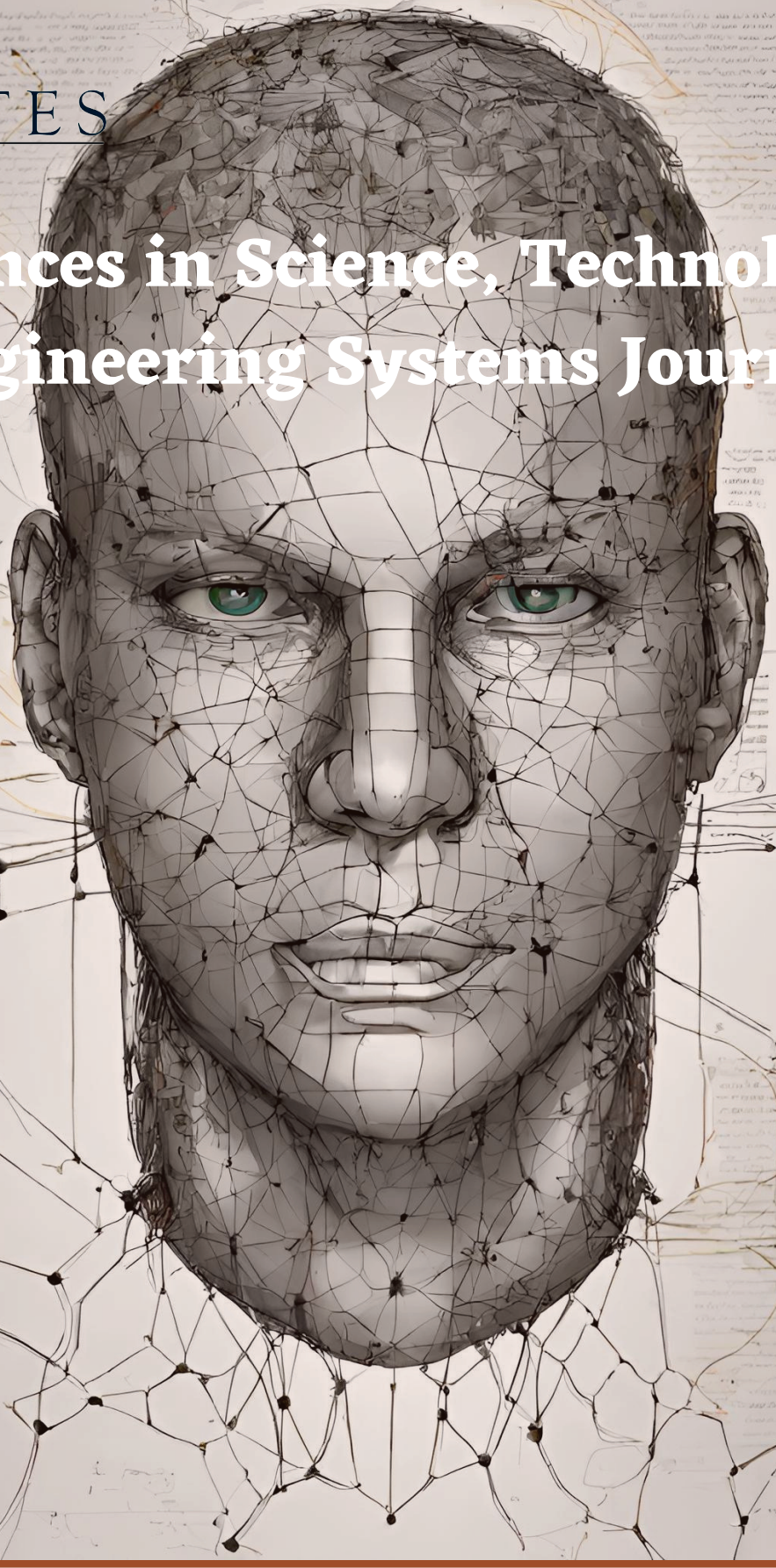




ASTES

# Advances in Science, Technology & Engineering Systems Journal



VOLUME 9-ISSUE 1|JAN-FEB 2024

[www.astesj.com](http://www.astesj.com)  
ISSN: 2415-6698

## EDITORIAL BOARD

### Editor-in-Chief

**Prof. Passerini Kazmerski**  
University of Chicago, USA

### Editorial Board Members

**Dr. Jiantao Shi**  
Nanjing Research Institute  
of Electronic Technology,  
China

**Dr. Tariq Kamal**  
University of Nottingham, UK  
Sakarya University, Turkey

**Dr. Hongbo Du**  
Prairie View A&M University, USA

**Dr. Nguyen Tung Linh**  
Electric Power University,  
Vietnam

**Prof. Majida Ali Abed  
Meshari**  
Tikrit University Campus,  
Iraq

**Dr. Mohmaed Abdel Fattah Ashabrawy**  
Prince Sattam bin Abdulaziz University,  
Saudi Arabia

**Mohamed Mohamed  
Abdel-Daim**  
Suez Canal University,  
Egypt

**Dr. Omeje Maxwell**  
Covenant University, Nigeria

**Mr. Muhammad Tanveer Riaz**  
School of Electrical Engineering,  
Chongqing University, P.R. China

**Dr. Heba Afify**  
MTI university, Cairo, Egypt

**Mr. Randhir Kumar**  
National University of  
Technology Raipur, India

**Dr. Serdar Sean Kalaycioglu**  
Toronto Metropolitan University, Canada

**Dr. Daniele Mestriner**  
University of Genoa, Italy

**Ms. Nasmin Jiwani**  
University of The Cumberland, USA

### Regional Editors

**Dr. Hung-Wei Wu**  
Kun Shan University,  
Taiwan

**Dr. Maryam Asghari**  
Shahid Ashrafi Esfahani,  
Iran

**Dr. Shakir Ali**  
Aligarh Muslim University, India

**Dr. Ahmet Kayabasi**  
Karamanoglu Mehmetbey  
University, Turkey

**Dr. Ebubekir Altuntas**  
Gaziosmanpasa University,  
Turkey

**Dr. Sabry Ali Abdallah El-Naggar**  
Tanta University, Egypt

**Mr. Aamir Nawaz**  
Gomal University, Pakistan

**Dr. Gomathi Periasamy**  
Mekelle University, Ethiopia

**Dr. Walid Wafik Mohamed Badawy**  
National Organization for Drug Control  
and Research, Egypt

**Dr. Abhishek Shukla**  
R.D. Engineering College,  
India

**Mr. Abdullah El-Bayoumi**  
Cairo University, Egypt

**Dr. Ayham Hassan Abazid** Jordan  
University of Science and Technology,  
Jordan

**Mr. Manu Mitra**  
University of Bridgeport, USA



## Editorial

In the ever-evolving landscape of science and technology, this issue brings together a collection of 16 accepted research papers that delve into diverse domains, ranging from robotics and healthcare to machine learning, cyber-physical systems, and environmental sustainability. Each paper represents a significant contribution to its respective field, offering novel methodologies, insights, and solutions to address contemporary challenges. The papers span a spectrum of cutting-edge topics, showcasing the breadth and depth of research endeavors undertaken by scholars and practitioners globally.

In the agricultural sector of Senegal, where climate change, environmental impacts, and traditional methods pose challenges, the paper proposes a smart farming management system named MbaïMi that utilizes IoT technology. MbaïMi aims to improve Senegalese agriculture by providing real-time decision support tools, optimizing seed selection, and helping manage water resources, ultimately contributing to sustainable agriculture, environmental benefits, and food self-sufficiency [1].

Using Oura Rings, paper investigates left and right differences in sleep patterns of two healthy women in their 40s. The study reveals small differences in bed time, sleep time, rem sleep, sleep score, and heart rate. Notably, significant differences in sleep latency, awakening time, and sleep efficiency are observed, providing insights into individual variations in sleep indices and the potential applications of wearable technology in sleep science [2].

Addressing the challenges of traditional resume screening, paper introduces an innovative approach using machine learning (ML) and natural language processing (NLP) techniques. The study implements the RankSVM algorithm to rank resumes based on semantic attributes, outperforming traditional scoring methods. The proposed comparative ranking, relying on semantic descriptions, enhances the accuracy of candidate selection for job requirements, demonstrating the potential for automation in human resource departments [3].

Focused on IoT security, paper enhances the LoRaWAN server framework, strengthening it against cyber threats by migrating from RVC4 to AES encryption. The integration with Grafana's mapping plugin leverages geolocation data, addressing vital aspects of IoT security and geolocation integration. The research contributes to the understanding of LoRaWAN's technological evolution, emphasizing the implications for academic and practical domains in reinforcing IoT and network security [4].

The study introduces an advanced emotion analysis framework that recognizes dominant facial emotions, integrates gesture recognition, and includes text-to-speech recognition. Leveraging pre-trained models like Residual Masking Network and CNN architectures, the multidisciplinary framework achieves promising accuracy rates for Asian and European facial expressions, highlighting its potential for understanding and interpreting different facial expressions in diverse cultural settings [5].

Focusing on urban drainage in South Africa, the paper explores the use of pervious concrete to mitigate challenges related to flash floods. By developing various structural pervious concrete mixes, the research identifies effective combinations that satisfy road and bridge works standards, demonstrating the potential of pervious concrete for urban drainage in South African cities [6].

Addressing network anomaly detection, paper employs a hybrid model combining convolutional neural network and Bidirectional Long-Short Term Memory. The research optimizes the model

through extensive parameter tuning, achieving high accuracy for binary and multiclass datasets. The proposed approach contributes valuable insights into selecting hyperparameters for deep learning techniques, enhancing anomaly detection in network traffic [7].

Investigating the progress of Machine Translation (MT) systems, paper provides a historical perspective on MT's evolution, focusing on its development against the complex background of low-resource languages. The interdisciplinary methodology integrates linguistics, technology, and culture, contributing not only to the understanding of MT's evolutionary history but also supporting the conservation of endangered languages, such as the Wa language in the Myanmar-Wa corpus [8].

Augmented reality is employed for maintenance support and training of a railroad vehicle's air compressor in paper. The study introduces modeling techniques for fast animation and visualization algorithms for fluid flow in air compressors. The developed augmented reality content is evaluated for usability, achieving good system usability and user experience scores, indicating effective application for maintenance support and training [9].

Recognizing the cultural nuances in healthcare, paper introduces a mobile app designed for Native Americans with diabetes. Developed through participatory design principles and direct input from stakeholders, the app features culturally tailored nutrition plans, community-based support systems, and engagement with tribal health resources. The innovative integration of technology and cultural heritage aims to improve engagement, self-efficacy, and health outcomes for Native Americans with diabetes [10].

The paper introduces a smart-agent architecture for effective direct load control, focusing on thermostatically controlled appliances. By employing a mathematical model to forecast aggregated power consumption behavior, the research demonstrates adaptability to changes, self-retraining capability, and continuous improvement in predicting aggregated power consumption, providing insights for demand-side management [11].

The integration of a Diffuse Kalman Filter (DKF) is vital for reliably determining the state of an autonomous vehicle. In three essential parts—Kalman filter, fuzzy control, and simulation of a GPS sensor signal—the research compares the efficiency of DKF against a conventional Kalman filter. The study highlights DKF's superior reliability in position estimation and its potential for optimal trajectory control in autonomous vehicles [12].

Investigating the stability of explainable Artificial Intelligence (XAI) algorithms, paper introduces two intuitive methods for assessing the stability of XAI algorithms. The study develops a taxonomy to categorize evaluation criteria and proposes an objective metric to classify XAI algorithms based on their explanation stability. The research contributes valuable insights into the determinism and stability of XAI methods, addressing challenges in generating consistent explanations [13].

The study explores the feasibility of differentiating between the heart rate patterns of *Macaca fascicularis* and human infants, evaluating the suitability of *Macaca fascicularis* heart rate data for privacy safeguarding methodologies. The research verifies the similarity of cynomolgus monkey heart rate data to human heart rate data, suggesting the plausible use of non-human subjects for privacy-related experiments and raising considerations for the ethical use of health-related data [14].

The paper introduces an automated Global System for Mobile Communication Signal Strength and Radio Climatological (GSM-RC) measuring device, integrating sensors for meteorological



parameters and GSM signal strength. The device offers synchronized measurements to reduce errors and demonstrates good correlation in signal strength compared to conventional devices. Field tests validate the device's reliability, making it a valuable tool for simultaneous measurements in diverse locations [15].

Focused on wind turbine simulation, paper presents a mathematical model of a five-phase permanent magnet generator supplying nonlinear load. The research explores the impact of nonlinear loads on generator performance, demonstrating advantages over conventional three-phase systems. The proposed system offers higher nonlinear load voltage and faster response, contributing valuable insights to the field of wind turbine simulation and nonlinear load analysis [16].

As we conclude this exploration into the diverse realms of research and innovation, it becomes evident that each paper contributes to the evolving narrative of technological progress. The findings and methodologies presented in these papers offer valuable insights that not only advance our understanding within specific domains but also collectively contribute to the broader tapestry of human knowledge. In this era of rapid technological evolution, the collaborative efforts of researchers showcased in this collection play a pivotal role in shaping the future landscape of science and technology. The journey continues, and with each new discovery, we move closer to unravelling the mysteries of the universe and enhancing the quality of human life.

## References:

- [1] A. Cisse, O. Diallo, E.H.M. Ndoye, M. Sy, O. Sene, J.J.P.C. Rodrigues, "A Smart Farming Management System based on IoT Technologies for Sustainable Agriculture," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 1–8, 2024, doi:10.25046/aj090101.
- [2] Y. Yoshida, E. Yuda, "Verify of Left and Right Differences in Sleep Index using the Ring-type Sensor," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 9–14, 2024, doi:10.25046/aj090102.
- [3] A.H. Alderham, E.S. Jaha, "Improved Candidate-Career Matching Using Comparative Semantic Resume Analysis," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 15–22, 2024, doi:10.25046/aj090103.
- [4] S.T. Ahmed, A. Annamalai, M. Chouikha, "Strengthening LoRaWAN Servers: A Comprehensive Update with AES Encryption and Grafana Mapping Solutions," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 33–41, 2024, doi:10.25046/aj090104.
- [5] A. Kulaglic, Z. Örpek, B. Kayı, S. Ozmen, "Analysis of Emotions and Movements of Asian and European Facial Expressions," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 42–48, 2024, doi:10.25046/aj090105.
- [6] P.P. Mokgatla, R.W. Salim, J. Ndambuki, "Enhancing Compressive Strength of Pervious Concrete for Use as Pavement Layer in Urban Roads Afer," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 49–66, 2024, doi:10.25046/aj090106.
- [7] T. Acharya, A. Annamalai, M.F. Chouikha, "Enhancing the Network Anomaly Detection using CNN-Bidirectional LSTM Hybrid Model and Sampling Strategies for Imbalanced Network Traffic Data," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 67–78, 2024, doi:10.25046/aj090107.
- [8] F. Yune, K.M. Soe, "Tracing the Evolution of Machine Translation: A Journey through the Myanmar (Burmese)-Wa (sub-group of the Austro-Asiatic language) Corpus," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 79–90, 2024, doi:10.25046/aj090108.
- [9] G.H. Kang, H.J. Kwon, I.S. Chung, C.S. Kim, "Development and Usability Evaluation of Mobile Augmented Reality Contents for Railway Vehicle Maintenance Training: Air Compressor Case," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 91–103, 2024, doi:10.25046/aj090109.

- [10] W.U. Hasan, K.T. Zaman, S. Alian, T. Liang, V. Pandey, J. Kong, C. Tao, J. Li, "Bridging Culture and Care: A Mobile App for Diabetes Self-Care Honoring Native American Cultural Practices," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 104–113, 2024, doi:10.25046/aj090110.
- [11] P. Yazdkhasti, J.L. Cárdenas–Barrera, C. Diduch, "Smart Agent-Based Direct Load Control of Air Conditioner Populations in Demand Side Management," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 114–123, 2024, doi:10.25046/aj090111.
- [12] M.G.T. Espinoza, "Comparing Kalman Filter and Diffuse Kalman Filter on a GPS Signal with Noise," *Advances in Science, Technology and Engineering Systems Journal*, **9**, 124–132, 2024, doi:10.25046/aj090112.
- [13] F. Gawantka, F. Just, M. Savelyeva, M. Wappler, J. Lässig, "A Novel Metric for Evaluating the Stability of XAI Explanations," *Advances in Science, Technology and Engineering Systems Journal*, **9**, 133–142, 2024, doi:10.25046/aj090113.
- [14] D. Hirahara, I. Kaneko, J. Nishino, J. Hayano, O.M. Mozos, E. Yuda, "Investigating Heart Rate Variability Index Classification in *Macaca fascicularis* and Humans: Exploring Applications for Personal Identification and Anonymization Studies," *Advances in Science, Technology and Engineering Systems Journal*, **9**, 143–148, 2024, doi:10.25046/aj090114.
- [15] G.A. Babatunde, E. Theophilus, O. Joseph. Sunday, A.K. David, O.G. Ayodele, "Development of a GSM-RC Automated Device for Measuring Mobile Communication Signal Strength and Meteorological Parameters," *Advances in Science, Technology and Engineering Systems Journal*, **9**(1), 149–164, 2024, doi:10.25046/aj090115.
- [16] P. Meesuk, V. Kinnares, "Mathematical Model of Wind Turbine Simulator Based Five-Phase Permanent Magnet Synchronous Generator with Nonlinear Loads and Harmonic Analysis," *Advances in Science, Technology and Engineering Systems Journal*, **9**, 165–174, 2024, doi:10.25046/aj090116.

**Editor-in-chief**

**Prof. Passerini Kazmersk**



# ADVANCES IN SCIENCE, TECHNOLOGY AND ENGINEERING SYSTEMS JOURNAL

Volume 9 Issue 1

January-February 2024

## CONTENTS

<i>A Smart Farming Management System based on IoT Technologies for Sustainable Agriculture</i>	01
Alioune Cisse, Ousmane Diallo, EL Hadji Malick Ndoeye, Mamadou Sy, Ousseynou Sene, Joel José Puga Coelho Rodrigues	
<i>Verify of Left and Right Differences in Sleep Index using the Ring-type Sensor</i>	09
Yutaka Yoshida, Emi Yuda	
<i>Improved Candidate-Career Matching Using Comparative Semantic Resume Analysis</i>	15
Asrar Hussain Alderham, Emad Sami Jaha	
<i>Strengthening LoRaWAN Servers: A Comprehensive Update with AES Encryption and Grafana Mapping Solutions</i>	33
Sheikh Tareq Ahmed, Annamalai Annamalai, Mohamed Chouikha	
<i>Analysis of Emotions and Movements of Asian and European Facial Expressions</i>	42
Ajla Kulaglic, Zeynep Örpek, Berk Kayı, Samet Ozmen	
<i>Enhancing Compressive Strength of Pervious Concrete for Use as Pavement Layer in Urban Roads</i>	49
Pontsho Penelope Mokgatla, Ramadhan Wanjala Salim, Julius Ndambuki	
<i>Enhancing the Network Anomaly Detection using CNN-Bidirectional LSTM Hybrid Model and Sampling Strategies for Imbalanced Network Traffic Data</i>	67
Toya Acharya, Annamalai Annamalai, Mohamed F Chouikha	
<i>Tracing the Evolution of Machine Translation: A Journey through the Myanmar (Burmese)-Wa (sub-group of the Austro-Asiatic language) Corpus</i>	79
Florance Yune, Khin Mar Soe	
<i>Development and Usability Evaluation of Mobile Augmented Reality Contents for Railway Vehicle Maintenance Training: Air Compressor Case</i>	91
Gil Hyun Kang, Hwi Jin Kwon, In Soo Chung, Chul Su Kim	
<i>Bridging Culture and Care: A Mobile App for Diabetes Self-Care Honoring Native American Cultural Practices</i>	104
Wordh UI Hasan, Kimia Tuz Zaman, Shadi Alian, Tianyi Liang, Vikram Pandey, Jun Kong, Cui Tao, Juan Li	
<i>Smart Agent-Based Direct Load Control of Air Conditioner Populations in Demand Side Management</i>	114
Pegah Yazdkhasti, Julian Luciano Cárdenas–Barrera, Chris Diduch	

<i>Comparing Kalman Filter and Diffuse Kalman Filter on a GPS Signal with Noise</i> Maximo Giovanni Tandazo Espinoza	124
<i>A Novel Metric for Evaluating the Stability of XAI Explanations</i> Falko Gawantka, Franz Just, Marina Savelyeva, Markus Wappler, Jörg Lässig	133
<i>Investigating Heart Rate Variability Index Classification in Macaca fascicularis and Humans: Exploring Applications for Personal Identification and Anonymization Studies</i> Daisuke Hirahara, Itaru Kaneko, Junji Nishino, Junichiro Hayano, Oscar Martinez Mozos, Emi Yuda	143
<i>Development of a GSM-RC Automated Device for Measuring Mobile Communication Signal Strength and Meteorological Parameters</i> Giwa Abdulgafar Babatunde, Ewetumo Theophilus, Ojo Joseph. Sunday, Adedayo Kayode David, Owolabi Gbenga Ayodele	149
<i>Mathematical Model of Wind Turbine Simulator Based Five-Phase Permanent Magnet Synchronous Generator with Nonlinear Loads and Harmonic Analysis</i> Peerawat Meesuk, Vijit Kinnares	165



## A Smart Farming Management System based on IoT Technologies for Sustainable Agriculture

Alioune Cisse<sup>1,\*</sup>, Ousmane Diallo<sup>1</sup>, EL Hadji Malick Ndoye<sup>1</sup>, Mamadou Sy<sup>1</sup>, Ousseynou Sene<sup>1</sup>, Joel José Puga Coelho Rodrigues<sup>2</sup>

<sup>1</sup>Laboratoire d'Informatique et d'Ingénierie pour l'Innovation (LI3), Department of Informatics, University of Assane Seck, B. P. 523 Ziguinchor, Senegal

<sup>2</sup>COPELABS, Lusófona University, Lisbon, Portugal

### ARTICLE INFO

Article history:

Received: 20 October, 2023

Revised: 19 December, 2023

Accepted: 19 December, 2023

Online: 20 January, 2024

Keywords:

Internet of Things

Smart farming

Smart agriculture

Prototype deployment

Sustainable agriculture

Food self-sufficiency

### ABSTRACT

Advances in Internet of Things (IoT) and wireless technologies are revolutionizing various sectors, including environment, education, healthcare, industry, etc. In the same dynamic, as the world population constantly evolves, solutions based on such technologies need to be proposed to improve the agricultural sector. Senegalese agriculture, primarily rain-fed and based on both cash crops and subsistence food crops, faces challenges from climate change, environmental impacts, and traditional methods. Lack of information on good practices, new efficient and low-cost farming methods, and access to seeds, along with the COVID-19 pandemic, further complicates the situation. This research work proposes a smart farming management system that uses IoT technology, named MbaiMi, whose main objective is to improve Senegalese agriculture through new digital technologies for sustainable agriculture, environmental benefits and food self-sufficiency. It provides a real-time decision support tool that offers farmers best seeds to choose at the beginning of work and helps optimize water resources. Additionally, the system includes features like sending SMS/advice, graphical visualizing of temperature and humidity changes for greenhouse cultivation control. The development of a web and mobile application, deployment of a prototype in southern Senegal, and validation case studies demonstrate the system's potential to improve farmers' yields and daily hard work to meet the challenges of sustainable agriculture, environmental benefits and food self-sufficiency.

### 1. Introduction

In general, IoT enables process automation by eliminating or reducing human interaction. In the automation process, IoT uses sensors to collect data, controllers to process the data, and actuators to complete the automation process [1]-[4].

To face the challenge of using IoT in various sectors of life in order to improve the living conditions of populations with sustainable development, researchers, practitioners and academics have proposed innovative solutions from the use of IoT technologies in several areas like healthcare, education, automation, industry and agriculture [5]-[7].

The term "smart agriculture" or "smart farming" has emerged as a result of the application of new technological innovations in the agricultural sector, including wireless sensors, Internet of Things (IoT), smart networks, location systems, data science, artificial intelligence, automation, robotics technologies, etc. [8],

[9]. Therefore, the main goal of smart agriculture and smart farming is to automate all aspects of farming and agricultural methods to make the process more efficient and effective. This is a global challenge with applications as solutions to solve various problems, including the increased need for improving the quality and quantity of crops, the increased demand for food self-sufficiency, labor force shortages, and the aging of farmers, as well as the expansion of cutting-edge agricultural technology and the optimization of human labor [7]-[10].

Smart farming may be used for remote and automated control of soil quality, precision irrigation, precision plant nutrition, climate change as well as monitoring and proper growth of crops and livestock by taking advantage of efficient management of sensor-generated data. [11]-[13]. These sensor's data are transmitted via wired or wireless communication technology in database servers or cloud platforms to be processed and analyzed by the algorithms in place.

\*Corresponding Author: Alioune Cisse, aliouneCisse1995@gmail.com

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj090101>

Senegalese agriculture is mainly rain-fed and seasonal, based on both cash crops and subsistence food crops. The vast majority of agricultural producers and livestock managers are smallholders that use traditional approaches with many inefficiencies such as higher human interaction and labor cost and face the challenges of climate change, environmental impact and rudimentary means. Added to this the lack of information on good practices, new efficient and low-cost farming methods and technologies, accessing good seeds, the recently economic effects of the COVID-19 pandemic and monetary changes in view (ECO in the UEMOA zone). Efforts must be devoted to the search of scale savings in the name of competitiveness, in particular by increasing agricultural production volumes [14].

This research work proposes the deployment of a smart farming management system that uses IoT Technology for sustainable agriculture, called MbaïMi. The proposed idea employs soil pH, humidity and ambient air temperature, and soil moisture level to provide a real-time categorization of the best crops for a specific soil. Additionally, the system uses this information to automatically manage irrigation as well as manage ambient temperature and humidity for greenhouse farming. The main contributions of this work are:

- The proposal of a smart farming management system architecture that uses IoT technology to remotely and automatically control and increase crop quality and quantity while optimizing human labor.
- The proposal of underlying algorithms of the system.
- The suggestion of a real-world, IoT-based smart farming management system prototype.
- The launch of the smart farming management web and mobile application called MbaïMi, which classifies the best crops for a specific soil in real-time based on factors including soil pH, humidity and ambient air temperature, and soil moisture level. Moreover, the proposal allows to automatically manage irrigation and ambient temperature and humidity for greenhouse cultivation.
- Prototype deployment in the southern region of Senegal and some case studies for evaluation and validation. This shows that the proposal can efficiently help farmers to improve their yields, facilitate their daily hard works to meet challenges of sustainable agriculture, environmental benefits and food self-sufficiency

The rest of this paper is structured as follows. Section 2 presents the related work addressing the most relevant solutions available in the literature. Section 3 describes the proposed system architecture with its underlying algorithms, while the technologies used for the system deployment are presented in Section 4. The

Validation of the proposal via a case study is presented in Section 5. Finally, Section 6 concludes the paper and presents future research works.

## **2. Related Work**

Nowadays, humanity is facing many phenomena such as climate change, floods, and threatening bush fires, which are exposing people, especially in parts of Africa, to famine. Hence, according to the arguments of the coordinator of the PIESAN (Project for Eco-Sustainable Intensification of Agriculture in the

Niayes (a locality in Senegal)), the natural degradation of the soil resulting from salinization and acidification, the lack of application of good sustainable agricultural practices (BPAD), unsuitable irrigation and water quality are among the constraints of the agricultural sector. In addition, it adds to these constraints, the lack of control of climatic hazards, the low adoption of proven technical and technological innovations [15].

To overcome this situation, researchers proposed solutions to address food self-sufficiency challenges, aiming to improve agricultural systems through new effective methods.

The work in this paper [10] proposed the design of a smart farm prototype in Thailand, featuring an automatic water control system using IoT devices and sensors. The system consists of a Raspberry Pi board installed in a control box to collect data from the field using a DHT22 sensor and a web-based application. Performance evaluation on a rice farm in Thailand demonstrated its usefulness for Agricultural 4.0 and increased productivity.

In Senegal, E-Tolbi has been introduced that is a mobile and web application aimed at improving agricultural output and field watering. It offers real-time decision-making, digital farmer profiles, irrigation water and fertilizer inputs management, and remote practices. Additionally, E-Tolbi offers farming organizations up-to-the-minute information on crop yields on fields and manufacturing facilities. However, the E-Tolbi's algorithm doesn't consider the knowledge of the best crops/seeds for a specific soil [16], [17].

A low-cost framework for food traceability is proposed in [18], using commercial and proprietary sensing devices to monitor air, water, soil parameters, and herbicide contamination during farming. The proposed framework includes an automated prototype for in situ glyphosate detection, monitoring indicators like temperature, humidity, and pH of the irrigation water, total Volatile Organic Compounds (VOCs) and equivalent CO<sub>2</sub>. The performance evaluation was conducted in an Italian olive farming company, proving its efficacy in food traceability.

The study in [19] proposed a low-cost intelligent system for smart irrigation using IoT technologies. It includes admin mode, onetime irrigation schedule estimation, neural-based decision making, and remote data monitoring. The system uses MQTT and HTTP to inform in real-time farmers about crop conditions and the performances evaluation is performed with a sample crop test-bed for presenting results that include irrigation schedule, neural net decision making and remote data viewing. The authors claim that their proposal proves beneficial with its intelligence, low cost and portability, making it suitable for farms and greenhouse.

AgroCares [20], a Dutch agricultural technology firm, has introduced a soil scanner in Kenya using NIR (near infrared technology) technology. The system provides farmers with immediate information on soil nutritional quality, enabling them to apply appropriate fertilizer. Agrocares offers various application services, including SoilCares, LeafCares, FeedCares, and InsectCares. The little Lab-in-a-Box is a portable detection laboratory for doing on-site analyses of micro and macronutrients. However, the solution requires daily presence on the property to update the samples and does not consider the knowledge of the best crops/seeds according to soil's condition.

FertiTogo, an interactive digital tool introduced during the 12th national Togolese peasant forum, helps farmers determine their agricultural soil fertility and provide precise fertilization recommendations. Using soil spectroscopy, it is a Togo-specific solution, lacking individual sensors and irrigation water control. However, updating the information requires constant implanting and pressing a button, which is not very practical [21].

The proposed smart agriculture model in [22] aims to develop a real-time monitoring system for soil properties, consisting of Farm, Server, and Client levels. It implements decision support advisory models for pest and disease forewarning, crop disease identification, and SMS alerts. The proposed network model implements connected objects with 802.15.4/Zigbee radio links, which therefore limits the range of the solution to small areas.

Based on the literature, the proposals share some similarities with the MbaïMi paper. However, as far as we know, none of the current proposed smart farming systems utilize soil pH, humidity, ambient air temperature, and soil moisture level measurements to provide large-scale farm managers with an effective real-time decision support tool for their agricultural practices, offering them best seed selection from the outset of their field work. Additionally, the proposal includes a prototype deployment, along with a web and mobile application.

### 3. Architecture of the Proposed Smart Farm Management System

The proposed architecture of the smart farm management system is depicted in Figure. 1. The goal is to facilitate communication between the various components of the system. This communication can be identified in several steps.

The key components of the system consist of an Arduino MEGA 2560 board, a water pump, sensors like soil pH sensor, soil moisture sensor, and temperature and humidity sensors (DHT11), a cloud service, and a web and mobile application.

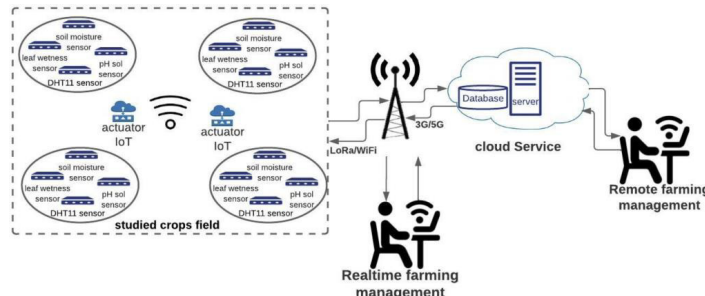


Figure 1: Illustration of the proposed architecture for a smart farm management system

#### 3.1. Description of the Architecture

**Studied crops field:** The crop farm being studied incorporates IoT devices, including an Arduino MEGA 2560 board (refer to Figure. 2) and a water pump. It also utilizes sensors such as a soil moisture sensor (refer to Figure. 3. A), a soil pH sensor (refer to Figure. 3. B), and DHT11 sensors for temperature and humidity (refer to Figure. 3. C). These sensors gather essential data on soil moisture, acidity, ambient temperature, and air humidity, which are then transmitted through the Gateway for real-time management or to cloud computing services for in-depth analysis.

The network technologies considered for transmitting this data include:

Wi-Fi : between farm equipment and the Gateway;

G/5G : between the Gateway and the Cloud service.



Figure 2: Arduino Mega 2560 board.



Figure 3: (A) The soil moisture sensor,



Figure 3: (B) The pH sensor,

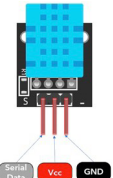


Figure 3: (C) The DHT11 sensor.

**Real-time farming management:** A web or mobile application allows farmers to access real-time information about the farm under study. The data is presented in different formats, including graphical representations of ambient humidity and temperature, with other data shown in percentages. Additionally, the algorithm proposed in this work provides a classification of the most suitable crops or seeds based on soil pH to increase yield. Furthermore, some tasks such as watering can be managed through the application.

**Cloud computing services:** It provides administration services for the proposed system's applications, database, and operating algorithm. Then, users may remotely access these data using a computer or a smartphone. The cloud service will enable additional analysis of the data gathered from the farm being studied.

**Remote farming management:** the farm management services of the proposed system can also be accessed remotely via a web or mobile application housed in the cloud and especially for further analysis of the data captured.

#### 3.2. Proposed Soil pH Check Algorithm

The proposed algorithm is based on the soil's pH level. The pH level captured is recorded in the variable denoted  $v$ . Therefore, the value in  $v$  undergoes verification in intervals of values corresponding to the acidity or alkalinity (basic) level of the soil (see Figure. 4).

After determining the level of alkalinity or acidity of the soil, the algorithm will check the adaptability of the crops or seeds selected in percentage based on the predefined ranges for each crop or seed.



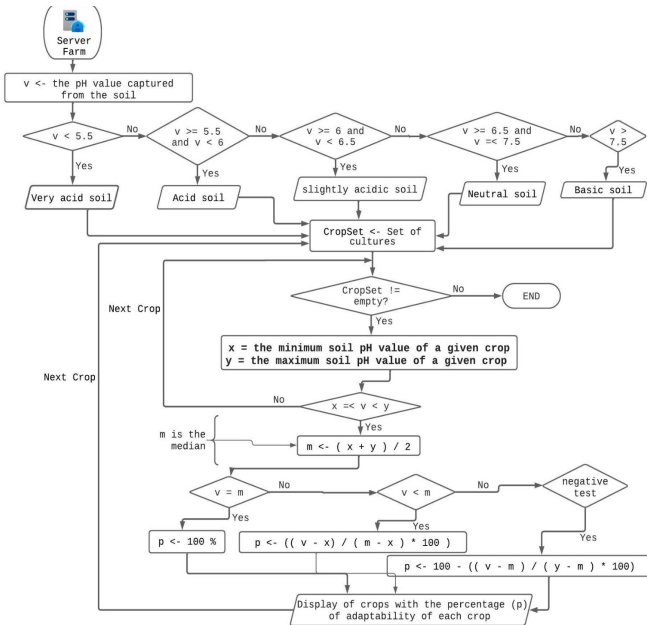


Figure 4: Algorithm for checking soil's pH.

### 3.3. Proposed Automatic Plant Watering Algorithm

Automatic plant watering is achieved by monitoring soil moisture levels and setting threshold values for each type of crop. The proposed algorithm, shown in Figure. 5), automates the watering process by comparing the soil humidity named “*hSol*” to the defined threshold values for a specific crop [*hSolMinC*; *hSolMaxC*]. “*hSolMinC*” represents the minimum soil moisture value, while “*hSolMaxC*” represents the favorable maximum soil moisture value for a particular crop.

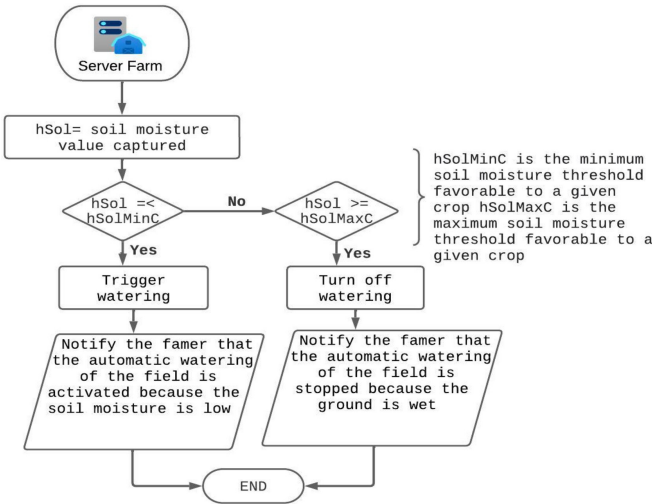


Figure 5: Automatic Plant Watering Algorithm.

### 3.4. Proposed Ambient Humidity and Temperature Management Algorithm

Humidity and temperature parameters affect the proper growth of plants in a farm. Each crop has its preferred climate range for optimum quality assurance in production. This is why threshold values are provided to study the frequency of optimum conditions and then alarm the farmer. The study of these parameters is especially useful for greenhouse agriculture. The proposed algorithm is shown in Figure. 6.

- Let *T* be the collected ambient temperature and [*Tx*; *Ty*] represents the defined range conducive to optimal crop growth;
- Let *H* be the collected ambient humidity and [*Hx*; *Hy*] represents the defined range conducive to good crop growth.

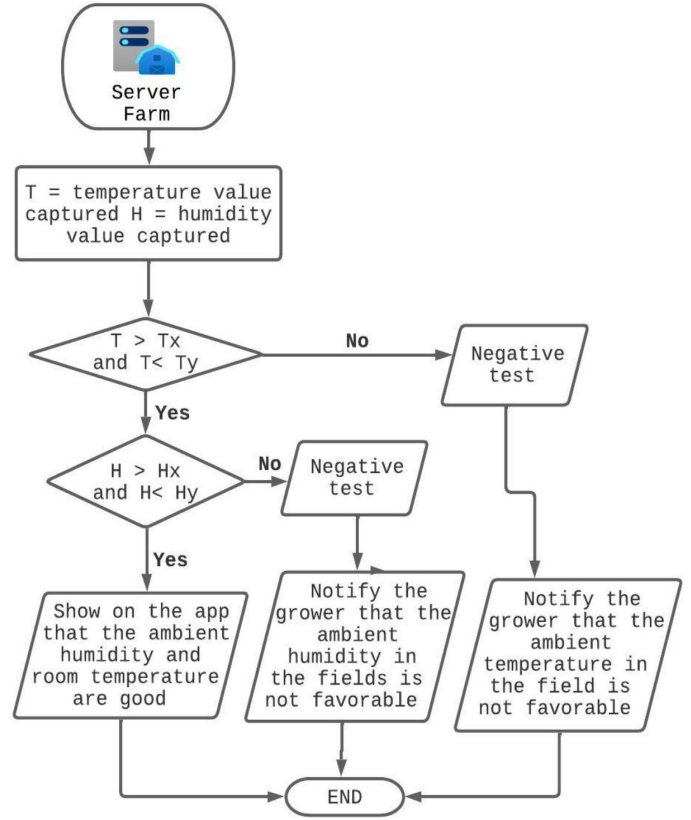


Figure 6. Ambient Humidity and Temperature Management Algorithm.

## 4. Used Technologies for the System Deployment

The paper’s work is based on a monolithic application that depicts a multi-level software application, where various components are integrated into a single program from a unified platform.

Choosing the right technology is a crucial step in developing an application, as its performance, durability, security, and time efficiency all depend on this choice. The Table 1 below provides a summary of the technologies used in this work.

Table 1: Summary of the used technologies.

Technologies	Target/Use	characteristics
Visual studio code frameworks [23]	implementation IDE of the web application	<ul style="list-style-type: none"> <li>- integrate a debugging console</li> <li>- integrate a terminal for launching server commands</li> <li>- features a robust code auto-completion engine (IntelliSense)</li> </ul>
ReactJS and AngularJS frameworks [24], [25]	front-end web application programming language	<ul style="list-style-type: none"> <li>- Free developed JavaScript library</li> <li>- develops native Android and iOS applications</li> </ul>

NestJS [25]	Language used for the back end.	<ul style="list-style-type: none"> <li>- Integrate TypeScript, object-oriented and reactive functional programming</li> <li>- server-side technology based on Node.js</li> <li>- API NEST access</li> </ul>
Arduino Software [26]	coding the microcontroller	<ul style="list-style-type: none"> <li>- open-source Arduino Software for writing and uploading code to Arduino boards</li> <li>- provides autocompletion, code navigation and live functionality</li> <li>- debugger</li> </ul>
MongoDB [27]	Manages smart farming data	<ul style="list-style-type: none"> <li>- version 6.0.1.</li> <li>- Manges NoSQL databases</li> </ul>

## 5. System Evaluation and Validation

### 5.1. Electronic Diagram of The Smart Farm System Prototype based on IoT

Figure. 7 displays the circuit diagram of the proposed prototype system, which was generated using Fritzing software [28]. In addition to the sensors mentioned in section 3.1, the system prototype includes:

#### MEGA 2560 microcontroller

The Arduino Mega 2560 board, powered by an ATmega2560 running at 16 MHz, offers 54 I/O, comprising 14 PWM, 16 analog, and 4 UARTs [29]. It includes ports for extra modules and a bootloader for programming alterations. While it is better suited for intricate projects than the Uno board [30], the Arduino MINI [31] is more suitable for farming use.

#### A Sprinkler Pump

In the prototyping process, a small water pump is utilized as a sprinkler pump. It is equipped with a small water output pipe and requires a 5V power supply.

#### A relay module

The Arduino board outputs 5V, necessitating a relay to alternate between the Arduino and a 12V or higher sprinkler pump, allowing for the pump to be powered on or off.

#### A SIM800L module

This mini-module supports 2G GSM and GPRS data, requires minimal power (3.4V to 4.4V), supports advanced AT instructions, and uses a serial communication interface, compatible with Arduino boards.

#### An ESP 32 module

The prototype employs a 38-pin ESP32 module, which is a development board featuring the SMD ESP32-WROOM-32 microcontroller, to test system functionalities [32]. This setup enables the control of sensors, modules, and actuators through Wi-Fi 802.11 b/g/n and dual mode Bluetooth version 4.2. The module is equipped with two cores, 38 GPIO pins, 16 PWM channels, and 4MB flash memory. It utilizes a micro USB Type B for power and ESP32 programming, integrating the USB controller to UART CP2102.

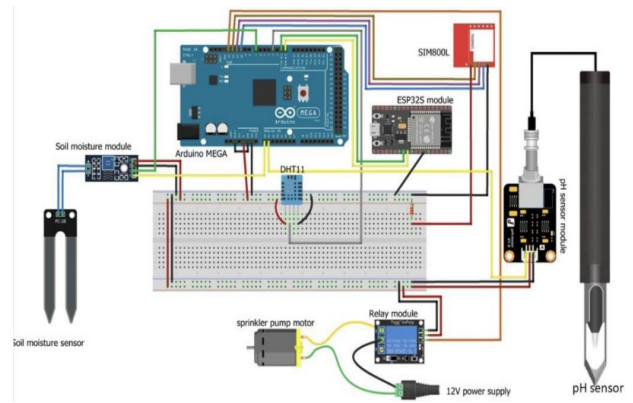


Figure 7: The electronic prototype diagram designed using Fritzing software [28].

### 5.2. Smart Farm Management System Prototype Deployment

The smart farm management system, utilizing IoT technologies, was created to offer real-time guidance to farmers, empowering them to make informed decisions in their agricultural practices. The system also assists in the management of water resources. Figure. 8 depicts the actual prototype's deployment at a farm in Senegal's southern area, serving as a case study for testing and validation.

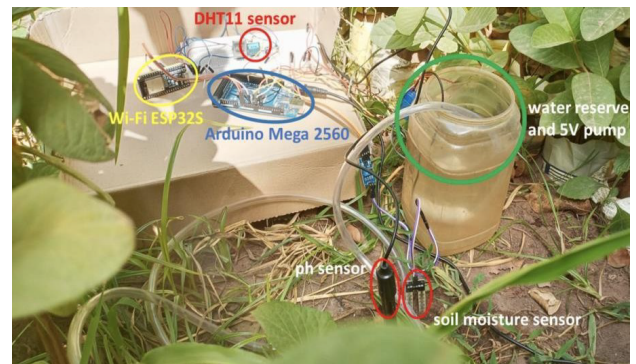


Figure 8. Deployment Capture of the actual prototype for testing and validation.

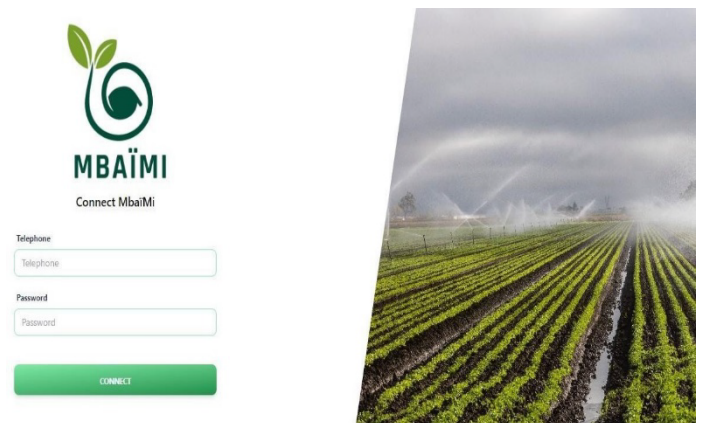


Figure 9: Web application access page.

### 5.3. The Smart Farm Web App

The MbaïMi smart agriculture Web application uses sensor data like readings of the pH of the soil, the humidity and temperature, and the soil moisture level to categorize in real-time crops for specific soils, assisting farmers in selecting seeds that are best suited to the soil, thereby increasing production.

#### 5.4. Web application access interface

The home page is user-friendly, but after opening the program, the administrator or farmer has to authenticate themselves to use its features (see Figure. 9).

#### 5.5. Interface for users setting

This *MbaiMi* system allows administrators and farmers to manage users by assigning roles, entering user information (first name, last name, phone number) and receiving an automatic login password via SMS or email (Figure. 10).

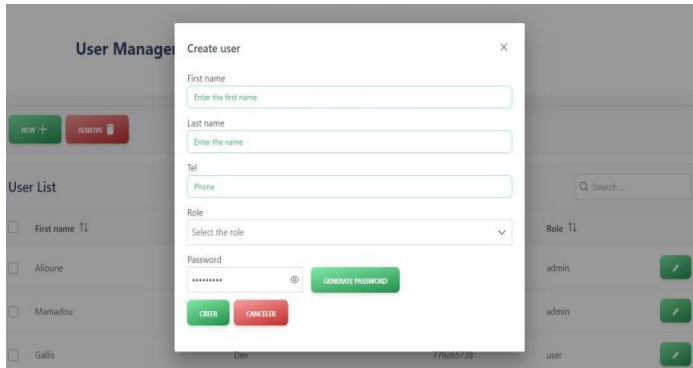


Figure 10: Interface for users setting.

#### 5.6. Interface for setting the studied farm

An administrator or farmer navigates to a menu-rich page to manage different farms (see Figure. 11), managing operations like creation, visualization, update, and deletion, and identifying farms by name, owner, and area.

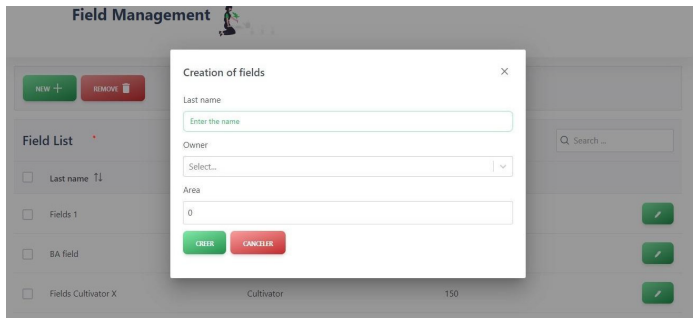


Figure 11: Interface for setting the studied farm.

#### 5.7. Interface for setting the used sensors

An administrator or farmer configures various sensors and actuators, providing information about their names, types, and connected nodes on a page like Figure. 12.

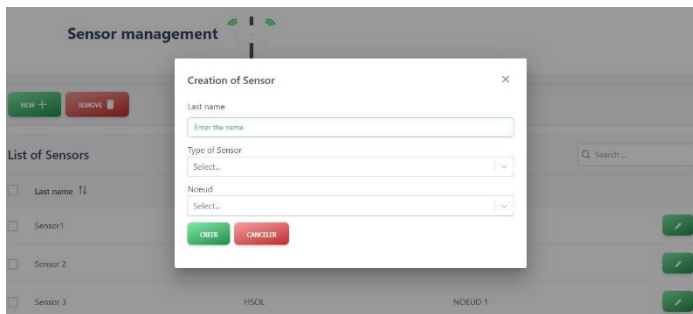


Figure 12: Interface for setting the sensors.

#### 5.8. Interface for crop settings

The interface below (see Figure. 13.) allows users to set thresholds for investigated factors like pH, soil moisture, and leaf humidity, and includes descriptions of each crop to improve seed information in the database.

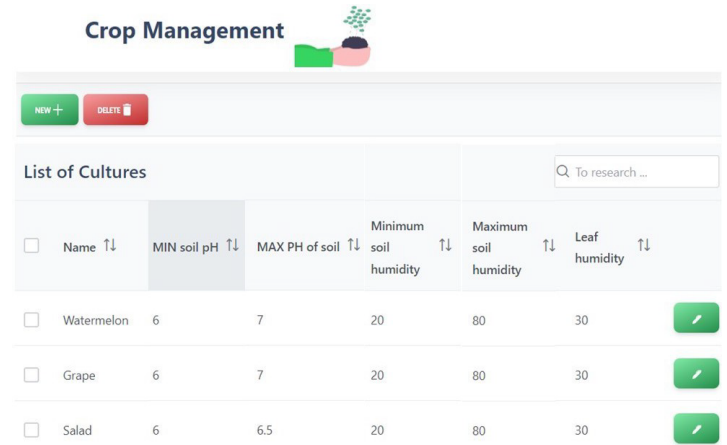


Figure 13: Interface for setting crops.

#### 5.9. Interface showing real-time categorization of best crops /seeds for a studied soil

In the web or mobile application, crops/seeds are displayed along with farm's adaptation percentages using ergonomic graduated color bands (see Figure. 14). For example, watermelon seeds are favorable for the studied farm with a percentage of 95.59%, so the green color band, whereas the passion fruit seeds are unfavorable with a percentage of 8.83%, hence the red band. By choosing the best crops /seeds for a specific soil one may reach challenge sustainable agriculture, reduces environmental degradation and waste seeds and products through more efficient use land, water, fertilizers and pesticides.

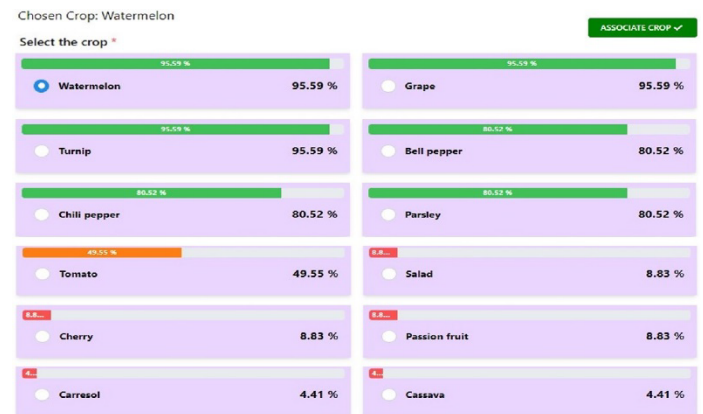


Figure 14: Real-time categorization of the best crops for a given soil.

#### 5.10. Interface for displaying the statistical data

Administrators can monitor sensor data collected on each farm node and visualize temperature and humidity evolution diagrams on the statistics page (see Figure. 15.).



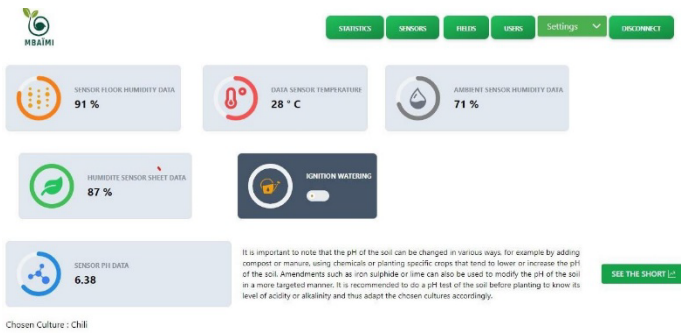


Figure 15: Display of the statistical data.

### 5.11. The Smart Farm Mobile Application

As shown in Figure. 16., the user initiates the mobile application and inputs their personal information to access their user space.

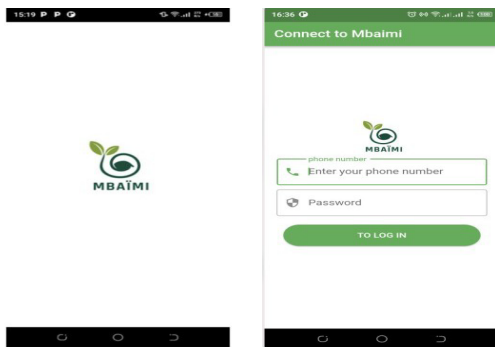


Figure 16: Mobile App Authentication.

The mobile application also assists farmers to consult adapted crops by monitoring sensor data collected every 10 minutes as well as the temperature and humidity diagrams of the farm (see Figure. 17). It also allows farmers to manually control watering of plants and send automatic SMS messages, since they may not have a constant connection to the internet.

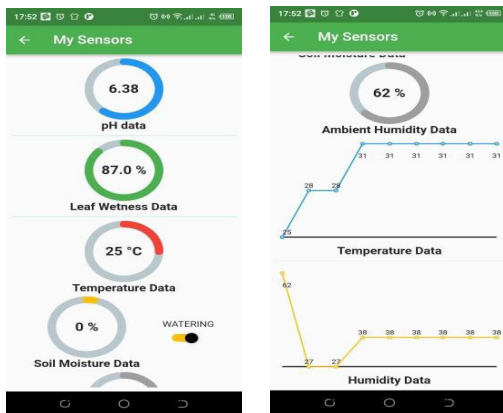


Figure 17: Display of the mobile application sensor data.

## 6. Conclusions and Future Works

This paper introduces *MbaiMi*, an implementation of a smart farm management web and mobile application with a real hardware prototype that uses IoT technology, which provides farmers an effective decision- support tool for their agricultural practices. The system uses soil pH, humidity, temperature, and moisture levels to determine the soil acidity used to provide a real-

time categorization of best crops for specific soils, aiding in seed selection and increasing production. The system also offers autonomous irrigation and temperature control, optimizing water resource utilization and encouraging greenhouse farming.

The *MbaiMi* system is a promising decision support tool for farmers in Senegal. The development of web and mobile

applications and the deployment of the prototype in southern

Senegal show that the system provides real-time categorization of the best crops /seeds, allows automatic irrigation and automatic management of ambient temperature and humidity of the studied soil. This demonstrates the system can improve farmer’s yields and daily hard work to meet the challenges of sustainable agriculture, environmental benefits and food self-sufficiency

LoRa technology is ideal for wireless communication in southern Senegal due to its 20 km transmission range and private mode configuration. Therefore, for future works we plan to include the LoRa technology for communication between the IoT devices in farms and the gateways.

## 7. Acknowledgements

This work is partially funded by Brazilian National Council for Scientific and Technological Development - CNPq, via Grant No. 306607/2023-9, and the LI3 laboratory of the Assane Seck University of Ziguinchor.

## References

- [1] S.R.J. Ramson, S. Vishnu, M. Shanmugam, “Applications of Internet of Things (IoT) – An Overview,” in IEEE 2020 5th International Conference on Devices, Circuits and Systems (ICDCS), 92–95, March 2020, doi:10.1109/ICDCS48716.2020.243556.
- [2] A.R. Madushanki, M.N. Halgamuge, W. A., A. Syed, “Adoption of the Internet of Things (IoT) in Agriculture and Smart Farming towards Urban Greening: A Review,” Int. Journal of Advanced Computer Science and Applications, **10**(4), 2019, doi:10.14569/IJACSA.2019.0100402.
- [3] Z. Eghbali, M.Z. Lighvan, “A hierarchical approach for accelerating IoT data management process based on SDN principles,” Journal of Network and Computer Applications, **181**, 103027, 2021, doi:10.1016/j.jnca.2021.103027.
- [4] A.A. Sadri, A.M. Rahmani, M. Saberikamarposhti, M. Hosseinzadeh, “Fog data management: A vision, challenges, and future directions,” Journal of Network and Computer Applications, **174**, 102882, 2021, doi:10.1016/j.jnca.2020.102882.
- [5] A. Khanna, S. Kaur, “Evolution of Internet of Things (IoT) and its significant impact in the field of Precision Agriculture,” Computers and Electronics in Agriculture, **157**(2), 218–231, 2019, doi:10.1016/j.compag.2018.12.039.
- [6] J. Muangprathub, N. Boonnam, S. Kajornkasirat, N. Lekbangpong, A. Wanichsombat, P. Nillaor, “IoT and agriculture data analysis for smart farm,” Computers and Electronics in Agriculture, **156**, 467–474, 2019, doi:10.1016/j.compag.2018.12.011.
- [7] M. Raj, S. Gupta, V. Chamola, A. Elhence, T. Garg, M. Atiqzaman, D. Niyato, “A survey on the role of Internet of Things for adopting and promoting Agriculture 4.0,” Journal of Network and Computer Applications, **187**, 103107, 2021, doi:10.1016/j.jnca.2021.103107.
- [8] T. Talaviya, D. Shah, N. Patel, H. Yagnik, M. Shah, “Implementation of artificial intelligence in agriculture for optimisation of irrigation and application of pesticides and herbicides,” Artificial Intelligence in Agriculture, **4**, 58–73, 2020, doi:10.1016/j.aiaa.2020.04.002.
- [9] E.H.M. Ndoye, O. Diallo, N. Hakem, E.N. Cabral, “Interference-Aware Nodes Deployment of a LoRa-Based Architecture for Smart Agriculture in the Southern Region of Senegal,” Advances in Science, Technology and Engineering Systems Journal, **7**(6), 248–255, 2022, doi:10.25046/aj070628.
- [10] P. Suampang, P. Jamjuntr, “A Smart Farm Prototype with an Internet of Things (IoT) Case Study: Thailand,” Journal of Advanced Agricultural Technologies, **6**(4), 241–245, 2019, doi:10.18178/joaat.6.4.241-245.
- [11] S. Kim, M. Lee, C. Shin, “IoT-Based Strawberry Disease Prediction System for Smart Farming,” Sensors, **18**(11), 4051, 2018, doi:10.3390/s18114051.

- [12] V.R. Pathmudi, N. Khatri, S. Kumar, A.S.H. Abdul-Qawy, A.K. Vyas, "A systematic review of IoT technologies and their constituents for smart and sustainable agriculture applications," *Scientific African*, **19**(11), e01577, 2023, doi:10.1016/j.sciaf.2023.e01577.
- [13] A. Mukherjee, S. Misra, N.S. Raghuvanshi, "A survey of unmanned aerial sensing solutions in precision agriculture," *Journal of Network and Computer Applications*, **148**, 102461, 2019, doi:10.1016/j.jnca.2019.102461.
- [14] S.S.A. Emira, K.Y. Youssef, M. Abouelatta, "Simulated IoT Based Sustainable Power System for Smart Agriculture Environments," *Advances in Science, Technology and Engineering Systems Journal*, **6**(1), 1030–1039, 2021, doi:10.25046/aj0601114.
- [15] PIESAN, Newsletter Project information of Eco-Sustainable Intensification of Agriculture in the Niayes, N°01, 2021.
- [16] S. Abdelkrim, L. Daniel, "Agritech : A la découverte de « Tolbi », la startup qui modernise l'agriculture sénégalaise," *StartupBRICS. StartupBRICS L'actu Tech & Startup Des Emergents*, 2022, <https://startupbrics.com/>.
- [17] TOLBI, <https://tolbico.com>, 2022.
- [18] A.V. Radogna, M.E. Latino, M. Menegoli, C.T. Prontera, G. Morgante, D. Mongelli, L. Giampetruzzi, A. Corallo, A. Bondavalli, L. Francioso, "A Monitoring Framework with Integrated Sensing Technologies for Enhanced Food Safety and Traceability," *Sensors*, **22**(17), 6509, 2022, doi:10.3390/s22176509.
- [19] N.K. Nawandar, V.R. Satpute, "IoT based low cost and intelligent module for smart irrigation system," *Computers and Electronics in Agriculture*, **162**, 979–990, 2019, doi:10.1016/j.compag.2019.05.027.
- [20] AgroCares. "Smart Farming | Nutrient Testing – AgroCares. AgroCares nutrient intelligence," 2022, <https://www.agrocares.com/>.
- [21] FertiTogo, "Plate-forme FertiTogo, système d'aide à la décision et à l'information," FertiTogo, 2022.
- [22] K.A. Patil, N.R. Kale, "A model for smart agriculture using IoT," in 2016 International Conference on Global Trends in Signal Processing, Information Computing and Communication (ICGTSPICC), IEEE: 543–545, 2016, doi:10.1109/ICGTSPICC.2016.7955360.
- [23] StudioVisual, Documentation for Visual Studio Code, 2022.
- [24] Ionic, <https://ionicframework.com/docs>, Introduction to Ionic, 2022.
- [25] NestJS, <https://docs.nestjs.com>, Documentation, 2022.
- [26] Arduino, <https://docs.arduino.cc>, Documentation, 2022.
- [27] MongoDB, <https://www.mongodb.com>, 2023.
- [28] Fritzing Software, <https://fritzing.org>, 2023.
- [29] Arduino Mega2560, "https://store.arduino.cc/products/arduinomega-2560-rev3," 2023.
- [30] Arduino UNO, <https://docs.arduino.cc/hardware/uno-rev3>, 2023.
- [31] Arduino Pro, <https://docs.arduino.cc/retired/boards/arduino-pro-mini>, 2023.
- [32] ESP32, "http://esp32.net/," 2023.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Verify of Left and Right Differences in Sleep Index using Ring-type Sensor

Yutaka Yoshida, Emi Yuda\*

Tohoku University, Graduate School of Information Sciences, Sendai, 980-8579, Japan

### ARTICLE INFO

Article history:

Received: 20 October, 2023

Revised: 19 December, 2023

Accepted: 19 December, 2023

Online: 20 January, 2024

Keywords:

Ring type wearable sensor

Sleep index

Left and right difference

### ABSTRACT

In this study, two healthy women in their 40s wore Oura Rings on both fingers and verified left and right differences in sleep index. Bed time, sleep time, rem sleep, sleep score, and heart rate had small left and right differences in both subjects, and a high correlation of 0.8 or more was observed (subject A:  $P < 0.001$ , subject B:  $P < 0.01$ ). Subject B showed a significant difference in the mean values of sleep latency, awakening time and sleep efficiency, and one finger was overestimated ( $P < 0.05$ ). Although the detection accuracy of sleep stages differs depending on the stage, it was shown that the influence of left and right differences is particularly small for bed time, sleep time, rem sleep, sleep score and heart rate.

## 1. Introduction

This paper is an extension of the 2022 World Automation Congress presentation [1]. Recent ring-shaped sensors can non-invasively measure heart rate, body surface temperature, and body acceleration. Indexes related to sleep time and sleep quality can also be calculated and visualized using a smartphone application. Among ring-shaped sensors, the Oura Ring has received particular attention in recent years and has been verified the accuracy of detecting HRV (heart rate variability) index and sleep stages during sleep using a combination of Electrocardiogram (ECG) and polysomnography (PSG) [2-5]. In addition, studies have been reported on accuracy comparisons with other wearable devices [6-9] and long-term tracking of physiological characteristics related to menstrual cycles and sleep [10].

In a study on accuracy evaluation regarding the wearing position of wearable sensors, it has become clear that the index obtained from wristwatch-type sensors varies depending on the wearing position [11], making it somewhat difficult to evaluate a stable index. However, there are no study reports that examine the left and right differences in the index when the Oura Ring is worn on both fingers.

Therefore, in this study, we wore Oura Ring to both fingers to record biological signals during sleep and verified left and right differences in sleep index calculated by the application.

## 2. Method

### 2.1. Experimental protocol

Subjects were two healthy women (subject A: 42 y.o., subject B: 41 y.o.). The gist of the study was explained to the two subjects, and consent was obtained. The studies involving human

participants were reviewed and approved by Center for Data-driven Science and Artificial Intelligence Tohoku University Institutional Review Board (No.2022-6, Approved 25 August, 2022).

Before going to bed, the subjects wore two ring type wearable sensor on both index fingers to record their sleep index for the night. The measurement period was from January 25, 2022 to October 24, 2022, with 172 days of data for Subject A and 10 days for Subject B. The measurements were not performed consecutively every day, but the subjects were allowed to decide for themselves, and the experiment was performed without straining while taking into account the prevention of skin irritation.

### 2.2. Ring type wearable sensor

The ring-type wearable sensor used Oura Ring Generation 3 (Oura, Inc., USA). Oura Ring has 15 sensors inside a titanium ring and can measure biological information such as heart rate, skin temperature, physical activity, and sleep index.



Figure 1: Oura Ring Generation 3

\*Corresponding Author: Emi Yuda, [yuda@ieee.org](mailto:yuda@ieee.org)

Pulse wave measurement uses an LED light sensor. Infrared and red LED run in dark places because they are more affected by ambient light, and green LED runs in bright places because they are less affected by ambient light. These sensors measure the amount of blood flow (change in blood vessel volume) that



Figure 2: State of charging

changes with the pulse of the heart. Relative temperature is measured with an NTC (negative temperature coefficient) thermistor. Physical activity and body motion are measured with the 3-axis accelerometer. The measured biological information is transferred to the Oura application via Bluetooth and output to a CSV file.

The size is 6 to 13 (ring width 7.8mm), and the weight is 4 to 6g. Battery life is 5-7 days, charging time is 20-80 minutes, and memory capacity is 16MB. The outside temperature range for use is -10°C to 52°C.

### 2.3. Sleep index and Statistical analysis

The sleep index is shown in Table 1. Heart rate is the average heart rate for 5 minutes, and respiratory rate is measured only once during one sleep. Mean values, standard deviations, and correlation coefficients of sleep indices obtained from both fingers were calculated. Mean values were compared by Paired t-test. The significance level was set at 5%, and significant at  $P < 0.05$ .

### 3. Results

Table 1 shows the statistical results of the sleep index of the two subjects. For subject A, there was no significant difference

between the left and right mean values, but significant correlations were observed in all indices ( $P < 0.001$ ). However, the strength of the correlation varies by index. In particular, bed time, sleep time, shallow non-rem, rem sleep, sleep score, heart rate and respiration rate showed a high correlation of 0.8 or more. On the other hand, subject B showed a significant difference in the mean values of sleep latency, awakening time and sleep efficiency, and one finger was overestimated ( $P < 0.05$ ). Bed time, sleep time, rem sleep, awakening time, sleep efficiency, sleep score and heart rate showed a significant correlation of 0.8 or more ( $P < 0.01$ ). Bed time, sleep time, rem sleep, sleep score, and heart rate were highly correlated with small left and right differences in both subjects.

Figure 3 and figure 4 are scatter plots of the left and right finger sleep indices for subject A and subject B, respectively. The vertical axis is the right finger and the horizontal axis is the left finger. In subject A, the distribution of sleep latency was such that the longer the right finger, the shorter the left finger, or the longer the left finger, the shorter the right finger. Awakening time and sleep efficiency had a square distribution. In subject B, the sleep latency was longer with the right finger than with the left. There is a high correlation of awakening time and sleep efficiency, but awakening time was longer with the right finger, and the sleep efficiency was greater with the left finger, resulting in a bias.

### 4. Discussion

There are many reports on the accuracy evaluation of HRV and sleep stages measured by Oura Ring during night sleep.

Reference [2] reported detection accuracy of 65%, 51%, and 61% for shallow sleep, deep sleep, and rem sleep, respectively. A study using machine learning combined accelerometers, body temperature, and HRV to detect four stages of light NREM sleep, deep NREM sleep, REM sleep, and wake with 79% accuracy[3]. Studies comparing HRV parameters during nocturnal sleep using ECG and Oura Ring have shown very high heart rate correlations during the night, indicating that HR can be measured with high accuracy [4,5]. A study comparing TST (total sleep time) between Samsung Gear Sport watches and Oura Ring showed a correlation of 0.86, and the average difference was within an acceptable range [7]. Based on the results of these previous studies, the cause of the difference between the left and right is thought to be the movement of the fingers and the detection accuracy of the device's algorithm. Since heart rate is detected as an average heart rate over a 5-minute period, there may be some fluctuations from beat to beat, but when looking at the baseline (average value) of heart rate, the difference between the left and right becomes smaller.

Table 1: Sleep index

Bedtime (min)	Awakening time (min)
Sleep time (min)	Sleep efficiency (%)
Sleep latency (min)	Sleep score
Shallow non-rem (min)	Heart rate (bpm)
Deep non-rem (min)	Respiration rate (frequency/min)
Rem sleep (min)	Relative body temperature (°C)



Table 2: Statistical results of the sleep index.

Sleep index	Subject A				Subject B			
	R-L	Mean±S.D.	Correlaion		R-L	Mean±S.D.	Correlaion	
			r	P-value			r	P-value
<b>Bed time (min)</b>	R L	403±89 404±82	0.843	<0.001	R L	461±107 437±97	0.832	0.003
<b>Sleep time (min)</b>	R L	339±68 339±67	0.917	<0.001	R L	375±68 373±68	0.815	0.004
<b>Sleep latency (min)</b>	R L	11±9 11±9	0.195	<0.001	R L	14±10* 7±4	0.529	<0.001
<b>Shallow non-rem (min)</b>	R L	194±48 197±45	0.812	<0.001	R L	218±63 211±47	0.577	0.081
<b>Deep non-rem (min)</b>	R L	40±21 38±20	0.590	<0.001	R L	33±17 42±21	0.324	0.362
<b>Rem sleep (min)</b>	R L	105±33 103±33	0.845	<0.001	R L	123±28 119±32	0.800	0.006
<b>Awakening time (min)</b>	R L	64±36 65±34	0.500	<0.001	R L	87±49* 64±50	0.837	0.003
<b>Sleep efficiency (%)</b>	R L	85±6 84±7	0.368	<0.001	R L	82±6* 87±8	0.824	0.003
<b>Sleep score</b>	R L	66±9 66±10	0.848	<0.001	R L	70±8 72±8	0.861	<0.001
<b>Heart rate (bpm)</b>	R L	63.38±3.63 63.39±3.67	0.994	<0.001	R L	64.97±4.50 63.63±4.05	0.870	0.001
<b>Respiration rate (frequency/min)</b>	R L	15.87±0.45 15.85±0.43	0.828	<0.001	R L	15.31±0.78 15.66±0.40	0.224	0.534
<b>Relative body temperature (°C)</b>	R L	-0.03±0.26 -0.03±0.27	0.735	<0.001	R L	0.05±0.25 -0.03±0.27	0.612	0.06

\*P<0.05(right finger vs. left finger)



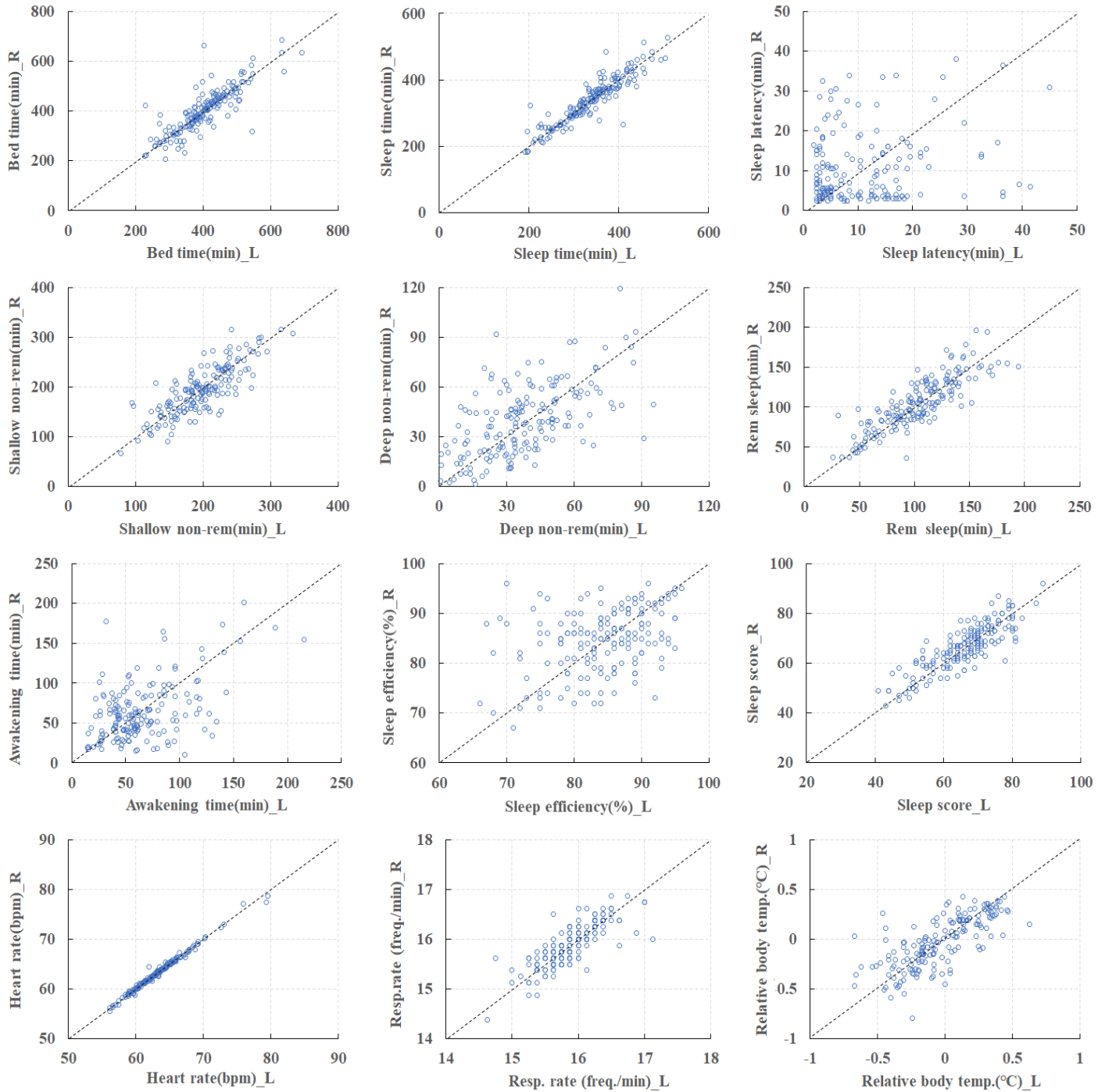


Figure 3: Scatter plots of the left and right finger sleep index for subject A.

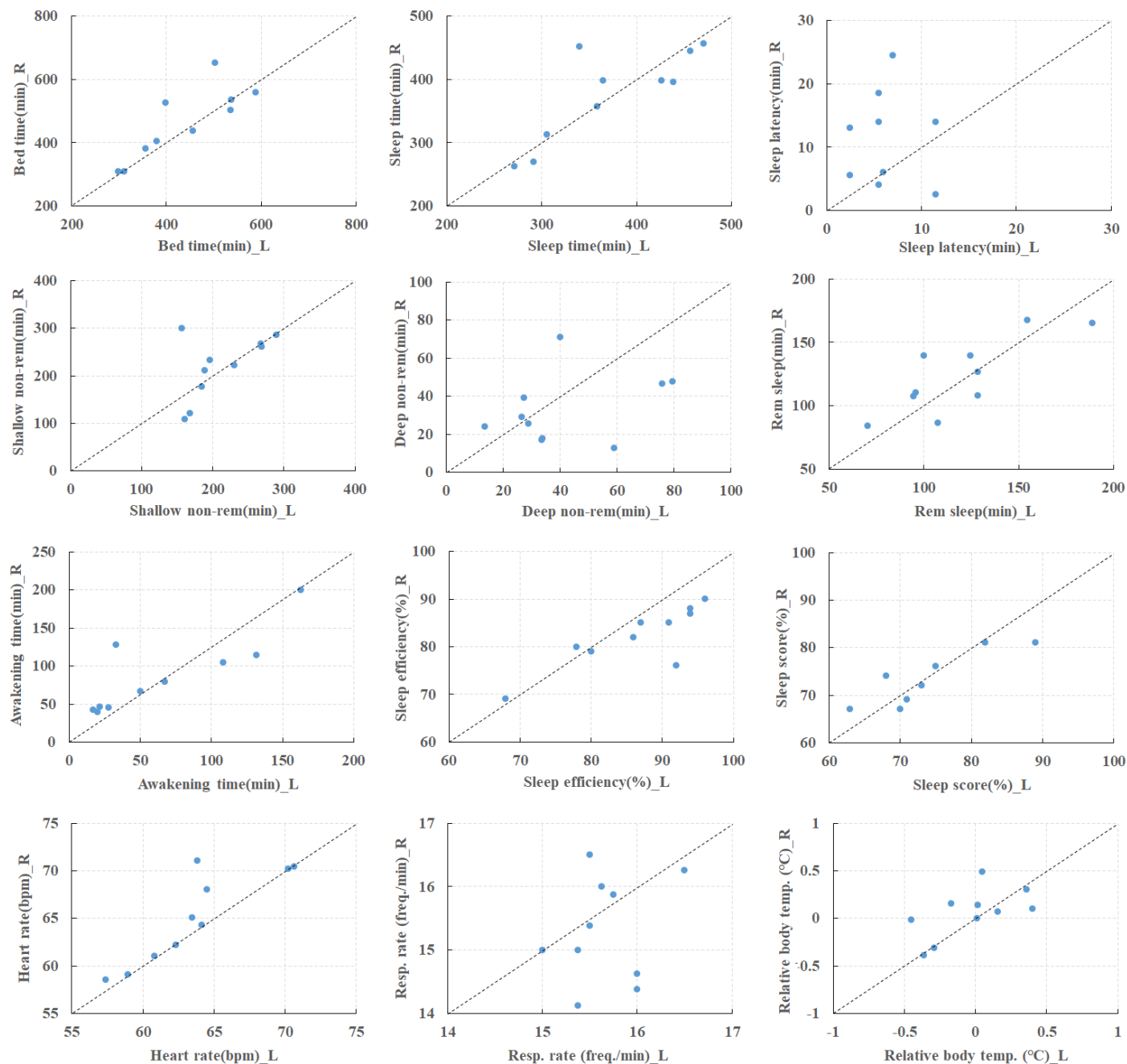


Figure 4: Scatter plots of the left and right finger sleep index for subject B.

In rem sleep, sympathetic nerve activity predominates, and the heart rate increases, whereas in non-rem sleep, parasympathetic nerve activity predominates and the heart rate decreases [12]. Therefore, the stages of rem sleep and non-rem sleep are easily detected, and the left and right differences in rem sleep are particularly small. The accuracy of detecting the shallow non-rem and deep non-rem stages is likely to decrease, and large differences between the left and right sides may occur. On the other hand, heart rate fluctuations are stable during non-rem sleep, but it is difficult to detect the difference between shallow non-rem

and deep non-rem. The left and right difference in non-rem sleep may also be due to the algorithm.

Awakening time is the time for turning over in bed, arousal during sleep, and although the Oura Ring's detection algorithm is unknown, it is thought that the acceleration sensor is involved in determining these times. Sleep efficiency is calculated as the ratio of total sleep time to bed time [13-15], and the longer the Awakening time, the more the sleep time is lost, and the lower the sleep efficiency. Therefore, bias occurs due to the difference in movement between the left and right fingers. In subject B, the

right finger had a longer awakening time, so it is thought that sleep efficiency was lower on the right and higher on the left.

Heart rate measured by the Oura Ring during night sleep is measured with high accuracy, and furthermore, it is suggested that wearing it on the left and right fingers does not affect the measurement accuracy. Although the detection accuracy of sleep stages differs depending on the stage, it was shown that the influence of left and right differences is particularly small for bed time, sleep time, rem sleep, sleep score and heart rate.

## 5. Conclusion

The left and right differences in sleep index obtained from the Oura Ring were verified. There were small left and right differences in bed time, sleep time, rem sleep, sleep score, and heart rate. It is suggested that indices can objectively and accurately evaluate sleep quality and fatigue recovery for health promotion. Other indicators are thought to differ between left and right depending on the individual, and it is necessary to take these differences into account when evaluating sleep. In addition, the subjects' health conditions, skin conditions such as wet skin or dry skin, and body positions during sleep were not recorded in this experiment, so indicators with large variations could not be categorized and analyzed. Considering these things, we will examine the effects of age and gender differences by increasing the number of subjects.

## Conflict of Interest

The authors declare no conflict of interest.

## Acknowledgment

This work was supported by a Grant-in-Aid for Strategic intellectual property/standardization, commercialization promotion support project for the Beyond 5G era from Ministry of Internal Affairs and Communications, Japan and the New Energy and Industrial Technology Development Organization (NEDO) Intensive Support for Young Promising Researchers Foundation.

## References

- [1] Y.Yoshida, H.Sakamoto, E.Yuda, " Precision of human workout-time detection using Random Forests and Wearable Sensor Data," 2022 World Automation Congress Proceedings, 429-432, 2022, doi:10.23919/WAC 55640.2022.9934354.
- [2] M.D. Zambotti, L.Rosas, I.M.Colrain, F.C. Baker, "The Sleep of the Ring: Comparison of the ÖURA Sleep Tracker Against Polysomnography," *Behavioral Sleep Medicine*, **17**(2), 124-136, 2019, doi:10.1080/15402002.2017.1300587.
- [3] M.Altini, H.Kinnunen, "The Promise of Sleep: A Multi-Sensor Approach for Accurate Sleep Stage Detection Using the Oura Ring," *Sensors (Basel)*, **21**(13), 4302, 2021, doi: 10.3390/s21134302.
- [4] R.Cao, I.Azimi, F.Sarhaddi, H.N.Vilen, A.Axelin, P.Liljeberg, A.M. Rahmani, "Accuracy Assessment of Oura Ring Nocturnal Heart Rate and Heart Rate Variability in Comparison With Electrocardiography in Time and Frequency Domains: Comprehensive Analysis," *Journal of Medical Internet Research*, **24**(1), e27487, 2022, doi:10.2196/27487.
- [5] H.Kinnunen, A.Rantanen, T.Kenttä, H.Koskimäki, "Feasible assessment of recovery and cardiovascular health: accuracy of nocturnal HR and HRV assessed via ring PPG in comparison to medical grade ECG,Physiological Measurement", **41**(4), 04NT01, doi: 10.1088/1361-6579/ab840a.
- [6] D.J.Miller, C.Sargent, G.D.Roach, "A Validation of Six Wearable Devices for Estimating Sleep, Heart Rate and Heart Rate Variability in Healthy Adults," *Sensors (Basel)*, **22**(16), 6317, 2022, doi:10.3390/s22166317.

- [7] M.A.Mehrabadi, I.Azimi, F.Sarhaddi, A.Axelin, H.N.Vilén, S.Myllyntausta, S.Stenholm , N.Dutt, P.Liljeberg , A.M.Rahmani, "Sleep Tracking of a Commercially Available Smart Ring and Smartwatch Against Medical-Grade Actigraphy in Everyday Settings: Instrument Validation Study," *JMIR Mhealth Uhealth*, **8**(10), e20465, 2020, doi:10.2196/20465.
- [8] E.D.Chinoy, J.A.Cuellar, J.T.Jameson, R.R.Markwald, "Performance of Four Commercial Wearable Sleep-Tracking Devices Tested Under Unrestricted Conditions at Home in Healthy Young Adults," *Nature and Science of Sleep*, **14**, 493-516, 2022, doi:10.2147/NSS.S348795.
- [9] L.Nauha, V.Farrahi, H.Jurvelin, T.Jämsä, M.Niemelä, M.Kangas, R.Korpelainen, "Comparison and agreement between device-estimated and self-reported sleep periods in adults," *Annals of Medicine*, **55**(1),2191001,2023,doi:10.1080/07853890.2023.2191001.
- [10] E.Alzuet, M.D.Zambotti, H.Javitz, T.Dulai, B.Albinni, K.C.Simon, N.Sattari, J.Zhang, A.Shuster, S.C.Mednick, F.C.Baker, "Tracking Sleep, Temperature, Heart Rate, and Daily Symptoms Across the Menstrual Cycle with the Oura Ring in Healthy Women," *International Journal of Women's Health*, **14**, 491-503,2022, doi:10.2147/IJWH.S341917.
- [11] E.Yuda, K.Yamamoto, Y.Yoshida, J.Hayano, "Differences in pulse rate variability with measurement site," *Journal of Physiological Anthropology*, **39**(1), 4, 2020, doi:10.1186/s40101-020-0214-1.
- [12] M.Yoshimoto, I.Yoshida, K.Miki, "Functional role of diverse changes in sympathetic nerve activity in regulating arterial pressure during REM sleep," *Sleep*, **34**(8),1093-1101,2011,doi:10.5665/SLEEP.1168.
- [13] B.V.Vaughn, S.R.Quint, J.A.Messenheimer, K.R.Robertson, "Heart period variability in sleep," *Electroencephalogr Clin Neurophysiol*, **94**(3), 155-162,1995, doi:10.1016/0013-4694(94)00270-u
- [14] P.Busek, J.Vanková, J.Opavský, J.Salinger, S.Nevsimalová, "Spectral analysis of the heart rate variability in sleep," *Physiological Research*, **54**(4), 369-376,2005.
- [15] D.L.Reed, W.P.Sacco, "Measuring Sleep Efficiency: What Should the Denominator Be? ," *Journal of Clinical Sleep Medicine*, **12**(2), 263-266, 2016, doi:10.5664/jcsm.5498.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

# Improved Candidate-Career Matching Using Comparative Semantic Resume Analysis

Asrar Hussain Alderham<sup>\*</sup>, Emad Sami Jaha

Faculty of Computer and Information Technology, Department of Computer Science, King Abdulaziz University, Jeddah, 22233, Saudi Arabia

## ARTICLE INFO

### Article history:

Received: 12 November, 2023

Revised: 27 December, 2023

Accepted: 27 December, 2023

Online: 20 January, 2024

### Keywords:

Semantic Resume Attributes

Comparative Description

Ranking

## ABSTRACT

A resume is a prevalent and generally employed method for individuals to showcase their proficiency and qualifications. It is typically composed using diverse customized, personalized methods in multiple inconsistent formats (such as pdf, txt, doc, etc.). Screening candidates based on the alignment of their resume with a set of job requirements is typically a laborious, challenging, time-intensive, and resource-intensive endeavor. This work is crucial for extracting pertinent information and valuable attributes that indicate acceptable applicants. This study aims to improve the candidate career-matching process for human resource (HR) departments by implementing automation and increasing efficiency. Machine learning (ML) and natural language processing (NLP) techniques are applied to infer and analyze comparative semantic resume attributes. Using semantic data comparisons, the ranking support vector machine (RankSVM) algorithm is subsequently applied to rank these resumes based on attributes. RankSVM detects tiny differences among candidates and assigns unique scores, resulting in an improved ranking of candidates based on their suitability for job requirements and from the best to worst match for the vacancy. The experiment and performance comparison results show that the proposed comparative ranking, which relies on semantic descriptions, outperforms the standard ranking based on regular scores in distinguishing candidates and distributing resumes across the ranks with an accuracy of up to 92%. Eventually, we obtained a list of the top ten candidates out of 228 technical specialists' resumes.

## 1 Introduction

This paper is an extension of work initially presented at the 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN) [1], which improves the candidate-career matching process. In the beginning, the Internet has played a leading role in several interactive communication fields. It has become crucial for job seekers and employers to find distinct candidates for their vacancies. Nowadays, job applications are predominantly conducted via online job portals or services. The job requirements are published by a human resources (HR) officer, while candidates upload their resumes via a specialized recruitment website.

Subsequently, the process of selecting an appropriate candidate for the job requirements requires time and human effort. Recruiters search available resume databases using a job-related keyword to select pertinent resumes for a specific job, but the traditional engine search is based on keyword filtering techniques without understanding the related semantic information of different resumes. This can

lead to similar resumes, which may not be closely related [2].

For example, in the technical domain (computer-related technologies), special information and common resume features such as skills, tools, and technologies may exist in the same single domain as computer science. Therefore, the capability of extracting all of these special skills and attributes can distinguish each resume from the others resumes, which is usually performed manually and is preferable to be partially/fully automated because it is an important, complex, and time-consuming task within any HR [3].

A resume is a document, either in printed or electronic form, that presents job-related knowledge and information to employers. The resume includes personal information, education information, work experience, technical skills under hard skills, leadership, creativity, time management, qualifications, and preference sections under soft skills. Each is written openly based on the author's personality and writing style. Furthermore, a resume typically comprises a document in which the information is presented in various file formats and structures. Therefore, it is probable that the format and content of numerous resumes submitted by applicants vary, necessitating a

<sup>\*</sup>Corresponding Author: Asrar Alderham, King Abdulaziz University, KSA, +966559813248 & [aalderham@stu.kau.edu.sa](mailto:aalderham@stu.kau.edu.sa)



review, evaluation, and filtration process to ensure improved candidate selection. HR job sections may, therefore, encounter difficulties and challenges when attempting to extract useful information and select potential candidates.

It is, therefore, imperative to keep pace with the astounding advances in machine learning (ML) technologies, as computing power becomes challenging when job requirements require further analysis and re-filtering of resumes by setting more precise criteria for extracting relevant resumes.

Much research has been done in resume analysis, including classification, summarization, and extracting information using various techniques such as ML, natural language processing (NLP), and Ontology. Such research efforts have faced many problems regarding the accuracy of selecting the most eligible candidates' resumes. Extensive resume analyses and further research studies are still needed to address these problems.

In this work, we aim to assist HR officers in improving the extraction, filtering, and selection of the most job-suited resumes in an automated and more accurate manner based on analyzing semantic resume attributes using NLP tools. Our contribution considers two groups of per-attribute semantic descriptions used for each resume as relative and comparative labels. This comparative description is more precise and informative than categorical descriptions. However, comparative labels must be derived from all possible comparisons between every two resumes per attribute. Moreover, using a ranking method like the ranking support vector machine (RankSVM) algorithm [4], results in discriminative ranking for all resumes per single attribute, per multiple attributes, or by all attributes.

The rest of this paper is organized as follows. The following sections 2, 3, 4, and 5 presents background in brief, the process of extracting information from resumes and candidate-career matching, relative and comparative descriptions. In addition, the motivations, the research paper question, and the work objectives. A reference to related works on resume analysis that employ various methodologies and strategies is provided in section 6. Section 7 subsequently provides a comprehensive outline of the methodology employed in this study and the proposed approach. After that, in section 8, the dataset, the experiments, and the results are described in detail. The paper is concluded in the 9 section, and in the 10 section, we suggest directions for future work.

## 2 Background

A brief background on candidate-career matching, resume parsing, and the information extraction process will be given in this section. Then, the relative and comparative descriptions will be highlighted.

### 2.1 Candidate-Career Matching

Job/career matching is an operational issue that is ever more important in every society. A candidate-career (job-candidate) matching task aims to assign the right job to the right candidate (applicant). Candidate-career matching is a tedious process that HR officers carry out. There are so many resumes to analyze in a limited time, with a lot of work to pass through each candidate's application and identify the best matches for the job. Hence, unfortunately, any

possible negligence in this process may lead to the wrong candidate being put in the wrong job by a human mistake [5], which in turn can cause a loss for both employers and candidates.

There has been a rapid increase in jobs and employment in recent years through numerous internet recruitment platforms. Moreover, nowadays, because of the COVID-19 pandemic, millions of employers and job seekers would prefer or be obliged to conduct their hiring or job-seeking attempts besides job interviews through an online recruitment platform [6].

In [7], Russell stated, "Core personality is made up of traits that have been conditioned over many years. Such traits are critical in assessing a candidate's ability to perform virtually any aspect of any job". He also stated in the candidate's selection process, "The selection process is clearly the most critical and controllable variable in the development of a productive and successful work team."

The proper match between the candidate and the job is crucial for HR in hiring the right candidates, and it is also beneficial for those candidates to avoid involvement in work that is not right for them. As a result, increasing the matching process of candidates with occupations through proper resume selection and analysis is crucial for improving accurate applicant employability and developing a successful employment process, which in turn helps enhance employee job performance.

Three factors can affect a person's job performance capabilities as mentioned in [8], which can be valuable and insightful for this research work:

#### 1- Organization Match

The degree to which a candidate's attitudes, values, ethics, and grooming match those required by the job is called an organizational match. Face-to-face interviews are commonly used to assess these factors. However, relying solely on face-to-face interviews is risky. The halo effect is an issue that occurs when interviewers see a part of themselves in a job candidate. On the other hand, the opposite of the halo effect is another potential pitfall of the interview process: unconscious bias. The more a candidate differs from the interviewer, the more conscious effort the interviewer must make to view the candidate positively or neutrally.

#### 2- Skills Match

The degree to which a candidate's educational background, technical skills, previous job experience, and specific expertise match those required for the job is referred to as a skills match. Many job positions necessitate specific sets of knowledge or technical skills. According to research, those in charge of selecting these positions are frequently biased toward believing that expertise and intelligence are significant. However, more than expertise and skill intelligence are required. According to job matching research, people perform better when fully engaged in the challenges of the job.

#### 3- Job Match

Job match is a major component of a candidate's success on the job. It refers to how much a person's cognitive skills, such as how fast and effectively they learn. Interests include whether a person desires to work with people, data, or things and personality, such as the ability to be part of a team, make decisions, manage clients, etc.

All three factors or cornerstones of job performance are essential in matching the candidate to the job. Therefore, in this research, inspired by these pivotal factors, we focused on the applicant selection

process based on the job requirements by analyzing semantic attributes provided in their resumes and finding the differential aspects between them.

## 2.2 Resume Parsing and Information Extraction

Resume parsing is the automated extraction of information from websites or inconsistently formatted documents, such as resumes, using complex pattern matching or language analysis tools. This procedure aims to create a possible recruiting database in various formats [9].

Resumes can contain semi/unstructured text, vary in information type, order, and writing style, and represent various file formats (.txt, .pdf, .doc, .docx, .odt, .rtf, etc.). A resume is typically divided into sections reflecting the candidate's competencies. Due to this diversity, resumes are difficult to parse, and extracting useful information is challenging. The recruitment or HR staff spends significant time and effort parsing resumes and extracting relevant data [10]. So, it is crucial to have an accurate parsing system for the resume section. To effectively and efficiently analyze the data from various forms and structures of resumes, the automation model must not rely on orders or types of information [11].

Information extraction (IE) technologies, such as NLP, take natural language text as input and help analyze the text efficiently and effectively to discover valuable and relevant knowledge that can be used to produce usefully structured information.

Since resumes are written in human languages, the computer used to parse candidate resumes must be constantly trained and adapted to deal with human languages and different expressions in the writing of resumes. The capabilities of machine/deep learning techniques, considered among the most effective approaches under the artificial intelligence (AI) umbrella, allow a model to learn patterns in data without being programmed [12]. Furthermore, NLP can also be employed for such purposes and used to understand human language in resumes and extract useful information embedded in them.

NLP can be described as using computational methods to process free text in spoken or written form, which serves as a mode of communication commonly used by humans [13]. The main objective of NLP is to analyze unstructured text and represent its meaning using pipeline processing steps. These processes operate at two levels: the syntactic level, which deals with the structure and grammar of the text, and the semantic level, which focuses on the meaning and interpretation of the text. The syntactic level involves dividing the raw text of a document into sentences using a sentence segmenter. Each sentence or statement is then further divided into words and punctuation, known as tokens, using a tokenizer. Subsequently, every token is assigned part-of-speech tags (such as nouns, verbs, adjectives, adverbs, etc.), which will play a significant role in the named entity recognition (NER) process. This step identifies all occurrences of a specific entity type in the text. The final step involves using relation recognition to search for possible links between the different entities in the text at the semantic level, where each word is analyzed to determine the meaningful representation of the sentence [14]. Figure 1 shows the architecture for a simple information extraction system.

We have this background to use as a basis for our goal. To

improve the process of matching candidates to job requirements, we propose a new approach that differs from existing work on the topic of resume analysis, where many of these existing related works will be reviewed in detail in section 6. Unlike traditional approaches, our proposed approach aims to enhance resume analysis using comparative descriptions. This involves extracting relevant information from resumes using ML and NLP techniques. These comparative descriptions are expected to be more precise and informative in distinguishing between resumes based on their semantic attributes. Previous research in various domains [4, 15], and [16] has shown that such comparative descriptions are superior to other forms of description.

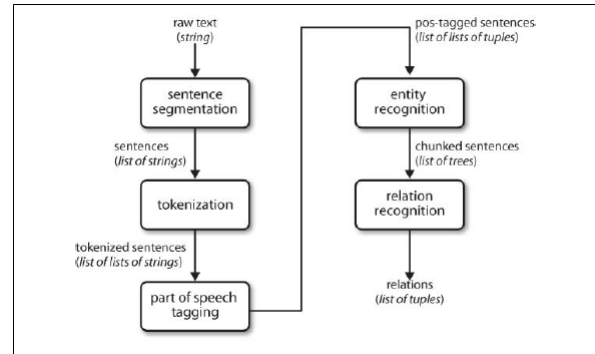


Figure 1: Simple pipeline architecture for an information extraction system [17]

## 2.3 Relative and Comparative Description

The parsing may fail and not extract high-quality data. If the search is more specific in describing job requirements, the search in most approaches is typically achieved based on an exact match of the search keyword or an explicit, predefined list of related words. However, other valuable implicit and semantic information is not analyzed or considered in the resulting retrieved data, which HR officers can usually infer and consider while manually searching and selecting candidates. Therefore, to fill this semantic gap between how humans and machines analyze resumes, we propose a new method to define several semantic attributes that describe different aspects of the process, like personal information, education level, experiences, certifications, etc. These semantic attributes may be explicitly or implicitly included in a resume. We need to parse them by searching keywords and giving each attribute a set of descriptive comparative labels demonstrating its significance.

Such semantic attributes can either be binary attributes associated with categorical labels or relative attributes associated with comparative labels [4]. Categorical (absolute) labels can be defined as nameable descriptions used to describe semantic attributes, such as education level, which can be labeled with any of the categories (BSc, MSc, Ph.D., etc.). Relative attributes are used to describe the degree of presence using labels such as "Very high," "High," "Average," "Low," "Very Low," and "None." Whereas comparative labels are named descriptions used to describe only relative attributes; for instance, they can be used to describe the education level of one resume based on the comparison with the education level in another resume. In other words, these labels describe the degree of comparison of relative attributes, using suitable labels such as



“Much higher,” “Higher,” “Same,” “Lower,” and “Much lower.” The proposed relative attributes are expressed using a bipolar five-point scale from 1 to 5. Corresponding comparative labels are assigned on a scale from 2 to -2. such that 2 is associated with the “Much higher” label, while -2 is associated with the “Much Lower” label. Thus, the relative attributes can represent the strength of the measured attributes but can also be comparable, making it easier to more accurately differentiate the minor underlying differences in the descriptive attributes between one resume and another [18]. As indicated in [19], there may be several expected advantages to describing the resume features in a relative (comparative) format:

1. It makes resume descriptions clearer and more meaningful (for example, resume A is better than resume B).
2. It enables comparisons with a reference resume object (e.g., resume A is higher than Asrar’s resume).
3. It makes attribute-based searching more efficient (e.g., search for a much higher resume with respect to skills).

### 3 Motivations

In our lives, getting a job that fits our ambition, qualifications, and abilities is very important, as matching our experience and knowledge with the job opportunities increases the efficiency of our production and progression in the labor market.

Since the advent of the Internet and the accelerating increase in online employment opportunities, matching candidates with job requirements is a tedious process that takes much time, effort, and cost to achieve. With many candidates seeking jobs, the need for a workforce increases to sort and analyze their resumes and choose the most suitable with high accuracy. For instance, when only two vacancies are required, and the applicants are more than a thousand candidates, each has his resume written differently to suit his thinking and how he presents his qualifications. It is often an inconsistently structured document that is difficult to analyze with the required accuracy automatically, ensuring distinguishing and the candidate matches the job from hundreds of others.

Proposing a comparative description-based approach adds to the existing resume analysis works a new viability expected to increase the current accuracy and efficiency of the selection of the candidates. It differs from merely comparing one resume with the other using traditional comparison methods. It compares the attributes of all resumes with each other based on a range that contains five points, making the selection and differentiation process for resumes more precise and informative.

### 4 Paper Research Question

Does the resume analysis based on semantic attributes described using comparative descriptions along with NLP and ML techniques can improve the process of candidate-career matching and selection?

### 5 Objectives

The main objectives of this work can be summarized as follows:

- To assist HR officers in improving the extraction, filtering, and selection of the most job-suited resumes in an automated, more accurate manner.
- To enhance existing work on resume analysis and candidate-career matching.
- To extract and organize the resume information related to a group of semantic attributes using relative /comparative semantic descriptions.
- To score the resume using initial relative labels and then compare them using comparative labels based on evaluation metrics for each semantic attribute.
- To rank resumes per attribute and then per all/some attributes based on comparative descriptions by matching the extracted semantic attributes from the resume to the job requirements.
- To improve the accuracy of candidate selection based on more accurate comparative descriptions, enabling the detection of tiny differences between candidates.

### 6 Related Work

Much work has been published on automatically extracting and analyzing information from resumes to improve candidate-career matching and selection. While many techniques and approaches are used in unstructured resume analysis, most researchers focus on ML and NLP to extract entities from resumes to ensure the best match between applicants and job requirements.

#### 6.1 Resume Analysis and Candidate Selection using NLP

Regarding resume analysis and the selection of candidates, in addition to the job domain allocation process and the classification of applicants in the field of ML and NLP techniques, several research studies have been summarized from 2018 to 2022 since they were interested in the current research scope, as will be illustrated about the studies [10], [20]–[43]. Started with [20], targeted at the challenge of resume data extraction using supervised and rule-based methods, which rely on hierarchical knowledge and large amounts of labeled data that are overly difficult to collect. They proposed a two-step of resume information extraction method. First, lines of raw resume text are segmented using a Naive Bayes (NB) classifier, which predicts each phase label in the resume. Second, after reducing text noise, term frequency-inverse document frequency (TF-IDF) cosine similarity is computed, and K-means clustering shows the resume attribute cluster. Writing style is a new feature proposed, whereas each line’s writing style comprises a word index, punctuation index, lexical attribute, and classifier prediction results. So, in the second step, the writing style is utilized to determine each semi-structured text relevant block and module components to get the structured resume text. The technique is practical and efficient, as demonstrated by experiments on a real-world dataset on information technology (IT) professional resumes.

To create a model that could extract valuable information from resumes, the authors of [21] used both Rule-based and Deep-learning approaches. Their work focuses on Vietnamese resume files in the IT domain. The approach of this study contains 4 phases: First, text segmentation; second, named entity recognition (NER); third, a combination of convolutional neural networks (CNN), bidirectional long short-term memory (Bi-LSTM), and conditional random field (CRF) to find name entities, and finally, text normalization. Above 81% accuracy in NER is a significant achievement for their model. This work demands a lot of time and people if they want an extensive enough data set.

In [22], the authors explained the proposed system used, which improves the decision-making process regarding candidates, as it consists of multiple modules: the section-based segmentation module, the filtration module, and the category-based matching module. They deal with the problem of selecting candidates who match the job requirements, and the human resources staff also exerts manual effort. They also addressed the issue of extracting information from unstructured resumes. The system uses an ML algorithm to build the applicant learning model and NLP to extract information. Note that this work covers resume analysis with all its sections and information like personal, education, experience, technical skills, internship, hobbies, etc. However, a limitation of the system is the absence of a research sample to evaluate the accuracy and efficacy.

In [23], they addressed two issues; the most significant was obtaining employer-interesting information from resume content. The second issue is structuring resumes automatically for database entry. They constructed a computer software that answers any question about resume content. A question is posed as a string of keywords that may not need to be grammatically correct. The response is supplied as a text block or a single resume clause. Their process of resume-text processing has four steps. Convert a resume from Word or PDF to text by Tika system, identify text keywords, create two directories, each with a  $\langle \text{key}, \text{value} \rangle$  pairs, and finally, text diapason definition. This paper gives examples of resume text extraction to answer arbitrary queries. Their approach allows for a structured resume through clustering and keyword labeling. The shortcoming of these methods is that they apply to process only any short text.

The study in [24] developed a web application that predicts the best resumes for a particular job description. The web app's design has two types of users: candidate and recruiter users. On the candidate side, the SpaCy model and regular expression (RegEx) of NLP extract different sections of data from the uploaded resume. On the recruiter side, Job descriptions, like resumes, must be scanned and parsed to extract essential data. Then, the TF-IDF cosine similarity compares a candidate's resume to the job description. Finally, the output is a ranked list of applicants based on total resume scores visible only to the recruiter. There were problems with parsing unstructured resumes; the candidate's education years were added to their job experience years, which were calculated. So, the applicant had to provide manual input for their job experience to count toward their overall resume score, in addition to the web application problems in working on .doc files only.

To avoid the time-consuming and redundant process of manually allocating projects to new hires by opening and analyzing their resumes one by one. In [25], they discussed developing and implementing a resume classifying application that uses NLP to

obtain only the essential and relevant information. In addition, an ensemble learning-based voting classifier consisting of 6 individual classifiers to classify a candidate profile into a suitable domain based on his interest, work experience, and expertise mentioned in the profile. The resume classifier application produces a bar graph showing candidates' suitability for various domains. The model uses Stack-Overflow REST API topic modeling techniques to add a new domain to the list of domains if the candidate profile fails to meet the confidence threshold value. The model results are encouraging, and their studies are distinguished by their focus on IT.

In [26], the authors offered a two-step process to rank applicants according to how well they match the requirements of a job vacancy. The first step is to create a resume parser that extracts complete information. This parser is a web app. In the second step, they employ bidirectional encoder representations from transformers (BERT) as sentence pair classification to rank job candidates based on matching the job descriptions by using a candidate's past resume experience descriptions. Their experiments first extracted LinkedIn resume text into predefined sections and achieved 100%

In [27], they created a resume analysis and position recommendation system that included competitiveness, personality traits, and job recommendation analysis with resume diagnosis. The diagnosis function uses ML and text mining to analyze the user's uploaded resume. It helps job applicants understand their competitiveness in the market and generates a list of recommended job openings based on their resumes. Job-seekers and companies immediately summarize their analysis results in reports and charts for job fairs. The job vacancy recommendation results were evaluated using a questionnaire survey on the applicant side. The experimental results confirmed that the job vacancies recommended by the created system met the expectations of job candidates.

One research work using NLP techniques for extracting information from legal documents is [28]. They have suggested using open information extraction from legal documents. Open information extraction involves extracting data from large datasets without a predefined data set. A sample case from the indiankanoon.com website is put into the system as a text document. As a first step, they performed data cleaning, including removing ambiguities, tagging different parts of speech (POS), chunking names, recognizing them using NER, and extracting relationships between them. No numerical results demonstrate the efficiency and accuracy of their proposed approach.

As the traditional recruitment process lacks speed and accuracy in selecting the best match of candidates for the job, a new model was proposed by [29]. They used the NLP method to summarize resumes in different formats by extracting important information for the job requirements focused on the technical skills and listing each summarized resume in a .txt file. The summarization model specification consists of four integrated modules implemented in Python: conversion, extraction, general, and resume list. The research concluded that the auto summarization model of resumes is not an efficient solution because it might not work well in all industries regarding the different skill sets needed in the different areas of recruitment.

In [30], they suggested an automated method for 'resume categorization and matching' to speed up the candidate selection and

decision-making process and further overcome obstacles to the proper screening of candidates. Their model is built on ML, which makes resume recommendations to HR based on job descriptions. This model works in two phases: extracting the features using NER of NLP and then classifying the resumes into the correct categories using LinearSVM—moreover, finding similarities between the resume and the job description using TF-IDF cosine similarity. The authors demonstrated the result of their model in the resume classification phase, with an accuracy of 78.53%.

Traditional hiring requires manual resume screening, which takes time and energy. Advanced machine learning-based NLP automates the resume screening and hiring process in this paper [31] by extracting entities from resumes using SpaCy and NER models and generating a graph showing each resume's score. The model is evaluated on 20 resumes, and the predicted summary resumes are saved as .txt files for each. The experiment achieved an F1-score of 91.35, a precision of 90.14, and a recall of 92.60. Their entity-recognized model achieved a 91.35 F1 score, 90.14 precision, and 92.60 recall.

In [32], they proposed a resume parser system using NER of Stanford CoreNLP and pattern matching to convert the resume data into a structured format and extract information about the candidates for the recruiting process. Based on skills extracted using TF-IDF and logistic regression classification techniques, the system predicted the candidate's genre, such as computer science, accounting, finance, statistics, business development, etc. The experiment was conducted with 100 resumes from students and job seekers from various universities. The result of the overall resume prediction accuracy of their system was 91.47%.

In this proposed system of [33], the authors made the entire enlistment process more practical and economical. As a result, they have put in place a system that compares applicants' resumes with job descriptions and displays the similarity percentage. Cosine similarity is the basis of the system under consideration. The NLP technology extracts the knowledge and vital abilities required for the specific employment position. It shows the recruiter the candidates' similarity results, which aids in selecting the best candidates. Additionally, candidates are ranked based on their similarity score to obtain the most cost-effective outcome possible. HR departments would reduce their burden if the system were implemented across various industries with a high demand for qualified workers.

Because resumes are unstructured, human-written documents, it is urgent to understand the context of the words. So, in [10], they addressed the problem of the time and effort expended to analyze and extract information, especially the parser of special skills from resumes. They proposed semantical and contextual rich IE using the advanced NLP library SpaCy, which has a feature called "Phrase Matcher". Their approach identifies a table or dictionary with the various skill sets, parses the resume to search all skill sets, and counts the frequency of those words of different categories. To select the appropriate candidates, they used the Matplotlib tool to represent the information visually after parsing. From 250 resumes and job requirements specified in the dictionary used in the result, they found two candidates' resumes satisfying the job requirements. Note that their research was limited to extracting special skills from the resume, which is only one of several other important parts such as experience, education, etc. Therefore, selecting candidates is

done visually, not using ranking algorithms.

In [34], they developed a job portal where employees and applicants can post their resumes, making recruiting easy and efficient. Their portal aids in the organization of resumes according to the needs of specific employers. Also, they used data from social media like LinkedIn to help recruit high-quality candidates worldwide while avoiding unfair and discriminatory practices. The result involves converting resumes from /pdf/doc/RTF to plain documents and tokenizing data entities using optical character recognition (OCR), NLP, and ranking algorithm tools. When comparing extracted entities and needed keywords, resumes were ranked based on their technical skills, and the results were supplied in the form of a pie chart and bar graph.

The research work of [35] was to develop a system that automates the eligibility examination and evaluation of candidates in recruiting students for job vacancies or higher education programs. This system handles the manual analysis of resumes with all its tedious tasks and provides accurate and practical evaluation results. They used ML approaches, NLP, and three classification algorithms to implement the system. The evaluation of this work was divided into several steps conducted on personal information, professional experience, academic background, and soft and technical skills. The results obtained in the classification are cross-validated by the results of online video interviews. This work used acceptance criteria for candidates based on the scores given after extracting the attributes, which, as a result, is an indicator in which resumes are classified between acceptable and rejected. Hence, this is what mainly differs from our proposed approach, which systematically ranks the candidates starting with the most matching the job requirements based on relative and comparative scoring instead of merely judging them with binary labels like (acceptable or rejected).

In [36], the authors address the problem of resume analysis and the complex use of the parser when supporting languages other than English. They used ML techniques in the context of NLP that achieved high accuracy in extracting information from resumes in an arbitrary format and five different languages. They created a system with many interconnected models, using state-of-the-art NLP models as a basis. Their approach used a new deep model architecture for sequential input data, called a transformer, with parallel processing of the input sequence in the form of the BERT language model. BERT is designed to learn deep bidirectional representations on unnamed text. Their models extract and categorize relevant resume sections (personal information, skills, education, previous job) and corresponding specific information at the lower hierarchical level (names, dates, addresses, competencies, etc.). Models were evaluated on a data set of 1,686 resumes. For Norwegian, Swedish, Finnish, Polish, and English languages, the system achieves F1-scores of 0.86, 0.88, 0.86, 0.87, and 0.82 at the section level, respectively, and 0.75, 0.80, 0.80, 0.81, and 0.83 at the item level, respectively.

One of the essential parts of a resume is the educational qualifications section, which captures the knowledge and skills relevant to the job. In [11], they solved the problem of a large amount of annotated data required to determine educational institutions' names and grades from a resume's education section. They proposed a semi-supervised model for accurately identifying degrees and insti-



tute names on a resume based on NER. It consists of CNN, word embeddings, and BI-LSTM. This model is used to predict unclassified education section entities and is corrected using a correction unit. This model was evaluated on 500 training and 50 test resumes, achieving an overall F1-score of 73.28 and an accuracy of 92.06%

Higher-level skills can be deduced from lower-level ones, and vice versa, rather than simply pulling out terms associated with competence. Extracting skills is a vital step in developing job recommendation systems. In [37], they proposed using CNN to create an explainable model that can extract high-level skills from resumes in their raw text format. The resulting model can predict the high-level skills mentioned in a resume and highlight the underlying low-level skills that led to that prediction. Experiments were conducted on anonymous IT resumes collected from various websites. The overall model achieves 98.79% recall and 91.34%, and more than 99% accuracy for specific skills.

In [38], they have proposed a two-stage embedding-based recommender system for matching available jobs with suitable candidates. A component for candidate retrieval using fused embedding and a module for fine-tuning and reranking candidates. Deep learning using CNN, representation learning, job-skill information graph, and geolocation calculator are fused for the job and candidate embedding. They have also implemented the Faiss index for clustering and compressing embeddings, which allows conducting runtime nearest neighbor searches. The final ranking score is calculated by a weighted linear equation that aggregates the first-stage relevancy score and contextual job and candidate features. Their job-to-candidate matching system has a satisfying title, description, requirement, and location-matching quality results. Their system has vastly improved the matching quality between jobs and applicants.

Discovering top candidates with few resources in a short time is the most pressing problem facing businesses today. In [39], the best candidates could be ranked using content-based suggestion, which employs NLP techniques to parse resumes. With an average parsing accuracy of 85% and a scoring accuracy of 92%, the system performs effectively. In addition, cosine similarity is used to discover the resumes most relevant to the job description provided, and the k-nearest neighbor (KNN) algorithm is used to select and rank Job applications.

People may have a better chance of finding jobs that will allow them to have a good life if they have the education level appropriate for the professional environment. In support of the impact of education on job applicants, the authors of [40] addressed gaps between the job market, job seeker, and educational institution skills. Recruiters want to check resumes for required skills automatically. Job applicants want to know what skills must be added to their resumes. Educational institutions recommend study programs and assist students in ensuring their courses cover job posting skills. Since all three of these users have skills in common, they made an app called "Skill Scanner," which uses NLP techniques to analyze, vectorize, cluster, and compare skills. It then makes reports with statistics and suggestions for all three users of which skills are covered and missing. They analyzed the master's program in data science and data scientist positions from Indeed.com and Kaggle.com. As part of a questionnaire, 108 representatives from their three parties were provided reports generated by Skill Scanner. Most users report that

their solution improves the efficiency, speed, fairness, explainability, autonomy, and support of skill-related processes. They proved that 89% of those who tried out their recommendation system had no problem using it.

Categorizing job applications received as resumes against open vacancies requires a significant amount of time and effort from an employer. This study [41] aims to develop an automated resume classification system (RCS) that classifies resumes according to their job categories. This study's main contribution is in preprocessing the resumes to a corpus and extracting vectorized representation using NLP techniques appropriate for classification tasks performed by ML algorithms. This work was evaluated on three extraction and representation techniques and nine ML classification models. The TF-IDF vectorizer was the best at extracting features and representation, and the SVM classes performed exceptionally well on parsed resume datasets with more than 96% accuracy.

In [42], authors introduce I-Recruiter, an intelligent decision support system (IDSS) for screening resumes and identifying the most qualified applicants for available IT sector vacancies. The semantic similarity between a resume and the job description determines an applicant's ranking in I-Recruiter. It provides information about the best applicants so that the hiring process can move forward. ML and NLP constitute the core of the system's functionalities. The essential components of this system are the training, matching, and extracting phases. The training block produces domain-specific word embeddings that have been trained. At the same time, the matching section identifies top candidates by comparing resumes and job postings for semantic similarity. In the last phase, It extracted some primary data on the top applicants. I-Recruiter showed excellent results, with an average accuracy of 96% and a short amount of working time.

Recently, In [43], they use NLP to extract useful information from resumes. The system will then go to the person's profiles on LinkedIn and GitHub, scrape information from those sites, and feed it into the ML Models to make a more accurate prediction. The first ML model, either KNN or SVM, predicts what kind of job role their resumes are best suited for it. The second model recommends improving their resumes by using cosine similarity, which compares the user's input to the model's prediction. They proved that the accuracy of the models ranges from 78% to 98% depending on the datasets utilized, the learning methods' complexity, and the dataset's size.

Several studies like [44, 45, 46, 47, 48, 49], and [50] suggest using ontology-based knowledge for resume parsing, an ontology is a knowledge base representation that can generate a semantic model of data connected with specific domain knowledge. Furthermore, ontology is utilized to define linkages between various types of semantic knowledge in a domain. Semantic web models use ontology to define the meaning of target data, information, or knowledge. The most fundamental comprehension of the semantic web is surpassed by ontology, which provides the ability for standard reasoning, typically based on the specification of inference rules. Ontologies let users organize information in taxonomies of concepts and their attributes and describe how these concepts relate to each other [51]. The authors of [52] say that there are two main reasons to use ontologies: to help people and software agents understand how information is structured and to make it possible to reuse existing



domain knowledge.

In [44], a resume ontology based on skills was created, which refers to the most necessary parts of a job to see how well the applicant matches the job description. Applicant and job profiles were annotated with a common vocabulary. The semantic concept similarity algorithm was modified to accurately compute and rank matching scores between profiles when a query was run. Based on the results of the system's experiments, the approach improved the search matching accuracy.

The author in [45] built an ontology that contains the most important factors to consider while hiring for IT positions. The primary purpose of the proposed approach is to reduce errors in the first stage of personnel selection by manually filtering hundreds of resumes/profiles to choose applicants for interviews. Their suggested system consists of a mobile application that automatically chooses online profiles from professional websites (like Indeed, LinkedIn, and Monster) and ranks them to eventually display the appropriate candidates for a given vacancy to the recruiter. They created an ontology focused on resume skills to match candidate resumes to job description requirements. GATE and Apache Tika tools were used to automatically extract skills from unstructured resumes in the implementation's first step. The specified information will be used to construct a structured RDF document, saved in a triple store, and queried for each job offer.

A supplement work that used the ontology for resume analysis is presented in the proposed research of [46]. The authors used text mining and ML tools to make an effective company recommender system (CRS) that could help recruiters find the best candidate for a given job title. Because the data to be extracted is unstructured, the authors used traditional and ontology-based information extraction strategies. Candidates for IT organizations were classified into three categories: low, average, and high, based on their rating score, which was interpreted as their competency level in the programming language mentioned in their project description, according to the model proposed. Finally, the organization can use the ranking to select the best possible candidates for the job openings.

The exploration in [47] also sought to address the issue of analyzing unstructured resumes. In addition to the limited availability of language dictionaries and the complexity of Polish linguistic dependencies, this drives their interest. The proposed prototype designed a hybrid resume parser service using text mining algorithms to extract information from Polish resume documents for IT recruitment. The resume parsing system combines three NER tools (Liner2, Nerf, Babelfy) with the anchor NER service and dictionary methods such as Fuzzy Dict and Competence Lexem. The proposed system also included a promising CRF method. Data from an IT recruiting firm was used to conduct the research, and the results showed that the hybrid approach proposed improves single results obtained for detecting educational institutions. The result of the combined tools (hybrid approach) was more than 60% better than the best single solution (30.92% vs. 19.23% for Babelfy).

The authors in [48] have considered the effects of ontology-based use in the resume search system for job applicants. They provided an ontology structure for representing resume and job description (JD) contents, extracting information from collected resumes and

JD using OCR, and classifying extracted information using an unsupervised approach called computer science ontology (CSO) classifier. This ontology searches resumes for job skills that match JD criteria. The challenge of candidate ranking in an automated recruitment system is solved using ML and NLP. This solution is based on an IT ontology called Job-Onto. Moreover, the planned system is implemented as a recruiting website that helps candidates and employers find opportunities and streamline their recruitment processes. The suggested recruitment system uses automated resume ranking based on criteria, with the recruiter controlling the weighting. Each time a job is posted, the system calculates the resume correlation score for each candidate. This score is the sum of their domain skill, general skill, and soft skill matching scores. Finally, a list of the best resumes for the front-end engineer job is shown. The proposed method proved effective in finding resumes that correspond with a particular JD in the IT industry, according to test results.

In [49], the authors proposed an approach to improve selecting the best candidates, including three primary processing steps: text extraction, text block classification, and resume facts identification via NER. A custom-built ontology is used to augment the extraction of technical skills. This research is unique because it focuses on many areas, including diversity and improving electronic recruitment re-usability. However, there are no numerical findings for the accuracy they intended to improve in extracting information from personal resumes.

Every job has distinct requirements: some require the most experience, education, or skills. Finding the proper applicants in a large resume pool is difficult, especially when we are looking for specific skills. So, in [50], their goal is to create a system to analyze resumes and job postings based on how well they match one another. Then, rank them in ascending order of score. They used ruled-based methods, developed their own ontology to facilitate matching and scoring, and employed NLP methods for information extraction. Their ontology contains 37493 distinct nodes, and 54632 relations were created. They did real case studies to evaluate the AI scoring and other features. For actual job vacancies, recruiters and the system simultaneously searched for candidates. The system's top candidates were ranked.

Finally, we highlighted a study that focused on sentiment analysis regarding personal interviews with candidates and used the SVM ranking algorithm that we will apply in our research. All users on social media can write their sentiments to express their emotions and opinions. Three processes comprise sentiment analysis: sentiment extraction, keyword preparation, review analysis, and classification. Moreover, Covid-19 has shifted all processes online, such as video conferencing. This study in [53] provides a job candidate rank (JCR) model for job candidates based on the interviewer's sentiment analysis. The candidate rank model has stages—first, sentiment text parsing, stopword removal, and stemming. Second, select features using document frequency (DF), expand using the WordNet database, and focus groups on ranking them. SVM and NB of ML classifications are included. Using the NB classifier, the model achieved 93% accuracy, while the SVM classifier achieved 89% accuracy.

## 6.2 Comparative Description Approach

This research was motivated by the successful use of comparative descriptions in other applications; therefore, we aim to apply and extend such effective capability as a novel contribution to resume ranking and selection processes for effective candidate-career matching. The ability of comparative descriptions regarding resume analysis has yet to be extensively investigated. Thus, we widen the exploration to embrace other prominent applications utilizing comparative descriptions, such as human identification by semantic/soft biometrics.

One of the challenges of human identification and semantic characterization is the quality of soft biometric description [19]. A search in soft biometrics showed better accuracy of comparative traits or attributes than the accuracy of categorical traits [4, 15, 16, 19].

In [16], the authors compared subject comparisons in the Soton gait database by comparing one target to multiple subjects. When comparing one target to more than one subject, they found that the comparative descriptions performed 17% better than the absolute/categorical descriptions.

Also, in [4], the authors used semantic clothing traits as soft biometrics for human identification. They further explored their validity and efficiency through corresponding comparative descriptions, allowing for more accurate differentiation.

The study in [15] analyzed a data set using gender as a comparative attribute and found that comparative annotations are more discriminatory than categorical labels. The study's approach on 100 annotated subject images showed correct-match reliability in the top 7% with ten comparisons or the top 13% with only five comparisons.

Additionally, in [19], the authors studied human identity through comparative facial soft biometrics using the labeled faces in the wild (LFW) dataset. Comparative soft biometrics allows each person to be uniquely identified in the database by creating a biosignature with their exact physical traits compared to others. Such comparative descriptions improved searches based on a given comparative trait (e.g., searching for someone younger).

The last study [54] focused on enhanced human-machine communication and demonstrated the benefits of relative attributes on four applications. The applications include active learning of discriminative classifiers, zero-shot learning from relative comparisons, automatic image description, and image search with interactive feedback. Relative attributes compare an image's attribute strength to others rather than predicting its presence. Relative attributes would be more natural and allow for richer communication, more detailed human supervision, higher recognition accuracy, and more informative descriptions of new images. The SVM ranking function is learned and used for each attribute to predict the attribute strength. They demonstrate that relative attributes lead to higher performance across all applications by comparing them to numerous strong baselines using image datasets of scenes and faces.

All previous research works aim to bridge the semantic gap between biometrics and human description by using relative and comparative descriptions of attributes. In particular, to the best of our knowledge, no study has used comparative descriptions for semantic attribute analysis of resume documents for improving job-candidate matching.

Table 1 presents our proposed approach compared with several previous studies.

## 7 Methodology

As in the literature, there is a need to improve the process of automatically extracting data from resumes and mining their information [55]. The improvement can assist in selecting candidates and ranking them according to the job's requirements. Recruiters may have subjective ideas about which aspects they want to emphasize, depending on the circumstances and nature of their company's requirements at that particular time. Therefore, in this research, we proposed a novel approach to improve the candidate-career matching process based on comparative semantic resume analysis. Our approach comprises three phases: the semantic attribute extraction phase using NLP techniques, the relative and comparative labeling phase, and the ranking phase using the ranking SVM algorithm. The proposed model can generally analyze the resume by finding the exact keywords for each resume section, which will be used to assign a suitable label for 13 semantic attributes each. Thus, our proposed model lets the recruiters modify the attribute-based searching as they want to depend on their own requirements. After constructing a table or dictionary, which covers various semantic sets of attributes, count the occurrences of the words belonging to different semantic attributes and assign scores for each attribute based on evaluation metrics, which are dynamic to change and determined by recruitment staff based on the job requirements. Then, aggregate all points for each attribute as regular scores and show the related normalization score. After that, we come to the comparative description part, to which we contribute to enhancing the resume analysis domain by inferring a suited relative descriptive label per attribute for each resume based on evaluating their given scores. Then, inferring corresponding comparative descriptive labels by describing a pairwise comparison per attribute for all of the attributes between each resume and all other resumes in the dataset. Finally, ranking all resumes per attribute using the Rank SVM algorithm through extracting usable relative measurements from those comparative descriptions as resume scoring per attribute and, consequently, leading to overall resume scoring concerning all sets or a selected subset of attributes in descending order. This method helps HR query the list of candidates ranked based on each semantic attribute in the resume or multiple semantic attributes.

Our study aims to provide an ML-based model that does not solely rely on training data to match candidate resumes to career requirements. However, it also enables dynamic learning and ranking per any combination of the proposed resume semantic attributes. We summarize all of the methodology phases and present the flowchart of the proposed approach of semantic resume analysis in the next subsection.

Table 1: Comparing our proposed approach with previous studies

Study	ML	NLP	Ontology	Scoring	Ranking	Comparative	Language
[10]	√	√	-	√	-	-	English
[11]	√	√	-	-	-	-	English
[20]	-	√	-	-	-	-	English
[21]	√	√	-	-	-	-	Vietnamese
[22]	√	√	-	√	√	-	English
[23]	√	√	-	-	-	-	English
[24]	√	√	-	√	√	-	English
[25]	√	√	-	-	-	-	English
[26]	√	√	-	√	√	-	English
[27]	√	-	-	-	-	-	English
[28]	√	√	-	-	-	-	English
[29]	√	√	-	-	-	-	English
[30]	√	√	-	√	√	-	English
[31]	√	√	-	√	√	-	English
[32]	√	√	-	-	-	-	English
[33]	√	√	-	√	√	-	English
[34]	√	√	-	√	√	-	English
[35]	√	√	-	√	-	-	English
[36]	√	√	-	-	-	-	English, Norwegian, Swedish, Finnish, Polish
[37]	√	√	-	-	-	-	English
[38]	√	√	-	√	√	-	English
[39]	√	√	-	√	√	-	English
[40]	√	√	-	-	-	-	English
[41]	√	√	-	-	-	-	English
[42]	√	√	-	√	√	-	English
[43]	√	√	-	√	√	-	English
[44]	√	-	√	√	√	-	English
[45]	√	-	√	√	√	-	English
[46]	√	√	√	√	√	-	English
[47]	√	√	√	-	-	-	Polish
[48]	√	√	√	√	√	-	English
[49]	√	√	√	√	√	-	English
[50]	√	√	√	√	√	-	English
[53]	√	√	-	√	√	-	English
<b>Our work</b>	√	√	-	√	√	√	<b>English</b>

### 7.1 Summary of Proposed Approach

Our approach is proposed to assist HR officers in improving the extraction, filtering, and selection of the most job-suited resumes in the technical domain (or specialties) in an automated and more accurate manner based on analyzing semantic resume attributes. For each resume, two groups of per-attribute semantic descriptions are used as relative and comparative labels, which are more precise and informative than categorical descriptions. The following scenario will be followed in this research study:

1. Construct a table or dictionary that covers various semantic sets of attributes with each keyword in the resume, such as (personal information, education level, technical skills level, professional experience, personal skills, additional qualification, etc.).
2. An NLP-based tool is used to parse and extract information from the entire resume to search the words in the table or dictionary.
3. Count the occurrences of the words belonging to different semantic attributes and assign scores for each attribute based on evaluation metrics determined by recruitment staff based on the job requirements.

4. Aggregates all points for each attribute as regular scores and shows the related normalization score.
5. Inferring a suited relative descriptive label per attribute for each resume based on evaluating their given scores.
6. Inferring corresponding comparative descriptive labels by describing a pairwise comparison per attribute for all of the attributes between each resume and all other resumes in the dataset.
7. Ranking all resumes per attribute using the SVM ranking algorithm through extracting usable relative measurements from those comparative descriptions as resume scoring per attribute and, consequently, leading to overall resume scoring concerning all sets or a selected subset of attributes in descending order.
8. Analyzing results based on four comparison aspects which compare the efficiency in distribution and discrimination of resumes rankings.

Figure 2 overviews the proposed semantic resume analysis approach, and we will explain each phase in detail in the following subsections.

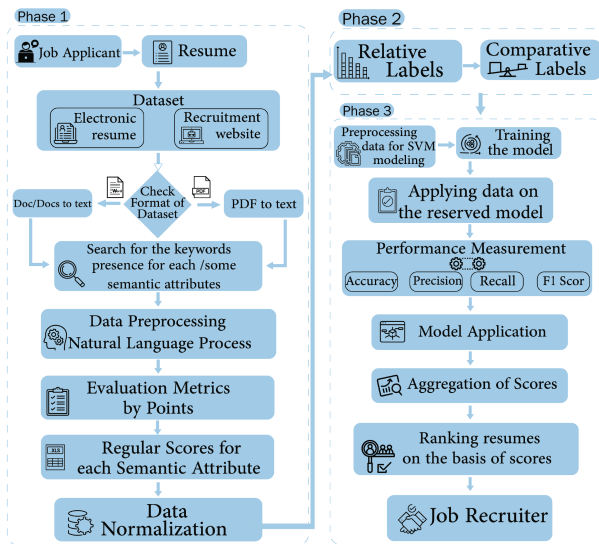


Figure 2: Overview of the proposed semantic resume analysis approach

## 7.2 Semantic Information Extraction Phase

An automated system for extracting information and selecting the potential candidates who best meet the position's requirements can significantly improve the efficiency of HR agencies [56]. However, resumes are unstructured documents with different file formats (pdf, txt, doc, etc.) and contain ambiguous and variable language. This heterogeneity makes it challenging to extract useful information.

To make the recruitment process simple, efficient, and automated, some advanced studies, like [10] and [33], solved the parsing problem with specific searches and unstructured data. They searched the whole document for words in predefined tables, dictionaries, or a BoW to list important words and count the number of detected words of interest. Therefore, we proposed a candidate selection system based on searching the keywords. If the keywords exist, it will count their presence and allocate the given points as evaluation metrics or as HR-specified job requirements.

Before we move to define the proposed semantic attributes, we will overview the NLP techniques in general and the related techniques we used in our approach:

### 7.2.1 NLP with Python

By "natural language," we refer to a human language used for everyday conversation, such as English, Arabic, Hindi, and Portuguese. In contrast to programming and formal languages such as computer and mathematical notations, natural languages have evolved as they are passed from generation to generation and are challenging to define with specific rules. NLP aims to "understand" complete human sentences so that appropriate replies can be provided. More and more technologies are appearing based mainly on natural language processing. In today's increasingly globalized and interconnected information world, language processing has assumed critical importance [17].

### 7.2.2 Proposed Semantic Attributes

Table 2 demonstrates the proposed semantic attributes that will be extracted from resumes and all relevant keywords and suggested evaluation matrices that will be defined and used in our approach.

### 7.2.3 Keywords Detection or Information Extraction

- Extracting files from the archive directory of resumes dataset and generating a list using `extractall()` python function.
- Checking the format of the input file from the dataset (pdf, doc/docx) and converting them to text using `str()` python function.

- Extracting semantic attribute information. We construct three functions (a function to check keywords, a function to check keywords and count their presence, and a function to check how many keywords are present). First, the personal information, name, email, and mobile number are parsed separately. Then, allocate the weights that the requirements illustrated using a simple resume parser existing in [57], which is used for extracting information from resumes by using the NLTK and SpaCy models of NLP. Since resumes have rather varied structures and formats, the keywords are similarly searched for the ten remaining semantic attributes. If the keywords exist, then the given points of evaluation metrics would be allocated. Aggregating all points for each attribute as regular scores and showing the related normalization score.

### 7.2.4 Data Normalization

For many procedures to work properly, the data must be normalized or standardized; since resumes have unbalanced scores after extracting attributes, we used normalization (MinMaxScaler) to transform features (the regular scores of the attributes) by scaling each attribute to a corresponding range (between zero and one). Then, we used the normalization score to get the relative scores, which range from 0 to 5, as explained in the next section.

The transformation imported using Scikit-learn of ML library is given by:

$$X\_std = \frac{(X - X_{\min}(axis=0))}{(X_{\max}(axis=0) - X_{\min}(axis=0))} \quad (1)$$

$$X\_scaled = X\_std * (max - min) + min$$

where  $min, max = feature\_range$ .

Table 3 shows an example of evaluation metrics by scores given for the personal information attribute, where the total is eight points for the whole personal information attribute and shows the related normalization score, which is calculated by "(1)" where  $feature\_range=(0,8)$ ,  $X=8$ ,  $min=1$ , and  $max=8$ .



Table 2: Proposed resume semantic attributes to be extracted and analyzed

Semantic attributes	Keywords
1- Personal information completeness	Name, Nationality, Mobile No., Email, Address, Phone, Contact No, Country, ...etc.
2- Education level	GPA, Bachelor, Master, PhD.
3- Technical skills level	Programming Languages, Tools, Framework, Databases, Operating systems, ...etc.
4- Professional experience	Experience, Participated, Worked, ...etc.
5- Personal skills (Soft skills)	Communication, Teamwork, Time Management, Problem-Solving, Creativity, Leadership, ...etc.
6- Awards	Awarded, Winner, ...etc.
7-Hobbies and interest	Play, Games, Surf, News, Design, Read, Video, Sports, ...etc.
8- Additional Qualification	Course, Program, Trained, Level, ...etc.
9- Professional Certificate	Certified, Cisco, Scrum, Professional, Management, Administrator, ...etc.
10- Suitable age requirement	Age, Years old, Birth, Birth year.
11- Career objective	Gain, Objective, Opportunity, Looking for, Looking forward, ...etc.
12- Project Experience	Project, Role, Client, Responsibilities, Environment.
13- Languages	Arabic, English, ILETS, STEP, ...etc.

Table 3: Example of evaluation metrics by scores, regular score, and normalization score for the personal information attribute

Personal information	Score
Name	2
Nationality	1
Address	1
Mobile No.	2
Email	2
Total (Regular score)	8
Normalized score	1

### 7.3 Relative and Comparative Labeling Phase

After we extract the important semantic attributes from the resume, we come to the comparative description part, to which we contribute to improving the resume analysis domain. In our work, apart from simple traditional evaluation metrics (a regular scoring system), we use the proposed comparative and relative labels to score or rate a resume more precisely.

After normalization, the points are totaled for each attribute in the regular scoring system to infer the relative scores. We multiply each point by five and then round it down. The labels of relative descriptions are defined as "Very high," "High," "Average," "Low," "Very low," and "None." Table 4 shows an example of relative labels and their corresponding scores.

Table 4: Relative labels and corresponding scores

Relative score system	Relative labels
5	Very high
4	High
3	Average
2	Low
1	Very low
0	None

After assigning the most appropriate scores based on relative labels for each semantic attribute in the resume, we will accordingly

use comparative labels to describe a pairwise comparison per attribute for all of the attributes between each resume and all other resumes in the dataset (one subject with all others). The comparative labels include "Much higher," "Higher," "Same," "Lower," and "Much lower." For example, when the relative score of personal information is 5 for resume A and 4 for resume B, then the comparative label for resume A will be labeled as "Higher" than resume B, with the new corresponding score equal to 1, as listed in Table 5, which shows an example of comparative labels and their corresponding scores.

Table 5: Comparative labels and corresponding scores

Comparative score system	Comparative labels
2	Much higher
1	Higher
0	Same
-1	Lower
-2	Much lower

We put the comparative scores as follows: the score will be "1" or "-1" if the difference between the relative scores of two resumes is 1 or 2. The score will be "2" or "-2" if the difference between the relative scores of the two resumes is more than 2, and the score will be "0" if the relative scores of the two resumes are the same. As a result, resume A will be compared to each of the other resumes, one by one, per attribute.

Note that when we compare resumes A and B, we do not need the opposite comparative label comparing B and A because we only compare the possible pairwise combinations, not the permutations. The combination relations refer to the combination of elements  $n$  taken  $k$ -combination at a time without repetition, and the order of elements selection does not matter here. Therefore, a  $k$ -combination of the set  $S$  is hence a subset of  $k$  different elements from  $S$ . If there are  $n$  elements in the set, the number of  $k$ -combinations is equal to:

$$\frac{n!}{k!(n-k)!} \quad (2)$$

whenever  $k \leq n$ , and which is zero when  $k > n$

## 7.4 Ranking SVM Phase

Using the ranking SVM algorithm, the final resume ranking per semantic attribute will be based on comparative labels concerning the other N-1 resumes. Rank SVM is a technique for sorting lists of objects using pairwise difference vectors to adaptively arrange given comparable peers based on a certain criterion. Rank SVM performs rankings using Standard SVM. The ranking SVM's objective is to sort the list of resumes for each attribute, and we can infer the ordering list of resumes for all attributes. This ranking will enable further search capabilities based on how relevant the retrieved objects are to a particular search query.

In our work, we used (**train\_test\_split**) function of sklearn that split input resumes and values data into train and test dataset according to **test\_size** ratio. The **test\_size** is equal to 0.33 mean, then 33% is test data, and 67% is train data. The output is **resume\_train**, **resume\_test**, which is (**x\_train x\_test**) and (**y\_train, y\_test**). All resume comparisons are placed as pairwise inputs into the SVM ranking algorithm for the training, with each attribute labeled. We used hyper-parameter optimization in SVM for both C and Gamma values to get their optimized values. Then, pass training data to SVM to predict the comparative labels against each attribute and resume pairs by using the three following functions [58, 59]:

- **svm = SVC(kernel = 'rbf', c=10, gamma=0.1)**
- **svm.fit(x\_train, y\_train)**
- **predict\_eachAttribute = svm.predict(x\_test).**

Rank resumes based on predicted ranks so that in each iteration, the resume with the highest rank is selected and removed for the next iteration. Ultimately, we will get the ranks of resumes based on scores against individual attributes. Finally, the output will be the corresponding usable relative measurements and the ordering list based on comparisons of all resumes.

This method helps HR query the list of candidates ranked based on single or multiple semantic attributes in their resumes. As such, we can apply a search to retrieve the resume that is "Much higher" than the others in a certain attribute. To rearrange the resumes based on the resulting rankings, we can count (sum) their nascent relative measurement scores based on multiple or all attributes for each resume. The ranking is given in descending order to present the resumes from the best to the worst. A soft-margin ranking SVM method is used for a given set of attributes  $A$  to learn a ranking linear function  $r_a$  for each attribute, similar to the way used in [4]:

$$r_a(x_i) = w_i^T x_i \quad (3)$$

Where  $w_a$  is the coefficient of the ranking function  $r_a$  and  $x_i$  is a feature vector of attributes of a resume being ranked. Rearranging a set of comparisons into two groups can be considered a representation of the pairwise relative constraints needed to learn a ranking function. The first group is a set of dissimilarity comparisons  $D_a$  of ordered pairs so that  $(i, j) \in D_a \Rightarrow i > j$ . The second group is a set of similarity comparisons  $S_a$  of non-ordered pairs so that  $(i, j) \in S_a \Rightarrow i = j$ . Then, the following formula is used to get the  $w_a$

coefficients of  $r_a$  from the  $D_a$  and  $S_a$  sets:

$$\begin{aligned} & \text{minimize} \quad \left( \frac{1}{2} \|w_a^t\|^2 + C \sum \xi_{ij}^2 \right) \\ & \text{subject to} \quad w_a^T(x_i - x_j) \geq 1 - \xi_{ij}; \quad \forall (i, j) \in D_a \\ & \quad \quad \quad |w_a^T(x_i - x_j)| \leq \xi_{ij}; \quad \forall (i, j) \in S_a \\ & \quad \quad \quad \xi_{ij} \geq 0 \end{aligned} \quad (4)$$

$\xi_{ij}$  is the misclassification bias, and  $C$  is the trade-off between maximization of margin and minimization error. The resulting optimal  $w_a$  function can then be used to (explicitly) rank all training samples according to  $a$ . "Equation(3) is used to map a feature vector  $x_i$  to a feature vector consisting of several real-value relative measurements."

## 7.5 Performance evaluation

For SVM performance evaluation, we assess the accuracy of the proposed model, along with a classification report that quantifies the quality of the predictions of each semantic attribute.

A classification report assesses the quality of a classification algorithm's predictions. How many predictions are correct, and how many are incorrect? True Positives (TP), False Positives (FP), True Negatives (TN), and False Negatives (FN), to be exacted, as in Figure 3 [60].

Actual	Positive	TP	FN
	Negative	FP	TN
		Positive	Negative
		Predicted	

Figure 3: Simple representation of confusion matrix [60]

The input of the classification report is ( $\mathbf{X}, \mathbf{Y}$ ) where  $X$  is predicted values and  $Y$  is actual values. We used (**classification\_report**) function of sklearn library to print precision, recall, and F1-score. Similarly, used (**accuracy\_score**) function of sklearn library to print the accuracy of the SVM model. Whereas the accuracy is the proportion of accurate predictions to the total number of data samples [61]. The precision shows how many predictions for a certain class are of the same class. The recall shows how many of certain class predictions were right, and the F1-score is the geometric mean of precision and recall [62]. The numeric implementation of the classification report is proven in the next section of the experiment results.

## 7.6 Practical Example of the Three Proposed Phases

The graph in Figure 4 shows our approach example with the following information:

- Inputs: 3 resumes (resume A, resume B, and resume C) with semantic education attributes extracted from resumes.
- Processing: Describing the relative and comparative labels with

corresponding scores for comparing between resumes.

-Output: Extract usable relative measurements in which these values are set and adjusted based on comparing each resume with the other resumes. Then, get the ordering list of all resumes using the SVM ranking algorithm.

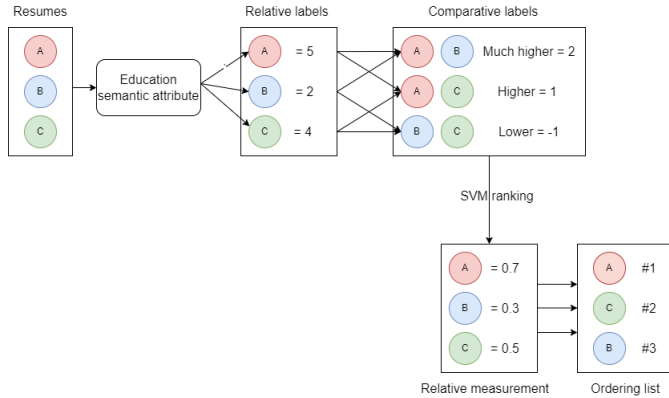


Figure 4: A simple example clarifies the main three phases of our proposed approach

## 8 Experiments and Results

### 8.1 Dataset

Our proposed approach’s experiment was conducted on a dataset containing 228 resumes in various text file formats, such as doc, docx, and pdf. We collected resumes from various technical (computer-related) specialists from various sources, including Github, Kaggle, LinkedIn, etc. Most resumes contain information about technical skills, such as programming in Python or Java, and other technical experience. We used the resume dataset for analysis, and each resume was assigned relative and comparative labels.

### 8.2 Implementation of three proposed phases

After the three phases of our approach, we generated 25,878 resume comparisons between resumes for each attribute. To select the candidate who matches the job requirements, we aggregated the scores and ranked the resumes according to each attribute and all attributes, as detailed in the methodology chapter. The resumes were ranked in descending order of total score, with the highest resume coming first and the lowest coming last. Next, we analyzed the results using four comparison-related aspects:

1. Ranking based on the regular scores. Figure 5 shows the histogram distribution of resume rankings for all attributes based on the regular scores.
2. Ranking based on the comparative scores. Figure 6 shows the histogram distribution of resume rankings based on the comparative scores for all attributes.
3. Ranking based on the comparative scores and regular scores. Figure 7 shows the histogram distribution of resume rankings for all attributes based on the comparative and regular scores.

4. Ranking based on the comparative scores and relative scores. Figure 8 shows the histogram distribution of resume rankings for all attributes based on the comparative and relative scores.

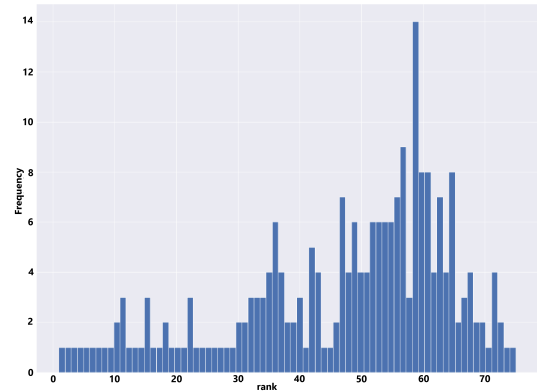


Figure 5: Resume rankings for all attributes based on the regular scores

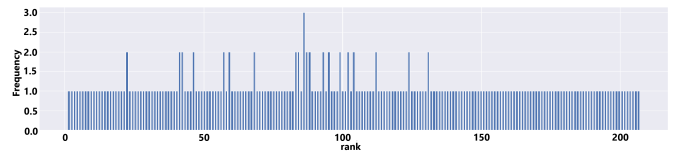


Figure 6: Resume rankings for all attributes based on the comparative scores

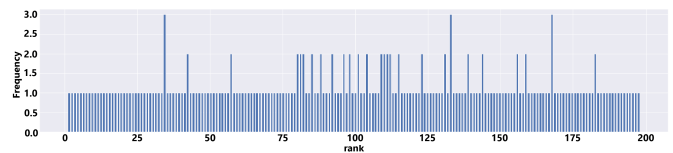


Figure 7: Resume rankings for all attributes based on the comparative scores and regular scores

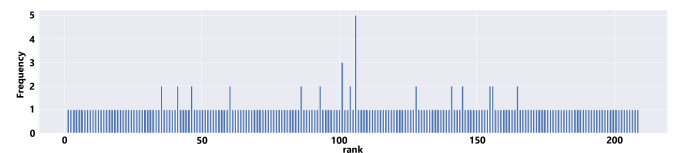


Figure 8: Resume rankings for all attributes based on the comparative scores and relative scores

### 8.3 Analysis of Results

Table 6 summarizes the four ranking results. The table compares the effectiveness of distributing and discriminating in the rankings of resumes with high unique values. It is observed that all three rankings based on comparative scores outperform the ranking based on regular scores, which offers very weak resume differentiation with high (undesired) redundant values and highly (confusable) similarities in resumes’ scores. It is important to observe that when it comes to resume distributions (as indicated in Table 6, Figure

6, Figure 8, and Figure 7), the more unique scores, the better the ranking result.

Table 6: Results of rankings with different four basis

Basis of rankings	Max Rank	No. of resumes that has unique scores	No. of resumes that has redundant scores
Regular scores	75	30	198
Comparative scores	207	187	41
Comparative and Regular	198	171	57
Comparative and Relative	<b>209</b>	<b>194</b>	<b>34</b>

Let us demonstrate the theoretical distribution of the duplicated rankings between resumes on each basis of rankings. Figure 9 shows the histogram distribution of the duplicated rankings between resumes for all attributes based on the regular scores. Figure 10 shows the histogram distribution of the duplicated rankings between resumes for all attributes based on the comparative scores. Figure 11 shows the histogram distribution of the duplicated rankings between resumes for all attributes based on the comparative and regular scores. Figure 12 shows the histogram distribution of the duplicated rankings between resumes for all attributes based on the comparative and relative scores.

As we proved in Table 6, the distribution of the duplicated rankings between resumes for all attributes based on the comparative and relative scores outperforms all other ranking bases.

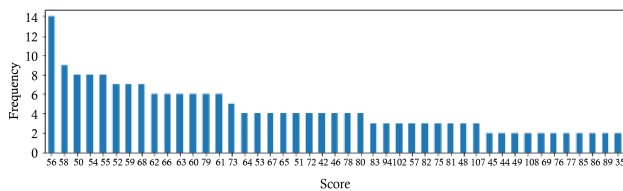


Figure 9: Duplicated rankings between resumes for all attributes based on the regular scores.

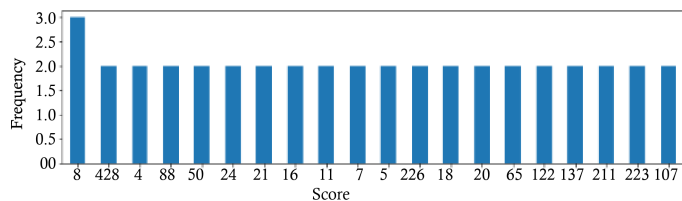


Figure 10: Duplicated rankings between resumes for all attributes based on the comparative scores.

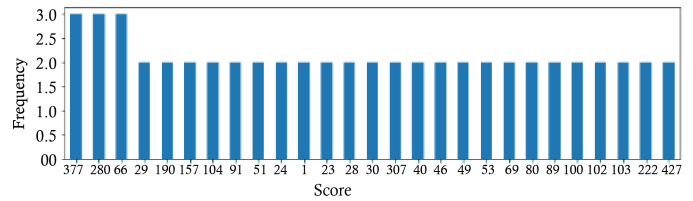


Figure 11: Duplicated rankings between resumes for all attributes based on the comparative and regular scores.

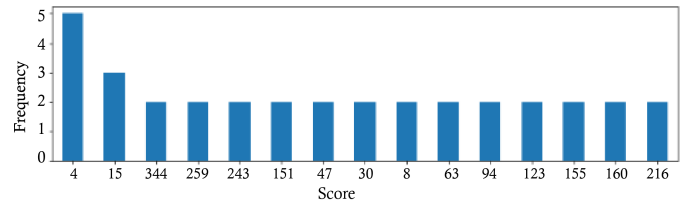


Figure 12: Duplicated rankings between resumes for all attributes based on the comparative and relative scores.

We proved that the method of ranking based on comparative and relative scores is very effective by comparing its accuracy with all other basis rankings as shown in Table 7. The ranking quality can further be measured as the accuracy of differentiating (as much as possible) between  $N$  of compared resumes. Resulting in as many as possible different unique scores  $S$ , which is, in the best-case scenario, equal to the number of compared resumes (i.e.,  $S = N$ ), implies that each resume has a unique (distinct) score. This approach is made by computing the percentage of the total sum of all unique and redundant scores concerning their frequencies using the following formula:

$$Accuracy = \frac{1}{N} \left( \sum_i^S \frac{1}{f_i} \right) * 100 \tag{5}$$

Where  $N$  is the number of resumes to be ranked.  $S$  is the total number of unique scores assigned to  $N$  number of resumes to be ranked, and  $f_i$  is the corresponding  $i$ th frequency of the  $i$ th unique score.

Table 7: Accuracy of rankings with different four basis

Basis of rankings	Accuracy
Regular scores	33%
Comparative scores	90%
Comparative and Regular scores	87%
Comparative and Relative scores	<b>92%</b>

As such, it is seen that all methods that rank resumes using comparative scores outperform those that rely on regular scores for ranking. The highest accuracy in ranking the resumes is achieved using ranks derived from comparative scores with the relative scores as the basis. This approach can be applied to each attribute. Figure 13 shows the rankings' histogram of the resumes for the personal information attribute, taking the comparative and relative scores as the basis.



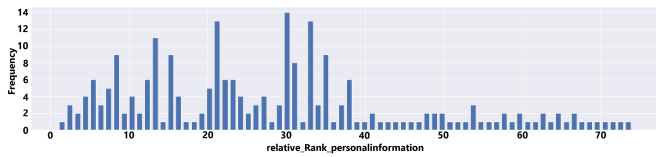


Figure 13: Resume rankings for personal information attribute based on the comparative scores and relative scores

Finally, Figure 14 depicts the top ten resumes’ (candidates) rankings with totaling scores on a comparative and relative basis.

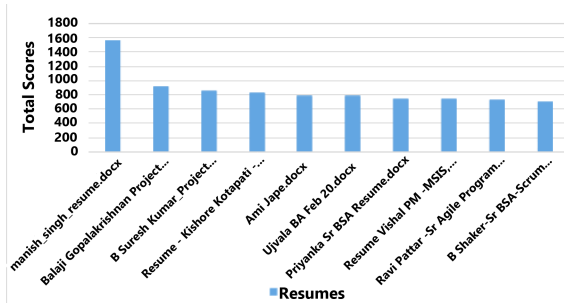


Figure 14: Top ten resume’ rankings

### 8.4 Evaluation of Resume Ranking per Attribute

Table 8 reports the accuracy [61] and classification report that contains a precision [63], recall [64], and F1 score [65] of the SVM ranking algorithm which quantifies the quality of the predictions of each of all semantic attributes.

Table 8: Evaluation metrics of the SVM ranking algorithm for each of all semantic attributes

Semantic Attribute	Accuracy	Precision	Recall	F1-Score
1- Personal information	99.168%	0.97751	0.97162	0.97456
2- Education level	99.496%	0.98852	0.99625	0.99238
3- Technical skills	99.461%	0.99106	0.98162	0.98634
4- Professional experience	99.754%	0.99516	0.99756	0.99636
5- Personal skills	99.841%	0.98810	0.99747	0.99278
6- Awards	99.660%	0.98212	0.99108	0.98660
7- Hobbies and interest	99.332%	0.98624	0.96987	0.97805
8- Additional Qualification	99.601%	0.98980	0.96697	0.97838
9- Professional Certificate	99.344%	0.97351	0.99453	0.98402
10- Suitable age requirement	99.941%	0.99099	0.99979	0.99539
11- Career objective	99.976%	0.99984	0.99963	0.99973
12- Project Experience	99.742%	0.98223	0.99774	0.98998
13- Languages	99.964%	0.99964	0.99975	0.99969
<b>Average score</b>	<b>99.652%</b>	<b>0.98805</b>	<b>0.98952</b>	<b>0.98878</b>

Through the SVM algorithm performance table, we note that the career objective attribute is higher than other attributes in accuracy, precision, and F1-score, and the suitable age requirement attribute is higher than other attributes in the recall. In general, the average values of all evaluation metrics represent high performance up to 99%. The table results prove the effectiveness of using an SVM algorithm to rank resumes for each semantic attribute separately.

## 9 Conclusions

This study presents a model to improve the candidate ranking and selection process for career matching. The model consists of three phases: firstly, the extraction of semantic information for 13 proposed attributes using NLP tools; secondly, inferred of relative and comparative descriptions by comparing every two resumes for each attribute; and finally, the use of the ranking SVM algorithm to order all resumes for each attribute. Implementing a comparison description in the candidates’ selection process has been suggested and proven more accurate and informative. It helps to distinguish the tiny differences between resumes. The experiment results were analyzed using four comparison aspects: ranking based on regular scores, comparative scores, comparative and regular scores, and comparative and relative scores. The histogram results prove that all three ranking methods based on comparative scores with high unique rankings values outperform the ranking using mere regular scores, which offers very weak resume differentiation with high redundant rankings values and high similarities in resume scores. Also, the ranking attained by the comparative scores with the relative scores as the basis achieves the highest accuracy of 92%. We reported evaluation metrics of the SVM ranking algorithm with an average accuracy of 99.652%, an average precision of 0.98805, an average recall of 0.98952, and an average F1-score of 0.98878, which quantifies the quality of the predictions of each of all semantic attributes. Finally, we have obtained a list of the top ten candidates out of 228 technical specialists’ resumes.

## 10 Future Work

In the future, a proposed model should consider keyword matching for 13 attributes and focus on the semantic relations and meanings of the resume language to account for variations in resume authoring. We will explore additional parts that individuals may incorporate into their resumes to enhance the evaluation and comparison process for more effective candidate-career matching. Furthermore, we will explore the possibility of applying the proposed methods to many domains beyond the technological field and incorporating languages other than English, such as Arabic. Finally, we will evolve our work to analyze the job description document to calculate its matching with the resume as another way different from calculating evaluation metrics of job requirements.

**Acknowledgment** This work is selected as a special issue of a conference paper published in the 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN) to be extended into a journal paper.

## References

[1] A. H. Alderham, E. S. Jaha, “Comparative Semantic Resume Analysis for Improving Candidate-Career Matching,” in 2022 14th International Conference on Computational Intelligence and Communication Networks (CICN), 313–321, IEEE, 2022, doi:10.1109/CICN56167.2022.10008255.

[2] Y. Lin, H. Lei, P. C. Addo, X. Li, “Machine learned resume-job matching solution,” arXiv preprint arXiv:1607.07657, 2016.

- [3] C. Saxena, "Enhancing productivity of recruitment process using data mining & text mining tools," 2011.
- [4] E. S. Jaha, M. S. Nixon, "Soft biometrics for subject identification using clothing attributes," in *IEEE international joint conference on biometrics*, 1–6, IEEE, 2014, doi:10.1109/BTAS.2014.6996278.
- [5] D. Lee, M. Kim, I. Na, "Artificial intelligence based career matching," *Journal of Intelligent & Fuzzy Systems*, **35**(6), 6061–6070, 2018, doi:10.3233/JIFS-169846.
- [6] L. Columbus, "Remote Recruiting In A Post COVID-19 World," <https://www.forbes.com/sites/louiscolumbus/2020/03/30/remote-recruiting-in-a-post-covid-19-world/?sh=7c4ce7e599fd>, 2020.
- [7] C. Russell, *Right Person–Right Job: Guess Or Know: the Breakthrough Technologies of Performance Information, Human Resource Development*, 2002.
- [8] A. Swenson, "The new art of hiring smart: Matching the right person to the right job," Technical report, 2000.
- [9] A. K. Sinha, M. A. K. Akhtar, A. Kumar, "Resume Screening Using Natural Language Processing and Machine Learning: A Systematic Review," *Machine Learning and Information Processing: Proceedings of ICMLIP 2020*, **1311**, 207, 2021.
- [10] Y. Suresh, A. M. Reddy, et al., "A Contextual Model for Information Extraction in Resume Analytics Using NLP's Spacy," in *Inventive Computation and Information Technologies*, 395–404, Springer, 2021, doi:10.1007/978-981-33-4305-4\_30.
- [11] B. Gaur, G. S. Saluja, H. B. Sivakumar, S. Singh, "Semi-supervised deep learning based named entity recognition model to parse education section of resumes," *Neural Computing and Applications*, **33**(11), 5705–5718, 2021, doi:10.1007/s00521-020-05351-2.
- [12] A. Kulkarni, A. Shivananda, "Advanced Natural Language Processing," in *Natural Language Processing Recipes*, 97–128, Springer, 2019.
- [13] H. Assal, J. Seng, F. Kurfess, E. Schwarz, K. Pohl, "Semantically-enhanced information extraction," in *2011 Aerospace Conference*, 1–14, IEEE, 2011.
- [14] S. Singh, "Natural language processing for information extraction," arXiv preprint arXiv:1807.02383, 2018.
- [15] D. Martinho-Corbishley, M. S. Nixon, J. N. Carter, "Analysing comparative soft biometrics from crowdsourced annotations," *IET Biometrics*, **5**(4), 276–283, 2016, doi:10.1049/iet-bmt.2015.0118.
- [16] D. A. Reid, M. S. Nixon, S. V. Stevenage, "Soft biometrics; human identification using comparative descriptions," *IEEE Transactions on pattern analysis and machine intelligence*, **36**(6), 1216–1228, 2013, doi:10.1109/TPAMI.2013.219.
- [17] S. Bird, E. Klein, E. Loper, *Natural language processing with Python: analyzing text with the natural language toolkit*, "O'Reilly Media, Inc.", 2009.
- [18] D. Parikh, K. Grauman, "Relative attributes," in *2011 International Conference on Computer Vision*, 503–510, IEEE, 2011.
- [19] N. Y. Almodhahka, M. S. Nixon, J. S. Hare, "Comparative face soft biometrics for human identification," in *Surveillance in Action*, 25–50, Springer, 2018, doi:10.1007/978-3-319-68533-5\_2.
- [20] J. Chen, C. Zhang, Z. Niu, "A two-step resume information extraction algorithm," *Mathematical Problems in Engineering*, **2018**, 2018, doi:10.1155/2018/5761287.
- [21] L. Pham Van, S. Vu Ngoc, V. Nguyen Van, "Study of Information Extraction in Resume," Conference, 2018.
- [22] R. Nimbekar, Y. Patil, R. Prabhu, S. Mulla, "Automated Resume Evaluation System using NLP," in *2019 International Conference on Advances in Computing, Communication and Control (ICAC3)*, 1–4, IEEE, 2019, doi:10.1109/ICAC347590.2019.9036842.
- [23] Y. O. German, O. V. German, S. Nasr, "Information extraction method from a resume (CV)," 2019.
- [24] S. Amin, N. Jayakar, M. Kiruthika, A. Gurjar, "Best Fit Resume Predictor," *International Research Journal of Engineering and Technology*, **6**(8), 813–820, 2019.
- [25] S. T. Gopalakrishna, V. Vijayaraghavan, "Automated Tool for Resume Classification Using Sematic Analysis," *International Journal of Artificial Intelligence and Applications (IJAIA)*, **10**(1), 2019.
- [26] V. Bhatia, P. Rawat, A. Kumar, R. R. Shah, "End-to-end resume parsing and finding candidates for a job description using bert," arXiv preprint arXiv:1910.03089, 2019.
- [27] Y.-C. Chou, H.-Y. Yu, "Based on the application of AI technology in resume analysis and job recommendation," in *2020 IEEE International Conference on Computational Electromagnetics (ICCEM)*, 291–296, IEEE, 2020, doi:10.1109/ICCEM47450.2020.9219491.
- [28] P. Kadu, V. Z. Ashwini, "Knowledge Extraction from Text Document Using Open Information Extraction Technique," *International Journal of Advanced Trends in Computer Science and Engineering (IJATCSE) Volume*, **9**, 2020.
- [29] F. O. Oladipo, A. A. Ayomikun, "A MODEL FOR AUTOMATIC RESUME SUMMARIZATION," *Journal of Information Systems & Operations Management*, **14**(2), 147–159, 2020.
- [30] P. K. Roy, S. S. Chowdhary, R. Bhatia, "A Machine Learning approach for automation of Resume Recommendation system," *Procedia Computer Science*, **167**, 2318–2327, 2020.
- [31] K. Satheesh, A. Jahnavi, L. Iswarya, K. Ayesha, G. Bhanusekhar, K. Hanisha, "Resume Ranking based on Job Description using SpaCy NER model," *International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056*, **7**(05), 2020.
- [32] V. Mittal, P. Mehta, D. Relan, G. Gabrani, "Methodology for resume parsing and job domain prediction," *Journal of Statistics and Management Systems*, **23**(7), 1265–1274, 2020, doi:10.1080/09720510.2020.1799583.
- [33] N. Khamker, "Resume Match System," *International Journal of Innovative Science and Research Technology*, 2021.
- [34] S. Bhor, H. Shinde, V. Gupta, V. Nair, M. Kulkarni, "Resume parser using natural language processing techniques," 2021.
- [35] R. Haddad, E. Mercier-Laurent, "Curriculum Vitae (CVs) Evaluation Using Machine Learning Approach," in *IFIP International Workshop on Artificial Intelligence for Knowledge Management*, 48–65, Springer, 2021.
- [36] D. Vukadin, A. S. Kurdija, G. Delač, M. Šilić, "Information Extraction from Free-Form CV Documents in Multiple Languages," *IEEE Access*, 2021.
- [37] K. F. F. Jiechieu, N. Tsope, "Skills prediction based on multi-label resume classification using CNN with model predictions explanation," *Neural Computing and Applications*, **33**(10), 5069–5087, 2021.
- [38] J. Zhao, J. Wang, M. Sigdel, B. Zhang, P. Hoang, M. Liu, M. Korayem, "Embedding-based Recommender System for Job to Candidate Matching on Scale," arXiv preprint arXiv:2107.00221, 2021.
- [39] K. Tejaswini, V. Umadevi, S. M. Kadiwal, S. Revanna, "Design and Development of Machine Learning based Resume Ranking System," *Global Transitions Proceedings*, 2021.
- [40] K. Bothmer, T. Schlippe, "Skill scanner: connecting and supporting employers, job seekers and educational institutions with an AI-based recommendation system," in *The Learning Ideas Conference*, 2022.
- [41] I. Ali, N. Mughal, Z. H. Khand, J. Ahmed, G. Mujtaba, "Resume classification system using natural language processing and machine learning techniques," *Mehran University Research Journal Of Engineering & Technology*, **41**(1), 65–79, 2022.
- [42] B. Amro, A. Najjar, M. Macido, "An Intelligent Decision Support System For Recruitment: Resumes Screening and Applicants Ranking," 2022.

- [43] B. Kinge, S. Mandhare, P. Chavan, S. Chaware, "Resume Screening Using Machine Learning and NLP: A Proposed System," 2022.
- [44] O. Dada, A. Kana, S. Abdullahi, "An ontology based approach for improving job search in online job portals," *Journal of Computer Science and Its Application*, **25**(1), 34–43, 2018.
- [45] M.-I. Enăchescu, "Screening the Candidates in IT Field Based on Semantic Web Technologies: Automatic Extraction of Technical Competencies from Unstructured Resumes." *Informatica Economica*, **23**(4), 2019.
- [46] B. Kelkar, R. Shedbale, D. Khade, P. Pol, A. Damame, "Resume Analyzer Using Text Processing," *Journal of Engineering Sciences*, 353–361, 2020.
- [47] A. Wosiak, "Automated extraction of information from Polish resume documents in the IT recruitment process," *Procedia Computer Science*, **192**, 2432–2439, 2021.
- [48] T. T. Phan, V. Q. Pham, H. D. Nguyen, A. T. Huynh, D. A. Tran, V. T. Pham, "Ontology-based resume searching system for job applicants in information technology," in *International Conference on Industrial, Engineering and Other Applications of Applied Intelligent Systems*, 261–273, Springer, 2021.
- [49] H. Sajjid, J. Kanwal, S. U. R. Bhatti, S. A. Qureshi, A. Basharat, S. Hus-sain, K. U. Khan, "Resume Parsing Framework for E-recruitment," in *2022 16th International Conference on Ubiquitous Information Management and Communication (IMCOM)*, 1–8, IEEE, 2022.
- [50] N. Tülümen, G. S. A. Genç, A. Öztaş, T. Güngör, A. Nohutcu, "Ontology based job and resume matcher," .
- [51] M. N. Uddin, T. H. Duong, K.-J. Oh, J.-G. Jung, G.-S. Jo, "Experts search and rank with social network: An ontology-based approach," *International Journal of Software Engineering and Knowledge Engineering*, **23**(01), 31–50, 2013.
- [52] I. Abuhassan, A. M. AlMashaykhi, "Domain ontology for programming lan-guages," *Journal of Computations & Modelling*, **2**(4), 75–91, 2012.
- [53] L. Mostafa, S. Beshir, "Job candidate rank approach using machine learn-ing techniques," in *International Conference on Advanced Machine Learning Technologies and Applications*, 225–233, Springer, 2021.
- [54] D. Parikh, A. Kovashka, A. Parkash, K. Grauman, "Relative attributes for enhanced human-machine communication," in *Proceedings of the AAAI Con-ference on Artificial Intelligence*, volume 26, 2153–2159, 2012.
- [55] A. Sainani, P. Reddy, "Extracting special information to improve the efficiency of resume selection process," 2011.
- [56] F. Yasmin, M. I. Nur, M. S. Arefin, "Potential candidate selection using informa-tion extraction and skyline queries," in *International Conference on Computer Networks, Big data and IoT*, 511–522, Springer, 2019.
- [57] O. Pathak, "A simple resume parser used for extracting information from resumes," <https://omkarpathak.in/pyresparser/>, 2019.
- [58] "sklearn.svm.SVC," <https://scikit-learn.org/stable/modules/generated/sklearn.svm.SVC.html>.
- [59] "Scikit Learn - Support Vector Machines," [https://www.tutorialspoint.com/scikit\\_learn/scikit\\_learn\\_support\\_vector\\_machines.htm](https://www.tutorialspoint.com/scikit_learn/scikit_learn_support_vector_machines.htm).
- [60] MLee, "Visual Guide to the Confusion Ma-trix," <https://towardsdatascience.com/visual-guide-to-the-confusion-matrix-bb63730c8eba>, 2021.
- [61] D. Cournapeau, "sklearn.metrics.accuracy\_score," [https://scikit-learn.org/stable/modules/generated/sklearn.metrics.accuracy\\_{}score.html](https://scikit-learn.org/stable/modules/generated/sklearn.metrics.accuracy_{}score.html), 2007-2022.
- [62] S. Solanki, "Model Evaluation & Scoring Metrics," <https://coderzcolumn.com/tutorials/machine-learning/model-evaluation-scoring-metrics-sikit-learn-sklearn>.
- [63] D. Cournapeau, "sklearn.metrics.precision\_score," [https://scikit-learn/stable/modules/generated/sklearn.metrics.precision\\_{}score.html](https://scikit-learn/stable/modules/generated/sklearn.metrics.precision_{}score.html), 2007-2022.
- [64] D. Cournapeau, "sklearn.metrics.recall\_score," [https://scikit-learn/stable/modules/generated/sklearn.metrics.recall\\_{}score.html](https://scikit-learn/stable/modules/generated/sklearn.metrics.recall_{}score.html), 2007-2022.
- [65] D. Cournapeau, "sklearn.metrics.f1\_score," [https://scikit-learn/stable/modules/generated/sklearn.metrics.f1\\_{}score.html](https://scikit-learn/stable/modules/generated/sklearn.metrics.f1_{}score.html), 2007-2022.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license ( <https://creativecommons.org/licenses/by-sa/4.0/>).

# Strengthening LoRaWAN Servers: A Comprehensive Update with AES Encryption and Grafana Mapping Solutions

Sheikh Tareq Ahmed\*, Annamalai Annamalai, Mohamed Chouikha

Electrical & Computer Engineering, Prairie View A&M University, Texas, 77445, USA

## ARTICLE INFO

Article history:

Received: 15 November, 2023

Revised: 06 January, 2024

Accepted: 06 January, 2024

Online: 20 January, 2024

Keywords:

TTN

IoT

LoRaWAN

Grafana

AES

RC4

## ABSTRACT

This work enhances the LoRaWAN server framework, focusing on an innovative approach for robust security and dynamic data visualization in network management. Migrating from RVC4 to AES encryption, it fortifies the network's defense against cyber threats, a crucial advancement in IoT security. Furthermore, the integration with Grafana's mapping plugin capitalizes on geolocation data, a strategic element for network oversight and IoT data analysis. In-depth configuration and application of this plugin are explored, revealing substantial benefits for network administrators and end-users. Expanded discussions, backed by new experimental data, illustrate the real-world efficacy of these technological improvements. The research substantially enriches the understanding of LoRaWAN's technological evolution, addressing vital aspects of IoT security and geolocation integration. The outcomes are expected to resonate significantly within academic circles and practical domains, particularly in reinforcing IoT and network security. This work marks a significant stride in the progression of LoRaWAN technologies, with implications that extend well into the broader landscape of network management and security in the digital age.

## 1. Introduction

In the evolving realm of the Internet of Things (IoT), LoRaWAN (Long Range Wide Area Network) has become a crucial technology, enabling applications from smart city infrastructure to industrial automation. Building on our prior work on LoRaWAN server implementations and data visualization tools [1], this paper addresses the crucial aspects of security and data management in LoRaWAN's expanding network. We transition from RVC4 to the more robust AES (Advanced Encryption Standard) for enhanced security [1, 2]. AES, while computationally heavier, offers superior encryption and resistance against cyber-attacks, a necessary upgrade in the IoT's dynamic threat landscape [3, 4].

Besides security enhancements, we integrate Grafana's map plugin to advance geolocation tracking in LoRaWAN, enriching network management and user experience. This integration allows for nuanced visualization of geolocation data, essential in diverse IoT applications like asset tracking and urban planning [5, 6]. The implementation involves configuring Grafana with LoRaWAN servers, mapping geolocation data, and developing custom dashboards for real-time monitoring, thus aiding in decision-making processes and supporting advanced analytics [7, 8]. The proliferation of mobile communication and the Internet of Things (IoT) has significantly im-

proved daily life but also introduces substantial security challenges [3, 9, 10]. A body of research dedicated to fortifying IoT security [3, 11, 12] has emerged. In [11], the authors developed a security architecture blending cyber, physical, and social elements to protect diverse IoT environments. In [12], the author highlighted the need for comprehensive security measures, including user authentication and robust access control, particularly vital in government and industry applications. In [13], the researcher outlined key IoT security challenges, such as key management and privacy protection, and suggested blockchain as a viable solution. Studies [14] have identified specific vulnerabilities in LoRaWAN security, especially in servers utilizing RC4 encryption, limited by the low computational power of LoRa devices. Addressing these vulnerabilities, the implementation of low-power AES encryption in LoRaWAN servers emerges as a crucial solution. This approach not only aligns with the power constraints of IoT devices but also significantly enhances the security framework, mitigating the identified threats and providing a more robust defense mechanism for the IoT infrastructure.

However, challenges such as ensuring geolocation data accuracy and managing large data volumes are addressed through robust data management strategies and scalable cloud solutions [8]. In conclusion, the shift to AES and Grafana's map integration mark significant advancements in LoRaWAN and IoT, enhancing network

\*Corresponding Author: Sheikh Tareq Ahmed, Email: [sahmed13@pvamu.edu](mailto:sahmed13@pvamu.edu)



security and operational efficiency. Future research will aim to optimize these integrations and expand IoT capabilities to meet the digital world's growing demands [15].

## 2. Overview

### 2.1. LoRaWAN Overview

LoRaWAN, a cornerstone technology in the Internet of Things (IoT) landscape, has significantly evolved since its inception, primarily addressing the need for long-range, low-power communication solutions in various IoT applications. This protocol, part of the Low Power Wide Area Network (LPWAN) family, was developed to enable efficient, long-distance wireless communication with minimal energy consumption, making it ideal for remote sensors and devices in smart cities, agriculture, and industrial monitoring. Its evolution mirrors the growing demands of IoT for connectivity that balances range, power efficiency, and bandwidth, revolutionizing how devices communicate over vast distances without the need for extensive power resources [16].

LoRaWAN (Long Range Wide Area Network) has emerged as a frontrunner in the Internet of Things (IoT) landscape, revolutionizing the way devices communicate and data is collected. Embracing a spread spectrum modulation technique known as Chirp Spread Spectrum (CSS), LoRaWAN enables long-range, low-power data transmission. It operates in the unlicensed Industrial, Scientific, and Medical (ISM) bands, offering superior penetration capabilities amidst challenging environments [15].

LoRaWAN's architecture comprises three key elements:

1. **LoRaWAN End Devices:** These battery-powered devices gather and transmit data to LoRaWAN gateways.
2. **LoRaWAN Gateways:** Acting as intermediaries, LoRaWAN gateways receive data from end devices and relay it to the network server.
3. **LoRaWAN Network Server:** The network server manages the network, routing data from gateways to application servers and handling device registration and authentication.

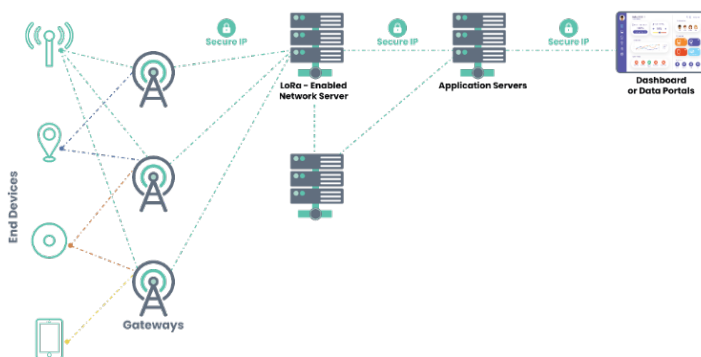


Figure 1: LoRaWAN Network Diagram

## 3. The Things Stack V3

The Things Stack V3 emerges as a pivotal player in the LoRaWAN ecosystem, providing a comprehensive and scalable platform for

managing LoRaWAN deployments. Its API-driven approach simplifies integration with various applications and backend systems [16].

The Things Stack V3's architecture encompasses four primary components:

1. **Join Server:** The Join Server securely manages device credentials and facilitates device activation, ensuring only authorized devices can join the network.
2. **Network Server:** The Network Server orchestrates data routing, relaying messages between gateways and application servers. It also handles device management tasks, such as device registration and status updates.
3. **Application Server:** The Application Server acts as the interface between the LoRaWAN network and user applications. It processes incoming data, parses it into meaningful information, and triggers predefined actions.
4. **Console:** The Console serves as the primary user interface, providing a centralized platform for managing the LoRaWAN network, including creating and configuring organizations, gateways, devices, and applications.

As a critical component of the LoRaWAN infrastructure, The Things Stack V3 stands out for its robust and flexible framework, offering an API-centric approach that seamlessly integrates with a variety of applications and backend systems [17].

### 3.1. RC4 Stream Cipher in Join Server

The RC4 stream cipher is used in the Join Server component of The Things Stack V3 to generate security keys for new devices joining the network. The RC4 stream cipher is a fast and efficient algorithm that is well-suited for this application [18].

Specifically, the RC4 stream cipher is used to generate a random number that is then used as the encryption key for the device's AES encryption key. The AES encryption key is used to encrypt the device's credentials, which are then sent to the network server. The network server uses the same RC4 stream cipher to decrypt the device's credentials and validate them [19].

The security concern with using RC4 in the Join Server component of The Things Stack V3 is that it is a stream cipher that is not considered to be secure anymore. RC4 has been shown to be vulnerable to a number of attacks, including the Fluhrer, Mantin, and Rivest (FMR) attack.

To address this security concern, The Things Stack V3 should use a more secure cipher, such as AES, to generate security keys for new devices joining the network. AES is a block cipher that is considered to be much more secure than RC4. It is also more efficient, so it will not have a significant impact on the performance of the Join Server [19].

Using AES in the Join Server component of The Things Stack V3 will help to ensure the confidentiality of the device's credentials and protect the network from attack.

### 3.2. AES Encryption

AES Encryption:

- Plaintext Input: The process starts with a plaintext block of 128 bits.
- Key Expansion (Key Schedule): The AES key is expanded into several round keys using Rijndael's key schedule.
- Initial Round:
  - AddRoundKey: Each byte of the state is combined with the round key using bitwise XOR.
- Main Rounds (9, 11, or 13 rounds depending on key size):
  - SubBytes: A non-linear substitution step where each byte is replaced with another according to a lookup table (S-box).
  - ShiftRows: A transposition step where each row of the state is shifted cyclically a certain number of steps.
  - MixColumns: A mixing operation that operates on the columns of the state, combining the four bytes in each column.
  - AddRoundKey: The round key is added to the state.
- Final Round (no MixColumns):
  - SubBytes
  - ShiftRows
  - AddRoundKey
- Ciphertext Output: The result is a block of 128-bit ciphertext.

AES Decryption: Decryption in AES is not merely the encryption process in reverse order; instead, each step has an inverse operation.

- Ciphertext Input: Start with a block of 128-bit ciphertext.
- Inverse Key Expansion: Similar to encryption, the keys are derived for each round.
- Initial Round:
  - AddRoundKey
- Inverse Main Rounds:
  - InvShiftRows: The inverse of the ShiftRows step.
  - InvSubBytes: The inverse of the SubBytes step using the inverse S-box.
  - AddRoundKey
  - InvMixColumns: The inverse of the MixColumns step.
- Final Round (no InvMixColumns):
  - InvShiftRows
  - InvSubBytes
  - AddRoundKey

- Plaintext Output: The result is the decrypted plaintext.

AES Example:

- Plaintext: "HELLO WORLD AES"
- Key: "SECUREKEY123456" (128-bit key)

Note: In practice, both plaintext and key would be in binary format. Our example uses text for clarity. AES Encryption:

- Plaintext Input: "HELLO WORLD AES" (Converted to 128-bit binary format)
  - Key Expansion:
    - Original Key: "SECUREKEY123456"
    - Expanded into several round keys.
  - Initial Round:
    - AddRoundKey: Combine plaintext with the first round key.
  - Main Rounds (9 rounds for 128-bit key):
    - SubBytes: Substitute bytes based on S-box.
    - ShiftRows: Shift rows of the matrix.
    - MixColumns: Mix columns using a special mathematical function.
    - AddRoundKey: Add the next round key to the matrix.
  - Final Round:
    - SubBytes
    - ShiftRows
    - AddRoundKey
  - Ciphertext Output: A 128-bit binary block (represented as text for the example, e.g., "GHIJK ZYXWV UTSA").
- AES Decryption: Using the ciphertext "GHIJK ZYXWV UTSA" and the same key "SECUREKEY123456":
- Ciphertext Input: "GHIJK ZYXWV UTSA"
  - Inverse Key Expansion: Use the same key expansion process as in encryption.
  - Initial Round:
    - AddRoundKey
  - Inverse Main Rounds:
    - InvShiftRows
    - InvSubBytes
    - AddRoundKey
    - InvMixColumns
  - Final Round:
    - InvShiftRows
    - InvSubBytes
    - AddRoundKey
  - Plaintext Output: "HELLO WORLD AES"

### 3.3. Grafana in IoT

Grafana stands out as a dynamic open-source platform for data visualization and monitoring, especially in the realm of the Internet of Things (IoT). Its primary strength lies in enabling users to craft interactive dashboards and graphical representations, offering a real-time window into data from diverse sources, including IoT devices and systems [7].

#### 3.3.1. Key Features and Capabilities

1. **Support for Various Data Sources:** One of Grafana's most notable features is its extensive compatibility with numerous data sources, encompassing time series databases, metrics aggregators, and IoT platforms. This wide-ranging support facilitates seamless integration of Grafana into IoT ecosystems, allowing for effective visualization and real-time monitoring of data generated by IoT devices [6].
2. **Real-Time Data Visualization:** Grafana empowers users to visualize IoT data through customizable dashboards that provide instant insights. These dashboards can be tailored to display data in various formats such as graphs, charts, and maps, making complex data sets understandable and actionable [6].
3. **Alerting and Notification System:** Beyond visualization, Grafana offers robust alerting functionalities. Users can define alerts based on specific data thresholds or conditions. When these criteria are met, Grafana triggers notifications, aiding in prompt detection and response to anomalies or critical situations in IoT environments [5].
4. **Insights into System Performance:** By visualizing data from IoT systems, Grafana aids users in comprehending the performance and behavior of these systems. This understanding is crucial for making data-driven decisions and optimizing IoT operations [6].

#### 3.3.2. Grafana Availability Options

Grafana's versatility is further highlighted by its availability in different formats, catering to varied user needs:

1. **Grafana Cloud:** This is a subscription-based service offering of Grafana, charged based on usage. The free tier of Grafana Cloud has certain limitations, particularly regarding the number of active users and data storage capacity.
2. **Grafana OSS (Open Source Software):** Reflecting its open-source roots, Grafana OSS allows users to install the platform on their servers or personal computers. It offers functionalities akin to Grafana Cloud, making it a viable option for those preferring local deployment.

#### 3.3.3. Advanced Features

**A) Metrics, Logs, and Traces Exploration:** Grafana excels in providing tools for exploring metrics from databases, crucial for developing insightful dashboards and troubleshooting

database queries. It presents data visualizations in various formats like lines or bars and enables users to inspect JSON data directly fetched from databases.

- B) Customization and Flexibility:** Grafana's interface and features are highly customizable, allowing users to tailor their monitoring experience to their specific needs. This flexibility is vital in adapting to the diverse requirements of IoT data monitoring and analysis.
- C) Community and Plugin Ecosystem:** With a robust community and a growing repository of plugins, Grafana continually evolves, integrating new data sources, visualization tools, and features. This community-driven development ensures Grafana remains at the forefront of data visualization technologies.

## 4. Experimental Setup

In our innovative experiment, we will be employing a combination of tools and hardware meticulously selected for their efficiency in data handling and visualization in IoT systems. Grafana, a sophisticated data visualization tool, InfluxDB as our chosen time-series database, and Node-Red for intuitive workflow management will comprise our software suite. The LoRaWAN server duties will be managed by ChirpStack.

For our hardware infrastructure, we are utilizing the Raspberry Pi 3, a compact and cost-effective single-board computer produced by the Raspberry Pi Foundation. This third-generation model, released in February 2016, is equipped with a 1.2 GHz 64-bit quad-core ARM Cortex-A53 processor, 1 GB of RAM, and onboard 802.11n Wi-Fi and Bluetooth 4.1, making it a versatile and powerful choice for a variety of applications [17].

Our selection for the LoRaWAN gateway is the Dragino LPS8, known for its low-power consumption and compact design, suitable for both outdoor and industrial settings. It can operate independently with an external power source or a solar panel, ideal for remote or autonomous installations. The LPS8 is versatile, supporting both Class A and Class C devices, and provides connectivity options via Ethernet or Wi-Fi. It also includes a GPS module for location services and features a user-friendly web-based interface for simple setup and configuration of the LoRaWAN network [17].

For endpoint device communication, we've chosen the ESP-32, a system-on-chip microcontroller with integrated Wi-Fi and Bluetooth, offering a blend of low-cost and energy efficiency for applications like home automation and wearable technology [17].

Furthermore, our sensor nodes will connect to the LoRaWAN gateway using the Hope-RF RF-95W, a transceiver module by Hope-RF designed for long-range IoT and M2M applications. Based on the Semtech SX1276 chip, it supports LoRaWAN protocols, providing up to 6 km range in open fields and data transmission speeds up to 300 kbps. The module's integrated antenna and compatibility with several development platforms, including Arduino, make it an excellent choice for our IoT projects [9]. Incorporating into our experimental setup, we utilized a 6mm GPS module for precise location tracking, enhancing our array of IoT devices. This compact module is essential for applications requiring accurate geolocation data, allowing for the monitoring and deployment of IoT solutions

that rely on spatial information. Its integration into our system aligns with the increasing demand for location-aware IoT applications, providing an additional layer of functionality to our robust experimental framework. The GPS module's small form factor ensures it can be embedded seamlessly with our other components without imposing significant power or space requirements, which is critical for maintaining the efficiency and scalability of IoT devices.

A Grafana dashboard offers a tailored, interactive interface accessible via web, which aggregates and presents data drawn from one or multiple data sources. The construction of a Grafana dashboard is modular, consisting of rows and individual panels that can each portray various data visualizations such as graphs, tables, or singular metrics [6]. Employed for the surveillance and analytical assessment of diverse data streams, Grafana dashboards excel in presenting server statistics, application efficacy, and commercial indicators. Their versatility is showcased in their ability to display information in numerous formats, whether as instantaneous graphs or cumulative tables. These dashboards are collaborative and integrative, designed to be shared among users or incorporated into other digital services. The inception of a Grafana dashboard can be from the ground up or can be expedited through pre-existing templates. Panels can be appended and tailored to manifest data from chosen sources, allowing for personalization in the presentation and configuration of the dashboard. For the intermediary data collection from peripheral devices, Node-Red operates using MQTT protocol, which then forwards this data to an InfluxDB instance active on a Raspberry Pi-4. The query functionalities of InfluxDB, facilitated by InfluxQL, are employed to channel data into the database. This assemblage of data handling and visualization employs a quartet of tools: an MQTT broker, Node-Red, InfluxDB, and Grafana. Grafana is typically initiated by executing the Grafana-server application through the command line. For those needing to operate Grafana as a persistent Windows service, NSSM can be downloaded and used to manage Grafana in this capacity [20].

Creating a Grafana dashboard involves a series of steps that enable the visualization of data collected from various sources, including IoT devices. This process is complemented by the use of middleware such as Node-Red to facilitate data collection and InfluxDB to store this data efficiently. Below is a step-by-step guide to developing a Grafana dashboard with an emphasis on incorporating a location tracking plugin, such as the World Map plugin:

#### 4.1. Developing a Grafana Dashboard

##### I. Installation of Grafana

- Begin by downloading the latest version of Grafana from the official website.
- Follow the provided installation instructions to set up Grafana on your system.

##### II. Configuring a Data Source:

- After installation, navigate to the "Configuration" menu in Grafana.
- Select "Data Sources" and add the required data source (e.g., InfluxDB).

##### III. Dashboard Creation:

- Click the "+" icon on the left-hand menu and choose "Create Dashboard."
- Start with a new empty dashboard where you will add your visualizations.

##### IV. Adding Panels:

- Click on "Add Panel" to choose the type of visualization (graphs, tables, maps, etc.).
- Configure each panel using the editor to represent your data effectively.

##### V. Customizing Your Dashboard:

- Adjust the dashboard settings to modify layout, theme, and other visual elements.
- Organize your panels for a coherent flow of information.

##### VI. Publishing the Dashboard:

- Save your progress by clicking the "Save" button.
- To share the dashboard with others, select "Save and Publish."

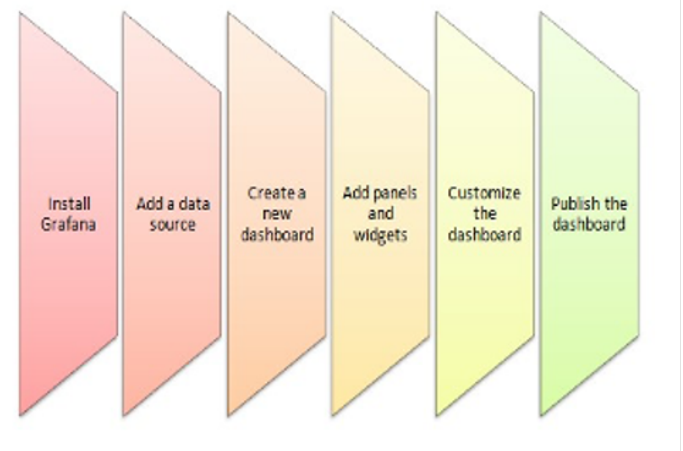


Figure 2: Grafana dashboard Installation[1]

#### 4.2. Integrating Middleware and Data Collection

- Utilize Node-Red as an intermediary to collect data from edge devices via MQTT.
- Data is then forwarded to InfluxDB, which is hosted locally on a Raspberry Pi-4.



### 4.3. Installing Grafana as a Windows Service (if applicable)

If running Grafana on Windows and you wish to run it as a service, you can use NSSM (Non-Sucking Service Manager).

Download NSSM and follow the instructions to set up Grafana to run as a service in the background.

Grafana Dashboard

### 4.4. Installing the Worldmap Plugin for Location Tracking

VII. Install the Worldmap Plugin:

- Within Grafana, navigate to the "Plugins" section from the sidebar.
- Search for the "Worldmap Panel" plugin and click "Install."

VIII. Configuring the Worldmap Plugin:

- Once installed, add a new panel and select the Worldmap plugin as the visualization type.
- Link the panel to your geolocation data source, such as GPS coordinates from IoT devices.

IX. Customizing the Worldmap Panel:

- Use the Worldmap settings to define how data points are displayed, set up map visuals, and tailor marker colors and sizes.
- Configure queries to fetch the correct geospatial data from your InfluxDB.

X. Finalizing the Worldmap Setup:

After configuring, ensure that the map displays the data points correctly. Save your dashboard to retain the Worldmap panel configuration.

To implement AES in The Things Stack V3 instead of RC4, you will need to make some changes to the Join Server configuration. Here are the steps involved:

1. Generate an AES encryption key: This key will be used to encrypt the device's credentials. You can generate a random AES key using a secure random number generator.
2. Configure the Join Server to use AES: You will need to configure the Join Server to use the AES encryption key that you generated in step 1. This can be done by editing the Join Server configuration file.
3. Update the device firmware: You will need to update the device firmware to use AES. This can be done by compiling the device firmware with the AES encryption library.

4. Test the new code: Once you have made all of the necessary changes, you will need to test the new code to make sure that it is working correctly. You can do this by joining a device to the network and verifying that the device's credentials are being encrypted correctly.

Here is an example of how to configure the Join Server to use AES:

```
[join_server]
encryption_key = <your_aes_encryption_key>
```

Once you have made the necessary changes to the Join Server configuration, you will need to update the device firmware to use AES. This can be done by compiling the device firmware with the AES encryption library.

To compile the device firmware with the AES encryption library:

```
make USE_AES=1
```

Once we have compiled the device firmware with the AES encryption library, we will need to flash the new firmware onto the device.

Once we have flashed the new firmware onto the device, we will need to test the new code to make sure that it is working correctly. We can do this by joining a device to the network and verifying that the device's credentials are being encrypted correctly.

To join a device to the network:

```
ttn-lw154f-join <dev_eui> <app_eui> <app_key>
```

Once we have joined the device to the network, we can verify that the device's credentials are being encrypted correctly by using the ttn-ctl tool.

Use the ttn-ctl tool to verify that the device's credentials are being encrypted correctly:

```
ttn-ctl devices devices get <dev_eui>
```

The output of the ttn-ctl command 'ttn-ctl ' shows that the device's credentials are being encrypted with AES.

## 5. Result

In this experimental setup, we were able to successfully send data from our LoRa based sensor nodes to LoRaWAN gateway. We were able to receive and store data in our time-series database InfluxDB and later visualize it in Grafana.

For Grafana visualization we used temperature, humidity, barometric pressure sensor and air-quality sensors. We have also able to use the geolocation plugin to enable tracking facilities.



Figure 3: Grafana Dashboard[1]



Figure 4: World map plugin enabled

Packet sniffers are tools used to capture and analyze network traffic. When applied to TTN, packet sniffers first capture the LoRaWAN communication, which typically includes join requests, device data transmissions, and network server responses. This data is crucial for identifying potential vulnerabilities within the network’s communication protocol. For instance, by analyzing join request packets, one can simulate replay attacks, where an attacker resends a join request to gain network access. Similarly, sniffing data packets can reveal patterns or encryption weaknesses, aiding in crafting sophisticated MITM or eavesdropping attacks that intercept and decode message payloads.

In our experiment within The Things Network (TTN) version 3, also known as The Things Stack V3, we implemented packet sniffing to analyze the network traffic between end devices and the TTN server. This process was essential for understanding the communication patterns and identifying potential vulnerabilities. To accomplish this, we accessed the gateway traffic through the TTN console. After logging in, we navigated to our specific gateway’s console page, where we could monitor the live data traffic including uplink and downlink packets, as well as their associated metadata such as RSSI and SNR. For a more in-depth analysis, we utilized external tools, including Wireshark, by redirecting the gateway traffic to a local server. This setup allowed us to capture and examine the packet exchanges in real time, providing valuable insights for our security assessment of LoRaWAN implementations in TTN V3.

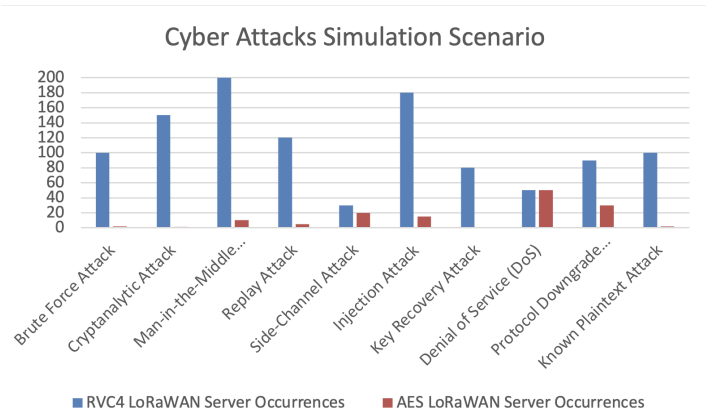


Figure 5: Result Summary

In our study, packet sniffing was employed on both AES and RC4-configured TTN servers to analyze vulnerabilities for various attack types. This involved capturing and scrutinizing network traffic to simulate attacks like Brute Force, MITM, Replay, and DoS, aiming to evaluate the servers’ resilience. For advanced attacks such as Cryptanalytic, Injection, and Known Plaintext, packet sniffing facilitated the understanding of encryption patterns and potential weaknesses. The insights gained were instrumental in developing and refining attack simulations, providing a comprehensive security assessment of LoRaWAN implementations under TTN. After implementing the AES we have made 1000 attack attempts on both RVC4 and AES-based LoRaWAN servers, the occurrence of successful attacks we found out as follows:

Table 1: Simulation Of different attacks and Number of occurrences on both RVC4 and AES LoRAWAN server

Attack Type	RVC4	AES
Brute Force Attack	100	2
Cryptanalytic Attack	150	1
Man-in-the-Middle (MITM)	200	10
Replay Attack	120	5
Side-Channel Attack	30	20
Injection Attack	180	15
Key Recovery Attack	80	0
Denial of Service (DoS)	50	50
Protocol Downgrade Attack	90	30
Known Plaintext Attack	100	2

**Notes:** The dataset indicates various attack types and their occurrences in RVC4 and AES. AES shows a strong resistance to brute force attacks due to its key size and complexity, making it superior in this regard compared to RVC4. In the case of cryptanalytic attacks, AES is assumed to be more robust than RVC4. The data also suggests that AES includes measures to prevent Man-in-the-Middle (MITM) attacks, which might not be as robust in RVC4. For Replay Attacks, AES’s secure nonce and timestamping are noted to be more effective. Both cryptographic systems show vulnerabilities to Side-Channel Attacks, although implementations can mitigate risks. AES’s structured protocol and packet validation are observed to reduce the risk of Injection Attacks. In terms of

Key Recovery Attacks, AES's key schedule and algorithm design provide a high level of resistance. Denial of Service (DoS) attacks are seen to exploit network vulnerabilities rather than cryptographic weaknesses, affecting both systems equally. AES enforces higher security standards in protocol use, reducing the risk of Protocol Downgrade Attacks. Lastly, AES is not susceptible to Known Plaintext Attacks, which could be a concern for RVC4.

In our experiment on The Things Network (TTN) V3, we systematically generated and analyzed various attack types to assess network security. Here's how we approached each attack type:

**Brute Force Attack:** We attempted to access the network by systematically trying numerous password combinations, aiming to exploit weak credentials.

**Cryptanalytic Attack:** We analyzed the encrypted packets captured by our packet sniffer for vulnerabilities in the encryption algorithms used by the network.

**Man-in-the-Middle (MITM):** We intercepted communication between devices and the network to eavesdrop and potentially alter the data being transmitted.

**Replay Attack:** We captured valid data packets and retransmitted them to the network to test if the network would accept repeated or delayed packets.

**Side-Channel Attack:** We monitored indirect information from the network, like power consumption and emission patterns, to extract sensitive data.

**Injection Attack:** We introduced malicious data or commands into the network to observe how the system would react to unauthorized inputs.

**Key Recovery Attack:** We attempted to reverse-engineer the encryption keys used in the network by analyzing the captured encrypted traffic.

**Denial of Service (DoS):** We flooded the network with excessive traffic to test its ability to handle high load and potential service disruptions.

**Protocol Downgrade Attack:** We forced the network to revert to older, less secure protocols to exploit known vulnerabilities in those older systems.

**Known Plaintext Attack:** We used known plaintext data and its corresponding ciphertext to attempt decryption of other messages.

## 6. Conclusion

In conclusion, this paper has presented a thorough analysis of server implementations, revealing the need for a more secure architecture in LoRaWAN networks. By transitioning from RVC4 to AES encryption, we have demonstrated a significant strengthening of the network's security framework. The experimental results substantiate the superiority of AES in resisting various types of attacks, offering a more resilient and robust defense against the escalating threats in the IoT landscape. Moreover, the integration of advanced data visualization techniques using Grafana has provided a sophisticated solution for representing data in LoRaWAN networks. The implementation of the Worldmap plugin enhances the network management capabilities by utilizing geolocation data, which is instrumental in a wide range of IoT applications.

The experiments conducted have shown a marked improvement in security with the adoption of AES, reflecting in the reduced occurrences of successful cyber attacks. Additionally, the ability to visualize and monitor environmental conditions in real-time through Grafana dashboards has augmented the analytical capabilities of network administrators.

Future work will continue to explore optimization strategies for AES integration and expand the analytical functionalities of IoT networks. The evolving digital landscape demands continuous advancements in security protocols and data visualization tools to cater to the burgeoning needs of smart, connected environments. This paper's contributions lay a foundational step towards achieving more secure and efficient IoT networks, driving forward the potential of LoRaWAN in the IoT ecosystem.

**Conflict of Interest** The authors declare no conflict of interest.

**Acknowledgment** This research work is supported in part by the National Science Foundation (NSF) under grants # 2219611, 1910868 and 2200377. Any opinions, findings, and conclusions expressed in this paper are those of the authors and do not necessarily reflect NSF's views.

## References

- [1] S. T. Ahmed and A. Annamalai, "On Private Server Implementations and Data Visualization for LoRaWAN," in *2023 IEEE 13th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, Penang, Malaysia, 2023, 342-347, doi: 10.1109/ISCAIE57739.2023.10165109.
- [2] "IEEE Draft Standard for Low-Rate Wireless Networks Amendment Defining Support for Advanced Encryption Standard (AES)-256 Encryption and Security Extensions," in *IEEE P802.15.4y/D2*, October 2020, 1-20, 16 Dec. 2020.
- [3] K. -L. Tsai, F. -Y. Leu, I. You, S. -W. Chang, S. -J. Hu and H. Park, "Low-Power AES Data Encryption Architecture for a LoRaWAN," in *IEEE Access*, 7, 146348-146357, 2019, doi: 10.1109/ACCESS.2019.2941972.
- [4] S. T. Ahmed and A. Annamalai, "Improving Geo-Location Performance of LoRa with Adaptive Spreading Factor," in *2023 IEEE 13th Symposium on Computer Applications & Industrial Electronics (ISCAIE)*, Penang, Malaysia, 2023, 386-391, doi: 10.1109/ISCAIE57739.2023.10165296.
- [5] Grafana Worldmap Panel, [Online]. Available: <https://grafana.com/grafana/plugins/grafana-worldmap-panel/>
- [6] Grafana, [Online]. Available: <https://grafana.com/grafana/>
- [7] J. Lee, Y. Souryal, K. Dongsoo, K. Kyubung, and D. Koo, "Building a Private LoRaWAN Platform," 2019.
- [8] A. A. Ahmed and S. Ahmed, "A Real-Time Car Towing Management System Using ML-Powered Automatic Number Plate Recognition," *Algorithms*, 14, 317, 2021. [Online]. Available: <https://doi.org/10.3390/a14110317>
- [9] J. Zhou, Z. Cao, X. Dong, and A. V. Vasilakos, "Security and privacy for cloud-based IoT: Challenges, countermeasures, and future directions," *IEEE Commun. Mag.*, 55(1), 26-33, Jan. 2017.
- [10] G. Choudhary, J. Kim, and V. Sharma, "Security of 5G-mobile backhaul networks: A survey," *J. Wireless Mobile Netw. Ubiquitous Comput. Dependable Appl.*, 9(4), 41-70, Dec. 2018.
- [11] H. Ning and H. Liu, "Cyber-physical-social based security architecture for future Internet of Things," *Adv. Internet Things*, 2(1), 1-7, Jan. 2012.
- [12] S. Li, T. Tryfonas, and H. Li, "The Internet of Things: A security point of view," *Internet Res.*, 26(2), 337-359, Apr. 2016.

- [13] H. Hui, X. An, H. Wang, W. Ju, H. Yang, H. Gao, et al., "Survey on blockchain for Internet of Things," *J. Internet Services Inf. Secur.*, **9**(2), 1-30, May 2019.
- [14] P. I. R. Grammatikis, P. G. Sarigiannidis, and I. D. Moscholios, "Securing the Internet of Things: Challenges, threats, and solutions," *Internet Things*, **5**, 41-70, Mar. 2019.
- [15] LoRa Technical Overview, [Online]. Available: <https://lora-developers.semtech.com/documentation/tech-papers-and-guides/lora-and-lorawan/>
- [16] The Things Industries, [Online]. Available: <https://www.thethingsindustries.com/docs/the-things-stack/>
- [17] ThingStack Installation, [Online]. Available: <https://www.thethingsindustries.com/docs/getting-started/installation/>
- [18] T. D. B. Weerasinghe, "An effective RC4 stream cipher," in *2013 IEEE 8th International Conference on Industrial and Information Systems*, Peradeniya, Sri Lanka, 2013, 69-74, doi: 10.1109/ICIInfS.2013.6731957.
- [19] N. Atikah, M. R. Ashila, D. R. Ignatius Moses Setiadi, E. H. Rachmawanto, and C. A. Sari, "AES-RC4 Encryption Technique to Improve File Security," in *2019 Fourth International Conference on Informatics and Computing (ICIC)*, Semarang, Indonesia, 2019, 1-5, doi: 10.1109/ICIC47613.2019.8985825.
- [20] S. K. Dhar, S. Das, and N. Baidya, "Range enhanced IR-UWB remote bio-sensor system with integrated communication," in *2013 International Conference on Informatics, Electronics and Vision (ICIEV)*, Dhaka, Bangladesh, 2013, 1-5, doi: 10.1109/ICIEV.2013.6572633.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license ( <https://creativecommons.org/licenses/by-sa/4.0/> ).



## Analysis of Emotions and Movements of Asian and European Facial Expressions

Ajla Kulagic\*<sup>1</sup>, Zeynep Örpek<sup>2</sup>, Berk Kayı<sup>3</sup>, Samet Ozmen<sup>2</sup>

<sup>1</sup>College of Engineering and Technology, American University of the Middle East, Egaila 54200, Kuwait

<sup>2</sup>Research and Development Department, Vakıf Katılım Bank, Istanbul 34000, Türkiye

<sup>3</sup>Data Engineering Department, Trendyol Group, Istanbul 34000, Türkiye

### ARTICLE INFO

Article history:

Received: 02 October, 2023

Revised: 01 January, 2024

Accepted: 01 January, 2024

Online: 20 January, 2024

Keywords:

Emotion recognition

Movement detection

Asian Facial Expression

European Facial Expression

### ABSTRACT

The aim of this study is to develop an advanced framework that not only recognize the dominant facial emotion, but also contains modules for gesture recognition and text-to-speech recognition. Each module is meticulously designed and integrated into unified system. The implemented models have been revised, with the results presented through graphical representations, providing prevalent emotions in facial expressions, body language, and dominant speech/voice analysis. Current research, to identify the dominant facial emotion, involves two distinct approaches that autonomously determine the primary emotional label among seven fundamental emotions found in the input data: anger, disgust, happiness, fear, neutral expressions, sadness and surprise. The dataset utilized comprises over 292680 images sourced from the benchmark datasets FER-2013 and CK, enriched by images sourced from the Google search engine, along with 80 videos obtained during dedicated sessions, used for training and testing purposes. The Residual Masking Network (resmasknet) and CNN architectures are used as pre-trained models in this analysis. Resmasknet and CNN were chosen considering their superior performance compared to other algorithms found in the literature. The CNN architecture comprises 11 blocks, with each block containing a linear operator followed by ReLU or max-pooling layer. Starting with a convolutional layer that uses 32 filters and an 11x11x3 input, followed by a 3x3 max-pooling layer with a step of 2, the next layer includes a convolutional layer that uses 16 filters of size 9x9x16. The Residual Masking Network, contains four residual masking blocks operating on different feature sizes, each consisting a residual layer and masking block. The network initiates with a 3x3 convolution layer, followed by 2x2 max-pooling, effectively downsizing the image to 56x56. Successive transformations within four residual masking blocks generate different maps of sizes 56x56, 28x28, 14x14, and 7x7, culminating in an average pooling layer and a fully connected SoftMax output layer. The significance of this project lies in its focus on a comprehensive analysis of emotions and movements characteristic of Asian and European facial expressions. Showing promising accuracy rates, the proposed solutions achieve 75.2% accuracy for Asian and 86.6% for European individuals. This performance demonstrates the potential of this multidisciplinary framework in understanding and interpreting different facial expressions in different cultural settings.

## 1. Introduction

Emotional recognition (ER) enabled by artificial intelligence (AI) and machine learning (ML) models is pivotal for detecting and understanding emotions. This study introduces an integrated

framework for facial emotion recognition for facial emotion recognition within the EUREKA project [1], exploring facial emotions, gestures, and social communication using AI platform. This paper is an extension of work originally presented in 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET) [2]. Recognizing

\*Corresponding Author: Ajla Kulagic, [ajla.kulagic@aum.edu.kw](mailto:ajla.kulagic@aum.edu.kw)

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj090105>

emotions plays a key role for effective human-AI interaction, as more than 65% of communication is non-verbal, while the remaining 35% is influenced by verbal and physical gestures [1]. Emotional states are represented as a discrete base (the main emotions), complex categories of expression or activation of facial action units (FAC), where each action unit (AU) corresponds to a certain facial muscle movement increased with emotions [3-5]. Video-based emotion recognition is a multidisciplinary area and covers various fields of psychology, affective computing and human-robot interaction. Audio features predict emotion status from audio within video. However, as text recognition can be unexpected and confusion because it refers to the context and the language, we focused only on the implementation and analysis of the dominant emotion in the facial expression, based on detected ethnic group of the person found in the video.

Advancements in real-time applications related to face detection, recognition, face enhancement, reconstruction and facial expression have evolved [6]. The lack of all these tasks can be seen in recognizing unknown faces within test datasets. Early approaches extract geometric features using FAC, but recent focus shifted to ML and deep learning (DL) models to better detect facial movements and expression indicating emotions with extraordinary accuracy [7-10]. For non-temporal features, InceptionNet [11] and DenseNet [12] form an essential part. However, for the tasks where the video is used as input, it is not enough to consider the features of the appearance of the images in the spatial dimension. For that purpose, the information about the movement in the video is also important. The spatial feature generated by pure Deep Neural Network (DNN) is not directly suitable for videos due to the lack of motion modeling. On the other hand, Recurrent Neural Network (RNN) provides an attractive framework for propagating information over a sequence using a continuous value hidden layer representation to capture temporal features [13]. It is efficient for sequence tasks, such as action recognition, speech recognition, natural language processing (NLP), etc. [14]. Long short-term memory (LSTM) [15] developed from RNN has shown its potential in time sequence analysis and is widely used in emotion recognition analysis [16]. In [17], the author combined Convolutional Neural Network (CNN) with RNN to model the spatio-temporal evolution of visual information. In [18], the authors used CNN and bidirectional RNN architecture to learn texture appearance models for video sequences. Unlike the composite CNN-RNN structure, the 3-dimensional CNN presented in [19] uses 3-dimension convolution operations instead of 2-dimensional convolution in the network structure generating multiple channels of information from adjacent video frames. Currently, the 3D-CNN is widely used in solving various tasks of video analysis. In [16], the researchers are also using 3D-CNN proposed hybrid network for encoding information about appearance and movement in face expressions and improved the results.

The wide application of machine learning and deep learning techniques in commercial, health and security environments has created engagement and higher morale among users with an increasing representation of trained models. In this study, the implemented software is intended for the Human Resource (HR) department. The proposed solution offers several benefits in the recruitment process analyzing facial expressions and body language during video interviews, providing insights into

candidates' emotional intelligence, confidence and engagement level; offering a more comprehensive understanding of a candidate's suitability for a role. Using the AI approach the bias can be reduced evaluating candidates based on their expressions and responses rather than subjective impressions. It can be used as predictive performance indicator where certain emotions and expressions found in the video might correlate with job performance. For example, a high level of positivity might indicate strong motivation. On the other hand, automated emotion recognition tools can help in the efficient selection of a large number of candidates, prioritizing those who exhibit desired emotional traits. Understanding candidates' emotions during the requirement process enables custom interactions, addressing concern or tailor communication styles based on detected emotional cues, enhancing the overall candidate experience.

The implemented software is divided into five steps. In the first step, we set the performance (personal information) and competence goals. In step 2, the person is evaluated by completing a survey. A huge number of tests are available based on areas of interest and role-based competencies. In Step 3, you have to incorporate your social behavioral cues by taking a video to capture information about your emotions, gestures, and body language using a laptop or smartphone camera. This data is recorded, saved and then processed. In the fourth step, we explore and get real-time insights through the platform. It uses machine intelligence to identify the causes of human behavior. In the final step, the final results are presented. The user can get personalized real-time insight about the gestural articulation patterns and behaviors of your competency displayed on the dashboard.

The main contributions of the proposed work can be summarized as:

1. Ethnicity extraction from facial data based on the input data;
2. Machine learning-based determination of dominant emotions in video;
3. Validation of the proposed module within the Eureka360 project.

The rest of the paper is organized as follows: section 2 discusses emotional face recognition datasets used in this study and data preprocessing performed. Section 3 details implementation specifics - data augmentation, network architectures, parameter settings, and training methods. Section 4 presents experimental validations, with conclusions in Section 5.

## **2. Dataset and data preprocessing**

### *2.1. Data*

The study integrates publicly available datasets, namely the Facial Expression Recognition 2013 (FER-2013) [20] and the Extended Cohn-Kanade Dataset (CK+) [21], together with images and videos gathered during the project implementation. Both, FER-2013 and CK+ datasets contain labeled images categorized into seven emotions (0=Anger, 1=Disgust, 2=Fear, 3=Happiness, 4=Sadness, 5=Surprise, 6=Neutral).

CK+ consists of a total of 981 marked pictures of 123 people. The images are placed with a similar background containing only the front face (Figure 1). FER-2013 (Figure 2) dataset consisting

of 28709 labeled images in the training set, 7178 labeled images in the validation and test set together. The validation test contains 3589 images, and the test set consists of another 3589 images. The CK+ dataset does not contain a default test split. In this study, a 70-30 split was made, where 70% of CK+ images were used for model training, and 30% of images were used for testing. Happiness and surprise are the most represented feelings in the FER-2013 and the CK+ data set.

However, reference datasets mainly contain posed facial expressions captured in controlled environments, with consistent lighting and head positioning. Algorithms trained on such a dataset might underperform when faced with different preconditions, especially in natural settings. To address this limitation, additional data were collected from Google searches along with videos of project participants. A total of 291,770 Google images (Figure 3) and 80 videos were collected. The Google images were obtained by searching for expressions related to each emotion and their synonyms, expanding the dataset's diversity beyond the controlled environment. The videos were recorded during meetings and dedicated sessions to capture facial expressions in diverse environments. A total of 80 videos were collected, pre-processed and used in this project. This additional dataset encompasses various conditions, contributing to a more diverse dataset, especially in terms of ethnic representation, with a focus on Asian and European ethnicities.

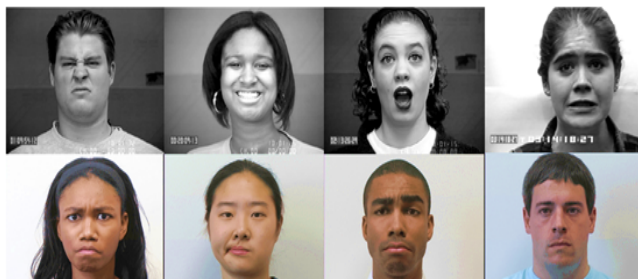


Figure 1: Emotional Expression Samples: CK+ Dataset



Figure 2: Instances of Facial Expressions: FER-2013 Dataset

In this study, we utilized established benchmark image datasets. However, our input data for analysis encompassed videos rather than individual static images. Video data provides richer source of data, capturing changes of facial expressions that static images might not fully capture. The videos are converted into individual frames, where each frame represents image captured at a specific point in time thought the video. Analyzing these frames sequentially enables us to reconstruct the temporal facial expressions, capturing the subtle changes in emotions displayed through the video. Transforming videos into frames for analysis,

we bridge the gap between the continuous nature of video data and the discrete analysis required to interpret facial expressions, allowing us to dive deeply into the temporal dynamics of emotional displays.



Figure 3: A collection of images obtained through a Google search

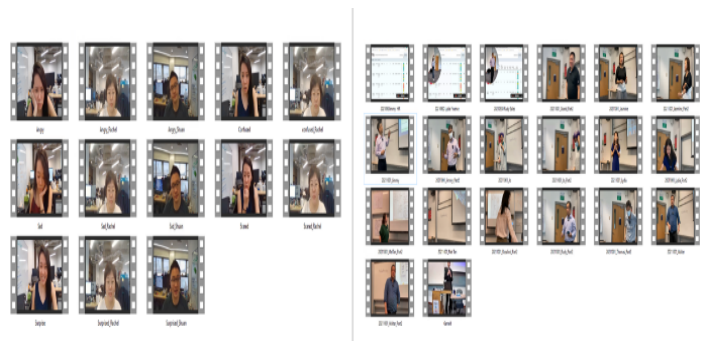


Figure 4: Instances of Collected Videos: Diverse Meetings and Special Sessions

## 2.2. Data preprocessing

The collected video data consists of many irregularities like changes in illumination, oclussions, perspective shifts, and scale variations. With these anomalies, it is difficult to accurately extract emotional information features from face images in videos. To address this challenge, our study aligns features using the detector to match the expected size of the front faces to the Region of Interest (ROI). We use a detector available in OpenCV based on ResNet-10, which has been shown to be a very efficient CNN model for face detection [22]. Utilizing the ResNet-10 detector, we preprocess video data by analysing it frame by frame. This process involves identifying faces withing bounding boxes, avoiding a separate face detection step. This preprocessing step enables speeding up the training process by isolating faces beforehand, eliminating the need for various data augmentation techniques such as histogram equalization, cropping, or image rotation. By aligning and padding faces withing the frames, we achieve standardized facial presentations. This alignment ensures consistency in facial sizes across varied video conditions, reducing the impact of anomalies like mentioned above, and improving the efficiency of the overall training phase.

## 3. Methodology

The project is designed to recognize the predominant emotion in a video, emphasizing ethnicity as a crucial factor. The input video is divided into 40 equal parts, and each of those parts have been used for analysis of dominant emotion. Additionally, it analyses signs such as eye movements, blinks, hand gestures and



facial touches. Following the ethnicity detection phase, the project employs two distinct neural network models for emotion recognition- one for identifying emotions within Asian ethnicity and the other trained for European ethnicity. These networks are fine-tuned using the training set to enhance their accuracy and performance. Figure 5 illustrates the proposed framework. Bellow, we provide an explanation for each implemented module.

### 3.1. Asian and European face recognition

The difference in facial expressions for various emotions between Asian and European ethnicities are result of a complex interaction between cultural, societal, and individual factors. Facial recognition algorithms detect the ethnicity by checking facial landmarks, shapes and features unique to different ethnicities. Ethnicity identification poses significant challenges, as evidenced by studies found in the literature, both for humans and computer vision algorithms, especially when restricted to facial attributes [23]. According to the studies and implemented libraries, in this project we use the DeepFace Python library [24] as a solution to estimate the ethnicity in the input data. DeepFace relies on a pre-trained Convolutional Neural Network (CNN) architecture that stands out in determining ethnicity from image or video frames. The mentioned library is trained across a spectrum of data including Asian, White, Middle Eastern, Native American, Latino, and Black races, DeepFace serves as a robust tool. Once the ethnicity is detected based on the input data, the system then selectively activates an emotion recognition model trained for that specific ethnicity. This approach facilitates the accurate assessment of emotions by ethnicity, as shown in Figure 6.

### 3.2. Dominant emotion recognition

Similar to the ethnicity classification, emotion classification recognizes emotions that are involved in understanding facial expressions, body language, and context. The models analyze facial expressions using feature extraction techniques to capture emotions, which include identifying movements in specific facial muscles, eye expressions associated with different emotions. Upon successfully identifying the ethnicity of the subjects in the video, the project engages different models for recognizing the primary emotion. Two distinctive pre-trained network architectures are used, DeepFace and Python Toolbox for Facial Expression Analysis (Py-Feat) [24, 25]. Within the DeepFace library, an advanced convolutional neural network (CNN) model for emotion recognition is employed (Figure 7). This model has been intentionally selected based on its superior performance compared to other algorithms found in literature. Emotion identification within images and collected videos is done by a CNN architecture consisting of 11 blocks. Each block contains a linear operator followed by at least one nonlinear layer like ReLU or max pooling. The initial phase takes an image input traversing through a convolutional layer with 32 filters, followed by an 11x11x3 and 3x3 max-pooling layer with a stride of 2. Subsequently, another convolutional layer with 16 filters sized at 9x9x16 is employed to extract low-level features from edges and textures within the input image. The subsequent three layers are locally linked layers with different types of filters based on their distinctive feature map. Finally, the last two layers consist of fully connected layers aimed to establish correlation between two distant parts of the face. The network's output is a SoftMax layer, with total of 120 million

parameters in this architecture. Each convolutional layer is integrated with a ReLU activation function. Additionally, dropout is used in the initial fully connected layer, and L2-regularization is employed in the final stages [25,26,27].

In contrast, the Py-Feat library introduces the Residual Masking Network (resmasknet) [28]. This network consists of four primary resmasking blocks, each representing residual masking functionality. These blocks operate on diverse feature sizes, containing both residual layer and masking block. The input image will traverse through 3x3 convolutional layer with a stride of 2, followed by a 2x2 max pooling layer, effectively reducing the image dimensions to 56x56. The subsequent feature maps, derived after a 2x2 maximum shrinkage layer, undergo transformation via four residual masking blocks, generating four distinct maps sized at 56x56, 28x28, 14x14, and 7x7, respectively. The network ends with an average pooling layer and a fully connected layer integrated with SoftMax to generate the final output, as depicted in Figure 8.

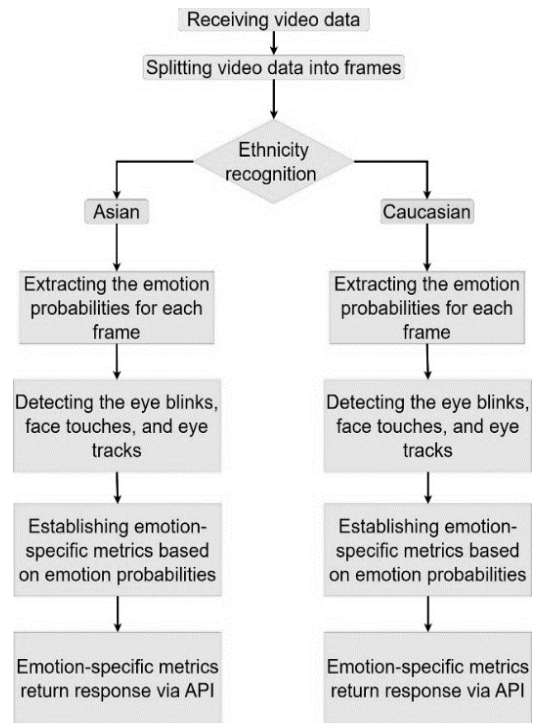


Figure 5: Emotion recognition framework for Asian and European facial expressions

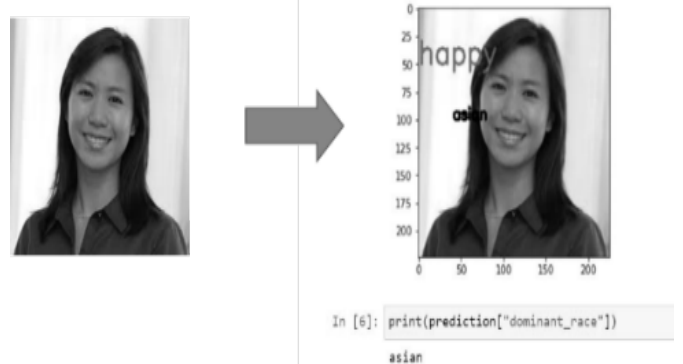


Figure 6: Example of dominant race recognition



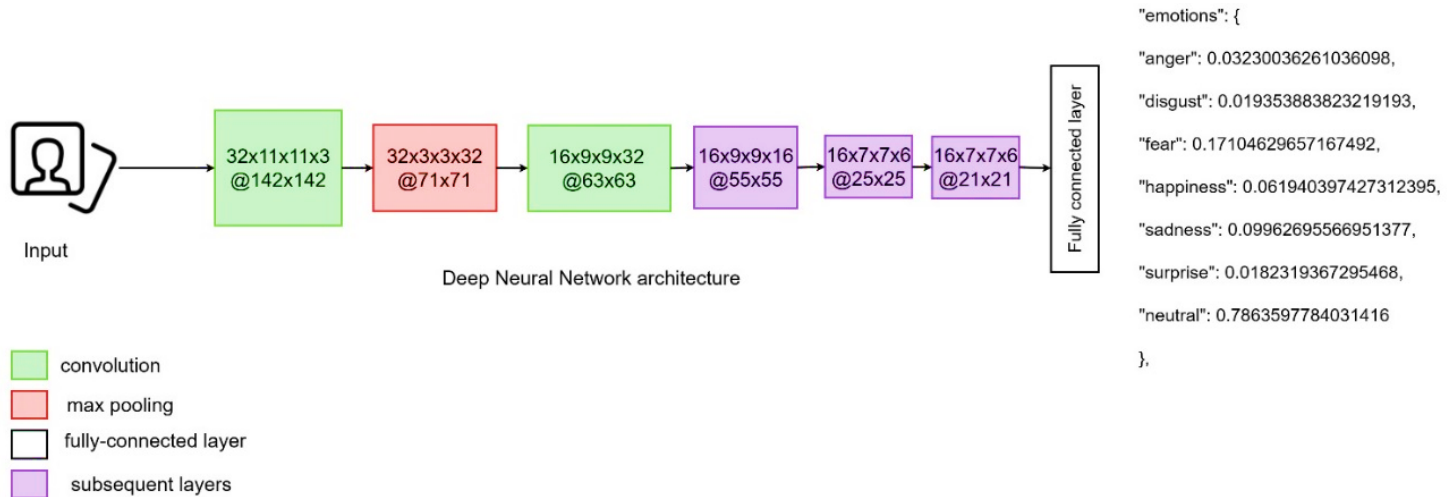


Figure 7: Schematic representation of DeepFace – Deep Neural Network architecture

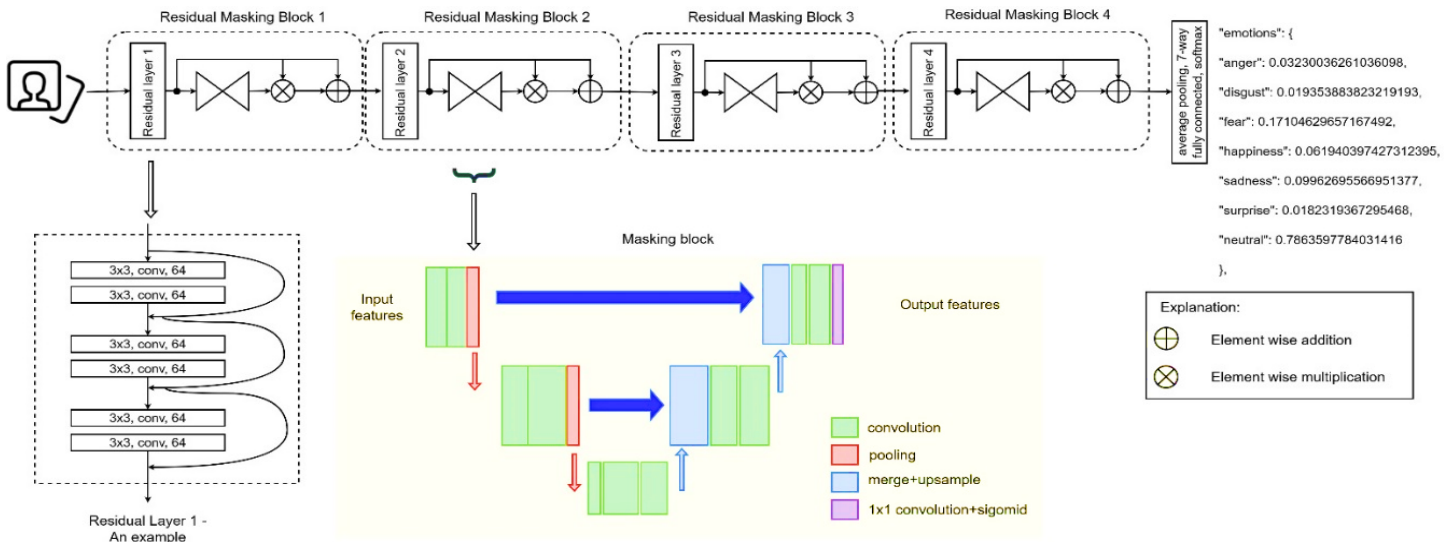


Figure 8: Schematic representation of Residual Masking Network (resmasknet) architecture

### 3.3. Eye tracking, blinking, and face touching detectors

Beyond emotion analysis, this project includes examination of eye movements, blinking frequencies, and facial touches. These non-verbal signs are considered expressive elements that contribute to a better understanding of emotional states. For instance, increased eye blink rates might signify nervousness or stress. Conversely, a reduction in blink frequency might indicate focus or intense concentration. Tracking the movement of eyes can offer valuable information about attention, interest, or cognitive processing related to emotions. People use their hands during conversations to further highlight some of the emotions. Touching face can indicate anxiety, discomfort, or even deception in some context. These non-verbal cues complement facial expressions and body language in emotion recognition. Integrating eye movements, blinking and facial touches into analysis enhance the accuracy and depth of understanding emotional states.

```
{'errorCode': 0,
'datetime': '2022-06-27 12:58:21.638802',
'emotions': {'average_eye_blink_time': 7,
'average_of_emotion_results': {'anger': 0.02,
'disgust': 0.02,
'fear': 0.19,
'happiness': 0.04,
'neutral': 0.83,
'sadness': 0.11,
'surprise': 0.0},
'average_of_face_touching': 2,
'dominant_emotions_in_difference_time_intervals': {'First': 'neutral',
'First-Mid': 'neutral',
'Last': 'neutral',
'Mid': 'neutral',
'Mid-Last': 'neutral'},
'longest_observed_emotions': 'neutral',
'number_of_hard_emotions_transitions_by_minutes': 0.0,
'number_of_observed_emotions_from_video': 7.0,
'number_of_soft_emotions_transitions_by_minutes': 14.0,
'number_of_transition_by_minutes': 14.0,
'shortest_observed_emotions': 'surprise'},
'ethnicity': 'white',
'filename': '720p_Medium.mp4',
'response_time': '50.89487981796265 second'}
```

Figure 9: Probability results of dominant emotions in the analyzed video

The project employs a fixed time window to count occurrences of blinks and facial touches. Moreover, the project extracts diverse features including blink duration, eyelid closure time, blink speed, and hand movement speed. It is important to note that the nearness of individuals to the camera significantly influences

the frequency of these observed events, resulting in higher event frequency for closer subjects.

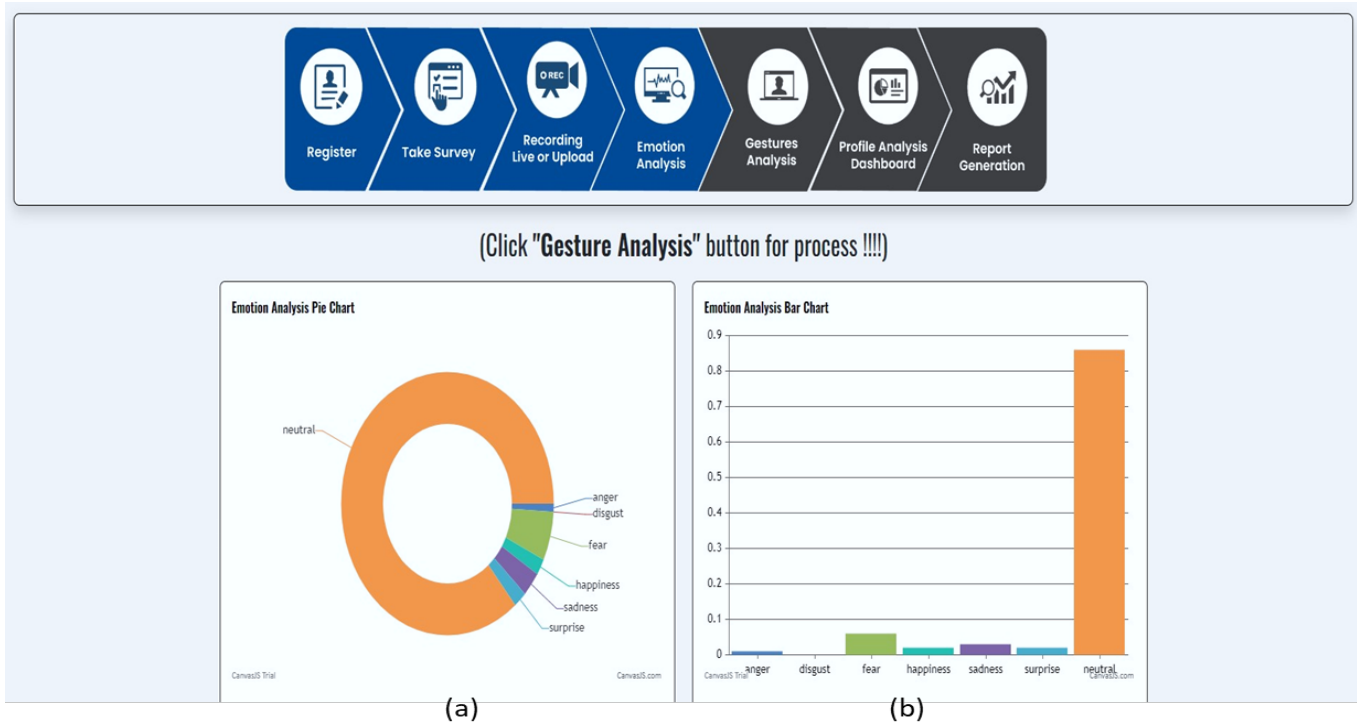


Figure 10: Visualized Result Representation. Emotion Analysis Pie Chart (a) represents the results in form of pie chart. Emotion Analysis Bar Chart (b) shows seven emotions (0=Anger, 1=Disgust, 2=Fear, 3=Happiness, 4=Sadness, 5=Surprise, 6=Neutral) and the numerical value (probability) of each emotion detected in the video (y - axis).

#### 4. Results

The Eureka Framework comprises several components, including registration, surveys/tests, video recording or uploading, emotional analysis, gesture analysis, a profile analysis dashboard, and a comprehensive report derived from the collected and analyzed data. This paper focuses exclusively on outcomes related to emotional analysis. The initial phase involves preprocessing the collected data and utilizing pre-trained models. Each data set used was analyzed and prepared separately, and at the end all data were used together for training and testing purposes. The reference data sets are recorded in controlled environments, while the collected Google images and videos contain different environments and conditions, which contributes to the diversity of the data set used in addition to ethnic diversity. The 80 videos collected and preprocessed, converting them into data frames. The first step is ethnicity detection within the input frames. A model is deployed to ascertain whether the individual in the input data is Asian or European person. Subsequently, depending on the ethnicity detected, specific models intended for different ethnic groups - Asian or European - are invoked. The accuracy of each model is separately calculated. Based on the input data in this work, the accuracy for Asian individuals stands at 75.2%, while for European individuals it reaches 86.6%. One reason for this may be due to the dataset's imbalance, where significantly more data for European individuals was available for training the models. Notably, the previously trained models from Py-Feat and DeepFace libraries, undergo tuning, involving adjustments in periods and parameters (learning rates, activation functions, etc.) to optimize their accuracy. The results for video recorded on June [www.astesj.com](http://www.astesj.com)

27, 2022 are shown in Figures 9 and 10. The person appearing in the video is identified as European, or white. Moreover, the analysis detected 7 instances of eye blinks, 2 facial touches and a dominant neutral emotion. The overall emotion analysis, based on seven emotions (0=Anger, 1=Disgust, 2=Fear, 3=Happiness, 4=Sadness, 5=Surprise, 6=Neutral), reveals on 83% of neutrality in the video. The video segments are analyzed separately, where the dominant emotions are calculated within each segment (first, first-middle, middle, middle-last, and last). Within segments the neutral emotion was detected. The results are graphically presented in the Figure 10, offering a view of dominant emotion through the analyzed video.

#### 5. Conclusion

The presented study demonstrates the implementation of a dominant emotion recognition model applicable across diverse sectors, notably crucial in HR for candidate assessment. Through questionnaires, competency-related tests and recorded or uploaded video introductions, an analysis based on collected data offers insights into candidate suitability. The proposed project empowers individuals to refine their skills and advance in their careers, simultaneously enabling corporations to maximize their investment in employees while exploring new opportunities. The main drawback of the above methods lies in their reliance on detecting overt facial movements, however, individuals might intentionally try to mask their emotions.

Employing AI in candidate interviews ensures unbiased selection. Leveraging unbiased datasets and interview videos,

well-trained models exhibit remarkable precision in detecting dominant emotions, eye blinks, facial touches, and eye tracking. Training models with varied datasets and interview videos resulted in high accuracy in recognizing emotions. This study highlights the model's efficacy not only in research but also in practical applications across various fields.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

This study was implemented thanks to the support provided by Vakif Katılım Bank R&D Center. The opportunities offered by Vakif Katılım R&D Center have added a huge impact on our scientific work. On this occasion, we would like to express our sincere gratitude to Vakif Katılım Bank R&D Center for their valuable support.

### References

- [1] <https://vimeo.com/775825642>, accessed:28.12.2022
- [2] B. Kayı, Z. Erbaşı, S. Özmen and A. Kulaglic, "Emotion and Movement Analysis Study from Asian and European Facial Expressions," 2023 3rd International Conference on Innovative Research in Applied Science, Engineering and Technology (IRASET), Mohammedia, Morocco, 2023, 1-5, doi: 10.1109/IRASET57153.2023.10152995.
- [3] DeepFace - Most Popular Deep Face Recognition in 2022 (Guide) - viso.ai, accessed:19.12.2022
- [4] Y. Taigman, M., Yang, M.A., Ranzato, & L. Wolf, L., "Deepface: Closing the gap to human-level performance in face verification", In Proceedings of the IEEE conference on computer vision and pattern recognition, 1701-1708, 2014, doi:10.1109/cvpr.2014.220
- [5] G. E. Dahl, T. N. Sainath, and G. E. Hinton. "Improving deep neural networks for LVCSR using rectified linear units and dropout". In ICASSP, 2013., doi:10.1109/ICASSP.2013.6639346
- [6] P. M. Ashok Kumar, Jeevan Babu Maddala & K. Martin Sagayam, "Enhanced Facial Emotion Recognition by Optimal Descriptor Selection with Neural Network", IETE Journal of Research, 2021, DOI: 10.1080/03772063.2021.1902868
- [7] S. Li and W. Deng, "Deep Facial Expression Recognition: A Survey," in IEEE Transactions on Affective Computing, **13**(3), 1195-1215, 1 July-Sept. 2022, doi: 10.1109/TAFFC.2020.2981446.
- [8] D. Kollias, M. A., Nicolaou, I., Kotsia, G., Zhao, & S. Zafeiriou, "Recognition of affect in the wild using deep neural networks", In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops ( 26-33), 2017, DOI: 10.1109/CVPRW.2017.247
- [9] H., Yang, U., Ciftci, & L., Yin, "Facial expression recognition by de-expression residue learning". In Proceedings of the IEEE conference on computer vision and pattern recognition (2168-2177), 2018, DOI: 10.1109/CVPR.2018.00231
- [10] S., Xie, & H. Hu, "Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks". IEEE Transactions on Multimedia, **21**(1), 211-220., 2018, DOI: 10.1109/TMM.2018.2844085
- [11] C. Szegedy, V. Vanhoucke, S. Ioffe, J. Shlens and Z. Wojna, "Rethinking the Inception Architecture for Computer Vision," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016, 2818-2826, doi: 10.1109/CVPR.2016.308.
- [12] G. Huang, Z. Liu, L. Van Der Maaten and K. Q. Weinberger, "Densely Connected Convolutional Networks," 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Honolulu, HI, USA, 2017, 2261-2269, doi: 10.1109/CVPR.2017.243.
- [13] P. Scovanner, S. Ali, and M. Shah. "A 3-dimensional sift descriptor and its application to action recognition", In Proceedings of the 15th ACM international conference on Multimedia (MM '07). Association for Computing Machinery, New York, NY, USA, 357-360., 2007 <https://doi.org/10.1145/1291233.1291311>
- [14] T. Mikolov, M. Karafiát, L. Burget, J. H. Černocký and S. Khudanpur. "Recurrent neural network-based language model." Interspeech (2010), doi: 10.21437/Interspeech.2010-343
- [15] S. Hochreiter and J. Schmidhuber. "Long Short-Term Memory". Neural Comput. **9**, 8 (November 15, 1997), 1735-1780. <https://doi.org/10.1162/neco.1997.9.8.1735>
- [16] Y. Fan, X. Lu, D. Li, and Y. Liu. "Video-based emotion recognition using CNN-RNN and C3D hybrid networks". In Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI '16). Association for Computing Machinery, New York, NY, USA, 445-450. 2016, <https://doi.org/10.1145/2993148.2997632>
- [17] S. E. Kahou, V. Michalski, K. Konda, R. Memisevic, and C. Pal. "Recurrent Neural Networks for Emotion Recognition in Video". In Proceedings of the 2015 ACM on International Conference on Multimodal Interaction (ICMI '15). Association for Computing Machinery, New York, NY, USA, 467-474. 2015. <https://doi.org/10.1145/2818346.2830596>
- [18] J. Yan, W. Zheng, Z. Cui, C. Tang, T. Zhang, Y. Zong, and N. Sun. "Multi-clue fusion for emotion recognition in the wild". In Proceedings of the 18th ACM International Conference on Multimodal Interaction (ICMI '16). Association for Computing Machinery, New York, NY, USA, 458-463. 2016, <https://doi.org/10.1145/2993148.2997630>
- [19] D. Tran, L. Bourdev, R. Fergus, L. Torresani and M. Paluri, "Learning Spatiotemporal Features with 3D Convolutional Networks," 2015 IEEE International Conference on Computer Vision (ICCV), Santiago, Chile, 2015, 4489-4497, doi: 10.1109/ICCV.2015.510.
- [20] FER2013 Dataset | Papers With Code, accessed:24.09.2023
- [21] CK+ Dataset | Papers With Code, accessed:24.09.2023
- [22] H., Kaushik, T., Kumar, & K. Bhalla, K. "iSecureHome: A deep fusion framework for surveillance of smart homes using real-time emotion recognition". Applied Soft Computing, **122**, 108788. 2022. <https://doi.org/10.1016/j.asoc.2022.108788>.
- [23] S. Fu, H. He and Z. -G. Hou, "Learning Race from Face: A Survey," in IEEE Transactions on Pattern Analysis and Machine Intelligence, **36**(12), 2483-2509, 1 Dec. 2014, doi: 10.1109/TPAMI.2014.2321570.
- [24] <https://viso.ai/computer-vision/deepface/>, accessed:25.09.2023
- [25] Y. Taigman, M. Yang, M. Ranzato and L. Wolf, "DeepFace: Closing the Gap to Human-Level Performance in Face Verification," 2014 IEEE Conference on Computer Vision and Pattern Recognition, Columbus, OH, USA, 2014, 1701-1708, doi: 10.1109/CVPR.2014.220.
- [26] A. Krizhevsky, I. Sutskever, and G. E. Hinton. "ImageNet classification with deep convolutional neural networks". In Proceedings of the 25th International Conference on Neural Information Processing Systems - **1** (NIPS'12). Curran Associates Inc., Red Hook, NY, USA, 1097-1105. 2012. DOI: 10.1109/ACPR.2015.7486599
- [27] G. E. Dahl, T. N. Sainath and G. E. Hinton, "Improving deep neural networks for LVCSR using rectified linear units and dropout," 2013 IEEE International Conference on Acoustics, Speech and Signal Processing, Vancouver, BC, Canada, 2013, 8609-8613, doi: 10.1109/ICASSP.2013.6639346.
- [28] L. Pham, T. H. Vu and T. A. Tran, "Facial Expression Recognition Using Residual Masking Network," 2020 25th International Conference on Pattern Recognition (ICPR), Milan, Italy, 2021, 4513-4519, doi: 10.1109/ICPR48806.2021.9411919.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Enhancing Compressive Strength of Pervious Concrete for Use as Pavement Layer in Urban Roads

Pontsho Penelope Mokgatla\*, Ramadhan Wanjala Salim, Julius Ndambuki

Civil Engineering Department, Tshwane University of Technology, Tshwane, 0001, South Africa

### ARTICLE INFO

Article history:

Received: 26 July, 2023

Revised: 19 November, 2023

Accepted: 19 November, 2023

Online: 20 January, 2024

Keywords:

Pervious Concrete

Permeable Concrete

Dolomite Rock Aggregates

Void Content

### ABSTRACT

South African Drainage and Stormwater Systems in urban roads has been of great concern, more so with recent flash floods in Gauteng and KwaZulu Natal Province in South Africa. Pervious concrete can be used to mitigate these challenges for urban roads. Pervious concrete is a concrete that contains no fines or only a small amount is added for binding. It consists of single sized rounded aggregates between 9.5mm to 12.5mm and bonded by cement and water at points of contact to create a high porosity system that drains water. The research focused on developing different structural pervious concrete mixes, with varying (a) aggregate sizes (7mm, 9.5mm and 19mm) (b) including (river fine aggregates, rubber flakes, fly ash and steel reinforced fiber) and combinations thereof. Results indicated that strong pervious concrete mixes can be developed using a combination of 9.5mm aggregate, river fine aggregates and fly ash, with a recorded maximum yield of 25MPa, and an accompanying Voids Content of 16%. Furthermore, against this performance, it was found that final compressive yield strength of 25MPa falls shy of the 35MPa COTO standard specifications for road and bridge works for south African road authorities (2020 edition). The derived pervious concrete mix was found to satisfy Class 4b Minor Collector Arterial in urban zones, with capacity to withstand up to 10 000 Average Daily Traffic

### 1. Introduction

Urbanization, which is concentration of population and its activities (social, cultural, and economic), has resulted in people moving from rural areas to urban in the quest for better opportunities as well as to develop the economy. Other urban benefits include quality health care and services such as (water, sanitary and transport). Unfortunately, with all these benefits some environmental aspects are affecting the very quality of life that was pursued due to congestion, pollution and a huge demand on water, electricity, and infrastructure. Urban areas are extensively covered by concrete, asphalt and infra-structure which makes it difficult for water to infiltrate the ground resulting in increased floods, run-off and contaminated groundwater [1]. The use of pervious concrete has increased with time to reduce some of these environmental concerns. The history of pervious concrete dates to 1800s when Europe was experiencing scarcity in cement as well as cost increase associated with sourcing cement, during this period it was used for pavement surfacing and for load bearing walls. Post WWII in the 1920s the use of pervious concrete increased in both Scotland and Europe due to then again

scarcity of cement, in this period pervious concrete was also utilized for slabs in double storey homes.

Pervious concrete is a concrete that contains no fines or only a small amount is added for binding. It consists of single sized rounded aggregates between 9.5mm to 12.5mm and bonded by cement and water at points of contact to create a high porosity system that drains water quickly. Its water to cement ratio is between 0.28 to 0.40 with the voids space between 15-30% where Conventional concrete has 3-5% voids spaces, and it has between 0.18-0.23 ratio of cement to aggregates. This voids space promotes water infiltration whilst recharging ground water with cleaner quality water. Even though it has high voids when placed properly a strength between 3.5MPa-20.5MPa are achievable. Pervious concrete has been developed to be both, firm and permeable, allowing vehicles to travel over it and water to infiltrate. Pervious concrete's characteristic of high drainage capacity unfortunately does not correspond with stability, durability and strength thus limiting its application to areas that have minimum to no traffic. Pervious concrete mixture has a very low slump because of its low mortar content as well as small number of fines if present at all. This research aimed to modify the standard pervious materials to achieve required compressive

\*Corresponding Author: Pontsho Penelope Mokgatla, pmokgatlap@gmail.com



strength whilst maintain its pervious concrete benefits of draining water quickly for a wide usage by modifying materials with the information provided by previous studies [2]

### 1.1. Pervious Concrete Background

Pervious Concrete (PC) is a concrete that contains no fines and only a small amount can be added for binding. It is a zero-slump concrete, consists of single sized aggregates between 9.5mm to 12.5mm and bonded by cement and water at points of contact to create a high porosity system that drains water quickly. Its water to cement ratio is between 0.28 to 0.40 [3] with the voids space between 15-30% [4] where conventional concrete has 3-5% voids space, and it has between 1:4-1:4,5 [5] ratio of cement to aggregates. This voids space promotes water infiltration whilst recharging ground water with cleaner quality water.

The use of Pervious Concrete on pavement in urban areas has very favourable environmental outcomes which forms part of Best Management Practice (BMP). The use of PC may drastically reduce drainage infrastructure required which will aid South African economy. As Engineers it is our responsibility to consider the future effects of our work and the human and environmental impacts. Pervious Concrete has been in use for decades but very limited use due to its strength and durability limitation. It has been listed as BMP because of its environmental benefits in other countries and it is important to enhance its materials based of previous studies to see how it may be used broader that its current limited use. Pervious Concrete limitations are mostly because of its strength and durability, studies have been ongoing to improve with various enhancing materials such as Fly ash, Fibers, waste rubber, Latex polymer, Styrene Butadiene rubber and Fine aggregates [6]-[8]

### Characteristics

As illustrated in Figure 1, Pervious Concrete also known as porous concrete, or permeable concrete, differs from conventional concrete. It is very permeable with void content range between 15% – 30 % which allows water to seep through the pavement layer directly infiltrating the ground. This water can be harvested, directed to the stormwater existing structures, or simply allowed to recharge the ground. Pervious Concrete’s permeability properties make it an environmentally friendly solution, which explains why it is being utilized in various ways such as sidewalks, driveways and low traffic roads and considered for medium to high traffic roads. The current asphalt roads are impermeable and water exposure is one of the risks especially when the drainage systems are poorly maintained or not available. After rain fall potholes are not visible which poses a risk on the drivers and pedestrians. The runoff caused by rain fall causes splashes on the roads which affects visibility. South African municipalities have limited budget which is always a project constrain and by opting for pervious concrete the amount of stormwater infrastructure will be reduced which thus reduce cost. The study focused on urban roads which are the most affected by floods, runoff, insufficient draining infrastructure which was the focus for the research to enhance Pervious Concrete with affordable and locally sourced materials.

The current road designs protect road layers from water however it is impossible to prevent water damages, water being in

solvent given time water destroys even the most permanent looking structures such as asphalt roads. Pervious Concrete as illustrates on Figure 2 allows water to infiltrate the ground and by design water doesn’t pond which all more reason to enhance the strength and durability of Pervious Concrete to benefit rural roads, urban roads even highways.

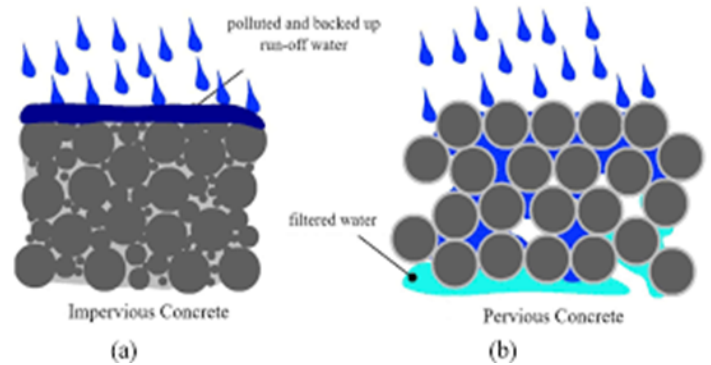


Figure 1: Pervious Concrete (American Society of Civil Engineers)

### Structural layout

Pervious concrete is commonly used for walkway and driveway however Pervious Concrete has a potential to reduce stormwater runoff, road splash and sprays, Pervious Concrete needs to be improved to meet the required strength and durability to be considered as a road pavement layer. This research focused on using Pervious Concrete as more than what is currently being utilized for by improving its mechanical properties. With all the pervious concrete has, it should not be limited to just walkways, driveways and so forth however be extended to urban roads.



Figure 2: Pervious Concrete Pavement Drain ability (Source: Quick Mix Website)

### a) Pervious Concrete layer

Pervious Concrete layer is the top layer that is exposed to the environment and traffic. The thickness of this layer ranges between 150 mm to 200mm see above Figure 3. Pervious Concrete layer must be strong and capable to withstand traffic and must be durable to all types of weather. Figure 3 Above shows the structure of Pervious Concrete pavement layers used for parking lots and sidewalks. Though the above structure is for no traffic areas, the similar structure enhanced will extend the use of Pervious Concrete and economy which is what the research aimed

to achieve. The pervious concrete layer has 15-30% void content and between 3.5-20.5Mpa compressive strength.

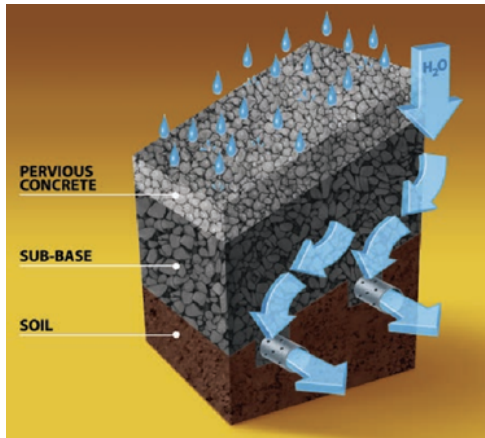


Figure 3: Pervious Concrete Pavement Structure (Image Courtesy of Quick Mix Website) and Sika USA

### b) Sub- base layer

The sub- base layer as shown in Figure 3 above has 20% to 40% voids and must be constructed uniformly at a correct elevation to provide support and it has to be sloped correctly to allow water flow to catchment areas. Other studies recommend compaction but not extreme levels following the natural drainage patterns. The designer can specify geotextile on the sub- base layer after compaction though not mandatory but preferred when sub-grade layer is of fine material.

### c) Sub-grade layer

Soil-layer which is the undisturbed soil beneath sub-base see Figure 3 which its purpose is to provide support and store infiltrated water, this layer too must be constructed evenly and to the correct level and desired slope. The soil layer may consist of fine material with low permeability, this does not affect construction of Pervious Concrete however a geo-textile may be required in between the sub-base and sub-grade to mitigate clogging and reduced retention capacity.

### 1.2. Problem Statement

Stormwater runoff is one of the greatest issues in South African urban roads which has caused environmental problems such as roads and embankments erosion whilst damaging properties and adding to pollution concerns. There are measures that have been put in place to mitigate stormwater runoff issues such as (surface drainage, open channels, stormwater pipe, culverts, and bridges) however none of the efforts has been sufficient to resolve this problem. Pervious Concrete is a highly permeable concrete which promotes water infiltration. The use of it would control surface runoff exposing damages on the road such as potholes during and after rainfall and reducing road ponding water sprays. Pervious Concrete doesn't just benefit the environmental but the economic too by saving cost on the stormwater infrastructures, reduce the risk of floods, help control stormwater runoff. Cities and Municipalities have budgetary constraints that can be addressed by constructing roads that have

stormwater management benefit reducing the requirements additional stormwater infrastructure.

### 1.3. Specific objectives

- To evaluate the influence of compaction on compressive strength of Pervious Concrete made with dolomitic rock coarse aggregates.
- To ascertain / establish the quantity of fine aggregates required to enhance the compressive strength of Pervious Concrete which usually doesn't contain fine aggregates or any fine aggregates.
- Determine the impact of waste tyre rubber when included in a Pervious Concrete mix and calculate the quantity required to enhance the mix.
- Derive a combination trial mix that will enhance the compressive strength of Pervious Concrete.
- To compare the achieved compressive strength results of the modified Pervious Concrete made with dolomitic rock aggregates against TRH 26 (South African Road Classification and Access Management Manual) and determine road class will the achieved mix satisfy.

To achieve maximum permeability large stones are required because large aggregates sizes create larger voids thus promoting permeability. However large voids are the reason that Pervious Concrete is low in compressive strength. Literature indicates the use of fine aggregates in small quantities has been shown to enhance and improve the strength of Pervious Concrete. In the quest of sustainable engineering ongoing studies to enhance Pervious Concrete by partially/fully replacing standard material with enhancing materials [10].

Though efforts have been taken to study and enhance Pervious Concrete, little is still known about Dolomite rock coarse aggregates Pervious Concrete which was the motivation of this research also being mindful of cost and availability of dolomite rock in South Africa. The objective of this research was to enhance Pervious Concrete from its typical compressive strength of 20Mpa to higher compressive strength that can be suitable for urban roads whilst maintaining sufficient void content (15-30%) to drain efficiently.

## 2. Materials and Methods

Coarse aggregates make up the largest material and provides the overall strength of Pervious Concrete mix, therefore the type of aggregate influences performance and strength. Pervious Concrete has high permeability due to high void content and low mortar content therefore the aggregates have to provide both structural integrity to withstand the traffic loads and environmental conditions whilst allowing water infiltration and drainage. Pervious Concrete uses the same materials as the conventional concrete however Pervious Concrete has no fines. Crushed natural stones are usually used, some studies experimented on river stones, limestones and few on dolomitic stones. The choice of coarse aggregates is thus very important as it plays an important role in forming both the critical mass and bonding component. The size, type and shape of coarse aggregates are also important because they have influence on permeability rate, compressive strength, and durability [11].

2.1. Literature Summary Table

Table 1: Literature Summary on Aggregate size, water to cement ratio and aggregate to cement ratio.

SN	Author	Year	Title	Aggregate Type	Aggregate Size (mm)	Cement to Aggregates ratio	water to cement ratio	Compressive Strength MPa	Additional Material	Conclusion
1	Sathish Kumar Veerappan; Hrvoje Gotovac	2022	Effect of Aggregate Size and Compaction on the Strength and Hydraulic Properties of Pervious Concrete	Limestone	4-8 8-16	1:4.5	0,32	24.4 11.73	N\A	In terms of compressive strength this study discovered that 25 drops of compaction resulted in the highest strength concrete, regardless of mix proportion, which further confirms that compaction affects compressive strength, Unlike conventional concrete, the study also indicated compressive after 7 days and 28days varied from 5% to 25% unlike conventional concrete with 30% to 35%. The concrete samples with larger aggregates produced lower compressive strength than those with smaller aggregates, even with the same cement and pore content. This may be due to the crushing behaviour of the larger aggregates surrounded by pores during testing
2	Fan Yu, Daquan Sun, Jue Wang, Minjun Hu	2019	Influence of aggregate size on compressive strength of pervious concrete	Basalt rock	2.36-4.75 4.75-6 4.75-9.5 6-8 8-9.5 10-12.5 12.5-15 10-15	1:4.5	0,31	19.86 21.73 29.33 30.19 27.39 29.44 26.30 32	N\A	Fan Yu discovered that though all groups experimented on had a close porosity of 20.9% though, the aggregates smaller than 7mm compressive strength didn't differ much and ranged between (19-21MPa) however a rapid strength increase was realised for aggregates larger than 7mm with the highest perming being between (27-32MPa)
3	L. K. Crouch, P.E ; Jordan Pitt and Ryan Hewitt	2015	Aggregate Effects on Pervious Portland Cement Concrete Static Modulus of Elasticity	Limestone	9,5 12,5 19,5	1:5.5	0,39		Sand and Fly Ash	Study found that, though it seemed like using large aggregates would result in greater void content however bigger aggregate size have less specific surface than smaller aggregates, smaller aggregates requires less paste and more paste required fills the void thus reducing void content some mixes were compacted and results were in accordance to the studies that compaction does enhance mechanical properties of pervious concrete. Using for uniform gradation results in higher effective void contents ( 23- 31%)
4	K. Cosic', L. Korat, V. Ducman, I. Netinger	2014	Influence of aggregate type and size on properties of pervious concrete	Dolomite	4-8 8-16	1:6.0	0,33	20-27Mpa	Steel slag, river sand	The study consisted of 5 experiment of dolomite or steel slag with 4-8mm to 8- 16mm aggregate fraction and (30:60 and 60:30 Proportions. The results indicated that dolomite aggregates had higher porosity compared to pervious concrete that included steel slag. 28days results didn't satisfy the required concrete strength for pavement concrete. highest compressive strength was achieved at 60% 8-16mm and 30% 4-8mm was mixed.
5	Benjamin Rehder, Kingsten Banh, Narayanan Neithalath	2013	Fracture behaviour of pervious concretes: The effects of pore structure and fibers	Limestone	4.75 9.5	1: 4.0	0,3		Polypropylene 1% or 2% by volume	the study discovered that 19% void content's compressive strength ranged between 18- 23MPa where as the higher porosity of 25% gave lower strength on 10-15MPa, the study discovered adding of fibers did not provide tangible benefits just a slight Improvement
6	Omkar Deo, Narayanan Neithalath	2010	Compressive behaviour of pervious concretes and a quantification of the influence of random pore structure features	Limestone	2.36 4.75 9.5	1:5.0	0,33		N\A	Increasing aggregate sizes resulted in increased compressive strengths for the pervious concrete specimens studied. The mixtures made with blended aggregate sizes were found to have their peak stresses in between those made with the single-sized aggregates that constitute the blend. An increase in pore volume fraction by approximately 10% was found to result in a reduction in the compressive strength by about 50%.
7	Manoj Chopra, Marty Wanielista and Ann Marie Mulligan	2006	Compressive Strength of Pervious Concrete Pavements	Limestone	6 8 10	1:6.0 1:8.0 1:10	0.38 0.41 0.45	14 21 10	N\A	study demonstrated an increase in aggregate to cement ratio results in decrease in strength and that compaction plays a big role in compressive strength. Best water to cement ration 0.35-0.39 which provided the highest compressive strength
8	M. Aamer Rafique Bhutta, K. Tsuruta,J. Mirza	2012	Evaluation of high-performance porous concrete properties	Limestone	2.5-5 5-13 13-20	1:5 1:7 1:5.5	0.35 0.30 0.30	20 24 16	SP (density: 1.06 g/cm3, and a thickening (cohesive) agent (water-soluble cellulose based polymer powder, density: 2.40 g/cm3	Combining SP and cohesive agent was beneficial and produced a acceptable pervious concrete in terms of workability and strength, the void content was between 18 to 28%.



Study by [12] discovered that single size 9,5mm natural Crushed stone aggregates yielded the highest compressive strength, whilst other studies discovered that aggregates lower than 7mm stone didn't show a significant change in compressive strength it ranged between 19- 21Mpa. Aggregates larger than 7mm showed a big increase in compressive strength with the increase in aggregate sizes. 6-8mm stone had a compressive strength of 29.33Mpa and 32Mpa for 10-15mm stone aggregate size indicated that the bigger the aggregate size doesn't necessarily result in higher compressive strength, what it seemed to increase was permeability due to larger void space [13].

Based on Table 1 of past literature on aggregate sizes of Pervious Concrete. This research focused on the dolomite rock coarse aggregates 7.1mm, 9,5mm and 19mm to determine if the dolomite rock will behave the same as limestone or river coarse aggregates. Dolomitic rock is affordable and readily available in Gauteng province.

Table 2: Coarse aggregates properties

Parameter	Dolomite Rock Coarse Aggregates			Units
	19mm Stone	9.5mm Stone	7.1mm Sand	
Loose Bulk Density (LBD)	1536	1532	1888	kg/m <sup>3</sup>
Compacted Bulk Density (CBD)	1664	1626	2127	kg/m <sup>3</sup>
Apparent Density (Relative Density)	2840	2810	2810	kg/m <sup>3</sup>
Water Absorption	0,76	0,8	1,32	%
Flakiness Index	5	-	-	-

### 2.2. Dolomitic Rock

Above Table 2 shows the properties of dolomite rock sourced at PPC Group. There are limited studies on dolomitic rock. Dolomitic rock has high surfaced hardness as well as density, it also has low water demand which is of benefit since Pervious Concrete has limited water. [14, 15] studies indicated improvement in compressive strength due to the dolomitic rock's high resistance to abrasion and to attrition another study indicated using Pervious Concrete with dolomitic rock as coarse aggregates resulted in improved durability and long-term performance in comparison to with other types of aggregates and it requires less energy to produce compared to other types of aggregates. Gauteng Province has 25% of its land covered in dolomitic rock [16], with upgrading, extensions and new infrastructure on the rise and this type of aggregates being locally discarded material due the risk related to foundation problems and groundwater, makes dolomitic rock very affordable.

There are limited studies on Pervious Concrete with dolomitic rock as coarse aggregates and no studies modifying dolomitic rock Pervious Concrete with either fine aggregates, fibers, or admixtures. Studies that have used dolomitic have shown impressive results.

### 2.3. Cementitious Binders

Portland cement is the most used in Pervious Concrete have a very important role in construction and just as important in forming Pervious Concrete, it is mainly used as an aggregate

binder that sets and hardens due to a chemical reaction between the dry ingredients and water it is rarely used on its own. Cement is an ingredient of concrete. It is a fine powder composed of crushed rock with burnt lime which behaves as a binder when mixed with water, coarse aggregates, and fine aggregates [17]. Table 3 shows the chemical and physical properties of cement used.

Table 3: Chemical and Physical Properties

Chemical Properties	Typical Results	SANS Requirements
Sulphur trioxide: % by mass	2,1	3,5% maximum
Chlorides: % by mass	≤ 0,01	1 0,1% maximum
Physical Properties	Typical results	SANS Requirements
<b>Setting times:</b>		
Initial: minutes	150	≥ 75 minutes No requirement
Final: hours	3,5	
Specific Area (Blaine): m <sup>2</sup> /kg	360	No requirement
<b>Soundness:</b>		
Le Chatelier Expansion (mm)	1	10 maximum
typical mortar prism results determined in accordance with EN 196-1:		
Compressive strength at 2 days (MPa)	± 17	≥ 10,0
Compressive strength at 28 days (MPa)	± 38	≥ 32,5 ≤ 52,5
<b>Densities:</b>		
Relative density	3,04	
Bulk density, aerated: kg/m <sup>3</sup>	1300	
Bulk density, as packed: kg/m <sup>3</sup>	1500	

Cement water ratio and cement to aggregate ratio are very important in Pervious Concrete, the ratio will determine the compressive strength. Water to cement influences workability, permeability, and strength, it is key to strike a balance. If the paste is too little the mix might not mix well or reach required consistency, when it is too high it settles at the bottom without providing adequate bonding. Pervious Concrete tends to fail at the interface between aggregates, water and cement which results in low compressive strength, studies have demonstrated that the higher the cement water ratio produces low compressive strength. It is important to note that by making paste stronger does not result in overall increased strength. The cement paste volume influences permeability, excessive paste fills the voids increases the strength however eliminates permeability which will take the number one benefit of pervious concrete being very porous. The use of admixtures by partially replacing cement has indicated, improved workability, shrinkage, reduced bleeding, improved long term durability and strength. To achieve optimal strength of the cement paste, proper curing has to be ensured to allow the paste to reach its full potential strength.

### 2.4. Fine aggregates - Fine Aggregates

Pervious Concrete also known as zero slump and no fines concrete as it usually doesn't contain fine aggregates to allow large void formation. In attempts to enhance Pervious Concrete studies indicated that the use of fine aggregates in small quantities improves strength, resistance, and durability. Above Table 4 shows fine aggregates properties. The reason adding fine aggregates has proven beneficial is by adding fine aggregates increases connecting points of pervious concrete thus providing a stronger bond [18] Fine aggregates inclusion increases mechanical properties of pervious concrete however reduces void



content, it important to add enough fine aggregates to balance both void and strength. Fine aggregates inclusion improved workability, when fine aggregates particles filled the voids between coarse aggregates provides better bonding and interlocking thus improving workability, placing as well as compaction.

Table 4: Properties of Fine Aggregates (Fine aggregates)

Parameter	Fine Aggregates	
	River Sand	Units
Loose Bulk Density (LBD)	1352	Kg/m <sup>3</sup>
Compacted Bulk Density (CBD)	1477	Kg/m <sup>3</sup>
Bulk Density	2,547	Kg/m <sup>3</sup>
Apparent Density	2,604	Kg/m <sup>3</sup>
Water Absorption	0,855	%
Flakiness Index	-	%

Pervious Concrete does not include fine aggregates or aggregates less than 3mm for large voids formation to create a highly permeable surface, previous studies have shown that adding small amount between (5-10%) of fine aggregates may balance Pervious Concrete strength and permeability. The research experiment aimed to enhance Pervious Concrete made with dolomitic rock coarse aggregates by varying different fine aggregates percentages in Pervious Concrete, results have shown that fine aggregates enhance mechanical properties of Pervious Concrete.

2.5. Fly Ash

Table 5 below shows properties of Fly Ash that was sourced from PPC Group

Table 5: Fly Ash Properties

<b>CLASSIFIED FLY ASH</b>	
Finer Ash which further selected/Split from the classifier	
<b>Class (S) Fly Ash (CFA)</b>	
Used as a partial cement replacement when making concrete or cementitious products, because of its Finer properties it offers an accelerated reaction to the solution.	
<ul style="list-style-type: none"> <li>Alumina - Silicates consisting mainly of SiO<sub>2</sub>, AL<sub>2</sub>O<sub>3</sub>, CaO, Fe<sub>2</sub>O<sub>3</sub></li> </ul>	

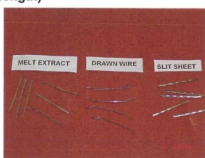
An admixture is defined as “a material other than water, aggregates, cementitious materials, used as an ingredient of a cementitious mixture to modify its freshly mixed, setting, or hardened properties and that is added to the batch before or during its mixing.

It is possible to produce concrete without concrete without cement and substitute it with fly ash as binder. Fly ash is a by-

product from coal combustion when generating electricity, this material is waste and cannot be used to generate power. The benefits of fly ash concrete such as improves workability on fresh concrete and accelerates ultimate strength on concrete and minimises thermal cracking [19].

Admixtures can be both natural and manufactured and benefits both fresh and hardened concrete. Previous studies managed to enhance and improve concrete strength significantly by adding (a) silica fume and (b) superplasticizer admixtures, further proving the significance of admixtures in Pervious Concrete. Studies summarized in Table 1 measured the effect of using fly ash and fine in Pervious Concrete and performance benefits and the study indicated that cement can be reduced up to 10% in Pervious Concrete without affecting or compromising permeability or strength. Fly ash inclusion reduced water demand which is a benefit due to limited water requirement in pervious concrete and fly ash improved workability since it acts as a pozzolan enhancing the cohesiveness of pervious concrete. Fly ash inclusion reduces the costs as because it is very cheap waste product. In [20], the authors demonstrated that partial replacement of cement with fly ash improved strength, resistance and durability by forming additional calcium silicate hydrates during hydration process and the pozzolanic reaction results in refined and denser microstructure.

Table 6: Steel Fibres Properties

Product Description: <b>STAINLESS STEEL FIBRES</b> (25 mm length)			
For re-inforcing castables and gunned concrete and refractory			
<b>Chemical Analysis:</b>		<b>430</b>	<b>446</b>
Cr	(%)	16-18	23-27
Ni	(%)		
Al	(%)		
<b>Physical Properties:</b>		<b>430</b>	<b>446</b>
Max. Melting Temperature	[°C]	1480/1530	1425/1510
Max. Service Temperature	[°C]	815	1100
Slit Sheet		Yes	
Melt Extract		Yes	Yes
Drawn Wire			
PACKAGING		20 kg Boxes	
STORAGE			
The above results are typical, subject to normal statistical variations and not for specification purposes.			

2.6. Steel Fiber

Steel fibre for reinforcing concrete is defined as short, discrete lengths of steel fibres with an aspect ratio (ratio of length to diameter) from about 20 to 100, with different cross-sections, and that are sufficiently small to be randomly dispersed in an unhardened concrete mixture using the usual mixing procedures. Table 5 provides information on the steel fibres used from ANVILS South Africa.

Steel fibers are very small and thin strands made of steel, the use of steel fibers in Pervious Concrete are ongoing and the available studies results indicates improvement on mechanical properties of Pervious Concrete by providing reinforcing to

Pervious Concrete which is important for this highly permeable pavement. literature results indicated following results (improved workability, increased strength and durability, crack resistance) when steel fibers are used [21]-[25].

Pervious Concrete's greatest benefits is being porous by having large voids, yet it is the reason that Pervious Concrete mechanical properties are compromised and require reinforcing without affecting to allow water to infiltrate the ground through the voids. Study review on the durability of Pervious Concrete reinforced with steel fibers indicated that adding fibres to Pervious Concrete improves both Tensile Strength and Permeability and this was due to fibres slightly increasing the void content in the mix. The study also indicated an increased void content, whilst splitting tensile and compressive strength in Pervious Concrete also improved due to fibres behaving as woven fabric in concrete which provides the structure with some reinforcement. In another separate research study, five (5) different types of fibres were used, three (3) which were non-structural for the reduction of shrinkage, and two (2) were of structural components to enhance mechanical performance. The compaction energies and w/c ratios differed for the various mixes, whilst aggregate size remained the same. The results showed an overall significant improvement in Pervious Concrete's strength. There is risk in using steel fiber in Pervious Concrete such as corrosion which may result in affecting the strength and durability hoped for when utilised, this however can be mitigated. The reason steel fiber rust is due to the exposure to environmental elements, air and water which results in the surface of the steel fiber corroding oxidisation. Ways to mitigate this or extend the lifespan of steel fiber reinforced Pervious Concrete are using stainless steel fibers as they are more resistant to rust, using zinc or aluminium coated steel fibers to prevent direct exposure on the steel fiber. Placing of fiber steel reinforced concrete is important, the distribution has to be even, and adequate curing achieved. like any other pavement layer, maintenance is required, and damages repaired. This study focused on the stainless steel fiber to mitigate corrosion.

**2.7. Waste Tyre Rubber**

Waste tyre rubber that was already in crumbs were collected from Dawhi rubber recycling in Germiston. To make sure they were free from oils and dirt, they were washed and properly dried ahead of time below Table 7 shows the waste tyre properties.

Table 7: Waste Tyre Properties

Parameter	Rubber Crumbs	Units
	Loose Bulk Density (LBD)	-
Compacted Bulk Density (CBD)	-	Kg/m <sup>3</sup>
Bulk Density	1,094	Kg/m <sup>3</sup>
Apparent Density	1,127	Kg/m <sup>3</sup>
Water Absorption	2,7	%
Flakiness Index	-	%

Waste tyres are amongst the biggest waste that is produced yearly. 100million waste tyres are generated each year with the projected 1200million by year 2030 Disposing waste tyre is a

challenge for cities. Studies were done with this waste material as an innovative solution to mitigate the waste tyres disposal challenges not only to recycle but also to enhance the strength of Pervious Concrete by partially replacing coarse aggregates with waste tyre rubber.

The inclusion of waste tyres has been explored however have contradicting information [26] discovered by partially including waste tyres to pervious concrete improved the mechanical properties such as resistance, cracking, ductility, and improved energy absorption. When waste tyres are partially included in a pervious concrete mix, a big reduction in mechanical properties such as compressive strength was experienced, the study indicated that this reduction in mechanical properties was due to inappropriate bonding between tyres and cement resulting in degradation and lack of bonding of the pervious concrete mix. In [27], the authors experimented on waste tyre crumbs and natural zeolite to enhance pervious concrete, findings were consistent with waste tyre reduced mechanical properties. A Study by [28] found that by replacing coarse aggregates with 9% of waste tyre improved compressive strength. There are other factors important to the strength and durability of pervious concrete such the following:

**2.8. Aggregates grading**

Stone aggregate's grading is just as important when considering the strength factor and permeability, it is important to keep a narrow grading. Poorly graded material requires excessive cement paste for filling up the voids which cannot be afforded on Pervious Concrete as it has a limited cement paste, gap graded is viable grading however is not optimal.

It is important to achieve well graded aggregates, which are aggregates that are proportioned and sized correctly to maximize aggregate volume and preserving workability, finishing and strength. Well graded aggregates provide the highest strength which was the study goal to enhance Pervious Concrete. Pervious Concrete commonly used gradation includes amongst others, most research used ASTM C 33 Sieves, namely (a) No. 67 (Sieves 19.0mm to 4.75mm), (b) No. 8 (Sieves 9.5mm to 2.36mm) and (c) No. 89 (Sieves 9.5mm to 1.18mm), and while up to 25mm single-sized aggregates have also been used. In this research SANS 3001-AG1 Grading Method was utilized which is a commonly used method in South Africa

**2.9. Aggregates size**

The aggregate size is important as it influences compressive strength, pore characteristic, paste thickness which all affects Pervious Concrete mechanical properties of pervious concrete Coarse aggregates generally consists of 9.5mm to 19mm stone sizes however studies have indicated an inclusion of smaller size aggregates enhances and improves mechanical properties of Pervious Concrete whilst maintaining acceptable water drainage for Pervious Concrete, large aggregates have larger voids which result in high permeability however compromises the concrete strength. Pervious Concrete is commonly used in parking lots, pavements, and low traffic roads where aesthetics matter therefore using the smallest size is useful. The thickness of the paste

influenced strength, the thicker the paste the higher the strength. Some indicated that larger aggregates had thicker coating thus indicated that with an increase in aggregate size resulted in higher compressive strength which is a contradiction to most studies see Table 1 which indicated compressive strength reduction with aggregate size increment. Including fine aggregates which are aggregates smaller than 7.1mm in small quantities indicated improvement in mechanical properties with lower yet acceptable draining capacity due to smaller void content. The size choice of Pervious Concrete is key as it will affect permeability rate as well as the strength of the concrete.

**2.10. Aggregates shape**

Aggregate shape equally plays an important role when considering compressive strength yield in Pervious Concrete. The shape of aggregates in Pervious Concrete directly influences the size of voids. Bearing in mind the shape of the aggregates, when void sizes in concrete increases, permeability increases. With increased permeability however, there is direct negative impact on mechanical strength parameters, leading to the reduced compressive strength of Pervious Concrete. The following are common types of aggregates shape:

- a) Angular aggregates: These are aggregates with irregular shape, characterized by their sharp edges, also known to provide good interlocking property with large void space and high permeability. The higher the void space means the lower the compressive strength. Angular aggregates also negatively affect workability.
- b) Rounded Aggregates: Rounded aggregates unlike angular aggregates offers good workability with less water demand. Compared to the angular, round aggregates performed better in terms of compressive strength and permeability. The interlocking of round aggregates is not that greatest which may result in weaker bond, they are also less inter-particle friction which results in reduced resistance.

The shape of aggregates in Pervious Concrete directly influences the size of voids, when void sizes in concrete increases, permeability increases. With increased permeability however, there is direct negative impact on mechanical strength parameters, leading to the reduced compressive strength of Pervious Concrete referred aggregates are the single sized, well graded aggregates which are proportioned correctly to promote maximum aggregate volume, workability, and strength.

**2.11. Cement To Aggregates Ratio**

Pervious Concrete is a unique concrete that is very permeable to allow water infiltration. Conventional concrete has high cement content in comparison to Pervious Concrete, the high cement to aggregates ratio of Pervious Concrete is important in creating a porous structure by creating large voids to allow water to pass through.

In the quest to improve Pervious Concrete different ratios of cement to aggregate have been experimented on see Table 1 and Table 8 where it shows the compressive strength results. The two

tables indicates that the highest compressive strength was mostly achieved when cement to aggregate ratio is 1:4 and between 9mm to 10mm stone aggregate size. Table 8 below shows compressive results of 34.3Mpa for 1:4 Cement to aggregate ratio and for cement to aggregate ratio of 1:7 the compressive strength was lower at 14.4Mpa, the results indicated that the higher the quantity of aggregates increased to one part of cement the compressive strength reduced. Consequently, this means the Aggregate Cement (A/C) ratio can be optimized thorough selecting the required ratio for optimum mechanical properties as well as physical properties (permeability). Research has shown that when higher aggregates amount is used, the concrete needs more water which affects concrete strength unless cement content is adjusted [29.30].

Table 8: Compressive strength of different Cement to Aggregate ratios [25].

S. No.	Concrete Mix	Water/Cement Ratio	Compressive Strength (N/mm <sup>2</sup> )			
			Day of Curing			
			1	3	7	28
1	A1(1:4)	0.3	6.26	14	22.75	34.3
2	A2(1:5)	0.3	5.87	11.2	18.2	27.78
3	A3(1:6)	0.35	2.56	6.5	10.4	16.02
4	A4(1:6.5)	0.35	2.4	6	9.8	15
5	A5(1:7)	0.4	2.31	5.78	8.7	14.44

**2.12. Water Cement Ratio**

Water-Cement (w/c) ratio which is the weight of water to weight of cement in pervious concrete mix. W/C plays an integral part in workability, permeability, strength, and durability in Pervious Concrete. Lower w/c ratio affects workability however results in higher compressive strength see Table 1. The study by [31] discovered that lower water to cement ratio doesn't achieve high compressive strength due to poor workability and poor hydration of the cement and poor cohesion between cement and aggregates. Excess amount of water improves workability but negatively affects durability, excess amount of water drains the paste which leads to clogging of the voids. Water content has to be closely monitored. Figure 4 below illustrates compressive strength versus varying water-cement ratio:

The water cement (w/c) ratio generally used in Pervious Concrete ranges from 0.3 to 0.4, and this is sufficient for viscosity, coating, and binding of aggregates. Previous studies have revealed that to achieve an optimum Pervious Concrete mixture, the water-cement (w/c) ratio has to range from 0.32 to 0.34 as demonstrated in Figure 4. Water to cement ratio, though important it does not determine the final concrete's properties there are more factors just as important such as aggregates quality and grading, mixing time, the use of admixture in the mix, the technique used, compaction and curing method are all key to the performance of pervious concrete. Studies have indicated that by increasing w/c ratio there's a direct increase in compressive strength, however, reduces permeability. Previous studies have assessed the hydrological and structural performance using the paste thickness

and discovered that the thicker the paste resulted in higher compressive strength.

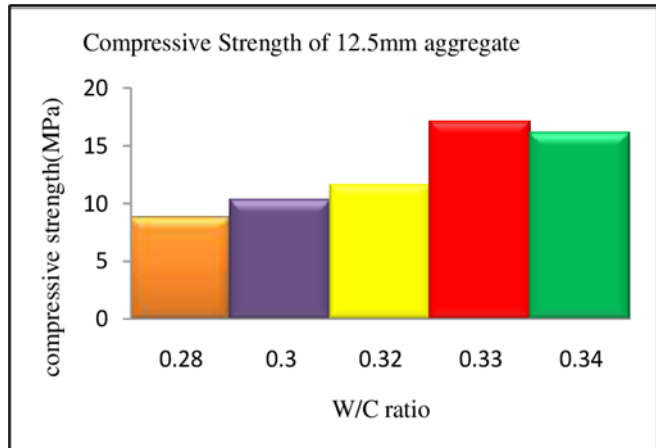


Figure 4: Variation in compressive strength of Pervious Concrete with various water-cement ratios [5].

Furthermore, Pervious Concrete tends to fail at the interface between aggregates, water, and cement. Striking a balance between water and cement is just as important since making the paste stronger does not necessarily result in the overall increased compressive strength. Figure 4 indicates that to achieve the highest compressive strength, water to cement has to be 0.33, this is contrary to the studies by [32] where for optimal strength w/c ratio varied between 0.30 - 0.35.

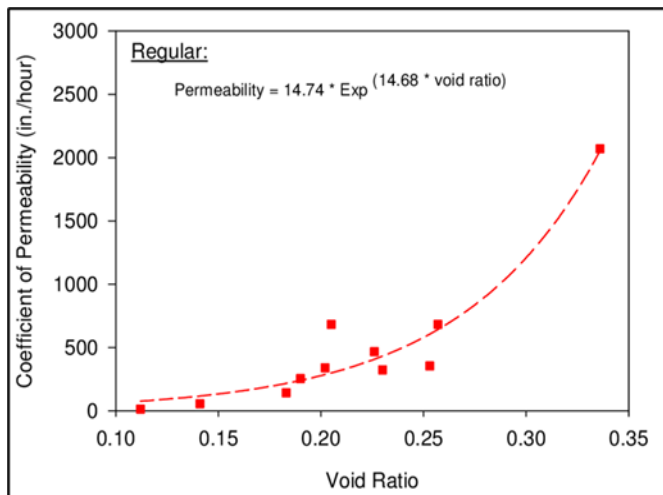


Figure 5: Relationship between pervious concrete void ratio and permeability for all mixes placed using regular compaction energy, (John T. Kevern)

### 2.13. Acceptable Void Ratios In Pervious Concrete

Figures 5 and 6 illustrate an exponential relationship between Permeability and Void Ratio, and both factors are directly related regarding Pervious Concrete. In other words, an increase in Void Ratio yields an increase in Permeability. In terms of conducted research, Pervious Concrete’s void ratio typically ranges from 15 to 30%. Previous studies have illustrated that a range of 15 to 30% in voids provides acceptable levels of compressive strength, hydraulic conductivity, and permeability results. When the void ratio is lower the unit weight is therefore higher which results in higher compressive strength. Subsequently, the higher the voids

content, the more likelihood the mix will pose a lower compressive strength. The following theoretical Figure 5 and 6 illustrates the above-mentioned concept.

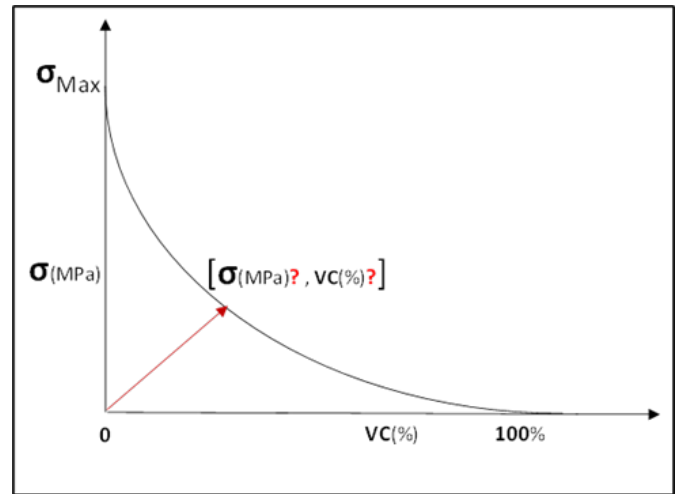


Figure 6: Theoretical Compression Strength (MPa) and Voids Content (VC%)

### Material Preparation and Handling

This section of the research methodology demonstrates how research material was handled during placing, mixing, and curing. The preparation process involved the following activities in a numerical sequence:

#### a) Sieving and Weighing

Material was batched according to the aggregate size. A large enough quantity was taken from the batch, cleaned to remove any impurities, material was dried and weighed to sample size then sieved with the shaker in descending size order. The largest sieve was placed at the top and smallest at the bottom, the results captured on each sieve were recorded. Percentages and cumulative percentages on each sieve were calculated and recorded.

#### b) Moulds Preparation

Mould sizes used were 150mm x 150mm x 150mm laboratory moulds that were cleaned with the wire brush and oiled.

#### c) Concrete mixing

Mixing was done according to the standard SANS 5861- (2006) using electric mixing drum. One adjustment was made where 5% of the cement was allowed to mix for a minute with coarse aggregate to all coating and improve bonding. The rest of the material was added and mixed for 3 minutes. A consistency check was performed using a ball method for every mix.

#### d) Concrete Casting and Curing

Specimens were casted according to SANS 5861-3 (2006). To determine the impact of compaction, samples were cast with and without compaction on the 3 sets of coarse aggregate sizes (7.1mm, 9.5mm and 19mm) dolomite rock. Different energies were administered in 3 layers (15, 20, 25) blows. Compacting Pervious Concrete is discouraged however there have been



studies that found improved strength with minimum compaction. A tamping rod was used for compacting and levelling.

The curing process was according to SANS 5861- 3 (2006), concrete moulds were placed in a curing bath with heat circulating pump, which kept the temperature at  $23 \pm 2^{\circ}\text{C}$  constant throughout the bath.



Figure 7: Laboratory sieved and weighed Material and labelled according to size.



Figure 8: Weighed and sorted material according to the mix design.



Figure 9: Electric Concrete Mixer



Figure 10: Ball Method (PC consistency check)



Figure 11: Cleaned and Oiled Mould.



Figure 12: Compacting with a rod.





Figure 13: Compacted and Levelled Samples



Figure 16: De-Moulded and Placed in a temperature-controlled bath.

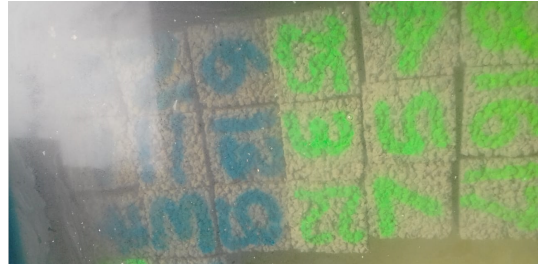


Figure 14: Labelled and Color Coordinated Samples I



Figure 17: Samples ready for Testing



Figure 15: Labelled and Color Coordinated Samples II



Figure 18: Sample Testing

## **Test Methods for Mechanical Properties of Aggregates**

### *2.14. Sieve Analysis*

The Standard SANS 201:2008 was utilized for sieve analysis for fine aggregates. Dry material was proportioned and placed in an oven at a temperature of 100 degrees Celsius, then weighed and recorded. Wet sample weight was weighed and recorded, extreme care was taken to wash and decanting until the water was clear. Fine material was dried and sieved again to calculate sieve analysis and fineness modulus.

Sieve analysis for coarse aggregates was performed in accordance with the standard SANS 201 (2008). Aggregates were washed and dried overnight. 19mm coarse aggregates was sieved through a couple of sieves bigger and smaller sieve sizes and only aggregates retained at 19mm sieve were utilized. 9.5mm coarse aggregates were sieved through a couple of sieves bigger and smaller sieve sizes and only aggregates retained at 9.5mm sieve were utilized. 7.5mm coarse aggregates was sieved through a couple of sieves bigger and smaller sieve sizes.

### *2.15. Bulk and Relative Density*

The standard SANS 3001: AG20;2014 for coarse aggregates bulk densities, relative densities and water absorption was used in determining Bulk and Relative Densities of the aggregates. The standard procedure was followed. Aggregates were first sorted according to size and proportioned according to the sample mass stated in the standard. The aggregates were then washed off to remove the dust, and then soaked in portable water for 24hrs. The visibly dried yet damp aggregates mass was recorded. Submerged in wire basket and unstrapped all bubbles and recorded the wet mass were recorded. The drained and dried aggregates were oven dried, and the mass recorded. From the above recordings the bulk and relative density were calculated.

### *2.16. Test Methods For Trial Mixes*

Several trial mixes were made so as to determine which materials will have an impact on Pervious Concrete strength. First a control Pervious Concrete was developed then additional materials were introduced one at a time in different dosages informed by the literature review chapter 2. Later a combination of the best-performing materials was assessed. To develop a strong Pervious Concrete mix, testing had to be classified into different testing phases, whereby different materials (fine aggregates, tyre rubber, steel fibers) discussed in literature review chapter, their content in terms of volumetric percentages, different combinations, and performance thereof, had to be explored. Consequently, materials were eliminated as they proved to weaken the Pervious Concrete mix. And on the contrary, admixtures enhancing strength were adopted and optimized further.

### *2.17. Fresh And Hardened Concrete Test*

Slump test is not effective for Pervious Concrete because Pervious Concrete has water to cement ratio restrictions that are so low as stated in the literature review. Pervious Concrete is zero slump concrete. Thus, to assess its workability the balling method was used see Figure 10.

### *2.18. Compressive test*

Compressive strength test of hardened concrete was done in accordance with Standard SANS 5861- 3-2006, The compressive strength test was performed for 7 days and 28 days.

### *2.19. Void content*

Voids Ratio, quite simply put, is the ratio between voids (or empty spaces) to the volume of solid concrete mass. The void content of Pervious Concrete is crucial as it directly affects permeability as well as compressive strength. To determine the total volume in the specimen was calculated as follows:

$$\text{Void content (\%)} = \frac{T-D}{T} \times 100$$

Where,  $D = (M_c - M_m) / V_m$  (Density)

$M_c$  = mass of measure filled with concrete

$M_m$  = net mass of concrete by subtracting mass of measure

$V_m$  = volume of measure

$T = M_s / V_s$  (Theoretical Density)

$M_s$  = total mass of materials batched

$V_s$  = total absolute volume of materials

### *2.20. Analysis Approach*

To ensure reliability and integrity of results, the experiments were carried out at SGS Matrolab, which is an independent SANAS accredited commercial laboratory located at Borksburg. The results were then compared to:

- I. Roads Asset Classification (TRH 26)
- II. Technical Design Specification (COTO)

There is currently no technical specification for Pervious Concrete. The current South African pavement design manuals were developed for conventional solid concrete, and not for pervious concrete. To this effect, all the material performance data, material behavior and algorithm modelling, were developed for conventional concrete, and not for any other "Green Solutions" such as Pervious Concrete. This critical limitation in research and development, particularly in South Africa requires that we rely on International Best Practice Research Guidelines. These guidelines place compressive strength criteria in determining which concrete class is applicable for which areas of pavement. The guidelines are derived from global research papers, which adopts the following synthesis in terms of International Best Practice:

- a) Sidewalks: 7 – 15Mpa
- b) Drainage Layers: 7 – 15Mpa
- c) Car Parks: 10 - 15Mpa
- d) Low Traffic Roads: 25 - 30Mpa

The above practice guidelines were used in this research in determining the suitability of the derived pervious concrete mix in one of the above pavement structures.

### **The experimentation was phased out in the following phases:**

- I. Phase I: A trial mix experiment to determine the appropriate aggregate size.
- II. Phase II: Determine the influence of compacting Pervious Concrete.



- III. Phase III: Investigate Pervious Concrete strength enhancing materials (Fine aggregates, Rubber, Fly Ash, and Steel Fibers)
- IV. Phase IV: Formulated a combination mix involving the best performing materials from Phase III

The material used to enhance the compressive strength of the Pervious Concrete in this study were based on the literature with the following parameters being kept constant according to the studies namely:

- I. In accordance to studies by [6; 29; 33] and Table 1 and 8, the Aggregates to Cement (A/C) was taken as 4:1.
- II. Water to Cement ratio (W/C) was taken as 0.32-0.35 as per [10,31,32] and Figure 4. The referenced studies achieved a balance between permeability and compressive strength with that water to cement ratio.
- III. Minimum acceptable void content was taken as 15% in accordance with the following studies.
- IV. Three coarse aggregates sizes namely 7.1mm; 9.5mm and 19mm were assessed with and without compaction in accordance with [4, 10, 34] studies. The best performing aggregate best on strength was kept constant throughout the experiment [1, 5, 35]
- V. The compaction method was kept constant on all samples that were compaction. Proctor Hammer Compaction Method was used, whereby Layer 1 (Roughly 1/3 of specimen height) was compacted at 15 Blows, Layer 2 (Adding roughly at 2/3 of specimen height) followed by 20 Blows and Layer 3 (Adding roughly at full height of specimen) followed by 25 Blows, in each instant applying a consistent weight of 2.5 Kilograms as per MOD AASHTO Compaction specification.

**Strength optimization**

Fly Ash, Fine aggregates, Rubber Crumbs, and Steel Fibers were material chosen for the study to enhance Pervious Concrete compressive strength; the range of determination was based on literature findings as stated below:

- I. Fly ash mix was added to the control mix 2% to 8% was in substitution of the proportioned cement weight.
- II. Fine aggregates mix was added to the control mix 0% to 7% was in substitution of the proportioned coarse aggregate weight.
- III. Waste tyre rubber was added to the control mix 0% to 7% was in substitution of the proportioned coarse aggregate weight (2,5 % to 10%) [8]
- IV. Stainless steel was added to the control mix 0% to 7% was in substitution of the proportioned coarse aggregate weight (1% - 2%) [22, 28].

With the best performing quantity from each additional material combination trial mix was developed create a Trial mix of (fly ash + Fine aggregates + Rubber) and (fly ash + Fine aggregates + Rubber+ Steel fiber) mixes. Table 10 indicates the mix design of the samples.

**3. Results**

**3.1. Aggregates Properties**

Aggregates properties of the dolomite rock coarse aggregates (7.1mm, 9.5mm and 19mm), river fine aggregates (fine aggregate) were determined, and aggregate grading analyzed. Table 9 below shows the aggregates properties of the aggregates used.

Fine and coarse aggregates were used, and a grading analysis was determined, material was sieved to determine the proportion using the standard sieve sizes. Figure 19 and Table 11 indicates good grading as it shows more than 90% passed a 4.75µm sieve and only between 5 and 25% passed thru a 150-µm sieve.

The fine aggregates were in accordance to SANS 1083: 2006 standard at 20% and coarse aggregates at 10% level. According to SANS 1083:2017, the acceptable FM lower limit is 1.2 and as presented below was acceptable which was important for particle size distribution of concrete fine aggregates. According to (Concrete-Institute, 2009) fine aggregates grading has major impact on workability, cohesiveness, and bleeding of fresh concrete.

Table 9: Aggregates Properties

Parameter	Dolomite Rock Coarse Aggregates			Units
	19mm Stone	9.5mm Stone	7.1mm Sand	
Loose Bulk Density (LBD)	1536	1532	1888	kg/m <sup>3</sup>
Compacted Bulk Density (CBD)	1664	1626	2127	kg/m <sup>3</sup>
Apparent Density (Relative Density)	2840	2810	2810	kg/m <sup>3</sup>
Water Absorption	0,76	0,8	1,32	%
Flakiness Index	5	-	-	-

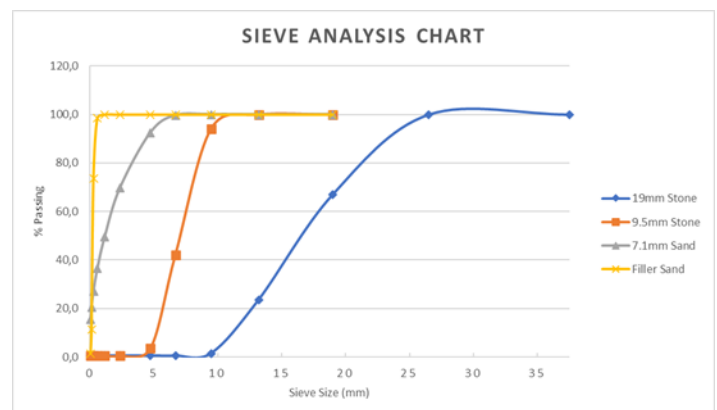


Figure 19: Sieve Analysis Chart

**3.2. Phase I Results**

Phase I determined aggregate size suitable to enhance and improve Pervious Concrete compressive strength and assess the behavior of dolomite rock aggregates of the chosen sizes.



Table 10: Mix Design

Mix Name	Aggregate	Water to Cement Ration	Aggregate to Cement	Aggregate Content (kg)	Cement Content (kg)	Water Content (kg)	Sand (kg)	Fly Ash (kg)	Tyre Rubber (kg)	Steel Fiber (kg)	Number of Cubes (150mm*150mm*150mm)					
PC 7.1 (1-3)	7.1mm	0.32-0.34	4 to 1	32	8	2,6	N/A	N/A	N/A	N/A	6					
PC 9.5 (1-3)	9.5mm										6					
PC 19 (1-3)	19mm										6					
PCC 7.1 (1-3)	7.1mm			32	8	2,6	N/A				N/A	6				
PCC 9.5 (1-3)	9.5mm											6				
PCC 19 (1-3)	19mm											6				
PCS/0 - 1	9.5mm			32	8	2,6	0				N/A	6				
PCS/2 - 1												31,36	8	2,6	0,64	6
PCS/5 - 1												30,4	8	2,6	1,6	6
PCS/7 - 1				29,76	8	2,6	2,24				6					
PC/2FA - 1				32	7,84	2,5	0,16				6					
PC/4FA - 1				32	7,68	2,5	0,32				6					
PC/8FA - 1				32	7,36	2,4	0,64				6					
PC/10FA - 1				32	7,2	2,3	0,8				6					
PC/2,5RB - 1				31,2	8	2,6	0,8				6					
PC/5RB - 1				30,4	8	2,6	1,6				6					
PC/7,5RB - 1				29,6	8	2,6	2,4				6					
PC/10RB - 1				28,8	8	2,6	3,2				6					
PC/SF-1				31,68	8	2,6	0,32				6					
PC/SF-2				31,36	8	2,6	0,64				6					
PC/C1 - 1				32	7,84	2,5	0				0,16	0,8	N/A	6		
PC/C2 - 2				31,36	7,68	2,5	0,64				0,32	1,6	N/A	6		
PC/C3 - 3				30,4	7,36	2,4	1,6				0,64	2,4	N/A	6		
PC/C4 - 4				29,76	7,2	2,3	2,24				0,8	3,2	N/A	6		
PC/C4 - 5				31,36	7,92	2,5	0,32				0,08	0,32	N/A	6		
PC/C4 - 6				30,72	7,84	2,5	0,64				0,16	0,64	N/A	6		
PC/C4 - 7				31,68	7,92	2,5	0,32				0,08	N/A	N/A	6		
PCC/C4 - 8				31,36	7,84	2,5	0,64				0,16	N/A	N/A	6		
				715,2	179,68	60,1	8,96	3,84	8,0	1,92	168					

Table 11: Sieve Analysis

19mm Stone Grading (Dolomite Stone)					9.5mm Stone Grading (Dolomite Stone)					7.1mm Sand Grading (Dolomite Super Sand)					Filler Sand Grading				
Sieve	Mass (g)	Mass Retained %	Cummulative Mass Retained %	Cummulative Mass Passing %	Sieve	Mass (g)	Mass Retained %	Cummulative Mass Retained %	Cummulative Mass Passing %	Sieve	Mass (g)	Mass Retained %	Cummulative Mass Retained %	Cummulative Mass Passing %	Sieve	Mass (g)	Mass Retained %	Cummulative Mass Retained %	Cummulative Mass Passing %
Dry Mass	2,525				Dry Mass	2,018				Dry Mass	1,012				Dry Mass	0,502			
37,5	0	0	0	100,0		0	0	0	100,0										
26,5	0	0	0	100,0		0	0	0	100,0										
19	0,835	33,1	33,1	67,0	19	0	0	0	100,0	19	0	0	0	100,0	19	0	0	0	100,0
13,2	1,098	43,5	76,5	23,5	13,2	0	0	0	100,0	13,2	0	0	0	100,0	13,2	0	0	0	100,0
9,5	0,555	22	98,5	1,5	9,5	0,118	5,8	5,8	94,2	9,5	0	0	0	100,0	9,5	0	0	0	100,0
6,7	0,025	1	99,5	0,5	6,7	1,05	52	57,9	42,1	6,7	0,004	0,4	0,4	99,6	6,7	0	0	0	100,0
4,75	0	0	99,5	0,5	4,75	0,778	38,6	96,4	3,6	4,75	0,072	7,1	7,5	92,5	4,75	0	0	0	100,0
2,36	0	0	99,5	0,5	2,36	0,062	3,1	99,5	0,5	2,36	0,232	22,9	30,4	69,6	2,36	0	0	0	100,0
1,18	0	0	99,5	0,5	1,18	0	0	99,5	0,5	1,18	0,202	20	50,4	49,6	1,18	0	0	0	100,0
0,6	0	0	99,5	0,5	0,6	0	0	99,5	0,5	0,6	0,134	13,2	63,6	36,4	0,6	0,008	1,6	1,6	98,4
0,3	0	0	99,5	0,5	0,3	0	0	99,5	0,5	0,3	0,094	9,3	72,9	27,1	0,3	0,125	24,9	26,5	73,5
0,15	0	0	99,5	0,5	0,15	0	0	99,5	0,5	0,15	0,066	6,5	79,4	20,6	0,15	0,312	62,2	88,6	11,4
0,075	0	0	99,5	0,5	0,075	0	0	99,5	0,5	0,075	0,05	4,9	84,4	15,6	0,075	0,05	10,0	98,6	1,4
Pan	0,005 + 0,008	0,5	100		Pan	0,000 + 0,001	0,5	100		Pan	0,01 + 0,14	15,6	100		Pan	0,003 + 0,00	1,4	100	
Total	2518				Total	2,008				Total	0,864				Total	0,498			
		FM	9,05				FM	6,58				FM	3,05				FM	1,17	

Aggregate size used (7.1mm, 9.5mm and 19mm). Figure 20 below presents compressive strength in the absence of compaction.

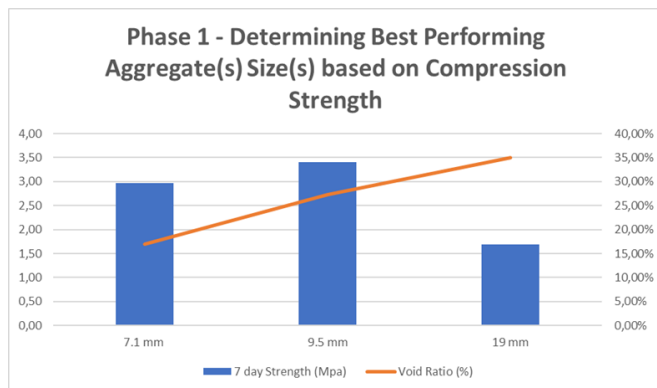


Figure 20: Determining best performing aggregate size(s), 0% Compaction.

Above Figure 20 indicates the relationship between the void content and compressive strength, that the smaller the coarse aggregate size the smaller the void space and the higher the compressive strength and 7.1mm, 9.5mm and 19mm aggregate size were proof with 2.97MPa, 3.4MPa, 1.7MPa respectively in 7 days. The results of 9.5mm aggregates highest strength within acceptable void content was consistent with literature. Smaller aggregate sizes (7.1mm and 9.5mm) yielded better compressive strength results when compared to the 19mm stone sizes, reason being the smaller aggregates have a larger number of contact points (cement fillets) per unit volume. In addition, these results correlate with reduced voids content, reconciling findings by other researchers in literature section.

### 3.3. Phase II Results

Phase II determined the influence of compaction on the chosen three aggregate sizes. The aim of this experiment was to determine the effect of compaction on compressive strength and the void content of Pervious Concrete. A trial mix experiment was carried out to establish the best performing aggregate size between the 7.1mm, 9.5mm and 19mm when compaction was included in all samples.

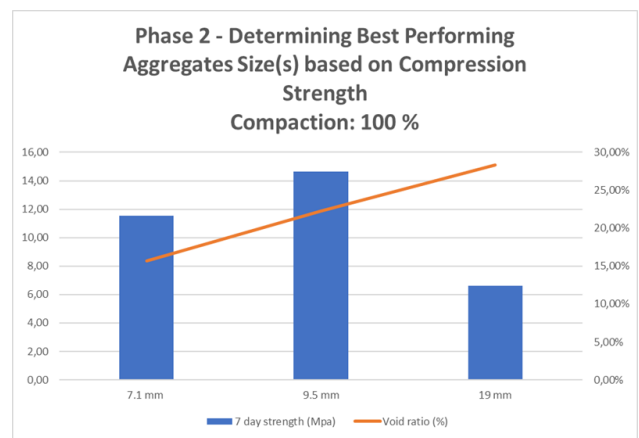


Figure 21: Determining best performing aggregate size(s), 100% Compaction.

The results in Figure 21 indicated an increase in strength parameter, particularly on the 9.5mm aggregate stone from 3.4MPa to 14.63MPa with void content of 27,26% uncompacted to 22,28% compacted. This indicates compaction effort creates an improved interlocking of aggregates, leading to improved cohesion properties and compression strength. The resultant Voids Content (%) is at 22,28%, sitting comfortably in the range of an acceptable Voids Content of 15 - 30%, as per published research mentioned in chapter 2.

### 3.4. Phase III Results

- Fine Aggregates

An experiment was carried out with only fine aggregates inclusion in varying quantities to determine exactly the quantity required to enhance a dolomitic rock Pervious Concrete. The fine aggregate mix contained 9.5mm stone aggregates and river fine aggregates, at a water/cement (W/C) ratio of 0.30 – 0.35 range, as per published research for Pervious Concrete. A total of 24 samples of Fine Aggregate Mix were mixed, with varying volumetric contents (0%, 2%, 5% and 7%) of River Fine aggregatess. Figure 22 below presents the 28days compressive strength results and void content.

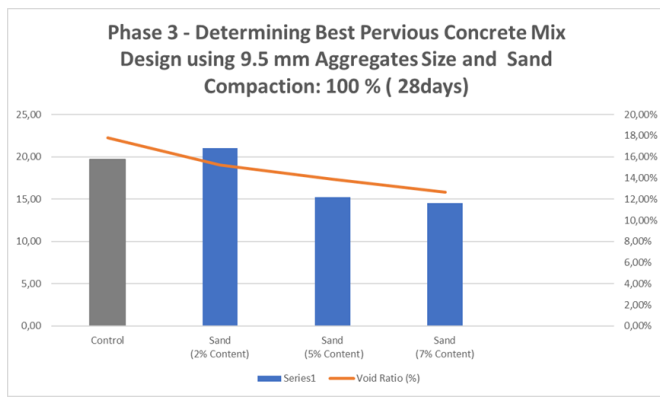


Figure 22: Fine aggregates Aggregate Mix (9.5mm Stone and River Fine aggregates Variations)

Partial inclusion of fine aggregates did increase compressive strength of Pervious Concrete from 14.6MPa without fine aggregates to 16.8MPa with 2% of river fine aggregates inclusion in 7 days and 19.7MPa to 21.5MPa respectively in 28day results, by including fine aggregates increases the amount of mortar which increases the contact area between aggregates, previous studies mentioned in chapter concurs with research findings.

The improvement in strength was also attributed to void reduction as discussed in Phase II that void content reduction results in increase compressive strength in Pervious Concrete. Another contributor to compressive strength increase was particle distribution, by adding fine aggregates it increased the surface area which improved bonding. Pervious Concrete comprises of large aggregates with limited fines, adding fine aggregates fills up the void spaces and fine aggregates blends with the cement paste which resulted in a well graded and interlocking mix. Figure 22 shows a decrease in compressive strength beyond 2%. At 5% partial fine aggregates compressive strength was 15.2MPa and 14.5MPa with 7% fine aggregates.

- Waste Rubber

A combination of 9.5mm aggregates and Rubber Crumbs were mixed, with varying volumetric contents (2.5%, 5%, 7.5% and 10%) of River Fine aggregates. Figure 23 represents the results.

Rubber crumbs are larger in size compared to fine aggregates and have irregular shapes which resulted with a mix that was not well graded, this resulted in unsatisfactory compressive strength results. The rubber crumbs affected the bonding between the coarse aggregates and the cement paste, compared to fine aggregates and coarse aggregates rubber has less stiffness by adding rubber crumbs to the Pervious Concrete mix reduces its stiffness therefore affecting its compressive strength. Evidently from the above results, an increase in rubber content in Pervious Concrete decreases compression strength. Scientifically, this makes sense as the higher the elasticity between the interlocking interface of aggregates, the lower the cohesion properties, leading to reduced compression strength. With concrete being classified as rigid pavement and granular materials being classified as flexible pavements, adding flexibility or elasticity to traditional classified rigid pavement (concrete), is a scientific contradiction in pavement design methodology. Elasticity thereof, speaks to flexibility in pavement concrete, and could possibly lead to further

research in the development of flexible pavements using Pervious Concrete.

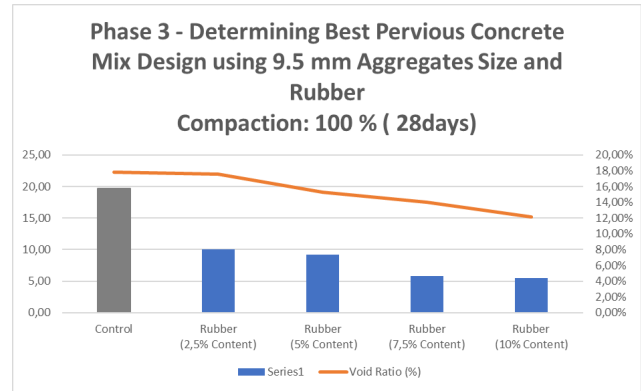


Figure 23: Fine aggregates Aggregate Mix (9.5mm Stone and Rubber Crumbs Variations).

- Fly Ash

Combination of 9.5mm aggregates and Fly Ash were mixed, with varying volumetric contents (2%, 4%, 8% and 10%) of River Fine aggregates. Figure 24 presents the 28days compressive strength results and the void contents.

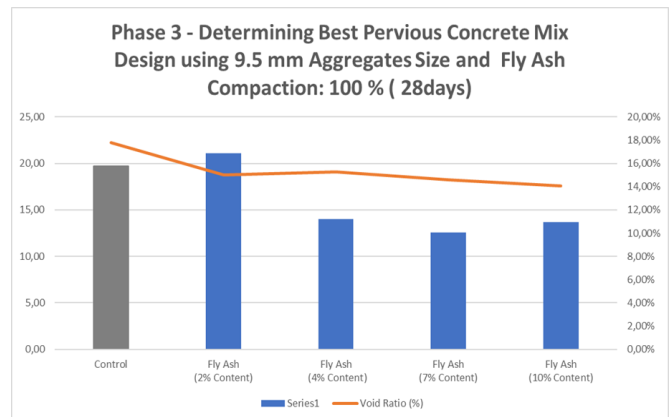


Figure 24: Cementitious Aggregate Mix (9.5 mm Stone and Fly Ash Variations)

Partially replacing cement with fly ash resulted in less water requirement and improved workability of the Pervious Concrete. The reduction in water led to a decrease in interconnectivity. Figure 24 results illustrated an increase in compressive strength when 2% of fly ash was included. Fly ash improving Pervious Concrete was a bonus due to cement cost hikes, though it was small quantities in the experiment batch, in road construction 2% will be a great cost saving. Pervious Concrete on its own was 19,69MPa in 28day when Fly Ash was included it slightly increased to 21.70MPa in 28 days. Beyond 2% a serious decline in compressive strength was noticed which didn't make sense. In [36], the authors discovered that the more fly ash added on the mix lowered compressive strength.

- Steel Mix

This experiment was to determine the effect of steel on material strength performance. Figure 25 below are the results of using steel fiber to enhance Pervious Concrete compressive

strength, with 19 MPa at 7 Days curing strength expected of 29 MPa after 28 Days of curing however that was not the case. Following the 28 Day strength, the 1% and 2% steel fibre reinforced concrete, has practically yield similar results, roughly at 23MPa compressive strength. This strength also represents a distorted concrete strength curve, with results somewhat lower than traditional solid concrete in strength yield, representing 74% concrete strength yield than traditional solid concrete, when following traditional concrete strength curve. Though the results indicated promising results for compressive strength its void content compared to the control dropped from 17.81% to 14.79% when 1% steel fiber was included

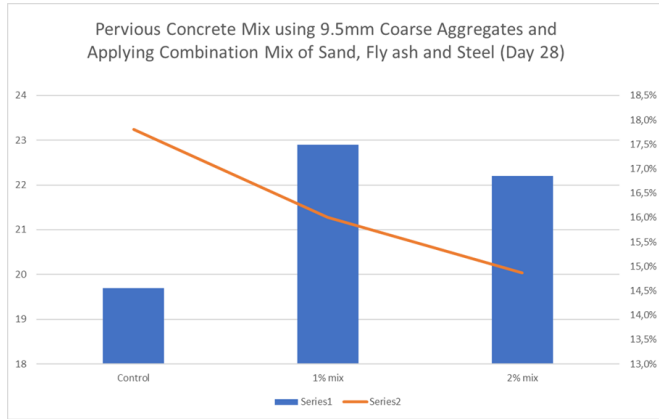


Figure 25: Fine aggregates Aggregate Mix (9.5mm Stone and Stainless Fiber Variations)

• **Strength Optimization**

Figure 26 below indicate the results of the highest performing material, fine aggregates and fly ash results were the highest individually, therefore it made sense to combine the two highest strength material when mixed with standard Pervious Concrete. It was by doing so that the highest compressing strength of 25MPa was achieved with an acceptable range of 15,61% of void content.

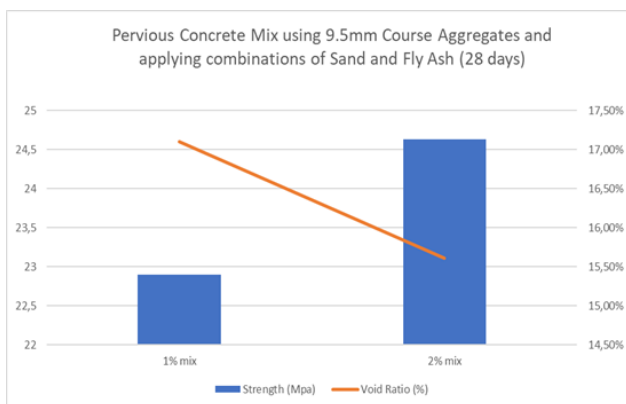


Figure 26: Pervious Concrete Mix using 9.5 mm Coarse Aggregates and applying combinations of River Fine aggregates, Fly Ash, and Steel 28 Days

**4. Conclusion**

The following results are based on the 28 days **compressive strength and void content** of different enhancing materials when using dolomitic rock:

- **Compaction** - Compaction has influence on Pervious Concrete, compressive strength improved though reducing the void content. A 7 day Trial was done with and without compaction which resulted with Non compacted pervious concrete strength of 3.4MPa and 14.6MPa compacted, this affected void content from 27% to 22%.
- **Fine Inclusion** - Inclusion of fines improved compressive strength; the highest strength was achieved when 2% fine aggregates volume of coarse aggregates as substituted. 28days compressive strength results were 19.7MPa without fines the increased to 21.5MPa with inclusion of fines. The void content was 18% to 15% this resulted from inclusion of both compaction and fines. Beyond 2% substitution compressive strength reduced which could have been associated with the restriction of water and cement.
- **Waste Tyre rubber** - waste tyre inclusion did not have positive results considering that pervious concrete compacted without any enhancers had the compressive strength of 19.6MPa however with the tyre it yielded 5.8MPa.
- **Fly Ash** - The following was drawn from fly ash inclusion, the compressive strength slightly increased from 19.7MPa to 21.1MPa, the void content was %18 then 17 % with fly ash inclusion.
- **Steel fiber** - Steel fiber did enhance pervious concrete compressive strength; this was studied by comparing the compressive strength results of the control specimen and steel fiber specimens. All samples were cured for 7 days and 28 days, the control compressive strength was 20MPa the steel fiber reinforced pervious concrete was 23MPa. The highest compressive strength was achieved when 1% partial replacement of coarse aggregates.
- **Combination Mix** - Based on the results obtained, best performing materials were combined which were fine aggregates and fly ash. The pervious concrete with a combination of 2% Fly Ash cement replacement and 2% Fine aggregates inclusion gave the highest compressive strength of 25MPa with a corresponding void content of 15%.

**5. Recommendations**

The following recommendation are derived from the current study.

- Prototypes, such as beams and scaled large slabs to determine flexural and or tensile strength, resistance, noise limits and cylinders to determine permeability are necessary.
- The study focused on improving compressive strength whilst remaining within acceptable void content however permeability as well as porosity are recommended.
- To investigate other locally available coarse aggregates, as the dolomite rock didn't reach the expected but has potential.
- Develop Technical Design Specification for Pervious Concrete in COTO
- Develop Design Criterion for the use of Pervious Concrete (Enforcement) and
- Align COTO Specification to TRH 26 Traffic Class Specification.



- Incorporate Pervious Concrete in the Drainage Manual of South Africa.
- Develop a Labour-Intensive COTO Specification for the Construction of Pervious Concrete, for the South African Context (Skills Development).
- Monitor performance of Pervious Concrete against Traffic and Drainage Demands and develop reliable database for continuous monitoring, research, and development.

## 6. Acknowledgement

The author wishes to acknowledge Tshwane University of Technology for the support.

## References

- [1] E Martinez, V Valeri, J Hernandez, D Fresno, Proposal of a New Porous Concrete Dosage Methodology for Pavements, *12(19) Materials*, 2019. <https://doi.org/10.3390%2Fma12193100>
- [2] M Offenberg, D Wade, C Weiss. Report on Pervious Concrete. United States of America: American Concrete Institute; 2010, <https://www.concrete.org/publications/internationalconcreteabstractsportal/m/details/id/51663557>
- [3] A Bonicelli, G Martinez, A. L Guillermo, F. Pumarejo, Improving Pervious Concrete Pavements for Achieving More Sustainable Urban Roads. *ScienceDirect*, **161**, 1568-1573, 2016, <https://doi.org/10.1016/j.proeng.2016.08.628>
- [4] F Yu, D Sun, & Hu. Study on the pores characteristics and permeability simulation of pervious concrete based on 2D/3D CT images. *Construction and Building Materials*, **200**, 687-702, 2019, <https://doi.org/10.1016/j.conbuildmat.2018.12.135>
- [5] M Seslija, J. D Radovic., & D Kukaras, Possibilities of Pervious Concrete Application in Road Construction. *International Standard*, 2017. <http://dx.doi.org/10.17559/TV-20160524162507>
- [6] P. D Tennis, M. L Leming, & D. J Akers, Pervious Concrete Pavements, Technical Report, Portland Cement Associates, 2004, <https://www.researchgate.net/publication/242722509>
- [7] N. Jin, 2010. Fly Ash Applicability in Pervious Concrete, The Ohio State University. <https://doi.org/10.1016/j.cscm.2021.e00502>
- [8] M Gesoglu, E Guneyisi, G Khoshnaw, S Ipek. Investigating properties of pervious concretes containing waste tire rubbers. *Construction and Building Materials*, **73**, 19-24, 2014. <https://doi.org/10.1016/j.conbuildmat.2014.12.073>
- [9] Y. J. Kim, A. Gaddafi, I. Yoshitake. Permeable concrete mixed with various admixtures. *Materials and Designs*, **100**, 110-119, 2016, <https://doi.org/10.1016/j.matdes.2016.03.109>
- [10] T Joshi, Evaluation of mechanical properties of pervious concrete with changing cement content. *Research Gate*, pp. 2-7, 2016, <http://dx.doi.org/10.1201/b20013-9>
- [11] K Cosic., L Korat, V Ducman, I. Netinger, Influence of aggregate type and size on properties of pervious concrete. *Construction Building Material*, **78**, 69-79, 2014. <https://doi.org/10.1016/j.conbuildmat.2014.12.073>
- [12] K. S Elango, & V Revathi, G Fal Binder Pervious Concrete. *Construction and Building Materials*, **140**, 91-99, 2017. <https://doi.org/10.1016/j.conbuildmat.2017.02.086>
- [13] A. K. Chandrappa, K. P. Biligiri, Pervious Concrete as a Sustainable Pavement Material- Resesrch Findings and Future Prospects: A state-of-the-art review. *Construction and Building Materials*, **111**, 262-274, 2016. <http://dx.doi.org/10.1016/j.conbuildmat.2016.02.054>
- [14] C Lian, & Y Zhunge, Optimum mix design of enhanced permeable concrete – An experimental investigation. *Construction and Building Materials*, **24**, 2664-2671, 2009, <https://doi.org/10.1016/j.conbuildmat.2010.04.057>.
- [15] A. Azada, S. F Mousavi, H. Karami, S. Farzin. Using Waste Vermiculite and Dolomite as Eco Friendly Additives for Improving the Performance of Porous Concrete. *Engineering Journal*, **22(5)**, 2018. <http://dx.doi.org/10.4186/ej.2018.22.5.87>.
- [16] Department of Public Works, appropriate development of infrastructure on dolomite guidelines for consultants, South Africa. 2003.
- [17] B Haung, H Wu, X Shu, G Burdette, Laboratory evaluation of permeability and strength of polymer- modified pervious concrete. *Construction and Building Materials*, **24**, 818-823, 2010. <https://doi.org/10.1016/j.conbuildmat.2009.10.025>.
- [18] J Yang, & G Jiang, Experimental study on properties of pervious concrete pavement materials. *Cement and Concrete Research*, **33**, 381-386, 2003, [https://doi.org/10.1016/S0008-8846\(02\)00966-3](https://doi.org/10.1016/S0008-8846(02)00966-3).
- [19] Y Aoki, R. S Ravindrarajah, H Khabbaz, Properties of pervious concrete. *Road Materials and Pavement Design*, 2012, doi:10.1080/14680629.2011.651834
- [20] H Wang, Investigation on the mechanical properties and environmental impacts of pervious concrete containing fly ash based cement aggregate. *Construction and building materials*, **202** 387-395, 2019. <http://dx.doi.org/10.1016/j.conbuildmat.2019.01.044>
- [21] B. Han, X. Yu, J. Ou. Self-Sensing Concrete in Smart Structures, 2014, <https://doi.org/10.1016/C2013-0-14456-X>.
- [22] K. H. Tan, E. Lim, T. F. Fwa, 2015. Steel fiber-reinforced pervious concrete for urban roads, Nepal: *Civil Engineering*, National University of Singapore. [tankh@nus.edu.sg](mailto:tankh@nus.edu.sg).
- [23] Lin, W. & Yoda, T., 2017. Bridge Engineering Classification, Design Loading, and Analysis Methods, 2017, [https://www.researchgate.net/publication/328019943\\_Bridge\\_engineering\\_Classifications\\_design\\_loading\\_and\\_analysis\\_methods](https://www.researchgate.net/publication/328019943_Bridge_engineering_Classifications_design_loading_and_analysis_methods).
- [24] Channappa: Reddy, P. K. Experimental Study On Steel Fiber-Reinforced Pervious Concrete, **6(8)**, 441-447, 2019, *International Research Journal of Engineering and Technology (IRJET)*.
- [25] Lee, M. et al., Mechanical Properties of High-Strength Pervious Concrete with Steel Fiber or Glass Fiber. *MDPI*, 1-18 2022. <http://dx.doi.org/10.3390/buildings12050620>.
- [26] C Ng, et al. Properties of Modified High Permeable Concrete with a Crumb Rubber. *The Open Civil Engineering Journal*, 2019. <http://dx.doi.org/10.2174/1874149501913010082>
- [27] J. Farnoosh, M. Khorram, G. Karimi, & N Hataf, Experimental investigation of mechanical properties of crumbed rubber concrete containing natural zeolite. *Construction and Building Materials*, **208**, 651-658, 2019. <http://dx.doi.org/10.1016/j.conbuildmat.2019.03.063>
- [28] K Chandramouli, S, N Chaitanya, A, M Mrunalini, G Hymavathi, Experimental studies on pervious concrete by using waste tires rubber as partial replacement of coarse aggregate, Chandramouli, K. et al., n.d. Experimental studies on pervious concrete by using waste tires rubber as partial replacement of coarse aggregate. | ISSN: **2456-3315**, 2022, <http://www.ijrti.org/>.
- [29] D. P. M Tripathi, S.M. Ashraf Hussain, P. Madhav. An Experimental Study on Pervious Concrete (Mix-Ratio, Strength and Porous Properties). *International Journal of Engineering Research & Technology (IJERT)*, **6(12)**, 2017, DOI:10.17577/IJERTV6IS120054
- [30] K Alalea, H. S. Wong, C. R Cheeseman. Clogging in permeable concrete: A review. *Journal of Environmental Management*, 2017 **193** 221-223, <http://dx.doi.org/10.1016/j.jenvman.2017.02.018>
- [31] Y. Joung, Z. Grasley, Evaluation and Optimization of Durable Pervious Concrete for Use in Urban Areas, Southwest Region University Transportation Center (U.S.), 2008, [https://rosap.ntl.bts.gov/view/dot/16416/dot\\_16416\\_DS1.pdf](https://rosap.ntl.bts.gov/view/dot/16416/dot_16416_DS1.pdf)
- [32] T. Mulyono, Anisha, 2019. Properties of pervious concrete with various types and sizes of aggregate, **276**, 2019, <https://doi.org/10.1051/mateconf/201927601025>.
- [33] T. Ahmed, S. Hoque, Study on pervious concrete pavement mix designs. **476(2020) 2020**, <http://dx.doi.org/10.1088/1755-1315/476/1/012062>
- [34] A. E. El-Maaty, Establishing a Balance Between Mechanical and Durability Properties of Pervious Concrete Pavement. *American Journal of Traffic and Transportation Engineering*, **13-15**. 2016, <http://dx.doi.org/10.11648/j.ajtte.20160102.11>
- [35] D Ramadhan, F Muslim, & L Sagita, Mechanical properties of porous concrete and design recommendation for low traffic road pavement in Indonesia. **794(2021)**, 2021, DOI 10.1088/1755-1315/794/1/012003
- [36] Sudhir Kumar Boddu, Investigating the Effect of Aggregate Size and Binder Material Proportion on Strength and Permeability of Pervious Concrete by Statistical Modeling. *Springer Nature Singapore*. 2022, [http://dx.doi.org/10.1007/978-981-19-2273-2\\_16](http://dx.doi.org/10.1007/978-981-19-2273-2_16)

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Enhancing the Network Anomaly Detection using CNN-Bidirectional LSTM Hybrid Model and Sampling Strategies for Imbalanced Network Traffic Data

Toya Acharya\*, Annamalai Annamalai, Mohamed F Chouikha

Electrical and Computer Engineering, Prairie View A & M University, Prairie View, Texas, 77446, USA

### ARTICLE INFO

#### Article history:

Received: 15 November, 2023

Revised: 17 January, 2024

Accepted: 17 January, 2024

Online: 20 January, 2024

#### Keywords:

Random Under Sampling

Random Over Sampling

Network Anomaly Detection

UNSW-NB15

NSL-KDD

Data imbalance

Convolutional Neural Network

CNN-BLSTM

Data Imbalance

### ABSTRACT

The cybercriminal utilized the skills and freely available tools to breach the networks of internet-connected devices by exploiting confidentiality, integrity, and availability. Network anomaly detection is crucial for ensuring the security of information resources. Detecting abnormal network behavior poses challenges because of the extensive data, imbalanced attack class nature, and the abundance of features in the dataset. Conventional machine learning approaches need more efficiency in addressing these issues. Deep learning has demonstrated greater effectiveness in identifying network anomalies. Specifically, a recurrent neural network model is created to recognize the serial data patterns for prediction. We optimized the hybrid model, the convolutional neural network combined with Bidirectional Long-Short Term Memory (BLSTM), to examine optimizers (Adam, Nadam, Adamax, RMSprop, SGD, Adagrad, Ftrl), number of epochs, size of the batch, learning rate, and the Neural Network (NN) architecture. Examining these hyperparameters yielded the highest accuracy in anomaly detection, reaching 98.27% for the binary class NSL-KDD and 99.87% for the binary class UNSW-NB15. Furthermore, recognizing the inherent class imbalance in network-based anomaly detection datasets, we explore the sampling techniques to address this issue and improve the model's overall performance. The data imbalance problem for the multiclass network anomaly detection dataset is addressed by using the sampling technique during the data preprocessing, where the random over-sampling methods combined with the CNN-based BLSTM model outperformed by producing the highest performance metrics, i.e., detection accuracy for multiclass NSL-KDD and multiclass UNSW-NB15 of 99.83% and 99.99% respectively. Evaluation of performance, considering accuracy and F1-score, indicated that the proposed CNN BLSTM hybrid network-based anomaly detection outperformed other existing methods for network traffic anomaly detection. Hence, this research contributes valuable insights into selecting hyperparameters of deep learning techniques for anomaly detection in imbalanced network datasets, providing practical guidance on choosing appropriate hyperparameters and sampling strategies to enhance model robustness in real-world scenarios.

### 1. Introduction

This study is extension version of the presented conference paper titled "Efficacy of CNN-Bidirectional LSTM Hybrid Model for Network-Based Anomaly Detection" [1] at the 2023 IEEE 13th Symposium on Computer Applications & Industrial Electronics (ISCAIE).

As technology undergoes rapid advancements, the transmission of information has transformed significantly,

adopting various methods such as wired, wireless, or guided networks. This evolution in network technology is pivotal to people's daily activities. Whether it's communicating with others, accessing online resources, or sharing information, the efficiency and security of these interactions depend heavily on the underlying network infrastructure. A system attains security when it effectively maintains the three essential notions of computer information security: Availability, Integrity, and Confidentiality (CIA). In essence, information security involves safeguarding information from unauthorized entities and protecting against

\*Corresponding Author: Toya Acharya, [tacharya@pvamu.edu](mailto:tacharya@pvamu.edu)

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj090107>

illegal access, use, disclosure, reformation, recording, or destruction of data. Confidentiality guarantees that the information is accessible only to individuals or systems. In information network technology, encryption methods and access controls prevent unauthorized users from gaining access to sensitive data during transmission. Information integrity guarantees that data remains unaltered during transmission. In the context of information network technology, this involves implementing mechanisms to detect and prevent unauthorized modifications to data, ensuring that the information received is the same as what was sent. Availability ensures that information and resources are available and accessible when needed. In an information resources and security environment, availability involves designing robust and reliable systems that can withstand potential disruptions, whether they are due to technical failures or malicious attacks. The overarching goal of information security is to safeguard information from unauthorized access and malicious activities. This includes preventing unauthorized individuals or systems from gaining access to sensitive data, ensuring that information remains unchanged and reliable during transmission, and guaranteeing that information and resources are available when needed.

Measures to achieve information security encompass a range of strategies, including encryption to protect data confidentiality, checksums or digital signatures to ensure data integrity, and redundancy and fault-tolerant systems to enhance availability. Additionally, access controls, firewalls, and intrusion detection are commonly utilized to fortify the security posture of networked systems, mitigating the risks associated with information resources.

A traditional network cannot be fully protected by relying solely on a firewall and antivirus software. These security measures identify predefined anomalous activities and establish the rule to prevent those unusual events by the cyber expert. In anomaly detection, outliers and anomalies are occasionally employed interchangeably. This approach finds extensive use across diverse domains, such as commercial, network attack detection, health systems monitoring, credit card fraud transaction detection, and identifying faults in mission-critical infrastructure systems. Anomaly detection is crucial in cybersecurity, providing robust protection against cyber adversaries. Ensuring safeguard network resources is essential to safeguard the organization from cyber threats.

Anomalies are categorized into point, contextual, and collective types based on the results generated by the detection method [2]. Point anomalies occur when a specific activity diverges from the typical rules or patterns. Contextual anomalies involve unusual patterns within a particular circumstance that consistently differ from numerous normal activities. Collective anomalies occur when a group of related instances exhibit anomalous behavior compared to the normal activity dataset.

Intrusion detection techniques can be broadly classified into two main types: Signature-based Intrusion Detection System (SIDS) and Anomaly-based Intrusion Detection System (AIDS). Anomaly detections, in contrast, are classified according to their origins, resulting in network-based and host-based intrusion/anomaly detection systems. Detecting anomalies in data is facilitated by employing labels to differentiate between normal and abnormal

occurrences. There are three fundamental approaches to detecting anomalies: supervised, semi-supervised, and unsupervised methods. In the supervised approach, the system is trained on labeled data, distinguishing between normal and anomalous instances. On the other hand, unsupervised methods detect anomalies without prior labeling, relying on deviations from established patterns. Semi-supervised techniques combine elements of both, using labeled and unlabeled data for training. AIDS overcomes the drawbacks of SIDS by utilizing ML, statistical-based, or knowledge-based methods to model normal behaviors. However, it's worth noting that anomaly-based detection may produce false results due to alterations in user habits.

Numerous traditional machine learning algorithms favor shallow learning methodologies, giving significant importance to feature engineering designed for smaller data. The feature engineering phase needs more processing time and domain expertise to create pertinent features and eliminate unrelated ones from anomaly detection algorithms. The effectiveness of anomaly detection is intricately tied to feature engineering and data preprocessing implementation. Traditional machine learning methods, characterized by simplicity, low resource consumption, and subpar performance in areas like vision, language processing, and image translations, underscore the limitations of these approaches.

CNN is predominantly employed for image signals, leveraging its architecture to effectively capture and analyze visual information. Individual neurons play a key role in reducing the dimensionality of the network's features in the lower layers of a CNN. These neurons are adept at identifying essential small-scale features within the images, including boundaries, corners, and variations in intensity. The CNN network links lower-level features to produce more complicated features in the upper layers, encompassing fundamental shapes, structures, and partial objects. The ultimate layer of the network amalgamates these lower features to generate the output results.

The functioning of a long short-term memory differs from that of a CNN due to its specific design to safeguard long-range info within a sequential order. Unlike CNNs, LSTMs are crafted to remember and store information over extended sequences, avoiding the loss of crucial details. In the case of BLSTM, an additional LSTM layer is incorporated, introducing a reversal in the information flow direction. This architectural enhancement addresses challenges related to vanishing gradients, ensuring more effective training by considering information from both forward and backward directions in the sequence.

Data imbalance, including network anomaly detection, is a common challenge in ML applications. In network traffic anomaly detection, data imbalance refers to the unequal distribution of normal and anomalous instances in the dataset for training the detection model. Anomalies in network traffic are typically rare incidents compared to normal activities, leading to imbalanced data.

The deep learning approach addresses issues found in conventional machine learning. The effectiveness of the deep learning-based anomaly detection algorithm relies on factors such as the NN architecture, #hidden layers, activation functions, batch



size, and the number of epochs utilized during DL model testing, training, and validation. The careful selection of these factors, including hyperparameters and the architecture of NN in deep learning, is crucial for enhancing the detection accuracy of network traffic anomaly detection. The essential selection of ML or DL models overcomes the class imbalance problem. The ensemble method, which combines more different individual models, requires longer training time and consumes more resources. The sampling method generates random data or deletes the random data based on the implemented sampling methods to create the balanced form of the final dataset, which is efficient in dealing with the imbalanced dataset.

## 2. Literature Review

The rapid increase of information and technology has led to widespread connectivity of numerous end terminals to the internet and networks. Those smart terminals contribute to generating substantial volumes of data, commonly referred to as big data. This huge flood of data is a valuable resource for analysis and insights. Machine learning and deep learning algorithms come into play to extract meaningful information from this vast data pool. The daily growth of big data presents difficulties for conventional machine learning algorithms, necessitating thorough feature extraction and discovery. DL substantially increases anomaly detection and model performance. Nevertheless, the dataset's attributes and features, hyperparameters in deep neural networks, and the structure of neural networks are pivotal elements that impact the efficacy of identifying anomalies in network-based IDS.

Conventional machine learning strongly relies on intricate and time-consuming feature engineering, often impractical for real-time applications. In the [3] study, the authors proposed an approach for payload classification utilizing CNN and RNN to detect attacks, achieving detection accuracies of 99.36% and 99.98% on the DARPA98 network data, respectively. CNN methods discern specific grouping patterns through convolution around input neighborhoods, while RNN works on sequences by calculating correlations between previous and current states. In another [4] study, class imbalance was handled utilizing a CNN with a Gated Recurrent Unit (GRU) hybrid model. To address the data class imbalance and feature redundancy, they used a hybrid sampling technique that integrates Pearson Correlation Analysis (PCA), repeated edited nearest neighbors, Random Forest (RF), and adaptive synthetic sampling. With the detection accuracies of 99.69%, 86.25%, 99.69%, and 99.65% on the NSL-KDD, UNSW\_NB15, and CIC-IDS2017 datasets, respectively, their CNN-GRU model performed better.

The research authors [5] proposed using an Adaptive Synthetic Sampling (ADASYN) technique in a DL-based network intrusion detection system to overcome dataset imbalance. On the NSL-KDD network data, they used an autoencoder to reduce dimensionality. The CNN-BLSTM hybrid DL method obtained the greatest F1 score (89.65%) and accuracy (90.73%). To address problems resulting from data in class imbalance and heterogeneous data distribution across various information sources, the research [6] used convolutional neural networks with federal transfer learning. The UNSW-NB15 multiclass network dataset produced an average detection accuracy was 86.85% for the model.

In [7], the researcher addressed data class imbalance on network datasets: NSL-KDD, KDD99, and UNSW-NB15 datasets using heterogeneous ensemble-assisted ML methods for binary and multi-class network intrusion detection. Using the NSL-KDD dataset, the model showed a 96.2% AUC and a true positive rate (TPR) of 94.5%.

The authors of [8] discovered that ML classifier performance increased with the decrease in target classes. Conventional ML approaches, such as Naïve Bayes, Random Forest, J48, Bagging, Adaboost, and BayesianNet, were used to investigate this idea on three network traffic-based intrusion datasets: KDD99, UNSW-NB15, and CIC-IDS2017\_Thursday.

In a study [9], the authors suggested a method for achieving network intrusion classification with low computing cost, creating a group of target classes based on the nature of network traffic. They created cluster characteristics for each group using K-means on the KDD99 network dataset, resulting in a detection accuracy of 98.84%. However, the intrusion detection model accuracy for user2root (U2R) is notably low at 21.92%, impacting overall performance. In [10], authors employed a hybrid method, combining CNN and LSTM, to enhance model classification accuracy, achieving 96.7% and 98.1% on CIC-IDS2017 and NSL-KDD network data, respectively.

In the study [11], CNN and LSTM combined to create a hybrid model was proposed to enhance network intrusion detection model facilities for advanced metering infrastructure through cross-layer features combination. This method achieved the highest network intrusion detection accuracy of 99.79% on NSL-KDD and 99.95% on KDD Cup99 but with limited user2root (U2R) detection capabilities. Similarly, in [12], authors employed a hybrid method of combining CNN and LSTM to improve model network intrusion detection by capturing additional network traffic data's spatial and temporal features.

In [13], the researchers implemented a hybrid technique based on the mean control of the CNN and BLSTM to address issues of conventional data pre-processing and imbalanced numerical distribution of class instances in the NSL-KDD, achieving the optimal detection accuracy of 99.10%. However, the accuracy for the minority traffic data class remains suboptimal. Using a different methodology, the authors [14] created a DL model that combined CNN and BLSTM to learn temporal and spatial characteristics. Accuracy levels on the binary class UNSW-NB15 were 93.84% and binary NSL-KDD of 99.30%.

Data was preprocessed using one-hot encoding and min-max normalization by authors in [15], which achieved an accuracy of 96.3% on CNN and Bi-LSTM hybrid methods on the multiclass NSL-KDD dataset. Using preprocessed on given NSL-KDD data, researchers in [16] applied the hybrid model using CNN and BLSTM algorithm with a 95.4% accuracy rate. A bidirectional LSTM model was used by the authors in their study [17] for the binary NSL-KDD dataset with the highest accuracy of 98.52%. Using a Bidirectional LSTM deep learning model, authors [18] got 99% accuracy on UNSW-NB15 and KDDCUP-99, which is an exceptional achievement. But a lot of the models that are now in use need help effectively identifying uncommon (rare) attack types, especially user2root (U2R) and remote2local (R2L) attacks,



which frequently have poorer detection accuracy as compared with other network attack types.

To overcome the difficulties found in the above literature review, authors in [19] presented a Bi-LSTM-based network intrusion detection system on the NSL-KDD dataset, which offered a binary classification accuracy of 94.26%. Furthermore, the authors proposed a Bi-directional GAN-based method [20] for the NSL-KDD and CIC-DDoS2019 datasets. The bidirectional GAN model demonstrated strong performance with an f1 score and detection accuracy of 92.68% and 91.12%, respectively, on the unbalanced NSL-KDD dataset.

In the research study [1], [17] the Authors used the hyperparameters tuning to obtain the best model performance on network intrusion detection datasets, including NSL-KDD and UNSW-NB15. In [21], the Authors implement the BLSTM model combined with random over-sampling strategies, which produces a high anomaly detection accuracy of 99.83% for multiclass imbalance network anomaly datasets NSL-KDD dataset.

The deep learning model discussed in [3] and [4] overcomes challenges traditional machine learning encounters in anomaly detection. While the CNN standalone model is unsuitable for sequential data preprocessing, and RNN requires complex data preprocessing, this model effectively addresses these issues. Data imbalance problems are tackled in [5]-[8]. Feature engineering emerges as a critical factor in enhancing the accuracy of both ML and DL models. Much research has been conducted on feature engineering, with studies focusing on attribute grouping found in [9]-[12]. The BLSTM, which brings together two distinct LSTMs to allow input processing in both directions (from the past to the future and vice versa), is implemented in [13]-[20] to improve the accuracy of network anomaly detection models.

Most of the researchers mentioned above concentrate on enhancing the detection accuracy of conventional or ML DL models and employ ensemble methods for feature engineering to address data imbalance. However, there needs to be more emphasis on exploring hyperparameter selection in DL-based models, determining the train-test split ratio, and defining the architecture of DNN. Some researchers need to elaborate on adopting these values in their studies. Subsequently, this research addressed these limitations in network traffic anomaly detection systems. We experimented using binary and multiclass versions of the UNSW-NB15 and NSL-KDD. Our focus includes investigating the performance comparison between random under-sampling and over-sampling to identify superior methods for imbalanced network data.

The contributions of our research effort in the area of network anomaly detection and imbalanced datasets are listed as:

- a. Examining the impact of CNN and BLSTM neural network architecture and performance for binary/multiclass datasets, specifically NSL-KDD and UNSW-NB15.
- b. Exploring the model performance of hyperparameters on binary and multi-class network datasets, namely UNSW-NB15 and NSL-KDD.
- c. Exploring the enhancement of CNN Bi-LSTM by varying memory elements and numbers of layers of NN.

- d. This study's interest is developing and implementing a CNN Bi-LSTM hybrid model for network anomaly detection, achieving high accuracy rates of 98.27% on NSL-KDD binary data and 99.87% on UNSW-NB15 binary data.
- e. Exploring the network anomaly detection model based on CNN Bi-LSTM using UNSW-NB15.
- f. Investigating the random sampling methods for imbalanced data with detection accuracy greater than 99.83% for NSL-KDD multiclass data and 99.99% for the UNSW-NB15 multiclass dataset.

The rest of the paper unfolds: Section 3 delineates the system model and individual blocks comprising our CNN Bi-LSTM hybrid approach. Section 4 elucidates the experimental setup, experimental results, and discussion of the findings, and section 5 encapsulates the conclusion of this research.

### 3. Network Anomaly Detection Model Description

The complete proposed model comprises the following steps:

1. Network traffic-based data collection
2. Data pre-processing and cleaning
3. Training and testing data preparation
4. CNN BLSTM model preparation
5. Train and test model
6. Evaluation of CNN BLSTM model
7. Compare the model and decision-making

The CNN BLSTM-based model's entire implementation schematic is displayed in Figure 1. The ensuing sections offer a thorough explanation of the approaches mentioned previously. The components of CNN and BLSTM layers and the intricate architecture of neural networks are seen in Figure 2.

#### 3.1. Network traffic-based data collection

Numerous datasets are accessible for research in network intrusion detection systems. Examples include the KDD Cup99, Kyoto 2006+, NSL-KDD, CICIDS2017, UNSW-NB15, and several others, providing valuable resources for intrusion detection research. During this research, the UNSW-NB15 and NSL-KDD datasets are specifically employed.

NSL-KDD KDDTrain+ [22] originates from the DARPA KDD99 dataset, with the elimination of noise and undesired data. This dataset encompasses the complete NSL-KDD training set, including labels denoting attack types and difficulty levels. Comprising 41 features, it delineates five different attack classes: "Normal," "DoS," "Probe," "R2L," and "U2R".

NSL-KDD represents a refined form of the KDD99 data, free from duplicate records in the training set and the test sets. Each entry in the dataset consists of 42 attributes, with 41 of them related to the input traffic and the final label indicating whether the traffic

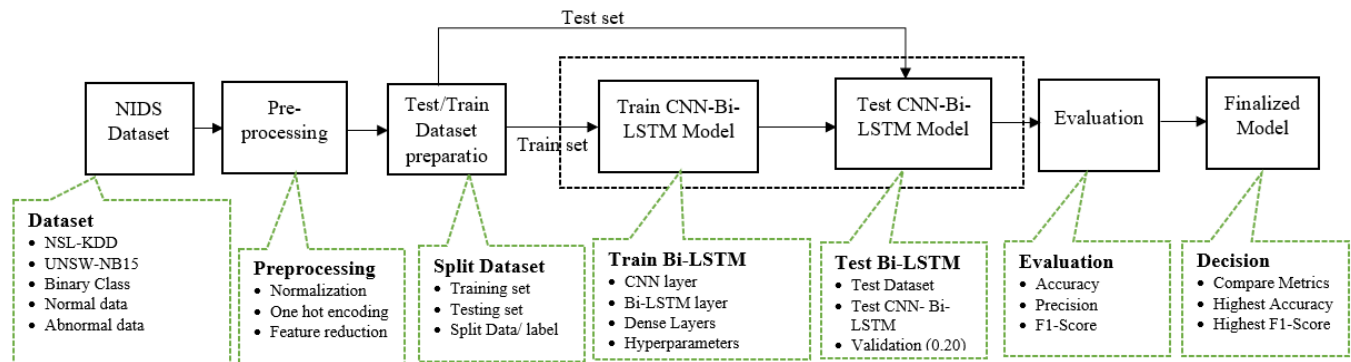


Figure 1: Pipeline for CNN BLSTM hybrid model

is normal or abnormal (target). The KDDTrain+ dataset encompasses 125,973 data entries, while the KDDTest+ dataset consists of 22,544 data entries utilized in this research work. Table 1 documents the detailed information regarding the traffic and data information [23].

Table 1: Details of NSL-KDD data

Traffic	KDDTrain+	KDDTest+
R2L	995	2,885
U2R	52	67
DoS	45,927	7,460
Normal	67,343	9,711
Probe	11,656	2,421
Total	125,973	22,544

Similarly, The Australian Centre for Cyber Security (ACCS) cybersecurity research team constructed the UNSW-NB15 dataset [24], unlike KDD99 and NSL-KDD, which is a recently developed network intrusion dataset created by IXIA PerfectStorm tools within the Cyber Range Lab of the ACCS, this dataset consists of approximately 100GB of PCAP files capturing raw network traffic flows between two hosts either server to client or vice versa. The Argus and Bro\_IDS tools and 12 other algorithms generated 49 features accompanied by class labels. Numerous records were utilized to construct the training and testing sets, where UNSW\_NB15\_training-set and UNSW\_NB15\_testing-set were used during this research work. The training set comprises 175,341 records, while the testing set comprises 82,332 records, encompassing various attacks and normal network activity. Table 2 shows detailed information regarding the attacks and normal traffic.

Table 2: Details of UNSW-NB15 data

Network Traffic	testing-set.csv	training-set.csv
Exploits	11,132	33,393
Generic	18,871	40,000
Worms	44	130
Fuzzers	6,062	18,184
DoS	4,089	12,264
Reconnaissance	3,496	10,491
Analysis	677	2,000
Backdoor	583	1,746
Shellcode	378	1,133
Normal	37,000	56,000
Total	82,332	175,341

The KDDTrain+ and KDDTest+ subsets of the NSL-KDD dataset were employed in our research experiment—likewise, experiments involved using training-set.csv and testing-set.csv from the UNSW-NB15 dataset.

### 3.2. Data pre-processing and cleaning

NSL\_KDD data is an improved version of the KDD99 dataset; minimum work is required for data preprocessing. The downloaded separate data files are used to test and train the model. The target class is initially isolated from the training and testing datasets to create the class labels. From the remaining attributes, numerical features and three categorical features—“protocol\_type”, “service”, and “flag” are extracted. The categorical features undergo conversion into numerical values using dummy one-hot encoding techniques, while the numerical attributes are standardized using standard Scalar methods. Afterwards, both types of feature sets are combined into a unified data frame, yielding the final data sets for training and testing. One hot encoding generates one binary variable for each individual categorical value. The dummy encoding is similar to one hot encoding and converts the categorical values into numeric binary values. The dummy encoding represents N categories using N-1 binary variable. Let’s say we have three categories of traffic “protocol\_type,” “service,” and “flag” that are going to be dummy encoded as [1 0], [0 1], and [0 0], respectively. The standard scalar converts the numeric values so that the data standard deviations become 1.

Since there are different types of services present in the KDDTrain+ dataset and KDDTest+ dataset, the one hot encoding produces unequal numbers of features. The KDDTrain+ dataset contains 126 features, while the KDDTest+ includes a total of 120 features after the implementation of one hot encoding. Those additional features “service\_aol,” “service\_harvest,” “service\_http\_2784”, “service\_http\_8001”, “service\_red\_i,” and “service\_urh\_i” are inserted into the KDDTest+ dataset after finding the exact location where those features reside into the KDDTrain+ dataset. We preserved the attacks\_types and difficulty\_level features because those features are highly relevant to the target class and increase the model's efficiency.

The UNSW-NB15 dataset was divided into two sets for training and testing purposes: UNSW\_NB15\_training-set and UNSW\_NB15\_testing-set. The UNSW\_NB15\_training-set comprises 175,341 entries, while the UNSW-NB15\_testing-set contains 82,332 entries, encompassing various attacks and normal data. Initially, the features on this dataset are 49. First, those categorical attributes are changed into numeric using dummy one hot encoding. All numerical attributes are applied to the standard scalar normalization method. After preprocessing the numeric and categorical features, 192 features for UNSW\_NB15\_testing-set

data and 196 features for UNSW\_NB15\_training-set data were generated. Again, here we are taking two sets of data: one we can use for training and the other for testing or vice versa. The categorical values of data entries are not the same for both datasets; hence, the one hot encoding produces unequal numbers of features on both data sets after preprocessing.

Some features generated from one hot encoding, such as state\_ACC and state\_CLO, are not included in the UNSW\_NB15\_training-set. Similarly, proto\_icmp, proto\_rtp, state\_ECO, state\_PAR, state\_URN, and state\_no features are not included on UNSW\_NB15\_testing-set. The empty features columns are added in the exact column location of those missing features on the respective dataset, generating 198 features plus one target class.

### 3.3. Training and testing data preparation

In experiments concerning the binary NSL-KDD dataset, the training and testing datasets were created using a split ratio. The train-test split approach assesses the performance of machine learning algorithms in making predictions from data that wasn't part of the training set. We opted for a 70:30 split ratio to generate the train and test dataset. For the CNN BLSTM hybrid model, 70% of KDDTrain+ was used to train, and the remaining data was used to test the model for binary NSL-KDD data.

A similar split percentage was employed in the binary class UNSW-NB15, using the "UNSW\_NB15\_training-set". In the case of multiclass experiments for UNSW-NB15 and NSL-KDD, two distinct files were selected—one subset for training the CNN-BLSTM model and another for testing. Detailed information regarding this split is provided in the respective experimental sections.

### 3.4. CNN BLSTM model

CNN is a forward DNN designed for image signal and classification. CNN comprises three primary layers: the convolutional, the pooling, and the fully connected layers. The convolutional layer is the main component of CNN and uses the convolutional operation to grab the various features from the image signal. Then, the number of pooling layers extracts features, and a fully connected layer employs the output from the preceding layer for classification. Combining convolutional layers with pooling layers is responsible for feature extraction, while the final fully connected dense layer is utilized for classification purposes. CNN also involves various hyperparameters, including the number of filters, stride, zero-padding, pooling layers, and others.

An RNN is an artificial NN designed to manage sequential data by integrating feedback loops into its structure. Diverging from conventional feedforward neural networks that linearly handle input data, RNNs feature connections forming loops, enabling them to retain a memory of past inputs and utilize that information to impact the current output. The memory in an RNN serves as a short-term storage, allowing the network to retain information about past events and use it to make predictions about future events. This is especially valuable in applications where context and temporal relationships are essential. Machine learning issues, including speech recognition, language processing, and picture categorization, have been resolved with RNN.

Yet, traditional RNNs encounter challenges, notably needing help with learning long-term dependencies attributed to the vanishing or exploding gradient problem. Advanced RNN versions such as gated recurrent units (GRUs) and long short-term memory (LSTM) networks have been devised in response to these constraints. These architectures include mechanisms for selectively storing and retrieving information across extended sequences, enhancing their effectiveness in tasks that demand capturing long-term dependencies.

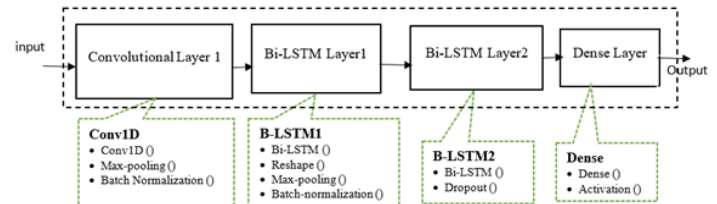


Figure 2: CNN BLSTM layer architecture

LSTM handles the vanishing gradient in RNN. There is a memory block and three multiplicative units in LSTM. The input corresponds to the write operation, output to read and forget gates corresponding to the reset operations for cells that make up the LSTM architecture. By allowing LSTM memory cells to keep and access data for longer periods. Those multiplicative gates mitigate the vanishing gradient.

To process input in both directions—from the future to the past and from the past to the future—bidirectional RNN combines two independent RNNs. Both forward and backward LSTM networks make up the Bi-LSTM. The features extracted by the forward LSTM hidden layer point forward, whereas those extracted by the reverse LSTM hidden layer point backward. By taking finite sequences into account about earlier and later items, the bidirectional LSTM can anticipate or tag the sequence of each element. Two LSTMs processed in series—one from left to right and the other from right to left—produce this. The CNN and BLSTM hybrid models have several layers, each with a set of hyperparameters. Figure 2 shows the CNN BLSTM's architectural layout.

### 3.5. CNN BLSTM model training

The CNN BLSTM model's neural network architecture is prepared for training. The datasets consist of two sets: one for training and the other for testing, or vice versa. The split percentage determines how much data is allocated for training and testing when a single data set is present. The selection of hyperparameters for model training is conducted through various experiments involving fine-tuning epochs and batch size to enhance detection efficiency. Within the training data, 20% is designated for validating the CNN Bi-LSTM model.

### 3.6. Test the CNN BLSTM hybrid model and evaluation.

Deep learning (DL) and machine learning (ML) models offer performance consistency. After the CNN BLSTM model is built, the model is trained using the training dataset with specified hyperparameter values. These chosen hyperparameter values influence the training duration. Following training, the model can assess the unseen dataset to evaluate its performance.



Hyperparameter selection lacks a predefined rule, allowing for random selection and subsequent fine-tuning through various experiments. After the model testing, performance metrics are determined based on the type of ML model employed. In the case of the supervised machine learning model, ground truth values are utilized to measure the performance metrics on the test dataset. Various metrics, such as detection accuracy, precision, F1-Score, recall, program execution time, and Area under the ROC, are available to compare the model efficiency. Confusion metrics from Karas generate True Positive (TP), True Negative (TN), False Positive (FP), and False Negative (FN) values. In the context of a classification report, the terms "weighted" and "macro" refer to different strategies for computing metrics such as precision, recall, and F1-score across multiple classes. Macro-averaging computes the metric for every class separately before averaging them. This means that each class is treated equally in the computation, regardless of size. Macro-averaging gives the same weight to each class, which can be useful when all classes are considered equally. Weighted averaging, on the other hand, takes the average of the metrics, but it weights each class's contribution based on its proportion in the dataset. In other words, classes with more samples have a greater impact on the average. Weighted averaging is especially helpful when working with unbalanced datasets—where certain classes may have substantially more instances than others.

The classification report provides a thorough summary of the model's performance metrics for the specified training and testing data sets. Lastly, to assess the performance of our CNN BLSTM hybrid model, the performance metrics are compared with the findings of earlier research publications.

### *3.7. Compare models and decision-making.*

Several sets of experiments were conducted to find the hyperparameter settings that yielded the best results. Following model testing and evaluation, choosing the best model pipeline from various options is part of the cognitive process of comparison and decision-making. Throughout this study, several sets of experiments are carried out to determine values for various CNN Bi-LSTM model hyperparameters to enhance the model's performance. To create an effective Bi-LSTM pipeline, it is necessary to decide on the hyperparameters, which include optimizers, number of epochs, batch, NN design, class size, and techniques of raw data preprocessing. This is achieved by evaluating performance metrics across multiple sets of experiments. The performance metrics of the Bi-LSTM model are then juxtaposed with previously published results for the binary/numerous class UNSW-NB15 and NSL-KDD. The class imbalance problem in the multiclass version of both NSL-KDD and UNSW-NB15 datasets was exposed with sampling data during the preprocessing stages. The sampling methods randomly deleted on down-sampling and randomly generated data samples in over-sampling. This resulted in the balanced form of datasets to compare the CNN Bi-LSTM model performance.

## **4. Results and Discussion**

To detect anomalies, intrusion detection uses a mix of DL and ML methods. The implementation of a network anomaly detection model is implied using Python script. Python has specialized packages for building machine learning models, including NumPy,

Pandas, Keras, and Scikit-learn. Additionally, commonly used tools like Java, C#, WEKA, Visual C++, and MATLAB play vital roles in network anomaly detection systems. On the Jupyter Notebook platform, seed values are fixed to guarantee consistency in outcomes over several runs. Plots and tables representing the results of experiments are analyzed using the Microsoft Office suite. Every experiment is run on a Windows machine with an i7 processor and 16GB of RAM.

Python and the packages it is linked with keep version information used in all experiments. For example, TensorFlow 2.9.1, Keras 2.6.0, and Python 3.7.12 are used. Hyperparameters will be determined, performance will be evaluated across class sizes, and the efficacy of various sampling approaches will be assessed about the CNN BLSTM model for the multi-class and binary-class UNSW-NB15 and NSL-KDD. Detailed explanations of these experiments are provided in subsequent sections.

The architecture shown in Figure 2 consists of a single 16-unit convolution layer that uses batch normalization and max-pooling. BLSTM neural network layer 1 contains 50 memory units; batch normalization, max-pooling, and reshaping come next. Bi-LSTM neural network layer 2 with 100 memory units and dropout is also available. The dense layer consists of a sigmoid activation, and the final output is obtained. The detection accuracy of the model is evaluated through a series of tests involving the adjustment of optimizers, learning rate (LR), number of epochs, batch size, and dropout rate. As explained below, the UNSW-NB15 and NSL-KDD binary/multiclass network traffic datasets are used for these investigations.

### *4.1. Experiment: Model performance Vs. Optimizers*

In the context of ML and DL, an optimizer is an algorithm or method used to adjust the parameters of a model to minimize or maximize a certain objective function. The performance of an optimizer is crucial in training machine learning models because it determines how well the model learns from the data. Choosing the optimizer is essential during the training of the CNN BLSTM model, as it significantly contributes to expediting results for the machine learning/deep learning model.

TensorFlow offers nine optimizers (Ftrl, Nadam, Adam, Adadelta, Adagrad, gradient descent, Adamax, RMSprop, and Stochastic Gradient Descent (SGD)) based on the optimizer's methods. The choice of optimizer can significantly impact the training performance of an ML model. Optimizers may converge at different rates or achieve different final accuracies on a given task. An optimizer's performance may be influenced by the model's architecture, the dataset, and the hyperparameters employed.

It is common practice to experiment with several optimizers to determine which combination of optimizers and hyperparameters is optimum for a given task. Additionally, some optimizers may perform better on certain types of neural network architectures or for specific data types. In summary, the relationship between the optimizer and machine learning performance is crucial, and choosing the right optimizer is an important part of the model training process. It often involves experimentation and tuning to find the optimal combination for a given task.

The model used in the experiment comparing Optimizers versus Accuracy has a 20% dropout rate and the Relu activation



function. To determine the best optimizer for our CNN-BLSTM model, seven optimizers, including Nadam, Ftrl, SGD, Adam, RMSprop, Adagrad, and Adamax, were tested. Based on the model performance metrics for UNSW-NB15 and NSL-KDD binary data, which are shown in Table 3, it was found that the Nadam optimizer performed best for NSL-KDD. In contrast, the Adam optimizer produced the best accuracy for the UNSW-NB15 dataset. Interestingly, although both optimizers used the same model architecture, they performed differently for both Network Intrusion Detection System (NIDS) datasets.

Table 3: Model performance Vs. Optimizer

Number of epochs = 10, Batch = 256, NSL-KDD C2 and UNSW-NB15 C2				
Optimizer	ACC-NSL	F1-NSL	ACC-UN	F1-UN
Ftrl	53.47	69.68	80.99	80.99
RMSprop	97.87	98.01	97.93	98.46
Adamax	97.65	97.78	95.33	96.51
<b>Adam</b>	<b>98.02</b>	<b>98.16</b>	<b>99.15</b>	<b>99.38</b>
Adagrad	96.98	97.21	94.04	95.62
SGD	97.74	97.91	99.14	99.37
<b>Nadam</b>	<b>98.13</b>	<b>98.26</b>	99.11	99.34

ACC: Accuracy in %, F1: F1Score in %, NSL: KDD-NSL, UN: UNSW NB

Table 4: CNN BLSTM performance Vs. Optimizer on NSL-KDD

Model Performance Vs. Optimizer on NSL-KDD Multiclass Datasets			
Epochs=10, Batch_size= 512, Training_data = KDDTrain+, Testing_data = KDDTest+, Multiclass=5			
Optimizers	Accuracy %	wt Precision %	wt F1score %
<b>Adam</b>	<b>88.46</b>	<b>88.87</b>	<b>88.23</b>
RMSprop	85.49	87.15	82.84
Nadam	84.79	86.97	82.45
SGD	82.86	84.99	77.12
Adamax	82.6	86.99	82.72
Adagrad	75.65	67.01	69.94
Ftrl	43.08	18.56	25.94

Table 5: CNN BLSTM performance Vs. optimizer on UNSW-NB15

CNN Bi-LSTM Performance Vs. Optimizer on UNSW-NB15 Multiclass Datasets			
Epochs=15, Batch_size= 512, Training_data=UNSW-NB15Train82332, Testing_data = UNSW-NB15Test175341, Multiclass=10			
Optimizers	Accuracy %	wt Precision %	wt F1score %
<b>SGD</b>	<b>89.84</b>	<b>87.49</b>	<b>88.01</b>
Adam	87.21	87.47	85.96
Nadam	84.59	84.48	83.38
RMSprop	79.3	75.71	76.85
Adamax	76.84	76.37	74.74
Adagrad	70.82	63.28	62.05
Ftrl	31.94	10.20	15.46

The selection of the optimizers depends on the combination of the different hyperparameters and NN architecture of the CNN BLSTM model. Popular optimization algorithm Adam combines concepts from RMSprop and momentum. It adapts the learning rates of individual parameters and is widely used in deep learning. An Adam extension that uses the Nesterov Accelerated Gradient (NAG). NAG involves looking ahead in the direction of the momentum before computing the gradient that combines the benefits of Adam and Nesterov momentum. Figure 3. shows the accuracy comparison for NSL-KDD and UNSW15.

Tables 4 and 5 shows the comparative performance metrics of the multi-class NSL-KDD and UNSW-NB15. The same optimizer does not provide the same performance for a similar dataset. The hyperparameters and datasets used to test and train the CNN-based BLSTM model are provided in Tables 4. and 5. For the NSL-KDD

multiclass dataset, Adam performed better than SGD, whereas for the UNSW-NB15 multiclass dataset, SGD performed better than other optimizers.

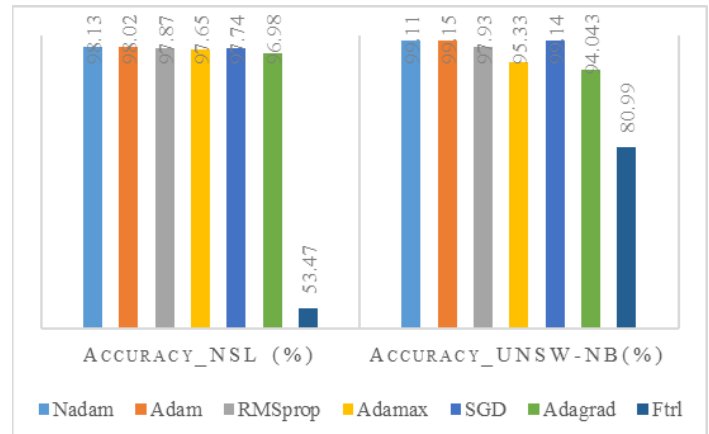


Figure 3: Optimizer Vs. Accuracy

#### 4.2. Experiment: Learning rate Vs. model performance

The learning rate, a positive scalar multiplied by gradient descent gradient, controls the step size in parameter space. A higher rate facilitates faster convergence but raises the risk of overshooting and oscillation. On the other hand, a lower rate ensures stability but may demand more iterations for convergence.

With optimizers chosen from the preceding Experiment 4.1, the same CNN BLSTM model neural network architecture is used to determine the ideal learning rate to enhance the model performance. The NSL-KDD binary data is preprocessed from the subset of the KDDTrain+ dataset, and the split ratio splits the data for training and testing. The learning rate determines the rate at which new weights are added to neural network models. The other hyperparameters remain constant throughout this experiment while the learning rates are adjusted to optimize the model's accuracy. Table 6 displays a comparison of learning rate with CNN BLSTM model performance. The model performs best on the UNSW-NB15 binary data and the NSL-KDD binary dataset, achieving a learning rate of 0.01 and 0.0002, respectively. The same learning rate provides different model performances.

Table 6: CNN BLSTM model Learning rate Vs. Performance metrics

Epochs size = 10, Batch = 256, KDD_C2 (Nadam), UNSW-NB15_C2 (adam)				
LR	ACC-NSL	F1-NSL	ACC-UN	F1-UN
<b>0.01</b>	97.49	97.67	<b>99.67</b>	<b>99.76</b>
0.001	98.16	98.29	99.54	99.66
0.0001	98.06	98.20	95.81	96.85
<b>0.0002</b>	<b>98.18</b>	<b>98.3</b>	97.9	98.44
0.0003	98.14	98.27	98.44	98.86
0.0004	97.97	98.11	99.13	99.35
0.0005	98.11	98.25	99.09	99.32

LR: Learning rate, ACC: Accuracy in %, F1: F1Score in %, UN: UNSW NB

#### 4.3. Experiment: Model dropout rate Vs. model performance

The phrase "dropout rate" in machine learning usually refers to a regularization method that neural networks employ to avoid overfitting. When a model becomes overfit, it can have poor generalization on new, unknown data because it has learned the training set too well, including its noise and outliers. During training, randomly selected neurons (units) in the neural network

are "dropped out" or omitted temporarily. This means these neurons do not contribute to the forward or backward pass during a particular iteration of training. The probability of a neuron being dropped out is called the dropout rate. The dropout rate is one hyperparameter that must be determined before training the model.

The CNN BLSTM model was tested and trained for both datasets using a batch size of 256 and 10 epochs. Different dropout rate values were used to evaluate the efficiency of the model. The model performed better than the others, with a 30% dropout rate on the UNSW-NB15 dataset; however, a 60% dropout rate worked better for the NSL-KDD. The hyperparameter values, dropout rates, and corresponding performance metrics are presented in Table 7. The experimental results highlight the varying dropout rates for distinct datasets despite the similarity between the two datasets.

Table 7: Dropout rate Vs. model performance

Epochs size = 10, batch = 256, KDD_C2 (madam), UNSW-NB15_C2 (adam)				
DropOut %	ACC-NSL	F1-NSL	ACC-UN	F1-UN
0.1	98.10	98.24	97.44	98.15
0.2	98.02	98.16	98.98	99.25
<b>0.3</b>	98.16	98.29	<b>99.87</b>	<b>99.9</b>
0.4	98.04	98.17	99.27	99.47
0.5	97.93	98.09	99.47	99.61
<b>0.6</b>	<b>98.21</b>	<b>98.33</b>	99.81	99.86
0.7	98.01	98.15	99.58	99.69
0.8	98.04	98.18	98.57	98.94

ACC: Accuracy in %, F1: F1Score in %, NSL: KDD-NSL, UN: UNSW\_NB KDDTrain+, UNSW-NB15 training.csv binary with test-train split

The batch size is a hyperparameter in machine learning that determines how many samples are used in a training iteration. The batch size represents the number of samples used in a single training iteration. Using a smaller batch size incorporates a limited number of data samples and results in a longer training time for the CNN Bi-LSTM model compared to a larger batch size. Throughout experimentation (Experiment A-C), the batch size is altered while maintaining other hyperparameters, such as a fixed number of epochs is 5, the learning rate of the optimizer, and the dropout rate values assigned to the model based on previous findings with the respective datasets.

Table 8: Model performance Vs. batch size

Number of Epochs = 5, KDD_C2 (Nadam), UNSW-NB15_C2(adam)				
Batch	ACC-NSL %	F1-NSL %	ACC-UN %	F1-UN %
<b>32</b>	97.89	98.04	<b>99.40</b>	<b>99.55</b>
64	97.95	98.10	99.35	99.52
<b>128</b>	<b>98.06</b>	<b>98.20</b>	99.33	99.50
256	97.64	97.79	96.36	97.26
512	97.92	98.08	96.90	97.70

The dataset size, the amount of computing power available, and the specifics of the optimization issue can all influence the batch size decision. Experimenting with various batch sizes is a frequent way to determine which is most effective for a certain task.

The experimental result in Table 8 demonstrates how the neural network's hyperparameter combinations affect performance. In this experiment, batch sizes of 128 for the binary NSL-KDD datasets and 32 for the binary UNSW-NB15 datasets for epochs 5 demonstrated the best performance of the CNN BLSTM model.

#### 4.4. Experiment: Epochs Vs. model performance

An "epoch" in machine learning is one whole iteration through the training dataset a model goes through while training. The learning method processes the complete dataset throughout each epoch, modifying the neural network weights and parameters to reduce the error or loss function. A hyperparameter called epoch count determines an algorithm's running frequency over the full training dataset. The integer between one to infinity can be used as the epoch. Selecting smaller epoch values results in a longer training time for the model and vice versa. Underfitting, the ML model cannot identify the original patterns in the data, which can be caused by using too few epochs. However, an excessive number of epochs might cause overfitting, in which case the model becomes inattentive to new data and underperforms on previously unknown data.

The CNN BLSTM hybrid model performance for binary KDD-NSL and binary UNSW-NB15 with the different values of epochs are documented in Table 9. The performance increases with large values of epochs but is different for a while. After 75 epochs, the model performance decreases. The amount of data utilized for training and testing, the size of the output class, and other hyperparameter combinations affect the epochs and performance of the machine learning/deep learning models.

Table 9: Epochs Vs. model performance

Batch size = 256, NSL- KDD_C2 (Nadam)		
Number of Epochs	Accuracy-NSL %	F1Score-NSL %
2	95.48	95.94
10	98.13	98.26
25	98.21	98.33
50	98.20	98.33
<b>75</b>	<b>98.27</b>	<b>98.39</b>
100	98.26	98.39

The selection of epoch size to produce a superior performance on an imbalanced dataset is challenging. The binary dataset is more balanced than the multiclass network-based intrusion dataset. The experimental results in Table 9 are not the determining experiment for the number of epochs on multiclass NSL-KDD and UNSW-NB15 datasets. Hence, we experimented with and documented multi-class experimental results to determine the values of epochs where we can produce higher accuracy on the provided dataset. Tables 10 and 11 show the experimental results for multiclass datasets to investigate the values of epochs to make superior detection accuracy. In summary, while epoch size and class size are conceptually different, they can influence each other indirectly, especially when dealing with imbalanced datasets. Selecting the right number of epochs for a given problem is crucial, as is keeping an eye on how class sizes affect model performance.

Table 10: Model performance Vs. Epochs on UNSW-NB15 multiclass data

Batch=512, Optimizer=SGD, Training=UNSW-NB15Train.csv82332 testing data= UNSW-NB15Test.csv175341, Multiclass=10				
Epochs	ACC	wt Prec	wt F1Score	Prg exe time
<b>10</b>	<b>93.10</b>	<b>91.10</b>	<b>91.94</b>	<b>0.64</b>
25	83.09	79.94	79.69	1.17
50	86.4	82.47	83.35	2.24
75	87.1	86.46	85.24	3.12
100	90.04	88.36	88.01	4.24
150	82.23	81.49	80.82	6.36

200	81.13	78.65	78.84	7.95
ACC: Accuracy in %, wt_Prec: weighted Precision in %, wt_F1Score: weighted F1Score in %, Prg_exe time: Program script run time in hr.				

Program execution time is the sum of the model's training and testing phases. The program execution time depends on various factors, such as the neural network's architecture, training and testing data size, class size, and combination of hyperparameters. We documented the performance of the CNN BLSTM hybrid model along with the program execution time. The higher the epochs result, the longer the program execution time. Table 10 and Table 11 show the multi-class NSL-KDD run for almost 8 Hrs. to complete training testing and evaluate the model for 200 epochs. A similar scenario for multiclass UNSW-NB15 dataset. Hence, selecting epoch and batch size is the trade-off with the model training, testing, and evaluation time. We found in Table 10 and Table 11 the different epoch sizes for NSL-KDD (outperform at epoch size 10) and UNSW-NB15 (outperform at epoch 100) during multi-class model performance.

Table 11: Performance Vs. epochs on NSL-KDD multiclass data

Batch=512, Optimizer= Adam, Data=KDDTrain+_125973, KDDTest+ 22544, Multiclass = 5				
Epochs	ACC	wt_Prec	wt_F1Score	Prg_exe time
10	86.21	88.13	83.85	0.45
25	86.64	88.01	84.1	1.08
50	86.64	88.7	84.78	2.15
75	87.11	88.39	84.84	3.31
<b>100</b>	<b>87.63</b>	<b>89.85</b>	<b>86.44</b>	<b>4.53</b>
150	87.22	89.66	85.81	7.03
200	86.84	90.61	86.41	8.76
ACC: Accuracy in %, wt_Prec: weighted Precision in %, wt_F1Score: weighted F1Score in %, Prg_exe time: Program script run time in hr.				

#### 4.5. Experiment: Imbalance data sampling Vs. performance

This experiment investigates the sampling techniques for imbalanced data to provide a high detection rate. The researcher employed various techniques, such as a critical selection of ML and DL algorithms, ensemble methods, data sampling, etc., to address the issue of data imbalance because there are fewer attacks than typical traffic data in the provided network intrusion detection dataset.

Sampling methods generate or delete random data from the dataset based on class data distribution. Random under-sampling and random over-sampling are two techniques used in imbalanced classification problems, where one class (usually the minority traffic class) is significantly under-represented compared to the other class(es). These methods are utilized to tackle class imbalance and enhance the efficacy of machine learning models. Random Under Sampling (RUS) involves randomly removing instances from the majority class until the distribution between the majority and minority classes is more evenly distributed. However, random over-sampling produces an equal distribution by randomly duplicating minority class instances or creating synthetic instances to increase the number of minority class instances.

Tables 12 and 13 provide the hyperparameter information and performance of this experiment's CNN BLSTM hybrid model. Table 12 compares the NSL-KDD multiclass dataset's performance when random over- and under-sampling is applied. After preprocessing multi-class NSL-KDD data, the training and testing datasets merge into a single file. Sampling is implemented

on merged data, and a 70:30 split ratio is used to split data into train and test datasets.

Table 12: CNN BLSTM performance Vs. Sampling on NSL-KDD

CNN BILSTM, Epochs= 25, Batch_size= 512, Data = combine (KDDTrain+KDDTest+) (sampling)				
Class	Recall_RUS	F1_RUS	Recall_ROS	F1_ROS
DoS	0	0	99.86	99.92
Probe	100	36.69	99.89	99.91
U2R	10	18.18	100	99.73
R2L	0	0	99.45	99.63
Normal	0	0	99.93	99.93
Wt Average →	22.15	10.07	99.83	99.83
Macro Avg→	22	10.97	99.83	99.83
Accuracy %	22.15		<b>99.83</b>	
F1:F1Score, RUS: Random Under Sampling, ROS: Random Over Sampling] %				

Similarly, preprocessed training data and testing data files are merged into a single file to implement the sampling method on the UNSW-NB15 multiclass dataset. The sampled data is then split into training and testing datasets using a 70:30 train-test split ratio. During Random under-sampling, data instances are randomly deleted from the majority class, resulting in significant information loss. Deleting samples from the majority class results in a smaller sample, unsuitable for the deep learning model and worsens the model performance, which is found in the experiment and documented in Tables 12 and Table 13. Random over-sampling (ROS) helps prevent information loss, as none of the minority class instances are removed. It can be more effective when the amount of data in the minority class is limited.

Table 13: Model performance Vs. sampling on UNSW-NB15

CNN BLSTM Epochs=25, Batch_size=512, data = combine (UNSW-NB15training-set 175341+UNSW-NB15testing-set 82332) sampling				
Class	Recall_RUS	F1_RUS	Recall_ROS	F1_ROS
Analysis	0	0	100	100
Backdoor	0	0	100	100
DoS	0	0	100	99.95
Exploits	100	17.51	99.91	99.95
Fuzzers	0	0	99.98	99.99
Generic	0	0	99.98	99.98
Normal	0	0	100	100
Reconnaissance	0	0	100	100
Shellcode	0	0	100	100
Worms	0	0	100	100
Weighted Average	9.2	1.61	99.99	99.99
Macro Average	10	1.75	99.99	99.99
Accuracy (%)	9.20		<b>99.99</b>	
F1:F1Score, RUS: Random Under Sampling, ROS: Random Over Sampling] %				

This method generates random data based on the data distribution in the dataset. The huge amount of data is always suitable for deep learning models. Regarding detection accuracy, our suggested CNN BLSTM hybrid model performs better than the random over-sampling technique, offering over 99%. Tables 12, 13, and Figure 4 above detailed the CNN BLSTM hybrid model's performance for the UNSW-NB15 imbalance dataset and the multiclass NSL-KDD.

## 5. Conclusion

The previous research from the literature reviews shows that while the detection accuracy for rarely occurring attack classes (U2R, R2L) is low, the average model accuracy for normal traffic in the UNSW-NB15 and NSL-KDD is roughly 99%. Regardless



of the type of attack, each poses a threat to network machines equally. To provide a comparative analysis, we juxtapose our results with existing findings of 91.12% [20] and 90.83% [5] detection accuracy for NSL-KDD binary, and 99.70% [18], 82.08% [14], 82.08% for UNSW-NB15 binary datasets. Our experiments enhance accuracy to 98.27% on NSL-KDD and 99.87% on UNSW-NB15 binary datasets by carefully selecting hyperparameters and conducting various experiments. We explored the CNN BLSTM hybrid model's hyperparameters (dropout, epochs, batch size, learning rate, and optimizer) to maximize detection accuracy for the binary NSL-KDD and UNSW-NB15.

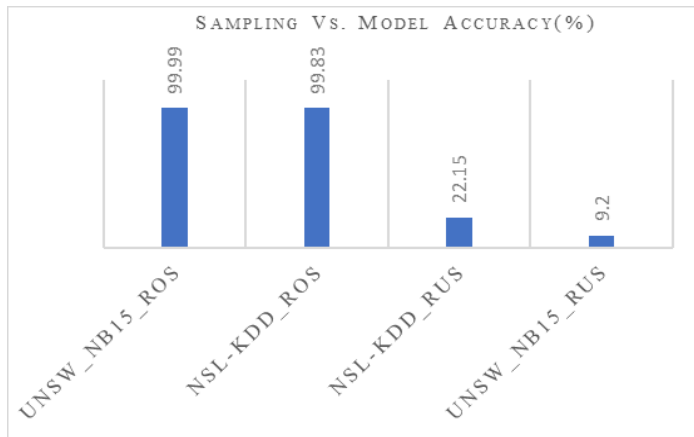


Figure 4: Sampling Vs. multiclass model accuracy (%)

The model performance depends on the combination of hyperparameters, the size of the dataset used to train/test the model, and the selection of the machine learning/ deep learning model. Our research provides information about the data size used during the experiments and the choice of hyperparameters. The suggested model uses random over-sampling techniques on a single set of data to provide 99.99% and 99.83% model accuracy for the multiclass UNSW-NB15 and NSL-KDD datasets, respectively (train and test data merge into a single file before sampling).

Selecting random over-sampling or under-sampling relies on the particulars of the dataset and the issue at hand. To achieve a balance, combining the two methods, a practice known as hybrid sampling may occasionally be necessary. It's crucial to remember that there are more sophisticated methods for dealing with class imbalance, such as SMOTE (Synthetic Minority Over-sampling Technique), which creates synthetic instances for the minority class instead of merely copying real instances. Thoroughly examining those approaches in various network intrusion detection multiclass datasets extends this research effort. The proper use of hyperparameters of neural networks, size of dataset used to train the model, and sampling methods for CNN BLSTM network anomaly model provide the highest detection accuracy for imbalance network data.

#### Conflict of Interest

The authors declare no conflict of interest.

#### Acknowledgment

The National Science Foundation (NSF) and Scholarship for Service CyberCorps (SFS CyberCorps) programs support this

research study. The award information of NSF and CyberCorps are #1910868 and #2219611, respectively. This cannot be completed without the continuous support of advisors and the Electrical and Computer Engineering Departments of Prairie View A&M University

#### References

- [1] T. Acharya, A. Annamalai, M.F. Chouikha, "Efficacy of CNN-Bidirectional LSTM Hybrid Model for Network-Based Anomaly Detection," in 13th IEEE Symposium on Computer Applications and Industrial Electronics, ISCAIE 2023, Institute of Electrical and Electronics Engineers Inc.: 348–353, 2023, doi:10.1109/ISCAIE57739.2023.10165088.
- [2] N. Moustafa, J. Hu, J. Slay, "A holistic review of Network Anomaly Detection Systems: A comprehensive survey," *Journal of Network and Computer Applications*, **128**, 33–55, 2019, doi:10.1016/j.jnca.2018.12.006.
- [3] H. Liu, B. Lang, M. Liu, H. Yan, "CNN and RNN based payload classification methods for attack detection," *Knowledge-Based Systems*, **163**, 332–341, 2019, doi:10.1016/j.knosys.2018.08.036.
- [4] B. Cao, C. Li, Y. Song, Y. Qin, C. Chen, "Network Intrusion Detection Model Based on CNN and GRU," *Applied Sciences (Switzerland)*, **12**(9), 2022, doi:10.3390/app12094184.
- [5] Y. Fu, Y. Du, Z. Cao, Q. Li, W. Xiang, "A Deep Learning Model for Network Intrusion Detection with Imbalanced Data," *Electronics (Switzerland)*, **11**(6), 2022, doi:10.3390/electronics11060898.
- [6] X. Ji, H. Zhang, X. Ma, "A Novel Method of Intrusion Detection Based on Federated Transfer Learning and Convolutional Neural Network," in IEEE Joint International Information Technology and Artificial Intelligence Conference (ITAIC), Institute of Electrical and Electronics Engineers Inc.: 338–343, 2022, doi:10.1109/ITAIC54216.2022.9836871.
- [7] T. Acharya, I. Khatri, A. Annamalai, M.F. Chouikha, "Efficacy of Heterogeneous Ensemble Assisted Machine Learning Model for Binary and Multi-Class Network Intrusion Detection," in 2021 IEEE International Conference on Automatic Control and Intelligent Systems, I2CACIS 2021 - Proceedings, Institute of Electrical and Electronics Engineers Inc.: 408–413, 2021, doi:10.1109/I2CACIS52118.2021.9495864.
- [8] T. Acharya, I. Khatri, A. Annamalai, M.F. Chouikha, "Efficacy of Machine Learning-Based Classifiers for Binary and Multi-Class Network Intrusion Detection," in 2021 IEEE International Conference on Automatic Control and Intelligent Systems, I2CACIS 2021 - Proceedings, Institute of Electrical and Electronics Engineers Inc.: 402–407, 2021, doi:10.1109/I2CACIS52118.2021.9495877.
- [9] M. Xiong, H. Ma, Z. Fang, D. Wang, Q. Wang, X. Wang, "Bi-LSTM: Finding Network Anomaly Based on Feature Grouping Clustering," in ACM International Conference Proceeding Series, Association for Computing Machinery: 88–94, 2020, doi:10.1145/3426826.3426843.
- [10] S.N. Pakanzad, H. Monkaresi, "Providing a hybrid approach for detecting malicious traffic on the computer networks using convolutional neural networks," in 2020 28th Iranian Conference on Electrical Engineering, ICEE 2020, Institute of Electrical and Electronics Engineers Inc., 2020, doi:10.1109/ICEE50131.2020.9260686.
- [11] R. Yao, N. Wang, Z. Liu, P. Chen, X. Sheng, "Intrusion detection system in the advanced metering infrastructure: A cross-layer feature-fusion CNN-LSTM-based approach," *Sensors (Switzerland)*, **21**(2), 1–17, 2021, doi:10.3390/s21020626.
- [12] P. Sun, P. Liu, Q. Li, C. Liu, X. Lu, R. Hao, J. Chen, "DL-IDS: Extracting features using CNN-LSTM hybrid network for intrusion detection system," *Security and Communication Networks*, **2020**, 2020, doi:10.1155/2020/8890306.
- [13] L. Zhang, J. Huang, Y. Zhang, G. Zhang, "Intrusion Detection Model of CNN-BiLSTM Algorithm Based on Mean Control," in Proceedings of the IEEE International Conference on Software Engineering and Service Sciences, ICSESS, IEEE Computer Society: 22–27, 2020, doi:10.1109/ICSESS49938.2020.9237656.
- [14] J. Sinha, M. Manollas, "Efficient Deep CNN-BiLSTM Model for Network Intrusion Detection," in ACM International Conference Proceeding Series,



- Association for Computing Machinery: 223–231, 2020, doi:10.1145/3430199.3430224.
- [15] A. Li, S. Yi, “Intelligent Intrusion Detection Method of Industrial Internet of Things Based on CNN-BiLSTM,” *Security and Communication Networks*, **2022**, 2022, doi:10.1155/2022/5448647.
- [16] J. Gao, “Network Intrusion Detection Method Combining CNN and BiLSTM in Cloud Computing Environment,” *Computational Intelligence and Neuroscience*, **2022**, 2022, doi:10.1155/2022/7272479.
- [17] T. Acharya, A. Annamalai, M.F. Chouikha, “Efficacy of Bidirectional LSTM Model for Network-Based Anomaly Detection,” in *13th IEEE Symposium on Computer Applications and Industrial Electronics, ISCAIE 2023*, Institute of Electrical and Electronics Engineers Inc.: 336–341, 2023, doi:10.1109/ISCAIE57739.2023.10165336.
- [18] P. TS, P. Shrinivasacharya, “Evaluating neural networks using Bi-Directional LSTM for network IDS (intrusion detection systems) in cyber security,” *Global Transitions Proceedings*, **2**(2), 448–454, 2021, doi:10.1016/j.gltp.2021.08.017.
- [19] Y. Imrana, Y. Xiang, L. Ali, Z. Abdul-Rauf, “A bidirectional LSTM deep learning approach for intrusion detection,” *Expert Systems with Applications*, **185**, 2021, doi:10.1016/j.eswa.2021.115524.
- [20] W. Xu, J. Jang-Jaccard, T. Liu, F. Sabrina, J. Kwak, “Improved Bidirectional GAN-Based Approach for Network Intrusion Detection Using One-Class Classifier,” *Computers*, **11**(6), 2022, doi:10.3390/computers11060085.
- [21] T. Acharya, A. Annamalai, M.F. Chouikha, “Optimizing the Performance of Network Anomaly Detection Using Bidirectional Long Short-Term Memory (Bi-LSTM) and Over-sampling for Imbalance Network Traffic Data,” *Advances in Science, Technology and Engineering Systems Journal*, **8**(6), 144–154, 2023, doi:10.25046/aj080614.
- [22] M. and B.E. and L.W. and G.A.A. Tavallae, “A detailed analysis of the KDD CUP 99 data set,” in *2009 IEEE Symposium on Computational Intelligence for Security and Defense Applications*, IEEE, 2009, doi:10.1109/CISDA.2009.5356528}.
- [23] L. Dhanabal, S.P. Shantharajah, “A Study on NSL-KDD Dataset for Intrusion Detection System Based on Classification Algorithms,” *International Journal of Advanced Research in Computer and Communication Engineering*, **4**, 2015, doi:10.17148/IJARCCCE.2015.4696.
- [24] N. Moustafa and J. Slay, “UNSW-NB15: a comprehensive data set for network intrusion detection systems (UNSW-NB15 network data set),” *2015 Military Communications and Information Systems Conference (MilCIS)*, Canberra, ACT, Australia, **2015**, 1-6, doi:10.1109/MilCIS.2015.7348942.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Tracing the Evolution of Machine Translation: A Journey through the Myanmar (Burmese)-Wa (sub-group of the Austro-Asiatic language) Corpus

Florance Yune\*, Khin Mar Soe

Natural Language Processing Lab, University of Computer Studies (Lecturer), Yangon, Myanmar

### ARTICLE INFO

Article history:

Received: 02 December, 2023

Revised: 06 January, 2024

Accepted: 06 January, 2024

Online: 20 January, 2024

Keywords:

Machine Translation (MT)

Low-resource Languages

Myanmar-Wa Corpus

Language Preservation

Transfer Learning

Long Short-Term Memory

Recurrent Neural Network

Pivot-Based SMT

### ABSTRACT

Machine Translation (MT) has come a long way toward reducing linguistic gaps. However, its progress in efficiently handling low-resource languages—such as the Wa language in the Myanmar-Wa corpus—has not received enough attention. This study begins with a thorough investigation of the historical development of MT systems, painstakingly following their development against the complex background of the Myanmar-Wa language region. Using an interdisciplinary methodology that integrates linguistics, technology, and culture, this investigation reveals the transformative journey of Machine Translation (MT) in its pursuit of overcoming linguistic barriers. It offers a thorough study that clarifies the opportunities and limitations present in MT's progress. More broadly, by clarifying the complex relationship between technology and linguistic diversity, our work not only advances our understanding of MT's evolutionary history but also supports the conservation of endangered languages, like Wa language. The research's conclusions have implications that go beyond machine translation to the larger conversation about language preservation and how technology development coexists harmoniously. Notably, this paper is an extension of work originally presented in "2023 IEEE Conference on Computer Applications (ICCA)", acknowledging its foundation and presenting substantial advancements.

### 1. Introduction

The critical role played by machine translation (MT) in eliminating linguistic barriers within the context of an expanding era of globalization cannot be overstated [1]. Essential instruments for fostering international collaboration, knowledge dissemination, and cross-cultural exchanges, MT systems have evolved to facilitate communication across linguistic and cultural boundaries. This paper aims to investigate the historical development of MT systems, with a specific focus on their adaptation to the intricate linguistic environment of the Myanmar-Wa corpus. The necessity for precise and efficient language translation is underscored by the accelerated pace of globalization, characterized by the seamless flow of ideas, information, and business across international borders [2]. Propelled by advancements in artificial intelligence and computational linguistics, machine translation has emerged as a pivotal element in the global communication paradigm. It has not only surmounted the constraints of traditional translation but has also enabled the bridging of languages previously considered challenging due to limited resources, as exemplified by the Wa language within the Myanmar-Wa corpus.

The ensuing methods are structured to fulfill the objectives of this study. Commencing with a historical overview of machine translation, encompassing its development and initial challenges, the analysis subsequently delves into the distinctive attributes of the Myanmar-Wa corpus, including its cultural significance and linguistic characteristics, along with the complexities associated with data acquisition. This research article proceeds to scrutinize early attempts at applying MT to the Myanmar-Wa corpus, elucidating their deficiencies and limitations. Moreover, it investigates innovative approaches and technological advancements that have enhanced MT for this specific language combination.

The subsequent section of the study involves linguistic analysis, dissecting the unique characteristics of the Wa language and examining the metrics and procedures employed to evaluate the efficacy of the MT system. Emphasizing the pivotal role that improved MT plays in cultural preservation and the broader objective of language revitalization, this analysis focuses on the practical applications and consequences of enhanced MT for the Myanmar-Wa language pair. Finally, the system's investigation concludes with an overview of major findings, a discussion of their

\*Corresponding Author: Florance Yune, [floranceyune@ucsy.edu.mm](mailto:floranceyune@ucsy.edu.mm)

implications, and concluding reflections on the dynamic field of machine translation and its critical role in bridging linguistic gaps.

The study begins with the acknowledgment that this paper is an extension of the work originally presented in the conference paper, "Myanmar-Wa Machine Translation based on LSTM-based Deep Learning Encoder-Decoder Model" [3].

## 2. Related Works

The evolution of machine translation (MT) stands as an enduring quest to surmount linguistic barriers, a journey that was inaugurated by the visionary insights of Warren Weaver. In [4] the author presented groundbreaking 1949 memo, Weaver introduced the concept of computer-assisted language translation, marking the genesis of MT as a distinct area of study and laying the foundation for subsequent developments. Early MT initiatives, notably the Georgetown-IBM experiment in the 1950s, played a pivotal role in establishing the rudiments of MT technology. Despite their nascent status, these early algorithms encountered formidable challenges, often yielding imprecise and unnatural translations. The intricacies of human language, encompassing idiosyncrasies, context-dependency, and cultural nuances, presented a significant hurdle. As linguistic theories matured and computational power increased, MT systems began to exhibit promise.

A pivotal advancement in machine translation materialized with the advent of rule-based MT in the 1960s and 1970s. These systems employed grammatical analysis and language rules to generate translations, exemplified by the Systran system initially designed for diplomatic and military communication [5].

The landscape of MT underwent transformative shifts with the introduction of statistical machine translation (SMT) systems in the 1980s, as seen in IBM's *Candide* and the European Union's EUROTRA project, as presented in [6]. Utilizing statistical models and extensive multilingual corpora, SMT systems revolutionized translation paradigms, ushering in a data-driven methodology that significantly improved translation accuracy and fluency. The 21st century heralded a new era in machine translation with the rise of neural machine translation (NMT) models, including recurrent neural networks (RNNs) and transformer models like the Transformer and BERT. Fueled by deep learning, these models exhibited remarkable proficiency in recognizing idiomatic expressions, context, and intricate linguistic patterns, with further enhancements through the integration of attention mechanisms [7].

In tandem with these advancements, MT systems grapple with challenges, particularly in handling low-resource languages like Wa in the Myanmar-Wa corpus. The absence of extensive bilingual corpora for low-resource languages hampers the development of precise and contextually appropriate translations. Additional complexities arise from linguistic nuances, such as tone systems, restricted written forms, and complex morphology. The lack of established grammatical rules and linguistic resources in these languages poses a challenge to the effective generalization of machine translation systems.

In response to these challenges, the field of machine translation (MT) has evolved with innovations such as transfer learning approaches, cross-lingual pre-training, and adaptations of existing models to low-resource languages. In [8] the author introduced a notable contribution to the field in "Cross-lingual Language Model Pretraining" (*Advances in Neural Information Processing Systems*, 2019). This concept emphasizes the pretraining of

language models in a cross-lingual manner, enhancing their ability to transfer knowledge across languages. To address the issues faced by low-resource languages and enhance inclusivity, researchers are exploring novel approaches. This study traces the historical development of machine translation (MT) systems, emphasizing their adaptation to the Myanmar-Wa corpus. Through an examination of technological advancements, the study sheds light on strategies and innovations aimed at achieving efficient Wa language translation. Crucially, this work extends beyond the realm of machine translation, contributing to broader domains such as linguistics and language preservation. By scrutinizing the challenges and solutions in adapting MT to the Myanmar-Wa corpus, the study contributes valuable insights to the growing body of information on underrepresented languages. This effort plays a vital role in ensuring the survival and resurgence of languages in a globalized environment [9].

Furthermore, in [10], the authors introduced a novel approach to NMT by incorporating syntax-directed attention mechanisms. This research, presented at the AAAI conference on artificial intelligence, delves into the utilization of syntactic structures to enhance the attention mechanism in NMT systems. By integrating syntax-directed attention, the researchers aimed to improve the model's comprehension of sentence structures and syntactic relationships, potentially leading to more accurate and contextually relevant translations. In [10] the author found the valuable insights to the ongoing advancements in NMT, particularly in addressing syntactic complexities for improved translation quality.

## 3. Myanmar-Wa Corpus

The extraordinary linguistic tapestry that is the Myanmar-Wa corpus derives its significance from the merging of two separate languages, Wa and Myanmar. This corpus provides insights into the difficulties and advancements in machine translation in addition to serving as a monument to linguistic diversity [11].

The Wa language, which is spoken by the Wa people, an ethnic group that live in the border regions of China and Myanmar, is stored in the Myanmar-Wa corpus, which is located inside this complicated linguistic terrain. MT systems face significant hurdles due to the unique linguistic complexities of the Myanmar-Wa corpus, which include a limited written form and a rich tone system [12]. These difficulties include maintaining cultural context, grammatical structures, and subtle semantic differences—all essential components of accurate translation.

### 3.1. Myanmar Language

Spoken largely in Myanmar, the Burmese language, also called the Myanmar language, is linguistically rich and culturally significant. It has distinct linguistic characteristics that present intriguing opportunities and problems for machine translation systems. This section gives a general introduction to Burmese, emphasizing its unique vowels and consonants, and discusses its use in the larger field of machine translation. Three basic character groups—vowels, consonants, and medial—combine to form Myanmar writing, and each adds to the language's complex structure. In addition, there are 33 primary consonants in the Myanmar language, each with a unique phoneme.

The tonal and analytical characteristics of the Myanmar language contribute to its linguistic complexity [13]. Its script,

which is based on the Indian (Brahmi) prototype, has changed over the course of a millennium to reflect Myanmar's rich linguistic past. Significantly, the Myanmar language has functioned as a symbol of cultural identity within the nation as well as a means of communication. Since it serves as the foundation for translation into the Wa language, it is imperative to comprehend its subtleties in the context of machine translation.

A wide variety of vowels and consonants make up Burmese. Voiced and voiceless nasals (m, n, ɲ, ŋ), stops and affricates (b, d, dʒ, ɡ, p, t, tʃ, k, ʔ), and fricatives (ð, z, θ, s, ʃ, sʰ, h) are among its consonantal inventory. Burmese also has voiceless approximants (l̥, m̥) and voiced approximants (l, j, w). Burmese consonants are peculiar compared to those in many other languages, which presents an interesting linguistic problem for machine translation systems.

Example Consonants:

Voiced Nasals: m (မ), n(န)

Voiceless Stops: p (ပ), t (တ), k (က)

Fricatives: s (စ), h (ဟ)

Voiced Approximants: w (ဝ), j (့)

There are several vowel sounds in the Burmese vowel system, such as diphthongs and monophthongs. Along with a variety of low and mid vowels, the monophthongs include the high vowels /i/ and /u/, as well as the middle vowels /ɪ/ and /ʊ/. These vowel sounds can be combined to generate diphthongs, which give languages distinctive phonetic patterns and increased complexity.

Example Vowels:

High Vowels: /i/ (အိ), /u/ (ဥ)

Central Vowels: /ɪ/ (အိး), /ʊ/ (ဥ့)

Low Vowels: /a/ (အ), /ɛ/ (ဧ)

### 3.2. Wa Language

The Wa language is still not as widely researched as the well-documented Myanmar language. Wa is a language belonging to the Mon-Khmer family that is spoken by about 950,000 people, mostly in northern Burma and some surrounding areas of Thailand and China. There are other varieties of Wa, such as Parauk, Vax, and Avax, each with a variety of dialects [14].

The Parauk language—also referred to as Phalok, Baroag, Praok, or Standard Wa—takes center stage in the Myanmar-Wa corpus. About 400,000 people in Burma speak it, mostly in the Shan States in the southeast, east, and northeast. Furthermore, the southern parts of China are home to the Parauk language. Wa is a language with unique linguistic characteristics, belonging to the Mon-Khmer family. An overview of the consonants in the Wa language is given in Table 1. Conversely, Table 2 displays the variety of vowels that add to the language's richness.

With limited English-language resources, the Wa language has not received much attention in grammatical research, despite its importance in the Myanmar-Wa corpus. The lack of thorough grammatical research emphasizes the importance of programs aimed at bridging the language divide between Wa and Myanmar [15].

Table 1: Consonants of Wa-Language

<b>K</b> က [k]	<b>KH</b> ခဲ [kʰ]	<b>NG</b> ငဲ [ŋ]	<b>S</b> စ [s]	<b>SH</b> ရှ [sʰ]	<b>NY</b> ည [ɲ]
<b>T</b> တ [t]	<b>TH</b> ထ [tʰ]	<b>N</b> န [n]	<b>P</b> ပ [p]	<b>PH</b> ဖ [pʰ]	<b>M</b> မ [m]
<b>Y</b> ယ [j]	<b>R</b> ရ [r]	<b>L</b> လ [l]	<b>V</b> ဗ [v]	<b>W</b> ဝဲ [ua]	<b>H</b> ဟ [h]
<b>X</b> အ [ʔ]	<b>C</b> ကျ [c]	<b>CH</b> ချ [tʃ]	<b>D</b> ဒဲ [d]	<b>G</b> ဂ [g]	<b>Q</b> ဂဲ [ŋg]
	<b>B</b> ဘ [b]	<b>F</b> ဖဲ [vʰ]	<b>J</b> ဂျ [ɲ]	<b>Z</b> ဧ [z]	

Table 2: Vowels of Wa-Language

<b>A</b> အာ [a]	<b>E</b> အေ [ei]	<b>IE</b> အဲ [ɛ]	<b>AW</b> အော် [ɔ]	<b>OI</b> အွဲ [oi~ɔi]
<b>AO</b> အော့ [au]	<b>AU</b> အော့ပ် [aʊ]	<b>AI</b> အိုင် [ai]	<b>I</b> အီ [i]	<b>O</b> အို [o]
	<b>U</b> အူ [u]	<b>EE</b> အေီး [ɛɛ]	<b>EU</b> အေး [ɛʊ]	

Tables 1 and 2 present Wa's phonological features, encompassing its vowels and consonants, whereas Figure 1 depicts the Wa classification (Mon-Khmer Family Tree).

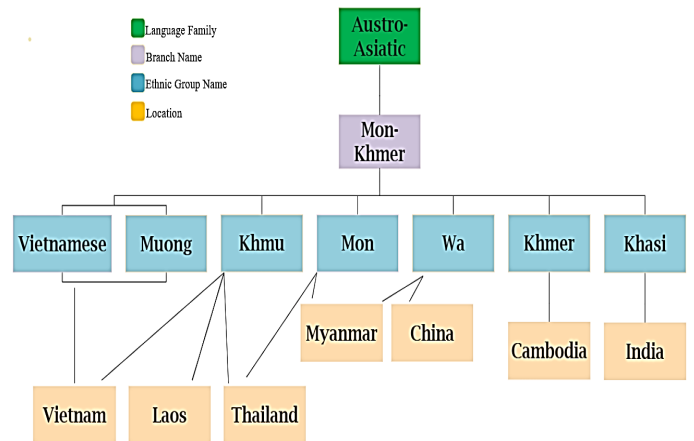


Figure 1: Mon-Khmer Family Tree [16]

### 3.3. Usage in Machine Translation

Burmese is becoming increasingly important in the worldwide environment, which makes it relevant in the field of machine translation. Accurate translation services between Burmese and other languages are in high demand due to Myanmar's recent political upheavals and growing involvement in regional and global issues.

Take, for instance, the translation of the straightforward phrase "Hello, how are you?" from English to Burmese. In this situation,



Burmese has unique phonological and grammatical patterns that machine translation systems must understand in order to provide a result that is accurate and suitable for the target culture.

Example Translation

"Hello, how are you?" (English)  
 "မင်္ဂလာ ပါ ၊ နေကောင်း ပါ သလား ။" (Burmese)

“Bawk yam pa mhawm, nyawm ot maix” (Wa)  
 (here “maix” means “you” in English and hidden word (နင်္ဂါ သင်္ဂါ) in myanmar)

Here, the Burmese translation effectively conveys the subtleties of the target language as well as the customs of polite conversation, highlighting the value of precise translation in cross-cultural communication. In conclusion, there are many interesting opportunities and problems for machine translation research because of the Burmese language's distinctive morphology, phonology, and cultural relevance. Accurate translation between Burmese and other languages is becoming more and more important as Myanmar continues to interact with the outside world, highlighting the need for improvements in machine translation technology catered to this unique language.

Table 3: VSO-SVO word order variation observed in Wa clauses

ကျွန်တော်တို့ နိုင်ငံ သည် နိုက်ပျိုးရေး နိုင်ငံ တစ်ခု ဖြစ်သည် ။	Hak tiex ex mawh hak tiex sum van tix mu heue	Our country is an Agro-based country.
ငါတို့ ၁၀ နာရီ မှာ အစည်းအဝေး လုပ် တယ် ။	Ex yuh bang kum yam kao taix	We had a meeting at 10 o'clock.
မင်း ဘယ် နေရာ မှာလဲ ။	Maix dee mawx	Where were you?
ဝမ်းနည်း ပါတယ်	Voi rhawm heue	I'm sorry.
ငါ မေ့ သွားတယ် ထင်တယ် ။	Pi awm aux tix yong	I guess I forget.
ဘာလို့ မေ့ တာလဲ ။	Jao tix pi awm lie	How could you forget?
ငါ မသိ တော့ဘူး ။	Ang aux lai tawng	I don't know.

3.4. An Academic Exploration of the Wa Language

Southeast Asia is where most Wa speakers are found. There are typological characteristics in this language that call for more investigation. Wa is regarded as a head-initial language, which means that modifiers like adjectives, relative clauses, and numbers come after nouns and verbs come before objects. Wa's subject and object markings are logical given where they fall in the sentence. Wa's particular grammatical structures are further demonstrated by the placement of the negation before the verb.

A detailed examination of the Wa language reveals a number of fascinating conclusions. For example, in Wa, several adverbs from other languages are articulated as verbs. Wa also has two alternate clause word orders, SVO and VSO as shown in Table 3, and allows components to be moved out of noun phrases. Negation frequently involves a supplementary negation particle. Interestingly, this shift in word order seems to depend on the kind of clause rather than verb transitivity or semantics.

The Wa language has received comparatively little scholarly study, despite its importance within the Mon-Khmer linguistic family. research now in existence primarily concentrate on phonological characteristics; no thorough grammatical research have been published in English. Especially, it explores the unique VSO-SVO word order variation found in Wa clauses.

3.5. Data Collection and Corpus Size Analysis for Myanmar-Wa Machine Translation

The establishment of the Myanmar-Wa corpus marked a significant milestone as it represented our inaugural venture into crafting a corpus tailored for machine translation. Throughout its development, meticulous attention was dedicated to linguistic nuances, with the Myanmar-Wa-Chinese dictionary serving as a valuable point of reference. Despite our commitment to precision, the absence of parallel data for the Myanmar-Wa language pair posed a formidable challenge during the data gathering phase.

The unique nature of low-resource languages, exemplified by Wa, presented limited opportunities for text alignment compared to more commonly spoken language pairs like English-Spanish or English-French, which benefit from extensive bilingual corpora. In response to this hurdle, we employed innovative approaches such as crowdsourcing, cross-lingual transfer learning, and data augmentation to augment the corpus's size and diversity.

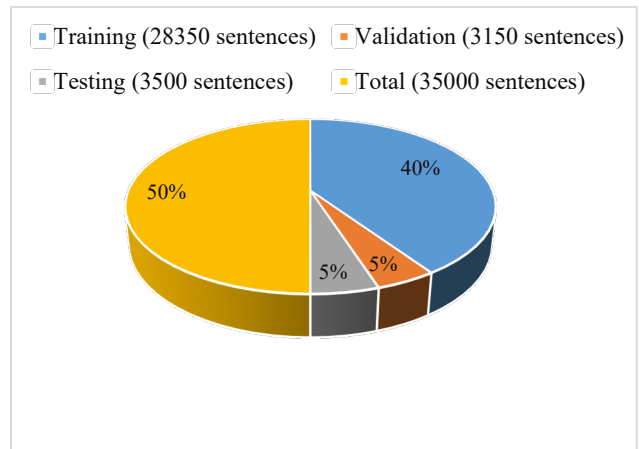


Figure 2: Corpus Size for Myanmar-Wa

It is crucial to note that the Myanmar-Wa corpus not only represents a pioneering effort in our machine translation endeavors but also holds immense value for future research initiatives. By addressing the specific data constraints inherent in low-resource languages, these initiatives aimed to lay the groundwork for the development of machine translation models characterized by heightened precision and contextual appropriateness.

Furthermore, the Figure 2 illustrates the corpus size chart provides a visual representation of the volume of data utilized for

training, validation, and testing phases in Myanmar-Wa language pair.

The Myanmar-Wa corpus, representing the pioneering effort in machine translation for this language pair, is strategically divided for training, validation, and testing. The training set comprises 28,350 instances, the validation set 3150, and the test set 3500, totaling 35,000 instances. This meticulous segmentation ensures a robust and comprehensive foundation for developing precise and contextually appropriate machine translation models for Myanmar-Wa.

#### 4. Methodologies of the Proposed System

The evolution of machine translation (MT) has witnessed remarkable success, driven by the availability of massive parallel corpora and advancements in Natural Language Processing (NLP). Despite these strides, adapting MT to the unique linguistic challenges posed by the Myanmar-Wa language pair has been met with considerable difficulties. Early attempts faced restrictions and flaws, primarily stemming from the scarcity of aligned sentences in the Myanmar-Wa corpus, hindering the development of reliable translation models. The absence of parallel corpora also impeded the progress of Statistical Machine Translation (SMT) techniques [17].

The low-resource nature of the Wa language, characterized by peculiar grammar and a small vocabulary, further compounded the challenges. Traditional machine translation models, designed for high-resource languages, struggled to capture the nuances of Wa [18]. Linguistic distinctions between Wa and Myanmar, encompassing vocabulary, syntax, and word order, posed additional hurdles, leading to suboptimal translation quality. Cultural allusions and subtleties, integral to accurate translation, were often overlooked in early efforts, underscoring the need for creative solutions.

To overcome these challenges, scholars and industry professionals explored innovative approaches, leveraging cutting-edge NLP methods and modifying existing models to suit the Myanmar-Wa language combination. The journey towards enhancing machine translation for Myanmar-Wa unfolded through various stages, each contributing to the refinement and improvement of translation quality.

##### 4.1. Statistical Machine Translation (SMT)

In the initial phases, Statistical Machine Translation (SMT) techniques were explored, but progress was hampered by the lack of parallel corpora. The limitations imposed by the absence of aligned sentences in the Myanmar-Wa corpus prompted the exploration of alternative methodologies.

##### 4.2. LSTM-based Machine Translation Models

As a response to the challenges faced by SMT, Long Short-Term Memory (LSTM) models were introduced. LSTMs, with their ability to capture sequential dependencies, showed promise in handling the complexities of the Wa language. The LSTM architecture, though an improvement, still grappled with the unique linguistic characteristics of Wa, necessitating further advancements. Figure 3 describes the LSTM encoder-decoder architecture.

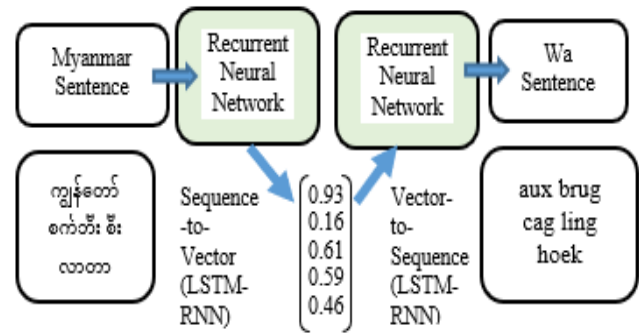


Figure 3: Architecture of LSTM Encoder-Decoder

##### 4.3. Transformer-based Machine Translation Models

The advent of transformer-based models marked a pivotal shift in the trajectory of machine translation for Myanmar-Wa. The transformer architecture, with its attention mechanisms and positional embeddings, introduced a novel approach to handling sequential dependencies. The dynamic representation of the transformer's output, dependent on the input sequence's length, addressed issues related to varying word orders and contextual nuances, contributing to improved translation quality.

Figure 4 illustrates the architecture of the Transformer Encoder-Decoder, highlighting the interconnected layers and the flow of information between the encoder and decoder. The transformer's innovative design, incorporating attention mechanisms and positional embeddings, marks a significant advancement in machine translation architectures.

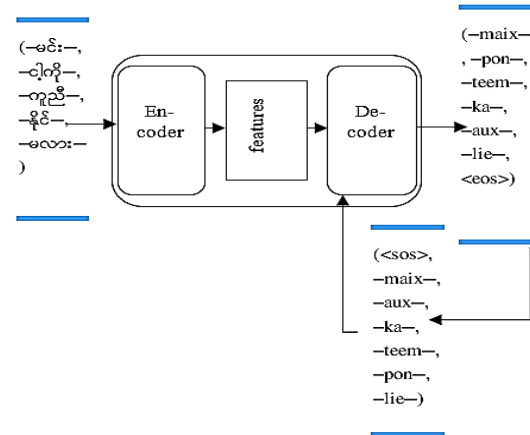


Figure 4: Architecture of Transformer Encoder-Decoder

##### 4.4. Traditional Transfer Learning on MT5 Pretrained Model

The concept behind traditional transfer learning is rooted in enhancing a machine's performance in a related task by leveraging knowledge from a pre-existing model. In Neural Machine Translation (NMT), this involves training a new model (child) with initial parameters from a previously trained model (parent) and a sparse parallel corpus. The Myanmar-Wa corpus, our pioneering effort in machine translation, showcases the utilization of transfer learning. Despite challenges in aligning low-resource languages like Wa, innovative methods, including crowdsourcing and data augmentation, were employed to augment the corpus's size and diversity.

The corpus size chart (refer to Section 3.5) illustrates the division of the Myanmar-Wa corpus into training (28,000 instances), validation (5,600 instances), and test (7,000 instances) sets, totaling 35,000 instances. Transfer learning benefits machine translation in several ways: reducing the need for extensive data, improving translation quality, accelerating training, and enabling cost-effective multilingual translation.

$$\theta_{fine-tuned} = \operatorname{argmin}_{\theta} (L(D_{Myanmar-Wa}, \theta) + \lambda L(D_{pretrained}, \theta))$$

where  $D_{Myanmar-Wa}$  represents the Myanmar-Wa parallel corpus,  $D_{pretrained}$  denotes the pre-trained multilingual corpus,  $\theta$  signifies model parameters, and  $\lambda$  is the hyperparameter governing the pre-trained model's influence.

The use of the MT5 model significantly enhances translation quality for Myanmar-Wa. Transfer learning, coupled with SMT, addresses data scarcity, and LSTM-based models capture language dynamics.

#### 4.5. A Pivotal Shift in Machine Translation with Myanmar and Wa Language Models

The pivotal role of machine translation (MT) in breaking linguistic barriers has been a relentless journey, marked by significant technological advancements. This section outlines the construction and training of a machine translation model, leveraging a pivot transfer learning approach for enhanced translation capabilities. The study focuses on the Myanmar (Burmese)-Wa corpus, representing a subgroup of the Austro-Asiatic language family.

The initial steps involve tokenization and encoding of the source and target sentences using state-of-the-art models. For the Myanmar language, a segmentation step is performed before tokenization. The source sentences are segmented and encoded using a Myanmar word segmentation tool [19], and then encoded using the Myanmar BERT model (UCSYNLP/MyanBERTa). The target sentences are encoded using the Myanmar word segmentation model, followed by encoding with the Facebook BART model.

Pivot transfer learning involves training an encoder-decoder architecture for machine translation. The Myanmar BERT model acts as the encoder. The Myanmar BERT model, initially designed for natural language understanding, serves as the encoder, extracting contextual information from the segmented source sentences. The language model, named MyanBERTa, is founded on the BERT architecture and tailored for the Myanmar language. MyanBERTa underwent a comprehensive pre-training phase encompassing 528,000 steps, utilizing a word-segmented dataset specific to Myanmar. This extensive corpus comprised 5,992,299 sentences, equivalent to 136 million words. The employed tokenizer is a byte-level Byte Pair Encoding (BPE) tokenizer with 30,522 subword units, learned subsequent to the application of word segmentation. The pre-trained Facebook BART model acts as the decoder, generating target translations based on the encoded information. Bidirectional AutoRegressive Transformer (BART), embodies a transformer architecture consisting of an encoder-decoder structure for sequence-to-sequence (seq2seq) tasks. The

model exhibits notable efficacy when fine-tuned for text generation applications such as summarization and translation [20].

This two-step process facilitates effective transfer of knowledge from one model to another, enhancing the overall translation performance.

The training process spans multiple epochs, with each epoch iterating through the training dataset. The combined model is fine-tuned using the AdamW optimizer, and a step-wise learning rate scheduler ensures optimal convergence. To mitigate GPU memory constraints, gradient accumulation steps are employed, allowing the model to accumulate gradients over several batches before updating the parameters.

Validation is a crucial step to assess the model's generalization on unseen data. The validation dataset, comprising the Myanmar-Wa corpus, is utilized to compute the validation loss and perplexity. Perplexity, a measure of translation quality, is calculated as the exponential of the average validation loss. This metric provides insights into how well the model captures the complexity of the language and context.

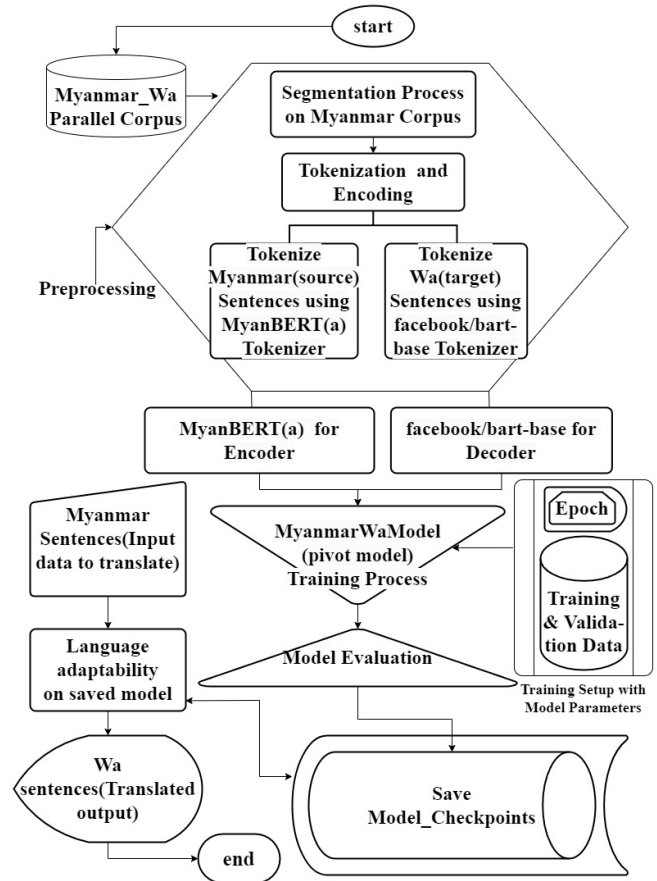
















Figure 5: Architecture of the Pivot Transfer Learning Model for Myanmar-Wa Machine Translation

To capture the model's progression, snapshots are saved at the end of each epoch. The saved models serve as checkpoints and facilitate model reproducibility. The model's evolution is crucial for understanding its learning trajectory and identifying potential areas for improvement.

This model-building process encapsulates the essence of pivot transfer learning, employing two specialized models for optimal translation performance. The journey through the Myanmar-Wa corpus showcases the adaptability and effectiveness of the proposed approach. The continual evolution of machine translation models, driven by innovative strategies and adaptability to diverse linguistic challenges, contributes significantly to the broader domains of linguistics and language preservation.

Figure 5 illustrates the architecture of the pivot transfer learning model employed for machine translation in the Myanmar (Burmese)-Wa corpus. The model leverages Myanmar BERT as the encoder and Facebook BART as the decoder, demonstrating the pivotal role of transfer learning in enhancing translation capabilities. The training process, validation steps, and model evolution are highlighted, showcasing the adaptability and effectiveness of the proposed approach in breaking linguistic barriers.

Table 4: Explanation of System Design Symbols

Symbol	Description
	Represents the start of the process or system.
	Denotes a database or data storage component.
	Signifies a step involving data preparation or preprocessing.
	Represents a component responsible for displaying information.
	Indicates a general process or computation step.
	Represents a point where data is merged or stored.
	Denotes a step involving data extraction or measurement.
	Represents manual input by a user or external entity.
	Signifies the end or termination point of the process or system.
	Represents direct data flow between components.
	Denotes the use of stored data in a process.
	Represents a predefined or standardized process.
	Indicates a loop limit or iteration constraint.
	Denotes a time delay or waiting period in the process.

In Table 4, the comprehensive index of symbols utilized in the system design is presented. Each symbol corresponds to specific

components or processes within the system, aiding in the visual representation and understanding of the system architecture.

#### 4.6. Pivot-Based Transfer Learning NMT Models with MT5

Pivot-based machine translation is a dual-stage process involving the translation from the source language (SL) to the pivot language (PL) and subsequently from the pivot language (PL) to the target language (TL). Due to the scarcity of parallel data for Myanmar-Wa, pivot-based NMT with MT5 models has been meticulously explored. The architecture is rooted in transfer learning, meticulously tailored for non-English language pairs, particularly for Myanmar-Wa.

The operational framework encompasses three crucial stages. First is the pretraining phase, leveraging source-pivot and pivot-target parallel corpora. This initial step allows the model to grasp the nuances of both source and target languages in the context of the pivot language. The subsequent stage involves the integration of a source-target parallel corpus for iterative refinement, fine-tuning the model to enhance its translation capabilities for the specific task at hand. The final stage revolves around leveraging the pivot language for pretraining both source encoders and target decoders, facilitating a holistic understanding of language dynamics.

In this context, the MT5 model stands as a robust alternative to the traditional MT5 model. By strategically selecting Myanmar-English and English-Fil (Philippines) as pivot language pairs, chosen for their analogous sentence structures, the system capitalizes on the shared linguistic characteristics between Wa and Filipino [21]. This strategic choice significantly contributes to improving translation quality and captures the synergies between these languages.

Drawing parallels to the Python code provided, the transfer learning process mirrors the combining of parameters from two distinct models ('model1' and 'model2') into a new model ('new\_model'). The encoder parameters from 'model1' are seamlessly integrated into the new model, and likewise, the decoder parameters from 'model2' are incorporated. This combination represents a fusion of linguistic knowledge from the two models into a unified architecture.

The use of pre-trained models, especially MT5, in this transfer learning paradigm underscores the flexibility and adaptability of this approach across diverse language pairs. This innovative methodology, akin to the Python code's model parameter merging, holds significant promise for advancing machine translation capabilities, particularly in low-resource language scenarios.

In this context, the MT5 model serves as a robust alternative to the MT5 model. By selecting Myanmar-English and English-Fil (Philippines) as pivot language pairs, which share similar sentence structures, the system capitalizes on the commonalities between Wa and Filipino, improving translation quality.

The utilization of pre-trained models, especially MT5, in transfer learning underscores the adaptability of this approach across diverse language pairs. This innovative methodology holds promise for advancing machine translation capabilities, particularly in low-resource language scenarios.

#### 4.7. Myanmar Word Segmentation Tool

To further refine the Myanmar to Wa machine translation, a specialized Myanmar word segmentation tool, developed by the



NLP Lab at the University of Computer Studies, Yangon, Myanmar [22], was employed. This tool, tailored for Myanmar language segmentation, enhanced the accuracy of the translation process, providing a solid foundation for subsequent advancements.

The step-by-step progression from SMT to LSTM, transformer-based models, traditional transfer learning, and pivot-based transfer learning culminated in the development of our proposed model. Each iteration addressed specific challenges posed by the Myanmar-Wa language pair, resulting in a more robust and accurate machine translation system. This comprehensive approach showcases the continual evolution of machine translation methodologies, driven by a commitment to overcome linguistic barriers and contribute to the broader domains of linguistics and language preservation.

## 5. Linguistic Analysis and Evaluation

Linguistic analysis and evaluation form a critical aspect in gauging the efficacy of machine translation (MT) systems, especially in the intricate context of Myanmar-Wa language translation. This section delves into the multifaceted linguistic properties of the Wa language and elucidates the methodologies and metrics employed to comprehensively assess the performance of MT systems on the Myanmar-Wa corpus.

### 5.1. Analyzing Linguistic Features

The Wa language, belonging to the Mon-Khmer language family, introduces numerous linguistic intricacies that offer both potential and challenges for MT. A meticulous examination of these distinctive features is indispensable for a nuanced understanding of Wa-Myanmar translation. Key linguistic characteristics that demand in-depth analysis include:

- **Word Order Flexibility:** Wa exhibits both Verb-Subject-Object (VSO) and Subject-Verb-Object (SVO) structures, introducing considerable diversity in word order. Successful translation necessitates a profound understanding of context [23].
- **Tonal Complexity:** As a tonal language, Wa relies on intonation and pitch variations for meaning. Preserving intended meaning during translation hinges on capturing tonal subtleties accurately [24].
- **Agglutinative Morphology:** Wa's agglutinative morphology involves conveying different grammatical functions through affixes added to root words. Fluent translation requires familiarity with these affixes [25].
- **Nominalization:** Verbs or adjectives in Wa commonly undergo nominalization, a process crucial for accurate translation.
- **Cultural Nuances:** The Wa language is intertwined with Wa culture and identity. To maintain cultural integrity in translations, special attention is required for terms and phrases holding cultural value.

### 5.2. Evaluation Metrics

The evaluation of MT systems for Myanmar-Wa encounters unique challenges due to limited parallel corpora and linguistic distinctions. Performance assessment utilizes various metrics and techniques, including:

- **BLEU Score:** Measures the similarity between reference translations and MT output, though it may not fully capture linguistic subtleties in low-resource language pairs like Myanmar-Wa.
- **METEOR Score:** Considers word order, stemming, and synonyms, providing a more comprehensive assessment of translation quality.
- **Manual Evaluation:** Expert linguists or native speakers evaluate translations, offering insights into cultural and contextual components.
- **Cross-lingual Embedding:** Measures semantic similarity between source and target sentences for additional information on translation quality.
- **Transfer Learning Assessment:** Evaluates how well models optimize Myanmar-Wa translation through the utilization of pre-trained knowledge from other languages.

Combining these quantitative and qualitative measurements provides a comprehensive understanding of the strengths and weaknesses of MT systems for Myanmar-Wa, guiding future advancements in the field.

### 5.3. Sentence Length Analysis

In machine translation, sentence length significantly influences model performance, especially in languages like Wa with diverse syntactic structures. This sub-section scrutinizes the impact of sentence length on Myanmar-Wa machine translation systems. The analysis involves:

- **Exploration of Distribution:** A detailed examination of sentence length distribution in both source and target languages, revealing potential correlations with translation quality.
- **Insights from Visualization:** Figure 6 illustrates the distribution of sentence lengths in the Myanmar-Wa corpus, offering insights into the nature of sentences and emphasizing the complexity of translation between Myanmar and Wa.

Longer sentences in the source language pose challenges for maintaining context and meaning in the target language. Understanding sentence length distribution aids in adapting machine translation models effectively. This visualization serves as an initial exploration, paving the way for further quantitative analysis and enhancements in models tailored for the Myanmar-Wa language pair. Each bar in the chart signifies a sentence, with the y-axis denoting sentence length and the x-axis representing the number of sentences.

Longer sentences in the source language may pose challenges for translation, particularly in preserving context and meaning in the target language. Understanding the sentence length distribution aids in adapting machine translation models to effectively handle the linguistic variations present in the Myanmar-Wa corpus. This visualization serves as a foundational exploration of sentence length characteristics, laying the groundwork for further analysis and improvements in machine translation models for the Myanmar-Wa language pair. Each bar in the chart represents a sentence, with the y-axis indicating the length of the sentence and the x-axis representing the number of sentences.

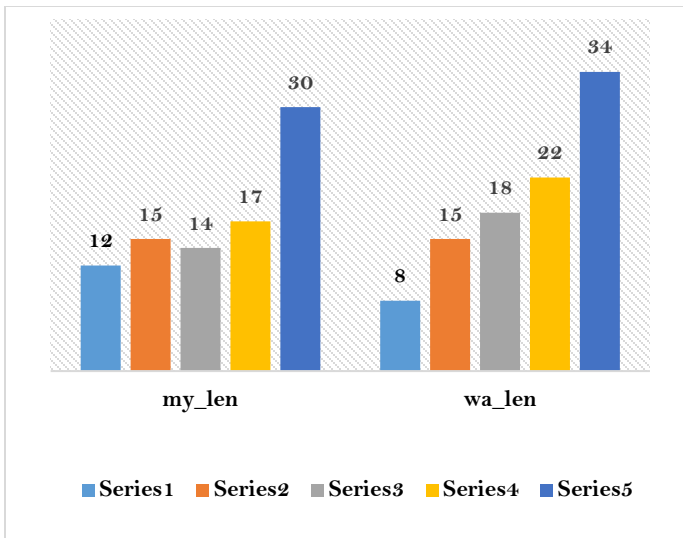


Figure 6: Sentence Length Analysis

## 6. Impacts and Future Directions

This section provides an in-depth exploration of the far-reaching implications and promising future avenues for studying improved Machine Translation (MT) in the Myanmar-Wa language pair. It examines the practical applications, language preservation aspects, and outlines key directions for future research.

### 6.1. Practical Applications

Improved MT for the Myanmar-Wa language pair has diverse practical applications, facilitating efficient communication, educational accessibility, and healthcare information translation. Additionally, it plays a crucial role in preserving Wa tradition and culture, as well as boosting tourism. Each application is detailed to highlight the significant implications.

### 6.2. Language Preservation

This subsection underscores the importance of enhanced MT as a tool for protecting and revitalizing endangered languages like Wa. It emphasizes the development of digital language resources, language learning support, and the preservation of dialectal variants and oral traditions. The broader cultural legacy of the Wa community is also discussed.

### 6.3. Future Research Directions

Identified research directions for MT in underrepresented languages like Wa are outlined, emphasizing the need for high-quality parallel corpora, domain-specific adaptations, and ethical considerations. The involvement of linguists, local communities, and technologists in collaborative projects is highlighted, along with potential enhancements through the incorporation of visual and audio modalities.

## 7. Comparative Analysis of Machine Translation Models

### 7.1. Previous Experiments on Myanmar-Wa MT System

The initial experiments on the Myanmar-Wa machine translation system utilized both Statistical Machine Translation (SMT) and Long Short-Term Memory (LSTM) models. The training data comprised 22,500 instances, with 700 instances for testing and 500 instances for validation. The analysis focused on the BLEU score, considering different batch sizes (64, 126, and

256). Results indicated that a higher batch size of 256 yielded the best accuracy in terms of BLEU score. These experiments, conducted with a limited dataset, underscored the need for a more extensive Myanmar-Wa parallel corpus in future endeavors. The SMT model achieved a BLEU score of 12.67.

### 7.2. Comparison with New Experiments and Added Corpus Size

Subsequent experiments saw the expansion of the Myanmar-Wa corpus, allowing for a comprehensive evaluation of machine translation models (Combined Model, T5, Transformer). The performance metrics for various models are summarized in Figure 7.

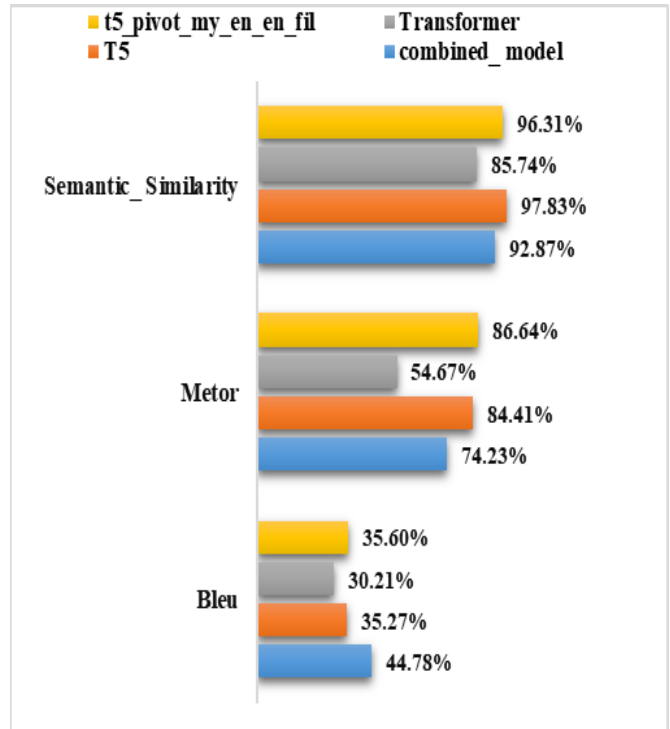


Figure 7: Performance Metrics for Various Models

### 7.3. Model Performance Analysis

- Combined Model (New):** Stands out with the highest Bleu score, indicating strong n-gram overlap. The Metor score is commendable, reflecting alignment with human reference translations. High Semantic Similarity emphasizes proficiency in maintaining semantic accuracy during translation.
- T5:** Demonstrates a strong Metor score, indicating correlation with human reference translations. Notably high Semantic Similarity suggests effective capture of semantic nuances. While the Bleu score is lower than the Combined Model, T5 excels across various metrics, showcasing versatility.
- Transformer:** Shows respectable performance in Bleu scores and Semantic Similarity. Lagging Metor score indicates potential challenges in maintaining fidelity to human reference translations. Further optimization may enhance performance.
- T5 Pivot My-En En-Fil (New):** Exhibits balanced performance across all metrics. Noteworthy Bleu and Metor scores, along with effective preservation of semantic content

indicated by Semantic Similarity. Strong contender in scenarios where a combination of metrics is crucial.

7.4. Observations and Insights

- Combined Model Outperforms: The newly introduced combined model excels across all metrics, with a Bleu score of 44.78, Metor score of 74.23, and high Semantic Similarity of 92.87.
- T5 Pivot My-En En-Fil: Demonstrates noteworthy improvements, particularly in Bleu (35.6) and Metor (86.64) scores. High Semantic Similarity (96.31) emphasizes effectiveness.
- T5 and Transformer Comparison: While T5 and Transformer models show respectable performance, T5 generally outperforms in terms of Bleu, Metor, and Semantic Similarity.
- Semantic Similarity Emphasis: Crucial for evaluating models' ability to capture meaning. The combined model excels, showcasing capability to preserve semantic nuances during translation.

7.5. Summary and Implications

The choice of the best model depends on specific translation task priorities. T5 and T5 Pivot My-En En-Fil exhibit robust performance in METEOR and semantic similarity, emphasizing versatility. Meanwhile, the Combined Model excels in Bleu scores and semantic similarity, showcasing linguistic precision. Notably, additional training data could further enhance models like the Combined Model.

These nuanced analyses should guide researchers in selecting models for machine translation tasks, recognizing that distinct aspects of performance are illuminated by different metrics. Strategies to enhance semantic preservation and leverage additional data for refining capabilities of all models could be explored in further investigations. The continuous evolution of machine translation models contributes to language understanding and communication.

It is imperative to note that the Myanmar-Wa corpus used in these experiments is the very first corpus created by the research team. The corpus size was expanded to enhance the machine translation system's performance and evaluation. Table 5 presents the sample result outputs on long sentences, middle sentences, and short sentences.

Table 5: Sample Result Outputs on Long Sentences, Middle Sentences, and Short Sentences

long sentences	
<b>combined_model</b>	source: ပြိုင်ပွဲ ဘယ်အချိန် စ ရ မလဲ ဆိုတာ မင်း ငါ့ ဆီ ဖုန်း ဆက် ပါ
	reference: you call me up when to start contest
	target: maix tah maox rhiem ka aux bang silah sang jah yam mawx
	prediction: maix tah maox rhiem ka aux sang pawk bawg yam mawx
<b>Transformer</b>	source: ပြိုင်ပွဲ ဘယ်အချိန် စ ရ မလဲ ဆိုတာ မင်း ငါ့ ဆီ ဖုန်း ဆက် ပါ

	reference: you call me up when to start contest
	target: maix tah maox rhiem ka aux bang silah sang jah yam mawx
	prediction: maix tah maox rhiem caw tix hoek yam mawx pa tah maox rhiem ka maix
<b>T5</b>	source: ပြိုင်ပွဲ ဘယ်အချိန် စ ရ မလဲ ဆိုတာ မင်း ငါ့ ဆီ ဖုန်း ဆက် ပါ
	reference: you call me up when to start contest
	target: maix tah maox rhiem ka aux bang silah sang jah yam mawx
	prediction: maix tah maox rhiem ka aux sang jah yam mawx
<b>T5 with pivot logic</b>	source: ပြိုင်ပွဲ ဘယ်အချိန် စ ရ မလဲ ဆိုတာ မင်း ငါ့ ဆီ ဖုန်း ဆက် ပါ
	reference: you call me up when to start contest
	target: maix tah maox rhiem ka aux bang silah sang jah yam mawx
	prediction: tah maox rhiem ka aux sang jah tix jah bang silah yam mawx
middle sentences	
<b>combined_model</b>	source: ဒါဟာ သိပ် ကောင်းတဲ့ စိတ်ကူး နဲ့ တူ ပါတယ်
	reference: that sounds like a good idea
	target: in awm krax keud pa kied sidaing mhawm
	prediction: in awm pa kied sidaing mhawm tete
<b>Transformer</b>	source: ဒါဟာ သိပ် ကောင်းတဲ့ စိတ်ကူး နဲ့ တူ ပါတယ်
	reference: that sounds like a good idea
	target: in awm krax keud pa kied sidaing mhawm
	prediction: in mawh grawng kaing pa sidaing mhawm
<b>T5</b>	source: ဒါဟာ သိပ် ကောင်းတဲ့ စိတ်ကူး နဲ့ တူ ပါတယ်
	reference: that sounds like a good idea
	target: in awm krax keud pa kied sidaing mhawm
	prediction: in awm mai krax keud pa kied mhawm
<b>T5 with pivot logic</b>	source: ဒါဟာ သိပ် ကောင်းတဲ့ စိတ်ကူး နဲ့ တူ ပါတယ်
	reference: that sounds like a good idea
	target: in awm krax keud pa kied sidaing mhawm
	prediction: in awm mai krax keud pa kied mhawm

short sentences	
<b>combined_model</b>	source: တို့ဘယ်မှာ ဆုံကြမလဲ
	reference: where shall we meet
	target: ex sang pup dee mawx
	prediction: ex sang pup dee mawx
<b>Transformer</b>	source: တို့ဘယ်မှာ ဆုံကြမလဲ
	reference: where shall we meet
	target: ex sang pup dee mawx
	prediction: ex sang sipup dee mawx
<b>T5</b>	source: တို့ဘယ်မှာ ဆုံကြမလဲ
	reference: where shall we meet
	target: ex sang pup dee mawx
	prediction: ex sang sipup dee mawx
<b>T5 with pivot logic</b>	source: တို့ဘယ်မှာ ဆုံကြမလဲ
	reference: where shall we meet
	target: ex sang pup dee mawx
	prediction: ex sang pup paox dee mawx

## 8. Conclusion

The journey through the evolution of machine translation within the context of the Myanmar-Wa corpus has been marked by iterative experimentation and continuous refinement. In the pursuit of refining translation capabilities, our study extensively evaluates various models, each contributing unique strengths and insights.

Among the models considered, the Combined Model (New) emerges as a standout performer, showcasing the highest Bleu score, a commendable Metor score, and exceptional Semantic Similarity. This comprehensive translation capability positions the Combined Model (New) as a leader across multiple metrics, highlighting its proficiency in maintaining semantic accuracy during translation.

T5, with its strong Metor score and notable Semantic Similarity, demonstrates versatility in translation tasks. While its Bleu score is slightly lower compared to the Combined Model, T5's performance remains commendable, underlining its effectiveness in capturing semantic nuances.

The Transformer model, although showing respectable performance in Bleu scores and Semantic Similarity, lags in Metor, suggesting potential challenges in fidelity to human reference translations. Further optimization may enhance its overall performance.

The T5 Pivot My-En En-Fil (New) variant exhibits balanced performance, particularly excelling in Bleu and Metor scores, along with high Semantic Similarity. This model stands as a strong contender, especially in scenarios where a combination of metrics is crucial.

In conclusion, our comprehensive evaluation indicates that the Combined Model (New) leads in overall performance, followed closely by T5 and T5 Pivot My-En En-Fil (New). These findings provide valuable insights for the advancement of machine translation in the Myanmar-Wa context, guiding future investigations and contributing to the broader landscape of machine translation research. The expanded corpus size, coupled with these model assessments, serves as a robust foundation for ongoing and future research endeavors.

## Acknowledgment

The Divine Creator, who gave me the permission to start the "Tracing the Evolution of Machine Translation: A Journey through the Myanmar (Burmese)-Wa (sub-group of the Austro-Asiatic language) Corpus" study, is the reason this research exists. I would like to express my sincere gratitude to Dr. Mie Mie Khin, the distinguished Rector of the University of Computer Studies, Yangon, for providing me with this exceptional chance to explore this exciting field of study. I am also grateful to Dr. Win Le' Le' Phyu, Professor; she is a guiding light from the University of Computer Studies, Yangon, and her strong support has fostered my academic endeavors and given them the time they need to grow. I also express my gratitude to my group members and UCSY NLP Lab colleagues, who have kindly offered their personal and professional perspectives, enhancing my comprehension of scientific research and life in general. My profound gratitude is extended to Dr. Khin Mar Soe, my committed research mentor, whose function as a teacher and advisor has given me immeasurable wisdom. Dr. Win Pa' Pa, a renowned professor at NLP LAB (University of Computer Studies, Yangon), deserves special commendation for her acts, which embody the traits of a model scientist and kind person. I would especially want to thank Drs. Yi Mon Shwe Sin and Aye Mya Hlaing, whose contributions to this effort were deeply appreciated. Mr. AI MHAWM KHAM is a valuable contributor who deserves special recognition for his expert translation from Burmese to WA. This translation was a critical factor in making this work possible. During the execution of this project, my family provided steadfast support. I am truly grateful to my parents and siblings for being the best role models in my life and for their unending love and support. Throughout this journey, their constant belief in me gave me more drive and determination. Finally, I would want to express my appreciation to Pastor AI NAP TAO and Associate Pastor AI KHWAT PAN for their unwavering prayers and words of spiritual support.

## References

- [1] G. Salton, "Automatic Text Processing: The Transformation, Analysis, and Retrieval of Information by Computer," IEEE Transactions on Audio and Electroacoustics, **16**(3), 366-366, Sep. 1968.



- [2] P. Hanks, "Computational Lexicography for Natural Language Processing," *IEEE Computational Intelligence Magazine*, **11**(3), 9-20, Aug. 2016.
- [3] F. Yune and K. M. Soe, "Myanmar-Wa Machine Translation using LSTM-based Encoder-Decoder Model," 2023 IEEE Conference on Computer Applications (ICCA), Yangon, Myanmar, 2023, 1-5, doi: 10.1109/ICCA51723.2023.10181692.
- [4] W. N. Locke and A. D. Booth, "Translation. Machine translation of languages", Fourteen essays, MIT Press, ISBN:9780262120029, 15 May, 1955.
- [5] W. J. Hutchins, "Machine translation: Past, present, future", Ellis Horwood, 2005.
- [6] P. Koehn, "Statistical machine translation", Cambridge University Press, 2010.
- [7] A. Vaswani and et al. "Attention is all you need", *Advances in neural information processing systems (NeurIPS)*, 2017.
- [8] G. Lample, and et al. "Cross-lingual Language Model Pretraining", *Advances in Neural Information Processing Systems*, 2019.
- [9] J. Smith, "Language Preservation in the Digital Age: Challenges and Strategies," *IEEE Computer*, **42**(10), 106-108, Oct. 2009.
- [10] K. Chen, R. Wang, M. Utiyama, E. Sumita, and T. Zhao, "Syntax-directed attention for neural machine translation," in *Proceedings of the AAAI conference on artificial intelligence*, **32**(1), New Orleans, LA, USA, April 2018.
- [11] Z. Zhizhi and Y. Qixiang. "A brief description of the Wa language", 1984.
- [12] D. Wu, "Machine Learning for Sino-Tibetan Language Translation: Challenges and Progress," *IEEE Transactions on Neural Networks and Learning Systems*, **29**(11), 5616-5631, Nov. 2018.
- [13] S. Duanmu, "The Phonology of Standard Chinese (2nd edition.)", Oxford University Press, 2007.
- [14] S. Mai. "A descriptive grammar of Wa." (Master's thesis, Payap University, 2012). Retrieved from [https://www.academia.edu/44390404/A\\_descriptive\\_grammar\\_of\\_wa\\_language](https://www.academia.edu/44390404/A_descriptive_grammar_of_wa_language).
- [15] Y. Xiang and J. Ke, "Linguistic characteristics of the Wa language", *A preliminary exploration. Language Documentation & Conservation*, **15**, 87-103, 2021.
- [16] Human Computation Institute, "Wa Dictionary Corpus". [Online]. Available: [https://www.humancomp.org/wadict/wa\\_corpus.html](https://www.humancomp.org/wadict/wa_corpus.html)
- [17] S. Hochreiter and J. Schmidhuber, "Long Short-Term Memory", *Neural Computation*, **9**(8), pp - 1735-1780, 1997.
- [18] T. Xie, "Burmese (Myanmar): An Introduction to the Script", Southeast Asian Language Resource Center, 2008.
- [19] Aye Mya Hlaing, Win Pa Pa, "MyanBERTa: A Pre-trained Language Model For Myanmar", In *Proceedings of 2022 International Conference on Communication and Computer Research (ICCR2022)*, November 2022, Seoul, Republic of Korea.
- [20] A. Lewis, Y. Liu, N. Goyal, M. Ghazvininejad, A. Mohamed, O. Levy, V. Stoyanov, L. Zettlemoyer. "BART: Denoising Sequence-to-Sequence Pre-training for Natural Language Generation, Translation, and Comprehension," *CoRR*, abs/1910.13461, 2019. [Online]. Available: <http://arxiv.org/abs/1910.13461>.
- [21] C. Callison-Burch and et al., "The META-SHARE Metadata Schema for the Description of Language Resources", *LREC*, 2008.
- [22] Win Pa Pa, Ni Lar Thein, "Myanmar Word Segmentation using Hybrid Approach", In *Proceedings of 6th International Conference on Computer Applications*, Yangon, Myanmar, 166-170, 2008.
- [23] K. Papineni and et al., "A Method for Automatic Evaluation of Machine Translation", *Association for Computational Linguistics*, 2002.
- [24] A. Lavie and A. Agarwal, "An Automatic Metric for MT Evaluation with Improved Correlation with Human Judgments", *Association for Computational Linguistics*, 2007.
- [25] G. Doddington, "Automatic Evaluation of Machine Translation Quality Using N-gram Co-occurrence Statistics", *Association for Computational Linguistics*, 2002.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Development and Usability Evaluation of Mobile Augmented Reality Contents for Railway Vehicle Maintenance Training: Air Compressor Case

Gil Hyun Kang<sup>1</sup>, Hwi Jin Kwon<sup>2</sup>, In Soo Chung<sup>1</sup>, Chul Su Kim\*<sup>1</sup>

<sup>1</sup>Industry-Academic Cooperation Foundation, Korea National University of Transportation Uiwang, 16101, Korea

<sup>2</sup>Department of Railroad Convergence System Engineering, Korea National University of Transportation, Uiwang, 16101, Korea

### ARTICLE INFO

#### Article history:

Received: 06 December, 2023

Revised: 11 January, 2024

Accepted: 11 January, 2024

Online: 20 January, 2024

#### Keywords:

Augmented reality

Railway vehicle maintenance

Air compressors

Usability evaluation

Mobile devices

### ABSTRACT

The air compressor of a railroad vehicle is an important equipment that produces compressed air used in braking systems. New visual interaction techniques were proposed and evaluated to develop effective augmented reality content for maintenance support and training of this device. To this end, modeling techniques capable of fast animation, storyboard production to support light maintenance, and visualization algorithms that implement fluid flow have been developed. In the case of air compressors using compressed air, 2D-3D line matching distance calculation algorithm and modified flow generation algorithm were proposed to implement fluid flow for visualization. If two algorithms are used at the same time, the educational effect can be enhanced by visualizing the air flow in the 3D object simultaneously on a 2D air pipe diagram. In addition, the use of flow generation algorithm alone can visualize fluid flow or quantity control, such as air discharge or lubricant supplementation. As a result of usability evaluation of 100 users of the developed air compressor augmented reality content, the system usability scale score was 76.65, which was good. Similarly, the user experience score using six questionnaires was very good with an average of 4.12. Therefore, it was found that this training content could be used very effectively for air compressor maintenance support and training at the site.

## 1. Introduction

Since railway vehicles are a means of mass transportation, it is important to perform reliable maintenance to avoid causing traffic congestion to many passengers if a failure occurs during train operation. Therefore, vehicle failures have been prevented through periodic preventive maintenance, and the lifespan of important parts has been managed and maintenance cycles have been extended through Reliability-Centered Maintenance (RCM) based on failure statistics analysis [1]. Modern railway vehicles have developed into a complex system in which various technologies are fused with the development of electronic and control technologies. In order to fully carry out the maintenance of railway vehicles with such modern advanced technology, not only the improvement of maintenance technology but also the educational techniques of maintenance personnel need to be innovated. Recently, with the development of sensor technology and communication technology and Artificial Intelligence (AI) technology, the maintenance of railway vehicles has rapidly

developed into predictive maintenance by Condition-Based Maintenance (CBM) and smart maintenance [2]-[4].

Most railway vehicles remained at the level where vehicle computers installed in intermediate vehicles and driver cabins, but event recording and information on major devices during train operation were displayed with driver front monitors and vehicle computers. Recently, however, more than 2,000 sensors have been installed and monitored to switch to a level that utilizes condition monitoring and AI technology during train operation to perform failure diagnosis as well as failure prediction and residual life prediction [5,6]. The core of this change is data-based analysis technology according to the development of computer and communication technology. Since these technologies essentially utilize AI and Internet of Things (IoT) technology using ground and onboard servers, it is expected that there will be a major change in the maintenance of the conventional human-dependent experience base. The Korean railway industry is also conducting a research project for smart maintenance of railway vehicles to overcome these technological changes as follows: (1) Predictive maintenance based on CBM by increasing the capacity

\*Corresponding Author: Chul-Su Kim, School of Railroad Engineering, Korea National University of Transportation, Uiwang, 16101, chalskim@ut.ac.kr

of new sensors and data communication, (2) Automation of the acquisition of railway vehicle light maintenance data when a vehicle in operation through a large scale vehicle scanner installed on the ground passes through a specific point such as a deposit entrance or terminal station, (3) Digitization and automation of inspection and repair records for digitization of maintenance data and connection with mobile devices, and (4) Augmented Reality(AR)-based maintenance training to replace existing computer-aided electronic manuals and mobile device-linked maintenance education contents. Research to apply AR techniques in the industry was initially a study related to the development and production of automobiles and aircraft, which was intended to be used for crash analysis, assembly optimization of batch products, and assembly production technology of complex systems [7,8]. In the automobile industry, driving support using head-up displays, assembly, design, assembly, maintenance, and manufacturing were reported as examples in the fields of in-car systems, design, maintenance, and manufacturing, such as marker-based or markerless graphic modeling using head mounted displays and AR glasses [9]-[12].

Research which applies AR technology to maintenance is also gradually expanding in the field of automobiles, starting with the maintenance of specific devices of aircraft and ships with high operational safety priorities, and railway vehicles [13]-[15]. These technological advances can be confirmed through various research survey papers [16,17].

For more than a decade, the railway sector has tried to apply AR technology to the maintenance and repair of railway vehicles using mobile devices, but it has remained at a very slow development due to the complexity of railway vehicles and limitations of IT technology [18]. The main reason for this was that the method of training many people at the same time in theoretical education and vehicles without a special education system was simple and the cost of education was low. However, this field practice is suitable for disassembling and assembling at the level of looking at and explaining the physical appearance, but there was a clear limit to education to understand the specific operating principles of complex equipment consisting of many wiring pipe checkpoints and to train complete maintenance. In addition, there is a problem that on-site classes have a high risk of safety-related accidents. Therefore, research on alternative education using video or virtual reality has been continuously conducted because the actual effect of education and training cannot be high except for education and training performed with expensive practical equipment. Recently, AR technology has also developed rapidly, such as virtual simulation of automobiles, visual design, virtual protection, and AR self-serving to reduce development costs, with the development of high-performance CPUs and graphic cards, 3D scanners, and the development of the game industry where animation technology is important [19,20]. In addition, active research is being conducted by expanding not only building infrastructure but also various manufacturing industries [21,22]. Nevertheless, the development of AR technology in the maintenance field of railway vehicles led by railway operators was very limited.

Recently, most of the AR application cases for maintenance in the railway vehicle sector have been developed mainly on main mechanical equipment such as boogie and braking devices

conducted by the research group of the National University of Transportation in Korea [23,24]. In the case of the air compressor conducted in this study, only some applications have been reported to manufacture industrial screw-type air compressors [25]. Therefore, this study aims to make it easy to use for train fleet maintenance by gradually converting important railway vehicle equipment, a collection of complex technologies such as mechanical engineering, electrical, communication, and electronic engineering, over several years to innovate existing manual-based maintenance methods in the railway industry.

Unlike previous studies such as object search and maintenance point search using head-ware devices and cameras, which are many AR research fields, this study aims to create a 3D AR model and connect this model with a database related to maintenance and a storyboard to enable reliable maintenance support and to effectively perform maintenance training with realistic content in 3D space with touch function on mobile devices [26]-[28]. This study aims to achieve the research objectives for supporting on-site maintenance of major railway vehicle equipment and effective education and training at low cost as follows: (1) Development of mobile-based maintenance support and educational training contents using AR technology, replacing the existing training methods of limited and inefficient document-based theory education and field practice, virtual reality video or Internet-based e-learning, (2) The development of contents that increase work efficiency by raising detailed information of major equipment and sub-components to the same level as actual field work. Therefore, the sub-component system is based on the part procurement BOM of the operator. Development of contents connected to tools for disassembly and assembly of parts, measuring devices or instruments for inspection, special equipment, and work procedures based on the maintenance manual, (3) Development of algorithms that enhance learning effect by implementing fluid or electric flow simultaneously with 3D AR models for operating principles and functions of equipment learned in 2D space from piping, wiring, and circuit diagrams to enhance maintenance support and education and training effect, (4) Analyze maintenance tasks to extend the user interface to inspect, measure, repair, and, Expanding practical education and training areas to provide a sense of immersion through physical equipment and interfaces, such as similar tools, etc., and (5) It increases the effectiveness of safety work through warnings related to safety work based on the operator's disaster database, such as serious injuries or deaths that may occur at the maintenance site.

The usability of AR contents developed for railway vehicle air compressor maintenance training was evaluated in two ways. The results of the System Usability Scale (SUS) survey and the User Experience (UX) experiment of 100 college students who are beginners in railway vehicle maintenance were as follows. The average SUS score of AR contents was 76.65, which was "Good." In addition, the results of the UX survey shows that the average usability score for six questionnaires was excellent at 4.12 out of 5. Therefore, AR content developed for Android-based mobile devices is expected to be conveniently used for maintenance training on a platform for maintenance efficiency for beginners in railway vehicle maintenance.

## 2. AR Contents and Developed Algorithm

### 2.1. Smart Maintenance Platform for Railway Vehicle

The Smart Maintenance Research Consortium Group, which includes the Korea Railroad Corporation, the Korea National University of Transportation, Hyundai Rotem, a railroad vehicle manufacturer, Seoul Metro, and AI software development companies, has been carrying out the railway vehicle smart maintenance research project with the support of the Korean government since 2019. Figure 1 shows the architecture of the final stage, the smart maintenance platform. The basic strategy is to install sensors on the main components of Electric Multiple Unit (EMU) and implement CBM and data analysis-based predictive maintenance in combination with AI. In addition, to improve maintenance efficiency, a maintenance decision-making system is established to support maintenance decisions in conjunction with CBM data by automating inspection and measurement and digitizing data. Finally, it is to establish a Mobile Maintenance Efficiency System (MMES) that promotes new innovation in automating maintenance records to support the implementation of new maintenance strategies and training employees using AR in mobile systems. This study introduces the development of AR contents that replace or support existing field training along with this MMES AR manual, and specifically relates to the process of developing AR contents for air compressors and the result of the user evaluation.

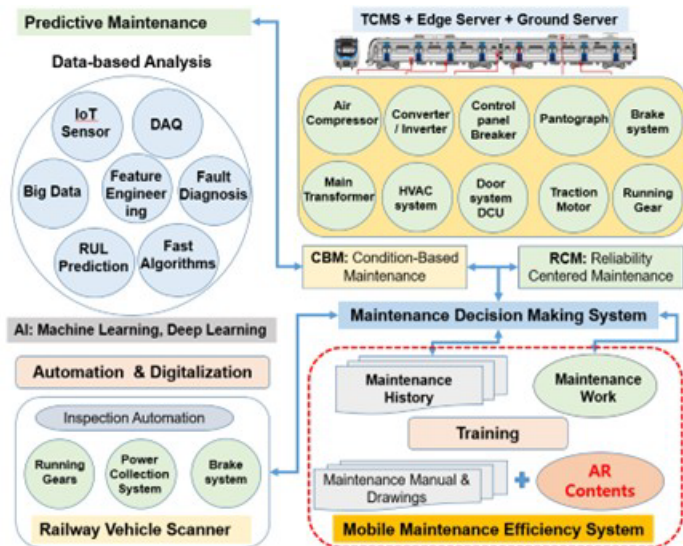


Figure 1: Architecture of Smart-maintenance platform for Railway vehicle

Train configurations vary from country to country due to population density, urban traffic characteristics, and traffic congestion. In the case of urban railroads in Seoul, a metropolitan city, 3 to 4 cars are built into one unit, and two units of 6 cars and 3 units of 10 car trainset is generally operated. Most of the major trunk lines are mass commuter transported by EMUs composed of 10 car trainset. The type of vehicle is also divided into a motor car with a propulsion device, a trailer car without power, a trailer car with cabin, and a motor car with cabin if there is a driver's room. However, most of the main equipment is installed in one unit, and when the number of unit increases, the device overlaps, but since it is the same equipment, AR contents can be manufactured for each major equipment. Most urban rail vehicles

are direct current EMUs using DC 1500V, while long-distance trains operated to the suburban lines are AC EMUs using 25kV AC [29,30]. In addition, in the case of AC-DC multiple EMU, a converter is added to the propulsion system, including the main transformer, making the vehicle complicated by installing ultra-high voltage power devices and switches. Table 1 shows the AR manuals and contents of major equipment for EMU developed as a Korean government-funded research project until recently.

Table 1: Developed AR manual and contents

Equipment	EMU	AR manual	AR training
Pantograph	AC/DC	○	○
Main transformer	AC	○	
Air compressor	common	○	○
Propulsion system	AC/DC	○	
Traction motor	common	○	
Auxiliary power	common	○	
Battery box	common	○	
Bogie	common	○	
Wheelset	common	○	
Axle box	common		○
Access door	common	○	○
HVAC	common	○	
Brake operating unit	common	○	○
Wheel brake system	common		○
Disk brake system	common		○
Air spring	common		○
Front connector	common		○
Inter connector	common		○

The reason why AR manual and AR contents differ on the MMES platform is that different small groups have different development objectives and different intellectual property owners. Therefore, some equipment is developed and compared with each other, and some equipment has been developed independently, but plans to optimize in the future through upgrades.

### 2.2. AR contents for Mobile Maintenance Efficiency System

Figure 2 shows the MMES installed on the mobile phone and the drive screen. The left side shows the app screen installed on the mobile phone and the screen consisting of maintenance work, maintenance manual, maintenance record, and AR content in the center. The picture on the right shows the developed AR manual screen by pushing the menu. The MMES platform was developed for the main purpose of digitizing maintenance records for smart maintenance, and when users perform maintenance and move to designated routes, and store maintenance history on digital workbooks according to maintenance guidelines, data is transmitted to ground servers, and AR manuals were produced to support maintenance in this process.



Figure 3 shows the screen of AR contents developed by the KNUT research team for educational and training purposes. It can be used by using a separate AR contents app on the upper left side of the left screen of Figure 2.

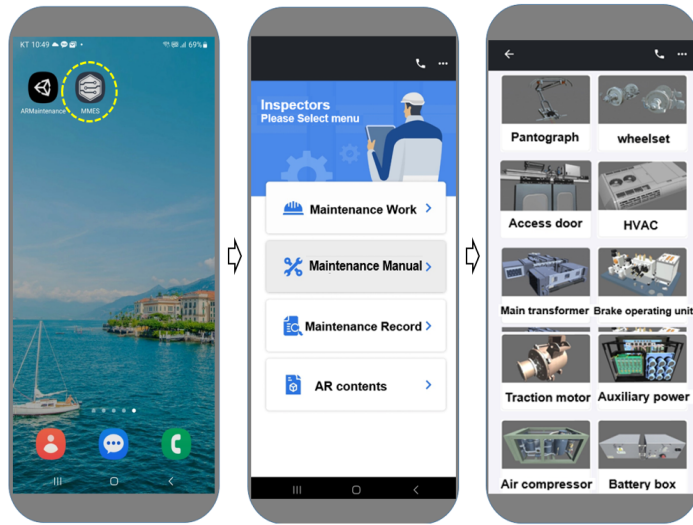


Figure 2: The AR manual screen for MMES

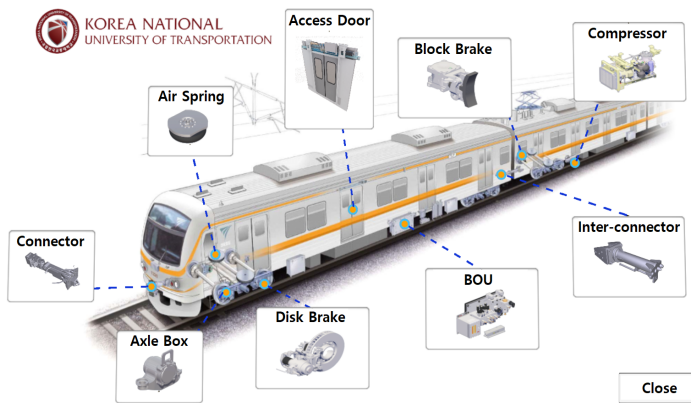


Figure 3: The AR training APP screen for MMES

### 3. Development of AR contents for Air Compressor

The air compressor of a railway vehicle is one of the essential equipment for starting, operating, and stopping trains. When starting an electric railway vehicle, compressed air is used to raise the current collecting pantograph to supply high-voltage electricity by contacting the power wire. During train operation, air is supplied to the braking device to stop the train at the station, and compressed air is supplied to the door system to enable passengers to get on and off the train. In the case of the 10-car electric vehicle of the Korea Urban Railroad, three sets are installed on the lower floor of the body, as shown in Figure 4. At least one air compressor is required for each train configuration unit. Modern air compressors are mainly reciprocated or screw types. In the operation, the motor is rotated to compress atmospheric pressure air to high pressure, store compressed air in the main air reservoir, and use compressed air to automate the brake system and opening and closing doors of the train. In the case of long-distance passenger trains, it is also used to supply water to the toilet or to provide water for handwashing. Thus, as with braking systems, air compressor maintenance training is important.



Figure 4: Air compressor of the Electric multiple unit

#### 3.1. AR Contents Development Process

To produce AR content for maintenance, it is necessary to select and model equipment for content and organize and design content. Figure 5 shows the planning and design procedure of AR content.

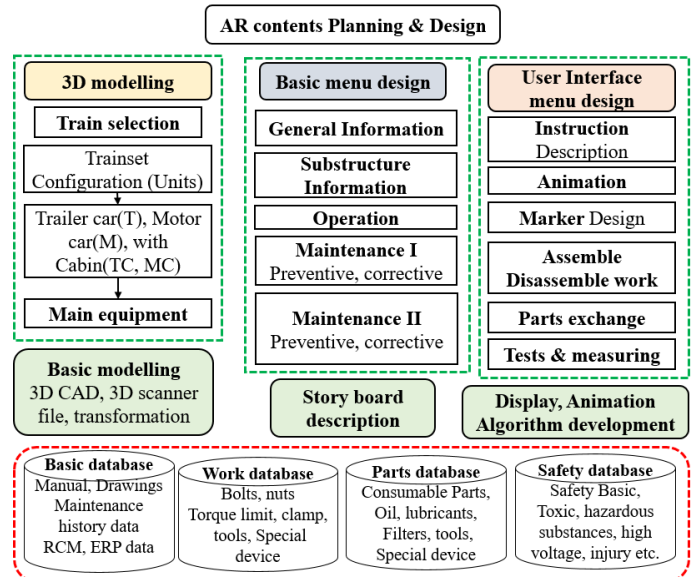


Figure 5: Procedures of AR contents planning and design

Modeling of an equipment assembly of a railway vehicle takes time and cost to model because the assembly is relatively large size and the number of parts is large, and there are typical methods as follows. (1) Using the parts manufacturer's 3D drawings to convert to a format for an AR production program: It uses 3D

CAD files, which is the most convenient method. (2) 3D scanning and measurement methods: 3D scanning parts are modeled by using scanners, and the rest are modeled by making cad drawings. (3) Model with similar dimensions: less precision, but with reference to drawings and real objects modeled with similar dimensions.

Figure 6 shows a 3D modeling technique that uses the Geometric Design X and Unity program to convert air compressor large CAD raw files into FBX to read reduced capacity files and convert them into APK format models, small files that run fast on mobile devices [31]. Depending on the type of equipment, 3D CAD format file formats are mostly large in capacity and cannot be used directly on mobile devices. It is also important to use a 3D scanner to create an initial device model file, since 3D CAD files are the most accurate models.

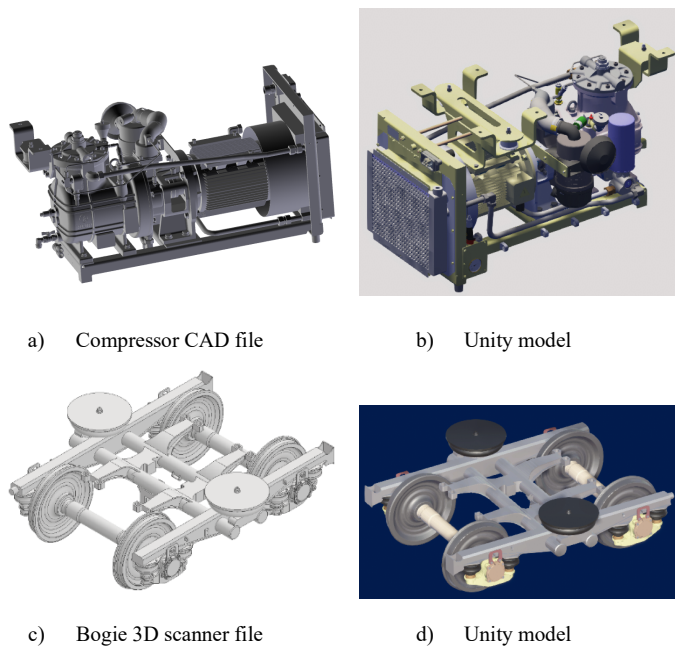


Figure 6: 3D modelling for AR contents

Figure 7 shows an example of modeling an air compressor by acquiring external coordinate data using a 3D scanner. It is essential to read and check it in the CAD program, then convert it to FBX format so that Unity programs can import and work on AR content.

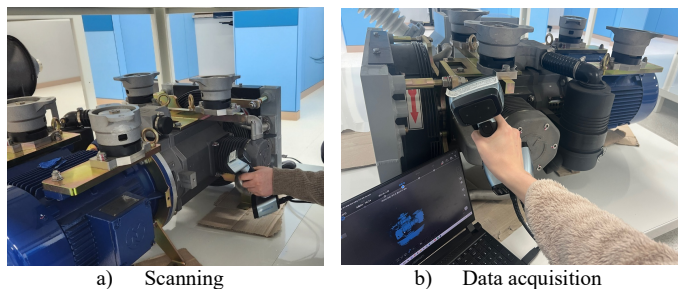


Figure 7: Process of using a 3D scanner

Table 2 shows a comparison of file size of among the air compressor, a railway vehicle bogie and brake operation unit. The bogie case is a small capacity CAD-compatible file STP format

using a 3D scanner, despite being a large structure. As shown in Table 2, 3D scanner files are only 6.8 M, which is advantageous for fast animation. Of course, the CAD program SolidWorks drawing file increases the file size to about 168MB only for the bogie frame.

Table 2: 3D modelling and file size

Equipment	cad file/3D scanner file size	Unity import file size (fbx)
Compressor	288MB (3D cad file)	68MB (77% reduced)
Bogie	168MB sldprt file (frame only) 6,497KB stp file/3D scanner	9,427KB
BOU	115MB (3D cad file)	22MB (81% reduced)

Therefore, for single structures that do not require disassemble, the dimensional information file of the 3D scanner is very efficient in creating AR models. This seems to be due to the fact that the welding structure of the truck frame is perceived as a part on the scanner, unlike the 3D CAD file, whereas the air compressor 3D CAD file consists of many subsystems, and the complexity of the structure increases the size of the CAD file.

### 3.2. Creating Storyboard for Maintenance

Storyboards are an important process of determining the overall structure of content and determining the hierarchy of objects to be implemented as unity, and determining the validity of training to determine information to be delivered to users. (1) Basic menus such as device general information, level determination of sub-components of the device, and maintenance classification are defined through the content major classification of the selected equipment assembly. (2) In order to avoid unity hierarchy complexity in content production, sub-menus are defined by integrating the middle and sub-classification levels. Design various levels of administration that describe objects and other functions in drawings of device sub-parts at the level required by the manual or implement the operation. At this stage, various scripts are prepared in conjunction with databases related to parts and sub-parts. (3) Maintenance defines the scope of development of the AR manual implemented by the operating company and MMES, and various scenarios such as assembly/disassembly, measurement and inspection taking into account other tools, equipment, and special instruments in the degree of exchange repair, and the contents related to training, maintenance support, and safety work.

In this way, storyboard creation determines the quality and system efficiency of the entire content, so it should be written in collaboration with maintenance experts to understand the level that can be implemented with development programs or algorithms in a systematic approach and implementation virtual space, and to have a high training effect. Figure 8 shows basic architecture of storyboard of air compressor.

Table 3 shows that classifies detailed tasks that are frequently repeated in railway vehicle maintenance work by category, and classifies training purposes by type and methods implemented as



AR contents. A storyboard should be created by focusing on such work content and reviewing the maintenance manual database and maintenance regulations.

Calculation (2D-3D LMDC) algorithm that calculates basic coordinates and travel distances to implement electricity or fluid flow at the same time. The original purpose of this algorithm was developed as a streamline matching variable calibration algorithm to make it easy to understand education and training of complex devices including circuit diagrams or pipes by displaying many 2D drawings widely used in railway vehicles in 3D space from real devices.

Table 3: Maintenance work definition for AR contents

Work	Detail task	Training target	Implementation
Assembly disassembly	Bolt, nut fastening loosening remove, insert pins seal Piping, wiring disconnecting	How to use tools, wrench and jig torque measurement assembling order procedure Connector fastener, How to use special tools, hazard understanding	Animation, scripts, button, movable window, Interaction algorithm requires Animation, Scripts, movable window
Cleaning washing Painting, oiling	Air filter, oil filter, case cleaning, painting, lubricating	How to clean the polluted filter or contaminated parts, How to select proper lubricant and control quantity etc. safety from electric shock or injury	Select the appropriate range of contents based on manuals due to the diversity of work
Repair exchange	PCB, switch, breaker, relay, brake shoes, consumable part	Remove dust, how to repair relay or switch contact how to exchange consumable parts	Select the appropriate range of contents based on manuals due to the diversity of work
Inspection Measurement	Gauges, display panel, warning lamp, gap or clearance Insulation resistance	How to read temperature, pressure gauges, how to measure voltage, current and to use tester, clearance measurement	Animation, Interaction algorithm requires
Drawings	Piping, wiring, electric and electronic circuit	Understanding piping, wiring and circuit diagrams	Interaction algorithm requires
Trouble shooting	Device function, operation manipulation of switch and reset switch	Understanding function and operation manipulation of the parts, the reason of the failures and faults of the equipment	Animation, scripts, markers, Interaction algorithm requires

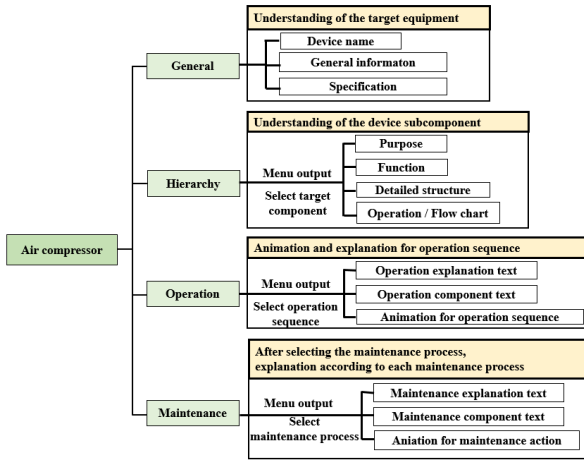
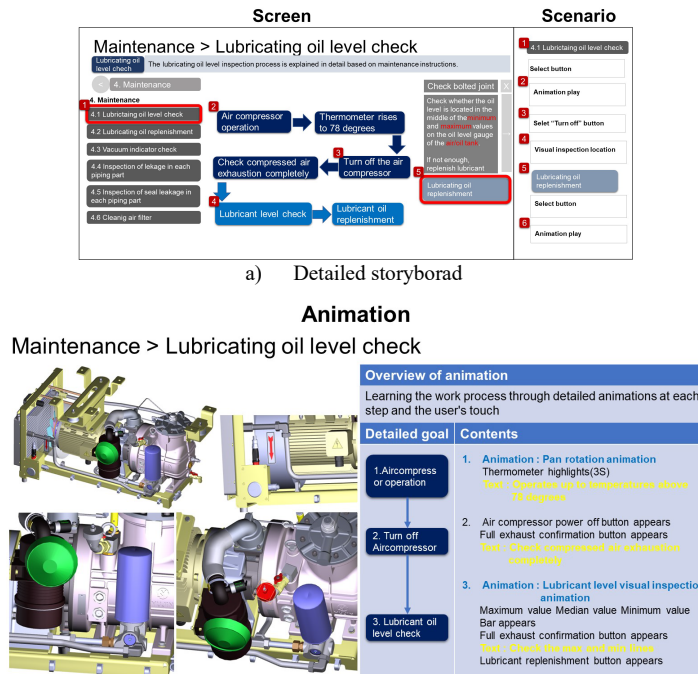


Figure 8: Storyboard creating for AR contents

Figure 9 is an example of a storyboard created during the development process. The left part of Figure 9(a) shows the screen configuration interface and text, and the right part shows the content scenario that the user needs to learn. Figure 9(b) is an example of a visualized storyboard, and the progress of the animation, the setting by time, and the additional functions that appear on the screen are written in detail. The detailed storyboard is continuously updated and managed if a change is needed.



b) Animation storyboard  
Figure 9: Example of storyboard

### 3.3. Algorithm Development for Flow Expression

The main algorithm used in the air compressor training content is as follows. Algorithm 1 is the 2D-3D Line Matching Distance

When used with Algorithm 2, Flow Generation (FG) algorithm, this algorithm can simultaneously display the electrical wiring diagram of the control circuit and the display of electrical flow from the actual wire, or the 2D air flow diagram of the brake system and the flow of compressed air or control air from the actual 3D pipe. An animation duration is obtained by calculating the distance from spatial coordinates with different dimensions and calculating the position that can display the same flow in space at the same time. Figure 10 shows the same flow matched to each other in the 2D drawing and 3D model. Because it passes through the same location value, it is easy for users to learn by comparing each other.

#### Algorithm 1: Code of the 2D-3D Line Matching Distance Calculation algorithm

$$\text{Result: } LD_i = \text{Animation Duration} \cdot \frac{\text{Distance}_i}{\text{TotalDistance}}$$

Animation Duration  $AD$ ;

Vector3 Position List  $P[P_1(x,y,z), P_2 \dots P_x]$ ;

Produce List of Distance  $D$ ;

for ( $i = 0, P - 1, i++$ )

//Calculate Distance and  $D_i(P_i - P_{i+1}) \dots D_i$

$$D_i = \text{Distance}(P_i, P_{i+1}) =$$

$$\sqrt{(P_{i+1x} - P_{ix})^2 + (P_{i+1y} - P_{iy})^2 + (P_{i+1z} - P_{iz})^2}$$

$$LD_i = AD \cdot \frac{D_i}{\text{TotalDistance}}$$

end

#### Algorithm 2: Code of the Flow Generation algorithm

Result: Particle System Quantity-Control

Object Prefab;

float Qcontrol; // Quantity-Control

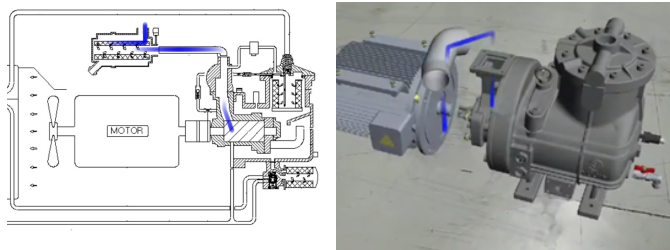
for ( $i = 0, P - 1, i++$ )

Instantiate Object Prefab As  $a$  on  $P_i$

Rotate  $a$  to look  $P_{i+1}$

Set a.Particle System. Lifetime, Speed,  
 Qcontrol, Emission.rate Over Time  
 Lifetime =  $L_i$   
 Speed =  $\frac{D_i}{L_i}$   
 Emission Rate over Time = Start Speed  
**Qcontrol += Qcontrol**

end



a) 2D drawing  
 b) 3D model  
 Figure 10: Flow displayed in 2D drawings and 3D models

The FG algorithm is an algorithm that makes it easy to express the flow of fluid in AR. This algorithm was originally developed to support the LMDC algorithm, allowing electricity or fluid to display the flow on the drawing by continuously flowing objects represented by dots. Since then, in the development of air compressor AR content, it has been expanded to display oil or air exhaust at the valve outlet. This called Prefab of Unity to enable the repetitive use of objects with specific attributes [32]. In order to express the pneumatic flow in 3D at the time of need, it is a method of instantiating the particle prefab with the properties of the particle system input. The module of the particle system was automatically modified once each attribute was entered to avoid the complexity of using a variable containing 24 properties that affect the entire system. When entered into the modified particle prefab, the prefab is instantiated, releasing the particles continuously, which has the effect of displaying the flow. It can also show arbitrary discharged airflow using random numbers and particles emitted by property changes. Figure 11 shows the detail process of the FG algorithm. It is released by each phase that increases the particles, giving a visual effect of creating a moving flow of electricity or compressed air.

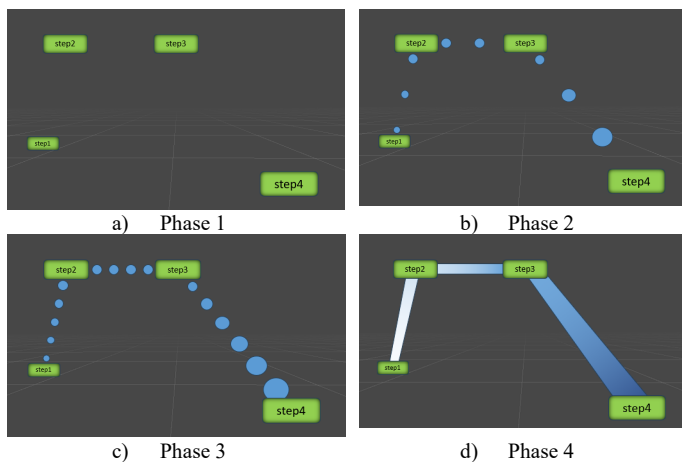


Figure 11: The process of generating a flow as particles increase using the FG algorithm

Figure 12 shows a flow chart of the algorithms that combine two developed algorithms to display electricity or fluid flows simultaneously in a 2D drawing and a real 3D object on one screen, or to control the flow of fluid using a single fluid flow generation algorithm to generate a screen such as air discharge.

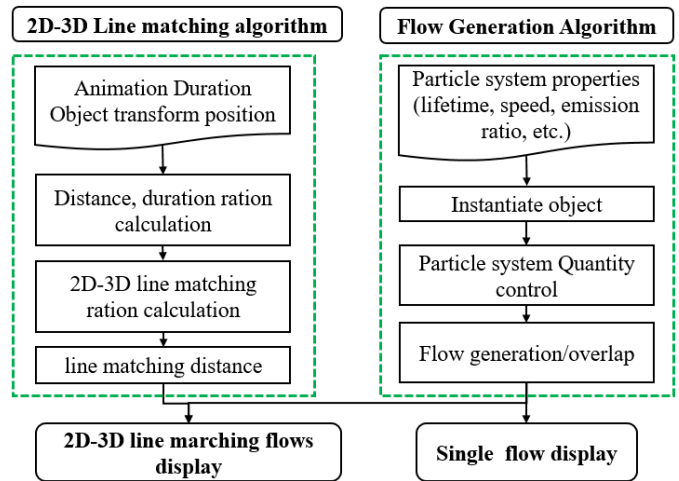


Figure 12: Flow chart of the algorithms for AR contents

Figure 13 shows the exhaust air at the outlet of the air compressor in a three-dimensional space using the FG algorithm alone. In this way, by showing the flow of fluid in various ways through Q-control, various functions of maintenance such as pressure control and exhaust air control can be actively shown to the trainee, thereby enhancing the training effect.

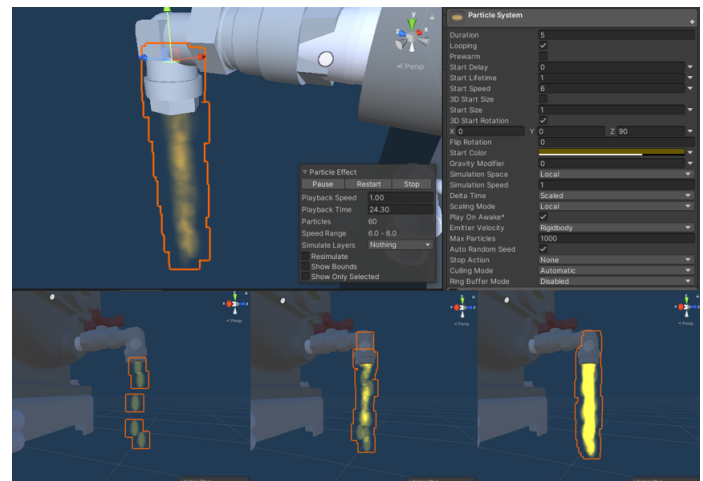


Figure 13: Fluid flow expression with quantity control for AR contents

### 3.4. Development of Menus for User Interaction

When the storyboard is completed, it is necessary to develop user menus to implement AR content on mobile device screens. In particular, menu arrangements that require using a limited narrow space were allocated in consideration of space use so that other menus were not pushed.

Figure 14 shows an example of menu configuration in a User Interface (UI) of air compressor AR content. The menu used for the content is located on the left and right sides of the screen so



that the device is displayed in the center. The menu consists of three types.



Figure 14: Menu design for sub-parts of air compressor

Figure 15 shows a detailed example of the three types of menus. The Global Navigation Bar (GNB) on the left, as shown in (a), is the main menu. This is a large category menu in GNB format, consisting of General, Hierarchy, Operation, and Maintenance, which are the highest levels of keywords for the maintenance process. When one of the GNB menus is touched, the Local Navigation Bar (LNB), a sub-menu of the GNB, appears, as shown in (b). Each LNB menu consists of maintenance process and necessary related information. The LNB is placed on the left side of the screen, and at the same time, a Side Navigation Bar (SNB) as shown in (c) is displayed on the right side. SNB is displayed with detailed information such as text and photos in a separate pop-up window [33]. In addition to this, there is "Hide menu" and "Open Scheme".

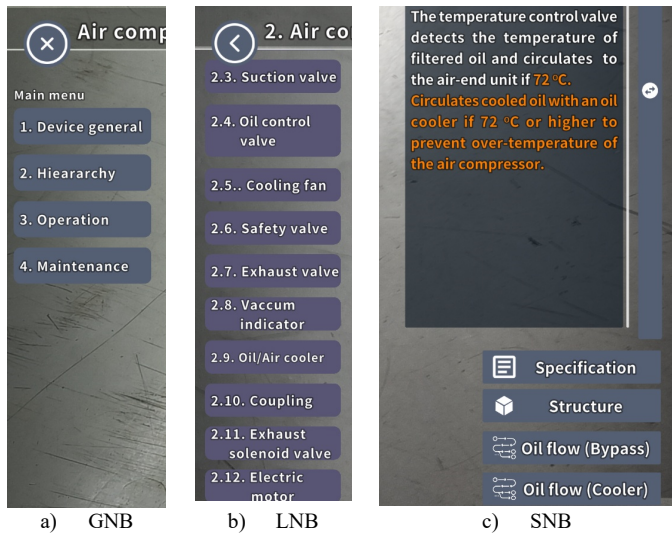


Figure 15: Detailed examples of menu types

In the early development version, GNB and LNB were placed on the left and right sides, but in the upgraded version, when the GNB is touched, the GNB menu disappears and the screen space can be used widely with a screen configuration replaced by LNB.

Figure 16 shows the enlarged screen space by moving the original right LNB to the left. Therefore, even a mobile phone with a small screen can use the screen widely for device screen.

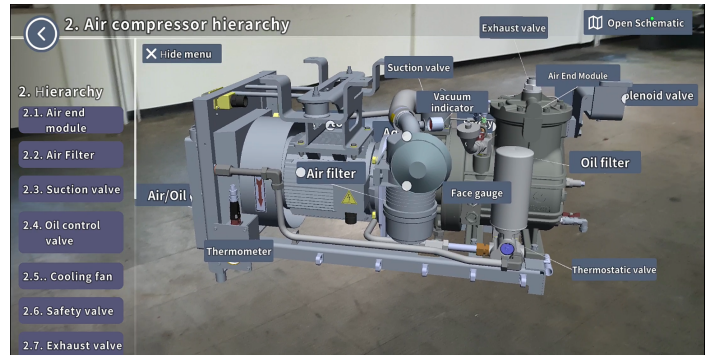


Figure 16: Enlarged screen space by moving the original right LNB to the left

### 3.5. Subcomponent Display Techniques

The developed model is assembled into sub-components, which are based on the maintenance manual and the replacement part procurement system using hierarchy and can identify individual parts from the assembly. It allows for a perceptive effect, and when a particular part is selected in the assembly, the assembly status can be easily identified by using a different color or by inserting an outline at the boundary between the equipment and sub-components. When the structure menu of SNB is touched at the component level, the sub-part is unfolded, and the name of the component is displayed with animation and buttons to make it easier for the user to understand the complex component, and if necessary, a separate SNB can be used to show the procurement product number of the component. This process is already detailed in the hierarchical design stage, but missing and insufficient functions or animations can be inserted additionally during use through content editing to increase completion. Therefore, through this content design, users can check the connection of device components and open the detailed structure between components in the center of the screen in detail, making it convenient to understand the entire device without disassembling the actual components one by one as shown in Figure 17.

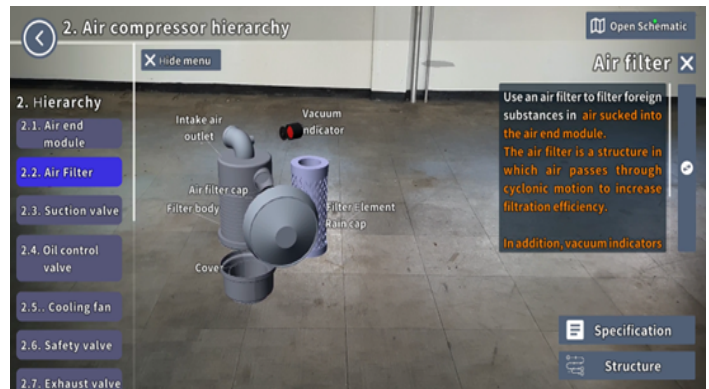


Figure 17: Subcomponent display screen of air filter assembly

The SNB-type menu on the right side of the window is mainly installed at the bottom of the movable window that can display detailed descriptions and other information about components. The SNB menu is used to show sub-parts in a disassembled structure, deliver operational information related to sub-parts, or animate compressed air, electricity, oil flow, etc. to help trainers understand and improve the convenience of maintenance training.

### 3.6. Maintenance Instruction Implementation for User Training

In the GNB menu, the maintenance menu mainly implements the maintenance items, contents, duration, procedures, inspection, repair and replacement methods of equipment by maintenance level required by the manufacturer in the manual approved by the vehicle operator, and the pre-learning support function for the replacement of spare parts and consumable parts. Sub-component contents such as oil amount, supplemental level, and exchange cycle related to the oil filter may be implemented as text in a pop-up window. In this way, the important information of the component is displayed visually to the user.

Figure 18 is a case screen showing lubricating oil filters and lubricating oil replenishment work among maintenance items. There is a maintenance support window built with a movable window that shows the process of replacing the lubricant filter. This content can train users by clearly presenting replacement of consumables, replacement of lubricating oil, supplementation procedures or criteria.

Maintenance AR content includes the ability to effectively pre-learn the replacement process of parts that have reached the replacement cycle. In the figure, the scene where the valve is opened and the lubricant is replenished from the lubricant plastic container to the refueling hole is realized as if it were real by algorithms that show fluid flow. The animation is also implemented with detailed procedures in the pop-up window, allowing users to effectively understand lubricating oil replenishment and learn how to do it.



Figure 18: Maintenance instruction for lubricating oil replenishment

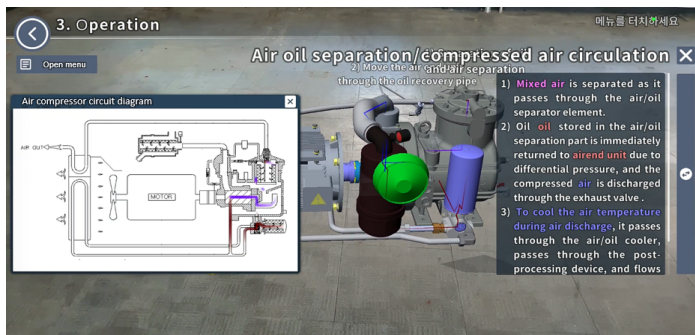


Figure 19: Implement fluid flow simultaneously in 2D pipe drawings and 3D model.

Figure 19 shows the function of implementing fluid flow simultaneously in a 2D pipe drawing and a 3D model. The algorithms described in the previous section are used to provide

effective theory education as well as to allow users to easily understand the operating principles of complex devices.

### 3.7. Safety Warning Function for Hazardous Maintenance Work

The maintenance work of railway vehicles includes various danger and hazard factors, which can lead to serious injury or even death of workers. Therefore, AR content needs to include the ability to proactively train these risks.

Figure 20 shows examples of safety warnings displayed in this content. (a) is a warning message to avoid burn hazard caused by touching hot areas during air compressor maintenance. (b) indicates a warning against body or other contact with cooling fans, coupling rotating parts, etc. (c) shows a safety warning to ensure that the air is exhausted in the pipe before starting work and to prepare for an unexpected situation where compressed air is ejected from the valve. This content reminds workers of things to pay attention to before and after starting work, and also makes workers aware of unexpected situations that may occur.

In addition, various risk factors such as electric shock caused by high-voltage electricity or serious injury caused by vehicle rolling are warned using appropriate warning text and sound effects. In particular, you can use red or yellow visually to highlight the hazard with highlights. Additionally, various pop-up windows can provide detailed explanations of maintenance hazards, such as handling toxic substances and pollutants.

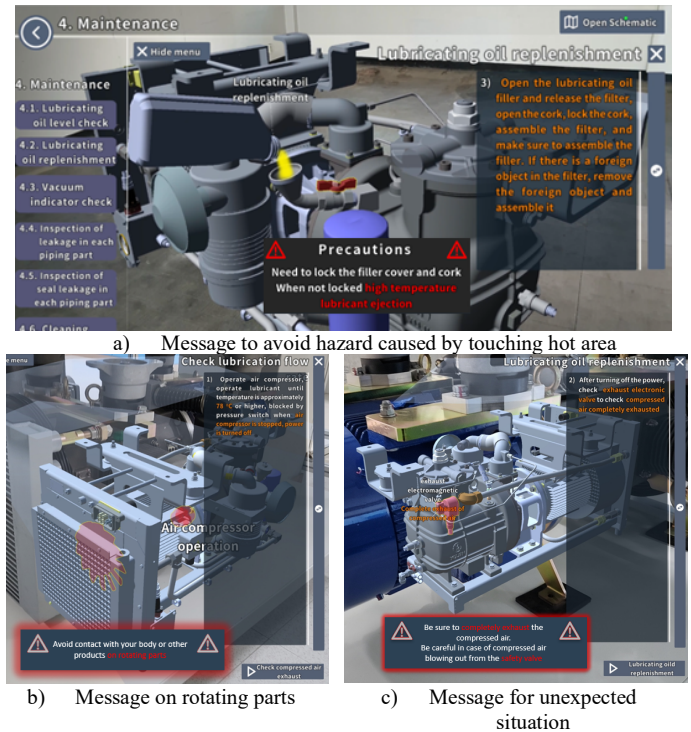


Figure 20: Caution and warning messages for safety

### 3.8. Finger Gestures on Mobile Device Touchscreen

This content can use a manipulation gesture that can control the location, size, and rotation of the 3D model using the mobile device's touch screen method. Table 4 shows the detailed functions of the 3D model manipulation interaction as an example [34]. Maintenance content users can open the app on the mobile device



screen, use the menu to load content on the screen, and use touch-type manipulation technology to perform training and maintenance work [35].

Table 4: Touch gesture for touchscreen manipulation




Value to change	Corresponding	Touch gesture
Position	swipe	
Scale	pinch/spread	
Rotation	twist	

Figure 21 is a photograph showing the inspection and work in front of the air compressor of a railway vehicle using the AR air compressor content of a mobile phone. The produced content is compatible with various screen sizes based on Android first. Using AR content, it overcomes the shortcomings of time, space, and manpower management, which are obstacles to existing document-based manual education, receives database information of a separately developed maintenance server, and is effectively digitized in conjunction with actual digital maintenance work records.



Figure 21: A photo of using mobile phone apps on the side of the vehicle at depot

#### 4. Usability Evaluation

The usability of the developed AR content can be performed using various methods. A commonly used technique is to use the content developed by the user group, then conduct a survey and analyze the results. These techniques include the SUS, the Computer System Usability Survey (CSUQ), and the Usability Indicator for User Experience (UMUX) [36]-[38].

In order to evaluate the usability of the developed AR content, a SUS survey and a survey on UX techniques were conducted on 100 college students who majored in railway vehicle systems, respectively. Table 5 show the participants of the usability evaluation experiments [39].

Table 5: The participants of the experiments

Experiment participants	100 students
Average age	23.8 years

Male	73
Female	27

#### 4.1. SUS

The SUS, as a quick and dirty usability scale is reliable tool for measuring the usability, originally created by John Brook in 1996 [40]. SUS survey consists of a 10 item questionnaire with 5 response options for participants of a scale of 1 to 5. Scale score 1 means strongly disagree and scale score 5 means strongly agree. Table 6 shows the modified SUS survey questions in this study [41]. The survey was carried out after using the AR contents.

Table 6: The modified SUS statements used in this study

No.	Questions
Q1	I think that I would like to use this AR content product frequently
Q2	I found the AR content unnecessarily complex
Q3	I thought the AR content was easy to use
Q4	I think that I would need the support of a technical person to be able to use this AR content
Q5	I found that the various functions in This AR content were well integrated
Q6	I thought that there was too much inconsistency in this AR content
Q7	I would imagine that most people would learn to use this AR content very quickly
Q8	I found the AR content very awkward to use
Q9	I felt very confident using the AR content.
Q10	I needed to learn a lot of things before I could get going with this AR content

The SUS score can be calculated using the following 3 equations [42].

$$Score_{Q1,Q3,Q5,Q7,Q9} = Scale\ Score - 1 \quad (1)$$

$$Score_{Q2,Q4,Q6,Q8,Q10} = 5 - Scale\ Score \quad (2)$$

$$SUS\ Scores = Sum\ of\ Scores \times 2.5 \quad (3)$$

Equation (1) is used to get the odd number question scores. The even number question scores are calculated by (2). Equation (3) is used to obtain the final value of the SUS survey. The scores have a range is from 0 to 100. It can also be mapped to the following adjective ratings. Figure 22 shows a comparison of mean SUS scores by quartile, adjective ratings and the acceptability of the overall SUS score [43].

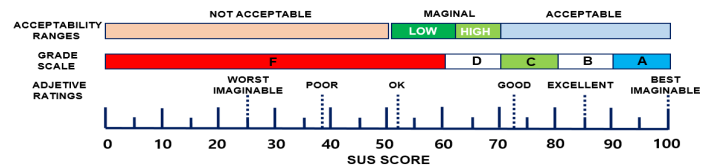


Figure 22: A comparison of mean SUS scores by quartile, adjective ratings, and the acceptability of the overall SUS score.

#### 4.2. User Experience Experiment

Another technique, UX assessment, was conducted on AR contents. Six items related to usability were selected that were

reviewed as suitable for content evaluation for maintenance of specific equipment, such as air compressors. The survey tool referred to the UX evaluation, and the detailed survey contents of the survey are shown in Table 7 [44]. An experience survey using a questionnaire tool developed after the user used AR content for air compressors was conducted and analyzed.

Table 7: Indicators of Usability assessment

Categorized Items	Questionnaire details
Simplicity	Appearance of AR content, function provision method, simplicity of procedures.
Accessibility	User should be able to access AR content smoothly.
Efficiency	It should be possible to derive accurate results without unnecessary time consumption.
Informativeness	User should be able to obtain the necessary information through appropriate means and methods.
Learnability	User should be able to handle the interface of AR content proficiently
User Support	User should be able to respond well to problems that arise when using AR content.

4.3. SUS Survey Results

Figure 23 is a graph showing the score for each SUS question of the AR contents of the air compressor. The threshold reference value for the usefulness of SUS is 68 points. The results of this survey are as follows. The items that were relatively better than other questions were Q3 "ease of using content" (87.2) and Q1 "intention to use content" (87). On the other hand, Q4 "needs for technical helpers" (68.3) and Q10 "needs for pre-learning" (69) questions scored relatively lower than other questions. This seems to be because survey participants are not familiar with the use of AR content in the screen touch manipulation method, as shown in Table 3. The average score of 10 SUS questions for this content was 76.65, which was higher than the overall threshold of 68 points for each question. In addition, the usefulness rating is "B(Good)".

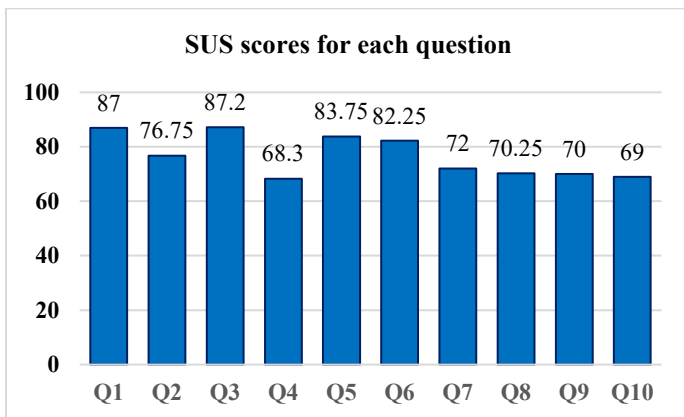


Figure 23: SUS score of the air compressor AR contents

Therefore, this content will be useful for beginners in railway vehicle maintenance. Therefore, it is considered suitable for use in the maintenance and repair training of air compressors for urban railway vehicles.

4.4. UX Survey Results

Figure 24 is a graph showing the UX evaluation results. This is the result of a survey after using AR contents developed four times. The overall average of the six questions was 4.12 points, and the score for the sense of reality item was the highest at 4.38 points. On the other hand, the score of the convenience item was 3.92, the lowest among the four items.

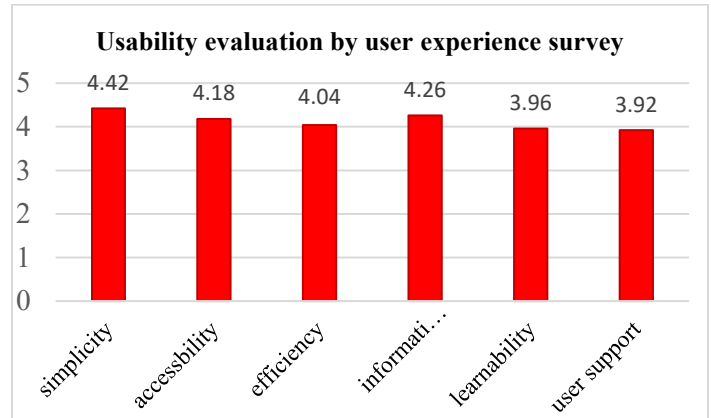


Figure 24: UX score of the air compressor AR contents

Table 8 shows the results of sub-items for each item of UX evaluation. The reason for the highest simplicity is that the maintenance procedure is simply and clearly structured to make it easier for even beginners to understand (4.58). It is presumed that various learning functions are also clearly organized based on keywords (4.41). The second highest informativeness item score is considered to be because the entire scenario of the content was systematically structured (4.38), and the text was written to make the maintenance procedure easier to understand (4.22). In comparison, the reason why the user support item is the lowest was that the description of the possible technical problems was insufficient (3.85), and the response method in case of an error was not described in detail (0.395).

Table 8: Scores of detail Questionnaire for UX

Categorized Items	Questionnaire details	score
Simplicity	Smooth output of content	4.27
	Simple configuration of various learning functions	4.41
	Simple and clear explanation of maintenance processes	4.58
Accessibility	Ease of running content	4.09
	Compatibility with mobile devices	4.20
	Clarity of description and usage	4.31
Efficiency	Comparison of efficiency with existing manuals	3.99
	Quickly acquire the information you want	4.13



Categorized Items	Questionnaire details	score
	Increased actual work efficiency	4
Informativeness	Provides accurate maintenance information	4.18
	Understand the entire maintenance process	4.22
	Systematicity of the entire scenario	4.38
Learnability	Seamless model touch availability	3.96
	Compatibility of UI and additional buttons	3.97
	Ease of screen transition	3.95
User Support	Description of response when error occurs	3.95
	Adequacy of help and error examples	3.96
	Examples of problems that may occur	3.85

#### 4.5. Discussion

Two evaluation techniques, SUS and UX, were used to identify problems and derive improvements of developed AR content. As a result of the two experiments, a common issue is that participants have some difficulty using the content and take time to learn how to use it. They experienced some errors in the content during the experiment and expressed lack of help or countermeasures. The content was produced in consideration of the use of novice maintenance workers in the design stage, so this part was highly evaluated, but it is also considered that the use of maintenance workers who are relatively unfamiliar with digital content should be considered. In future contents, it will be necessary to produce it by receiving feedback from maintenance workers who have not been exposed to AR technology in the design stage. If additional functions such as inserting help for each function, producing tutorial version, and voice guidance are produced, high-quality content can be produced.

#### 5. Conclusion

This study is a case of applying AR technology to the maintenance of railway vehicles and employee education and training as one of the smart maintenance research projects. AR contents that can be easily used in mobile devices for maintenance and training of air compressors for commuter electric vehicles have been developed. Two algorithms were developed and used to express fluid flows such as pressure air. In addition, in order to verify the usability of the developed content using a SUS survey and UX survey. The conclusions obtained from this study are as follows.

(1) By reducing the file size from 3D scanners or CAD original files, it is possible to implement fast animation on Android-based mobile devices. It was developed by linking training AR contents with the AR manual of the MMES, which is a high-level platform. Therefore, maintenance work and training are possible while watching actual equipment in the field depot, and maintenance support can be provided along with other equipment contents already developed or under development.

(2) The 2D-3D LMDC algorithm, which can implement flow simultaneously on a two-dimensional drawing or piping diagram and a three-dimensional equipment screen, was applied. In addition, the FG algorithm, which can control the flow rate of the fluid flow, was used to implement the movement or exhaust of compressed air, thereby enhancing the learning effect of the content.

(3) A simplified menu is placed on the left and right sides of the display screen, and the device is operated by user-friendly menu bar and touch screen operation technology on the central screen. In addition, through a systematic hierarchy structure when developing a storyboard, users can easily understand the operating principles and operations of the entire equipment, and perform reliable work by receiving detailed support for disassembly procedures and specific work methods

(4) The developed content can display components and sub-components according to the standardized BOM system of manufacturers and operators. This enables support for maintenance work using part hierarchies. In addition, through information on the replacement cycle and replacement process of spare parts based on the manufacturer's manual, it is possible to effectively train the replacement and process of spare parts and perform necessary inspections or repairs.

(5) Due to the nature of railway vehicle parts, there is a lot of heavy equipment, and there are maintenance tasks that involve workers' risks such as high voltage and high-pressure air. In order to eliminate safety accidents during work, safety items were visualized on a text display to generate a warning message, and a red highlight was applied to each inspection part for worker safety.

(6) The average SUS score for AR content of this air compressor content is 76.65, and the rating is "Good". In addition, the results of the UX test showed that the average usability score for the six-item questionnaire was excellent at 4.12 out of 5 score, and this content is estimated to be quite useful for training beginners in railway vehicle maintenance.

Future research will need to expand content that has not yet been developed in the vehicle system, further develop algorithms for efficient content production, and continue to improve usability in conjunction with the AR manual of the maintenance efficiency system.

#### Conflict of Interest

The authors declare no conflict of interest.

#### Acknowledgment

This work was supported by the Korea Agency for Infrastructure Technology Advancement (KAIA) by the Ministry of Land, Infrastructure and Transport under Grant RS-2021-KA162811.

#### References

- [1] F.A. Marten, "Reliability Centered Maintenance: A Case Study of Railway Transit Maintenance to Achieve Optimal Performance," Mineta Transportation Institute Report, 1006, 2010.
- [2] E. Quatrini, F. Costantino, G.D. Gravio, R. Patriarca, "Condition-Based Maintenance An Extensive Literature Review," *Machines*, 8(2), 31, 2020, doi:10.3390/machines8020031.
- [3] A.K.S. Jardine, D. Lin, D. Banjevic, "A review on machinery diagnostics and prognostics implementing condition-based maintenance," *Mechanical*

- systems and signal processing, **20**(7), 1483–1510, 2006, doi:10.1016/j.ymssp.2005.09.012.
- [4] J. Campos, "Development in the application of ICT in condition monitoring and maintenance," *Computers in Industry*, **60**(1), 1–20, 2009, doi:10.1016/j.compind.2008.09.007.
- [5] B. Zhang, Y. Sui, Q. Bu, X. He, "Remaining useful life estimation for micro switches of railway vehicles," *Control Engineering Practice*, **84**, 82–91, 2019, doi:10.1016/j.conengprac.2018.10.010.
- [6] C.P. Ward, P.F. Weston, E.J.C. Stewart, H. Li, R.M. Goodall, C. Roberts, T.X. Mei, G. Charles, R. Dixon, "Condition Monitoring Opportunities Using Vehicle-Based Sensors," *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*, **225**(2), 202–218, 2011, doi:10.1177/09544097jrrt406.
- [7] Z. Cujan, G. Fedorko, N. Mikusova, "Application of virtual and augmented reality in automotive," *Open Engineering*, **10**(1), 113–119, 2020, doi:10.1515/eng-2020-0022.
- [8] T. Masood, J. Egger, "Augmented reality in support of Industry 4.0–Implementation challenges and success factors," *Robotics and Computer-Integrated Manufacturing*, **58**, 181–195, 2019, doi:10.1016/j.rcim.2019.02.003.
- [9] L.F.D.S. Cardoso, F.C.M.Q. Mariano, E.R. Zorzal, "A Survey of Industrial Augmented Reality," *Computers & Industrial Engineering*, **139**, 106159, 2020, doi:10.1016/j.cie.2019.106159.
- [10] R.T. Azuma, "A Survey of Augmented Reality," *PRESENCE: Virtual and Augmented Reality Issues*, **6**(4), 355–385, 1997, doi:10.1162/pres.1997.6.4.355.
- [11] R. Palmirani, J.A. Erkoyuncu, R. Roy, H. Torabmostaedi, "A systematic review of augmented reality applications in maintenance," *Robotics and Computer-Integrated Manufacturing*, **49**, 215–228, 2018, doi:10.1016/j.rcim.2017.06.002.
- [12] R.G. Boboc, F. Girbacia, E.V. Butila, "The Application of Augmented Reality in the Automotive Industry: A Systematic Literature Review," *Applied Sciences*, **10**(12), 4259, 2020, doi:10.3390/app10124259.
- [13] H. Rios, E. Gonzalez, C. Rodriguez, H.R. Siller, M. Contero, "A Mobile Solution to Enhance Training and Execution of Troubleshooting Techniques of the Engine Air Bleed System on Boeing 737," *Procedia Computer Science*, **25**, 161–170, 2013, doi:10.1016/j.procs.2013.11.020.
- [14] H. Eschen, T. Kotter, R. Rodeck, M. Harnisch, T. Schuppstuhl, "Augmented and Virtual Reality for Inspection and Maintenance Processes in the Aviation Industry," *Procedia Manufacturing*, **19**, 156–163, 2018, doi:10.1016/j.profmfg.2018.01.022.
- [15] M. Kans, D. Galar, A. Thaduri, "Maintenance 4.0 in railway transportation industry," in *Proceedings of the 10th world congress on engineering asset management (WCEAM 2015)*, 317–331, 2016, doi:10.1007/978-3-319-27064-7\_30.
- [16] A. Muller, A.C. Marquez, B. Iung, "On the concept of e-maintenance: Review and current research," *Reliability Engineering & System Safety*, **93**(8), 1165–1187, 2008, doi:10.1016/j.res.2007.08.006.
- [17] P. Cipresso, I.A.C. Giglioli, M.A. Raya, G. Riva, "The Past, Present, and Future of Virtual and Augmented Reality Research: A Network and Cluster Analysis of the Literature," *Frontiers in Psychology*, **9**, 2086, 2018, doi:10.3389/fpsyg.2018.02086.
- [18] W. Zhang, B. Han, P. Hui, "On the networking challenges of mobile augmented reality," in *Proceedings of the Workshop on Virtual Reality and Augmented Reality Network*, 24–29, 2017, doi:10.1145/3097895.3097900.
- [19] J. Kiu, K. Zhang, "Design and Simulation Debugging of Automobile Connecting Rod Production Line Based on the Digital Twin," *Applied Science*, **13**(8), 4919, 2023, doi: 10.3390/app13084919.
- [20] V. P. Gangwar, D. Reddy, "Hospitality Industry 5.0: Emerging Trends in Guest Perception and Experiences," *Opportunities and Challenges of Business 5.0 in Emerging Markets*, 185–211, 2023, doi: 10.4018/978-1-6684-6403-8.ch010.
- [21] E.D.N. Orihuela, O.O.V. Villegas, V.G.C. Sanchez, R.I.B. Castillo, J.G.L. Solorzano, "Mobile Augmented Reality Prototype for the Manufacturing of an All-Terrain Vehicle," *Advanced Topics on Computer Vision, Control and Robotics in Mechatronics*, 49–75, 2018, doi:10.1007/978-3-319-77770-2\_3.
- [22] A. Chouchene, A.V. Carvalho, F.C. Santos, W. Barhoumi, "Augmented Reality-Based Framework Supporting Visual Inspection for Automotive Industry," *Applied System Innovation*, **5**(3), 48, 2022, doi:10.3390/asi5030048.
- [23] H.J. Kwon, K.S. Kim, C.S. Kim, "Development and Evaluation of Augmented Reality Learning Content for Pneumatic Flow: Case Study on Brake Operating Unit of Railway Vehicle," *IEEE Access*, **11**, 46173–46184, 2023, doi:10.1109/ACCESS.2023.3273605.
- [24] H.J. Kwon, S.I. Lee, J.H. Park, C.S. Kim, "Design of Augmented Reality Training Content for Railway Vehicle Maintenance Focusing on the Axle-Mounted Disc Brake System," *Applied Science*, **11**(19), 9090, 2021, doi:10.3390/app11199090.
- [25] S. Scheffer, A. Martinetti, R. Damgrave, S. Thiede, L.V. Dongen, "How to Make Augmented Reality a Tool for Railway Maintenance Operations: Operator 4.0 Perspective" *Applied Sciences*, **11**(6), 2656, 2021, doi:10.3390/app11062656.
- [26] N. Hagbi, O. Bergig, J.E. Sana, M. Billibghurst, "Shape Recognition and Pose Estimation for Mobile Augmented Reality," *IEEE transactions on visualization and computer graphics*, **17**(10), 1369–1379, 2010, doi:10.1109/TVCG.2010.241.
- [27] D.P. Cassent, "Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling," *Intelligent Robots and Computer Vision XV: Algorithms, Techniques, Active Vision, and Materials Handling*, 2904, 1996.
- [28] N. Wagner, G. Reitmayr, A. Mulloni, T. Drummond, D. Schmalstieg, "Real-time detection and tracking for augmented reality on mobile phones," *IEEE transactions on visualization and computer graphics*, **16**(3), 355–368, 2009, doi:10.1109/TVCG.2009.99.
- [29] M.S. Elbelkasi, E. Badaran, M.A. Rahman, "Overview of DC and AC Electric Railway Systems Considering Energy Efficiency Enhancement Methods," *Port-Said Engineering Research Journal*, **24**(1), 128–140, 2020, doi: 10.21608/pserj.2019.18206.1012.
- [30] M. Saeed, F. Briz, J.M. Guerrero, I. Larrazabal, D. Ortega, V. Lopez, J.J. Valera "Onboard energy storage systems for railway: present and trends," *IEEE Open Journal of Industry Applications*, **4**, 238–259, 2023, doi: 10.1109/OJIA.2023.3293059.
- [31] 3D Systems, *Geomagic Design X 2020 manual*, 2020.
- [32] Unity Asset Store, *Easy Performant Outline 2.0*, 2021.
- [33] H.J. Kwon, J.H. Cha, Y.K. Park, C.S. Kim, "A study on the realistic content for air compressor maintenance of the urban railway vehicle," *Journal of Korean Society for Urban Railway*, **10**(1), 1163–1175, 2022.
- [34] Unity Asset Store, *Lean Touch+*, 2021.
- [35] M. Kim, J.Y. Lee, "Touch and hand gesture-based interactions for directly manipulating 3D virtual objects in mobile augmented reality," *Multimedia Tools and Applications*, **75**, 16529–16550, 2016, doi:10.1007/s11042-016-3355-9.
- [36] J.R. Lewis, "IBM computer usability satisfaction questionnaires: Psychometric evaluation and instructions for use," *International Journal of Human-Computer Interaction*, **7**, 57–78, 1995, doi:10.1080/10447319509526110.
- [37] J.R. Lewis, "Measuring Perceived Usability: The CSUQ, SUS, and UMUX" *International Journal of Human-Computer Interaction*, **34**(12), 1148–1156, 2018, doi:10.1080/10447318.2017.1418805.
- [38] K. Finstad, "The usability metric for user experience," *Interacting with Computers*, **22**(5), 323–327, 2010, doi:10.1016/j.intcom.2010.04.004.
- [39] K. Hornbæk, M. Hertzum, "Technology Acceptance and User Experience," *ACM Transactions on Computer-Human Interaction (TOCHI)*, **24**(5), 1–30, 2017, doi:10.1145/3127358.
- [40] J. Brooke, "Sus: a "quick and dirty" usability," *Usability evaluation in industry*, **189**(3), 189–194, 1996, doi:10.1016/j.res.2007.08.006.
- [41] A. Bangor, P.T. Kortum, J.T. Miller, "An Empirical Evaluation of the System Usability Scale," *International Journal of Human-Computer Interaction*, **24**(6), 574–594, 2008, doi:10.1080/10447310802205776.
- [42] A.C. Wijaya, M.W.A. Munandar, F. Utaminigrum, "Usability Testing of Augmented Reality For Food Advertisement Based On Mobile Phone Using System Usability Scale," in *2019 International Conference on Sustainable Information Engineering and Technology (SIET)*, 266–269, 2019, doi:10.1109/SIET48054.2019.8986118.
- [43] A. Bangor, P.T. Kortum, J.T. Miller, "Determining what individual SUS scores mean: Adding an adjective rating scale," *Journal of usability studies*, **4**(3), 114–123, 2009, doi:10.1080/10447310802205776.
- [44] J. Shin, S.H. Lee, J.E. Yoo, C. Roh, J.A. Oh, S. Y. Kim, "Development and Evaluation of a Virtual Maintenance Training System for Thermal Power Plant Boiler," *Journal of Digital Contents Society*, **22**(5), 791–800, 2021.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Bridging Culture and Care: A Mobile App for Diabetes Self-Care Honoring Native American Cultural Practices

Wordh Ul Hasan<sup>1</sup>, Kimia Tuz Zaman<sup>1</sup>, Shadi Alian<sup>1</sup>, Tianyi Liang<sup>1</sup>, Vikram Pandey<sup>1</sup>, Jun Kong<sup>1</sup>, Cui Tao<sup>2</sup>, Juan Li<sup>1,\*</sup>

<sup>1</sup>Department of Computer Science, North Dakota State University, Fargo, 58102, USA

<sup>2</sup>School of Biomedical Informatics, University of Texas Health Science Center, Houston, 77030, USA

### ARTICLE INFO

Article history:

Received: 14 November, 2023

Revised: 06 January, 2024

Accepted: 06 January, 2024

Online: 26 January, 2024

Keywords:

Diabetes

Self-management

Native American

Cultural Sensitivity

Participatory Design

Mobile health

### ABSTRACT

Diabetes presents a significant public health issue for Native Americans, exacerbated by cultural nuances often ignored by conventional healthcare. To address this, we introduce a mobile app designed with the cultural context of Native American populations in mind. The app's development followed participatory design principles, with direct input from Native American stakeholders through focus groups and interviews. The app features culturally tailored nutrition plans incorporating traditional foods, community-based support systems, and engagement with tribal health resources. An interface highlighting Native American heritage, along with gamification using cultural storytelling, aims to enhance user engagement and educational content is provided within a culturally relevant framework. This innovative integration of technology and cultural heritage in health management is anticipated to improve engagement, self-efficacy, and health outcomes for Native Americans with diabetes, serving as a blueprint for culturally sensitive health interventions.

### 1. Introduction

The prevalence of diabetes among Native American (NA) populations is not only a pressing public health issue but a stark representation of existing health disparities in the United States. Data from the American Diabetes Association reveal that Native Americans endure the highest age-adjusted prevalence rates of diabetes across all racial and ethnic groups nationwide, with figures nearly twice as high as those observed in non-Hispanic whites [1]. This staggering statistic underscores the urgency for specialized healthcare approaches tailored to the distinct needs of NA communities.

One of the most pronounced obstacles in diabetes management within these populations is the limited accessibility to healthcare [2]. Geographic isolation compounds this challenge, with many NA communities situated in remote areas where healthcare facilities are sparse, under-resourced, and often miles away from those in need [3]. This situation is further complicated by socio-economic barriers such as poverty, which is prevalent in many NA communities, and lower rates of health literacy that impede the uptake and implementation of diabetes management practices [4].

Moreover, the cultural dimensions of NA life — encompassing food, lifestyle, and health beliefs — are deeply ingrained and central to their identity. However, these cultural aspects are

frequently overlooked in mainstream diabetes care strategies. The resultant disconnect renders many existing diabetes management applications ineffective for NA patients, as they fail to align with the patients' cultural contexts and lifestyle preferences [5].

Language and educational disparities between healthcare providers and NA patients can also lead to misunderstandings and misapplications of diabetes management plans, further inhibiting effective care. It is therefore imperative to embrace a diabetes management strategy that not only recognizes but integrates the cultural, socio-economic, and geographic particularities of NA patients.

Our study advocates for a diabetes self-management system that is culturally congruent with the lives and values of NA patients. This system is designed in close collaboration with the NA community, particularly engaging with members from the Lower Sioux Indian Community in Morton, MN, to gather requirements, iterate design solutions, and garner feedback on the prototype. Our methodology is inclusive, drawing upon the unique cultural traits of the NA population in our ontological framework, user interface, nutrition guidance, social networking features, and compilation of local diabetes care resources. This paper delineates the process of developing a mobile application that is not merely a tool but a culturally resonant companion for NA patients, offering tailored nutritional and lifestyle advice, fostering a strong sense of community, and enhancing the accessibility of local diabetes management resources.

\*Corresponding Author: Dr. Juan Li, Department of Computer Science, North Dakota State University Fargo 58102, USA, [j.li@ndsu.edu](mailto:j.li@ndsu.edu)

This work was supported by the National Science Foundation (NSF) with award numbers: 1722913 and 2218046.

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj090110>



By situating the lived experiences and cultural fabric of NA individuals at the forefront, we endeavor to contribute a meaningful and functional solution to the challenges posed by diabetes management in NA communities, thereby aiming to bridge the gap between healthcare availability and cultural appropriateness.

## **2. Related Work**

The concept of cultural appropriateness in healthcare interventions is not new; it has been thoroughly investigated by numerous scholars who advocate for the integration of cultural elements to foster positive and enduring health outcomes. The consensus is clear: to be effective, health interventions must resonate culturally with the target population. Airhihenbuwa [6] posits that cultural considerations should supersede behavioral ones, suggesting that aligning health interventions with cultural norms and values is paramount. Kreuter [7] echoes this stance, highlighting the enhanced efficacy of health programs that are culturally tailored to meet the specific needs of a given group. Further expanding on this idea, Barrera [8] devised a strategic framework for the development of culturally enriched interventions, demonstrating their superiority in bolstering health outcomes in contrast with conventional approaches.

The tripartite model of culture proposed by Asad [9], which classifies culture into knowledge, practice, and change, underscores the multifaceted influence of culture on health intervention success. This model suggests that understanding cultural knowledge, respecting cultural practices, and being open to cultural change are essential for the effectiveness of health interventions. The synthesis of these scholarly insights supports the premise that embedding cultural traditions and practices within health interventions is a pivotal strategy in diminishing health disparities and enhancing the acceptability and impact of such interventions.

Despite the known benefits of cultural tailoring, there exists a conspicuous deficiency in the availability and quality of diabetes management apps that cater to culturally diverse groups. Rodríguez [10] uncovered a scarcity of diabetes management applications available in local languages and noted a lack of credible certification or scientific backing within those that are available. Brzan [11] conducted an extensive analysis of 65 mobile self-management apps for diabetes and concluded that a mere nine met the criteria to be considered comprehensive and effective for diabetes management. This finding accentuates the need for a more rigorous and patient-centric approach to app development.

Highlighting the necessity for oversight, Sneha [12] advocates for regulation, standardization, and quality control within the burgeoning market of diabetes management applications. The current landscape is marked by an influx of apps with varying degrees of reliability and efficacy, necessitating a structured framework to assure their quality. Conversely, the research conducted by Kebede [13] offers a counterpoint, illustrating the positive impact of popular diabetes management applications like mySugr, as well as continuous glucose monitoring systems, on the self-care behaviors of individuals with type 1 and type 2 diabetes.

The intersection of culture and healthcare technology is a critical focal point in enhancing the efficacy of diabetes management applications. The influence of cultural and traditional

considerations on the success of such interventions is increasingly recognized in the literature. Petersen[14] underscores the profound effect of cultural practices on diabetes self-management, suggesting that these cultural influences often have a greater impact on patient behavior than the technological features of mobile applications themselves. This insight is pivotal, as it directs attention to the nuanced ways in which cultural contexts shape health-related behaviors and choices.

Echoing this sentiment, Millan-Ferro [15] advocates for the development of interventions that are culturally resonant, particularly for marginalized groups such as the Latino community in the United States. The call is for culturally sensitive tools that are not merely translations of existing resources but are built from the ground up with cultural relevance in mind. This approach ensures that interventions are more than just medically accurate—they are culturally meaningful and thus more likely to be embraced by the communities they intend to serve.

The work of Chacko [16] sheds light on the intricate ways individuals with Type 2 diabetes navigate their healthcare options, often blending modern biomedicine with traditional practices. In Kerala, India, for instance, patients incorporate Ayurvedic medicine and folk herbal remedies alongside conventional treatments, reflecting a holistic approach to health that honors cultural heritage. This amalgamation of healthcare practices highlights the necessity of acknowledging and integrating traditional health paradigms into modern healthcare solutions.

Furthermore, Goody [17] articulates the critical role that cultural food practices play in diabetes care and education. The application of frameworks such as the Campinha-Bacote Model of cultural competency is instrumental in enabling healthcare professionals to understand and respect the intricate relationship between cultural dietary habits and diabetes management. By recognizing and incorporating these cultural food practices into diabetes care, healthcare providers can develop more effective education and intervention strategies that align with patients' lifestyles and beliefs.

These studies collectively underscore the potential for culturally tailored diabetes management applications to not only enhance their practical utility but also to resonate deeply with the daily lives and values of diverse patient populations. It is evident that there is a pressing need for such culturally informed tools, yet the current landscape reveals a scarcity of interventions that successfully integrate these critical cultural and traditional elements.

Considering this, our research is poised to fill a significant void by designing a healthcare intervention that not only acknowledges but prioritizes the cultural and traditional aspects of the Native American community. Our aim is to develop a diabetes management app that is not just functionally robust but also culturally attuned, ensuring that the Native American traditions and perspectives are interwoven into every facet of the app's design and functionality.

## **3. Design**

The development of a culturally sensitive diabetes self-management application for Native Americans necessitates a methodological approach that is both systematic and adaptive to the unique needs of the community. This study employed a



community-based participatory research (CBPR) framework, which emphasizes collaboration and partnership with the target community at all stages of the app's development. This approach aligns with the principles of respect, relevance, and reciprocity, ensuring that the research is culturally grounded and community driven.

Our methodology integrates qualitative data collection methods to inform the design and functionality of the mobile application. The process was iterative and responsive, allowing for continuous feedback and refinement. This section details the comprehensive steps taken—from participant recruitment, design considerations to app development. Each phase was conducted with rigorous attention to maintaining cultural integrity and fostering a sense of ownership and validation within the Native American community.

The core objective of our methodological design was to create an application that is not only a technological tool for diabetes management but also a culturally resonant resource that empowers Native American users and acknowledges their unique perspectives on health and well-being.

Our engagement with the Lower Sioux Indian Community indicated a sufficient level of access to smartphones and internet services, making the deployment of a mobile health app a viable option. This was corroborated by community feedback during the initial phase of our study, where participants expressed familiarity with and access to modern mobile technology. Thus, while acknowledging the existing challenges, our study operates on the premise that a significant portion of the target population has the means to access and benefit from the app, making it a relevant tool for diabetes self-care in these communities.

### 3.1. Background on Native American Traditions and Perspectives

In developing our study and the subsequent mobile application for diabetes self-care, we delved deeply into the rich tapestry of Native American traditions and perspectives. The Native American communities, with their diverse and vibrant cultural practices, offer a unique lens through which healthcare solutions can be viewed and tailored.

At the heart of our narrative is an understanding that health and wellness in Native American communities are not merely physical states but are deeply intertwined with spiritual, community, and environmental well-being. These aspects, often overlooked in conventional healthcare approaches, are central to the Native American perspective on health. Food is much more than sustenance in Native American cultures; it carries historical significance and is a medium for community bonding and cultural expression. Traditional foods, preparation methods, and communal eating practices are not only nutritional choices but also a means to preserve heritage and identity. We integrated these aspects to ensure our app respects and supports these dietary traditions. The role of family and community is paramount in Native American societies. Healthcare interventions, to be effective, must recognize and leverage these social structures. Our design process involved consultations with community members to understand how family dynamics and community interactions influence health behaviors and diabetes management. Storytelling, a revered tradition in Native American cultures, serves as an educational and cultural preservation tool. By incorporating

elements of storytelling and native symbolism into our app, we aimed to connect users with their heritage, making the app not just a health tool but also a medium of cultural engagement.

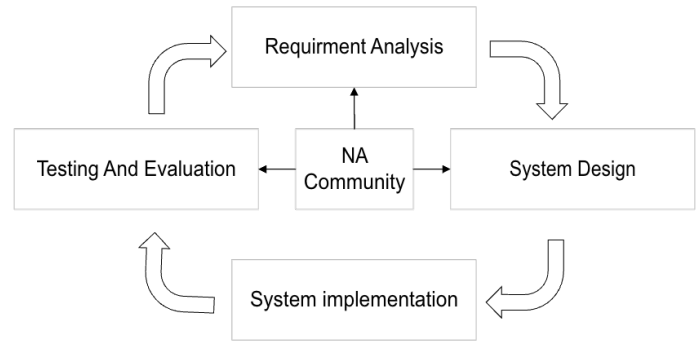


Figure 1: CBPR Process for Developing a Culturally Tailored Diabetes Management App

This narrative underscores our commitment to developing a diabetes self-management tool that is not only functionally effective but also culturally congruent. By incorporating these cultural insights, we aim to bridge the gap between technology and tradition, offering a healthcare solution that resonates deeply with Native American communities.

### 3.2. NA Community Based Participatory Research

Our research is deeply rooted in the principles of Community-Based Participatory Research (CBPR), an approach that ensures the development of our mobile application is closely aligned with the distinct needs and cultural context of the Native American (NA) community. Figure 1 illustrates the cyclical and collaborative nature of the CBPR process, which integrates the Lower Sioux Indian Community throughout the system development cycle.

This study has been granted ethical approval by Lower Sioux Indian Community Tribe's governance body under resolution number 16-153.

#### 3.2.1. Requirement Analysis

In the initial phase of our study, a detailed requirement analysis was conducted, as illustrated by the top box of Figure 1. This phase was twofold: it involved individual interviews with domain experts and focused group sessions with the community.

We engaged in knowledge elicitation sessions with an assembly of domain experts, including tribal leaders from the Native American reservation community in the upper Midwest, local diabetes physicians, nutritionists specializing in diabetes care, and public health faculty. Preliminary meetings were convened to gather coarse-grained, general knowledge and set the foundation for a more nuanced understanding of the domain. The insights from these meetings were meticulously documented, serving as a vital reference for developing a domain-relevant ontology and for informing subsequent, more detailed inquiries.

Parallel to expert interviews, focused group discussions were held with members of the Lower Sioux Indian Community to explore the lived experiences of managing diabetes within their socio-cultural context. These sessions were pivotal in extracting nuanced insights into the community's cultural traditions, resource accessibility, socio-economic factors, and the specific challenges encountered in diabetes self-management. The iterative nature of these discussions is signified by the two-way arrows in Figure 1, reflecting the dynamic exchange between the research team and

the community, ensuring that their voices directly influenced the requirement analysis.

The amalgamation of knowledge from domain experts and the experiential wisdom of the community formed a comprehensive foundation for our requirement analysis. This collaborative approach ensured that the resulting diabetes self-management app was not only grounded in expert knowledge but also deeply embedded in the cultural and practical realities of the Native American users it seeks to serve.

### 3.2.2. Cultural Diversity in App Design

Our mobile app embraces the cultural diversity of Native American tribes and communities through an adaptable and inclusive design. The following features have been integrated:

- **Dynamic Ontology-Based Foundation:** Built on an extensible ontology, the app's knowledge base is tailored to the Lower Sioux tribe but is designed to expand to include the cultural nuances of other tribes.
- **Modular Design:** This allows for the agile integration of tribe-specific content, ensuring that the app remains relevant across a spectrum of Native American cultures.
- **Community Involvement:** We leverage a community forum chatgroup function for direct input from users, ensuring the app's cultural accuracy and responsiveness.
- **Cultural Education and Sharing:** The app promotes inter-tribal cultural exchange, enhancing understanding and unity among its users.

By incorporating these elements, our app is designed to be a culturally intelligent tool for diabetes self-management, adaptable to the varied practices and traditions of Native American communities.

### 3.2.3. System Design

With the input garnered from the community, we progressed to the System Design phase, which is shown on the right side of Figure 1. In this stage, we employed a participatory design methodology to mold the system's interface, functionality, and features. Our design choices were culturally informed, integrating tribal logos and NA celebrity voices to ensure the application resonates with the community's cultural heritage and preferences. The participatory nature of this phase is illustrated by the continuous loop back to the central community collaboration element in the figure, signifying an ongoing exchange of ideas and feedback.

#### 3.2.3.1. Accessibility and Communication Strategy

Our mobile application has been meticulously crafted to facilitate effective communication and comprehension of health management strategies within the Native American community. During the design phase, input from tribal community members was sought to identify straightforward and easily understood terms, thereby avoiding intricate medical terminology.

Recognizing the potential obstacle that language can present, we have strategically reduced dependence on text. The application makes extensive use of intuitive visuals, symbols, and icons that are widely recognizable and can bridge linguistic gaps. These graphical elements are instrumental in effectively imparting health-related information.

In addition to visual aids, the application features an integrated natural voice assistant. This tool extends beyond mere navigation aid; it serves as an interactive medium for users who might favor auditory learning or those who find it more accessible. The voice assistant is programmed to communicate in a clear and comprehensible manner, catering to users with diverse literacy levels and language abilities.

In essence, our application overcomes language challenges by combining straightforward language, illustrative aids, and a user-friendly voice assistant, ensuring that all users, irrespective of their language proficiency, can easily access and understand health management strategies.

#### 3.2.3.2. Socioeconomic Factors in App Design

Our mobile app is strategically designed to mitigate the socioeconomic challenges faced by Native American populations in managing diabetes. It provides personalized nutrition plans that are sensitive to users' financial situations and the local availability of food, ensuring that the dietary recommendations are both affordable and accessible. The app respects and supports the inclusion of traditional Native American foods, fostering sustainable eating practices deeply rooted in cultural heritage. Recognizing the variety in health literacy levels, the app conveys information using straightforward language and visual aids to ensure comprehensibility for all users. A curated database within the app guides users to local, cost-effective options for diabetes-friendly meals and healthcare services. Additionally, community forums and groups within the app serve as platforms for sharing resources and strategies, reinforcing a network of support attuned to the users' socioeconomic realities. This holistic approach to app design underscores our commitment to providing a supportive and accessible diabetes self-management tool for Native American communities.

#### 3.2.3.3. Cultural User Centered Interface Design

Our approach adopts a Cultural-User-Centered Design (CUCD) approach to construct an interface meticulously crafted to resonate with Native American (NA) users. This design methodology is deeply rooted in cultural sensitivity, ensuring that the interface serves not just as a tool, but also as a cultural bridge. We've integrated a spectrum of culturally significant elements, each chosen for its profound connection to the NA community

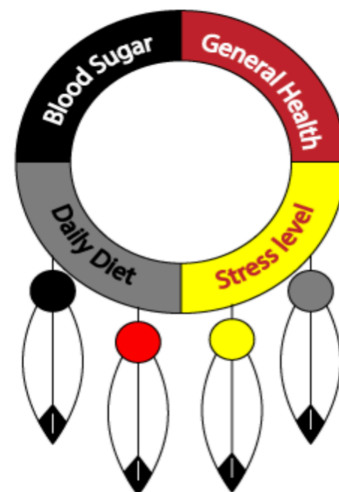


Figure 2: Tribal Emblem as UI Component

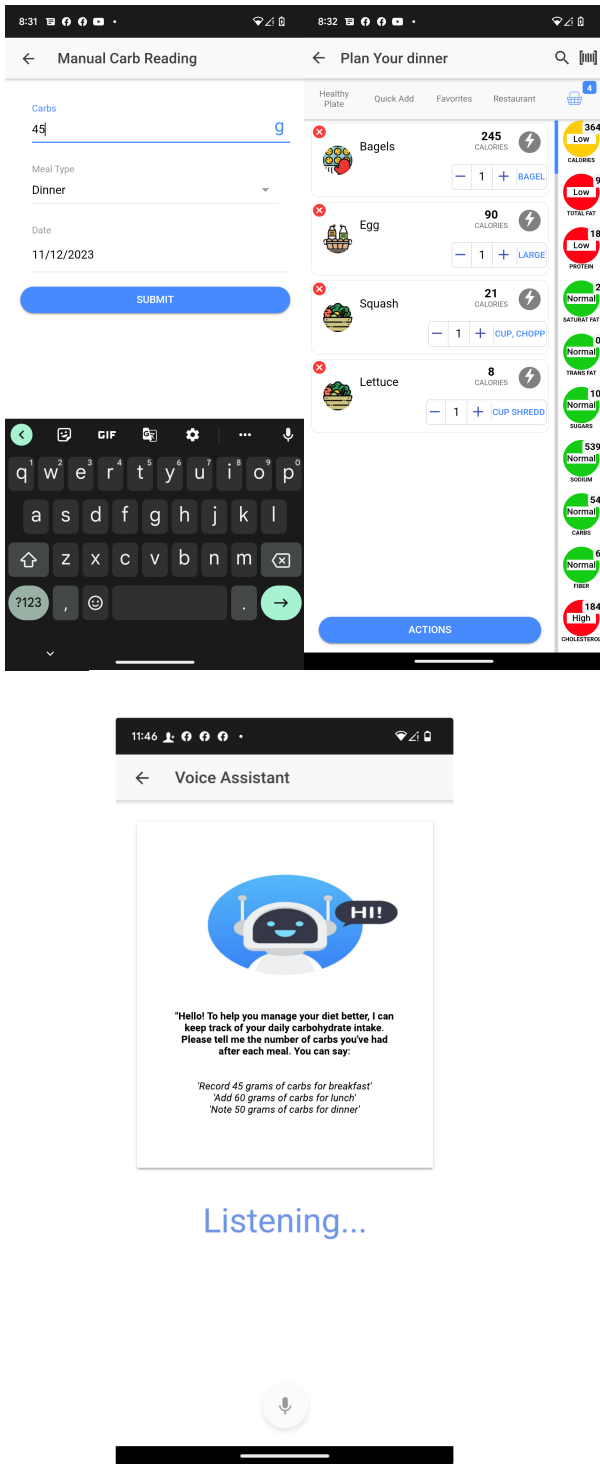


Figure 3: (a) Manual carb input Interface. (b) Calculated carb input through meal selection Interface (c) Voice assistant based Carb input

- **Tribal Emblem Integration**

The integration of the tribal emblem within our user interface serves as a cornerstone of our culturally tailored design strategy. This emblem shown in figure 2, segmented into four interactive quadrants, does more than provide aesthetic value; it functions as a dynamic interface element. Each quadrant is meticulously crafted

to represent an essential component of the system's functionality and to serve as a navigational aid, transforming into buttons that, upon interaction, reveal detailed logs and further information pertaining to the system's various modules. The emblem itself is deliberately chosen for its cultural significance, embodying the rich heritage and shared history of the community. It acts as a visual anchor, grounding the user experience in familiar iconography that resonates on a personal level with the community members. Its inclusion is a thoughtful acknowledgment of the community's values, traditions, and identity, intended to engender a sense of belonging and ownership over the technological experience.

This emblematic interface promotes accessibility and ease of use, with each interactive portion designed to guide users through the system's capabilities intuitively. For instance, selecting one quadrant might reveal insights into the daily diet analysis data, while another might delve into blood sugar logs and trends over

time. This method of organization not only simplifies navigation but also aligns with the community's cognitive patterns and cultural context, making the system's operation inherently intuitive for the intended NA users. It's a harmonious blend of function and tradition, ensuring that the system is not only efficient but also culturally congruent, facilitating a seamless and resonant user experience.

- **Celebrity Vocal Engagements**

To enrich the interface's auditory experience, we have procured voice recordings from Native American luminaries and venerated figures, like Evan Adams, known for his pivotal portrayal of Thomas Builds-the-Fire in the seminal film "Smoke Signals." His voice serves as a thread weaving through our interface, providing guidance and feedback, imbuing the user experience with authenticity and a deeper cultural linkage. This initiative goes beyond mere functionality—it is a homage to a voice that echoes the collective memories and experiences of the NA community.

- **Linguistic Tailoring**

Acknowledging the diverse linguistic capabilities within the NA communities, our interface offers customizable language options. It allows users to select their preferred language and adjust the complexity of the language to match their literacy level, ensuring accessibility and comprehension for all. Insights into this aspect of our design process are detailed in publication [18].

- **Multi Modal Interaction**

Our application offers a multifaceted interface designed to accommodate the distinctive communication styles within the community. As depicted in Figure 3(a), the app includes a straightforward graphical user interface for manual carbohydrate entry. This tactile method is uncomplicated, facilitating quick and direct input of carbohydrate consumption through simple screen interactions.

In another approach, demonstrated in Figure 3(b), the app leverages a sophisticated algorithm to estimate carbohydrate intake from meal selections. Users can describe their meals, and the app intelligently calculates and logs the carbohydrate content, streamlining the tracking process.

Additionally, Figure 3(c) introduces the convenience of a voice assistant. This modality allows users to verbally log their

carbohydrate intake, accommodating those who might find manual entry cumbersome or prefer the immediacy of speech.

These three modalities reflect our dedication to inclusive design, offering both visual and auditory options to manage health data in the most intuitive way possible. By embracing diverse methods of interaction, the app not only improves user experience but also respects the cultural and individual preferences of our wide-ranging user demographic. This inclusive approach positions our application as a comprehensive tool for managing diabetes through user-friendly carbohydrate tracking.

- *Cultural Narratives and Context*

The interface weaves in narratives, traditions, and practices that are the fabric of NA cultural identity. These elements are not mere adornments; they serve to create a digital experience that is culturally congruent, fostering a sense of community and relevance.

The CUCD methodology champions the idea that an interface should be a cultural artifact, one that NA users can see themselves reflected in. By weaving these culturally rich threads into the interface, we aim to deliver an experience that is not only functionally superior but also culturally enriching and empowering for NA users.

### 3.2.3.4. Cultural Tailored System Design

In the development of our system, we prioritize a nuanced approach tailored to Native American (NA) users, focusing on elements that cater to their unique dietary needs and cultural practices. This includes:

Our system is finely tuned to offer personalized nutrition advice, considering the distinct dietary preferences and needs of NA users. We understand the importance of incorporating traditional and culturally significant foods in our recommendations. This includes a special emphasis on game food, which is not only a dietary staple but also holds deep cultural significance in many NA communities.

Recognizing the diverse culinary landscape across NA communities, our app considers the availability of local foods. This involves integrating traditional ingredients and dishes that are integral to various NA cultures, ensuring that our nutritional guidance is not only health-focused but also culturally resonant.

The socioeconomic context of NA communities is a critical aspect of our system design. We strive to make our nutritional recommendations affordable, considering the varying economic backgrounds of our users. This ensures that our guidance is realistic and accessible, not just theoretically ideal.

Our app goes beyond basic nutritional advice by embracing the rich cultural heritage of NA communities. We respect and incorporate traditional food practices and preferences, ensuring that our recommendations do not conflict with cultural norms and practices. Our app also acknowledges the diverse cultural practices and beliefs of tribal members, offering personalized lifestyle recommendations. For instance, we suggest physical activities like traditional dancing or nature walks on ancestral trails, which are not only enjoyable but also culturally significant. These suggestions honor the traditions and enhance the physical well-being of NA users.

To bridge the gap between nutritional advice and practical meal options, we analyze local restaurant menus. This analysis is geared towards ensuring that the food options available to NA users at local eateries are in line with their dietary requirements and cultural preferences. Previously our work [19–21] on restaurant menu and personalized meal planning describes how health conditions can be used to recommend foods. This feature not only supports healthy eating habits but also promotes cultural integrity and relevance.

Understanding the centrality of family and community in NA culture, our app encourages social activities that strengthen these bonds. This could include family-based support systems or participation in cultural events and gatherings. By doing so, we aim to foster a supportive environment that is conducive to positive health outcomes and sustained user engagement.

Our social network feature is designed with a deep understanding of the importance of community connections in NA culture. It enables users to connect, share experiences, and support each other in their journey of diabetes self-management. This virtual community space is a platform for exchanging ideas, sharing cultural practices pertinent to diabetes management, and providing emotional support. It's more than a network; it's a communal space that strengthens the sense of belonging and encourages positive health behaviors.

We facilitate easy and meaningful communication between NA users and healthcare professionals who are attuned to their cultural context. This feature ensures that healthcare delivery is culturally competent, addressing the unique needs and preferences of NA users. It's a bridge connecting users with healthcare providers in a way that respects and understands their cultural background.

The app provides educational content specifically tailored to the cultural sensitivities and needs of NA users. Covering various aspects of diabetes management, including diet, physical activity, medication, and self-monitoring, this content is not just informative but also relatable. It's designed to be a culturally relevant guide in managing diabetes effectively.

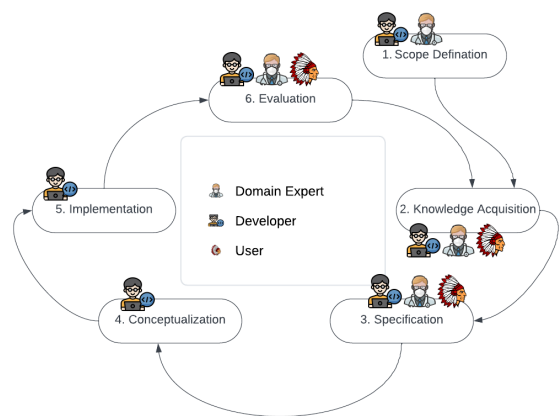


Figure 4: Ontology design and development life cycle

By incorporating features that facilitate effective communication and tailored education, our app plays a pivotal role in promoting culturally competent healthcare. It respects and acknowledges the influence of cultural context on health beliefs



and behaviors, aiming to bridge any gaps between healthcare providers and users.

Overall, our system is meticulously designed to provide a relevant, inclusive, and culturally sensitive user experience for NA users. It aims to promote healthy lifestyle changes, foster community support, and ultimately contribute to improved health outcomes within the NA community. Our approach is not just about managing diabetes; it's about embracing and integrating the rich cultural heritage of NA users into every aspect of their healthcare journey.

### 3.3. NA Biocultural Profile Ontology Development

In designing self-management solutions for diabetes within Native American (NA) populations, it is essential to acknowledge and integrate their distinct biocultural attributes. Our approach introduces a sophisticated biocultural user profile ontology, NAOnto, tailored to encapsulate the intricate tapestry of biological, cultural, socio-economic, and environmental factors that characterize NA diabetes patients. This ontology is a rich repository of data points, including but not limited to cultural influences, geographic locations, prevalent health beliefs and attitudes, varying literacy and numeracy levels, physical capabilities, familial support structures, and economic conditions. Together, these variables form a holistic view of a patient's well-being, firmly rooted in their unique biocultural landscape.

The creation of NAOnto was orchestrated through a structured six-phase process, as depicted in Figure 4. Beginning with scope definition, our team delineated the boundaries and objectives of the ontology, ensuring alignment with the specific self-management needs of NA diabetes patients. Subsequent knowledge acquisition involved meticulous gathering of information from a host of resources, including direct inputs from the community and domain experts. Specification followed, where the acquired knowledge was translated into a detailed framework, leading to the conceptualization phase, which established the ontology's foundational structures and relationships.

In the course of development, we have collaborated with the Community Health & Social Service Directors, the Clinic CEO, and two of their assistants from the Lower Sioux Indian Community. While these professionals did not participate in the broader user validation process, they played a crucial role during the design phase. Their insights and contributions were pivotal in tailoring our app's design to be culturally sensitive and to address the specific healthcare needs of the Native American community.

To ensure a wider relevance and applicability, a survey-based user study with nine participants was conducted across Native American communities in North Dakota and Minnesota.

The implementation phase saw the ontology being constructed and populated with data, while the evaluation phase subjected it to rigorous community and expert review, allowing for validation and calibration against real-world scenarios. This cyclical development process—iterative in nature—permitted ongoing refinement and updating of the ontology, ensuring that it remained dynamic and reflective of the evolving understanding of the NA community's

needs and the latest medical insights. The result is a dynamic, living model that is as responsive as it is comprehensive.

The utility of NAOnto extends beyond a mere database; it is a critical tool for crafting personalized diabetes self-management

strategies that honor the biocultural context of NA patients. Through NAOnto, we enable the creation of interventions that are not only medically precise but also culturally consonant, fostering greater acceptance and effectiveness.

Figure 5 presents a partial view of the NAOnto that is defined for this research. Further details of the ontology's structure and its application within our CBPR framework are elaborated in the accompanying publications [22, 23], where we delve into the granularities of NAOnto and its operational impact.

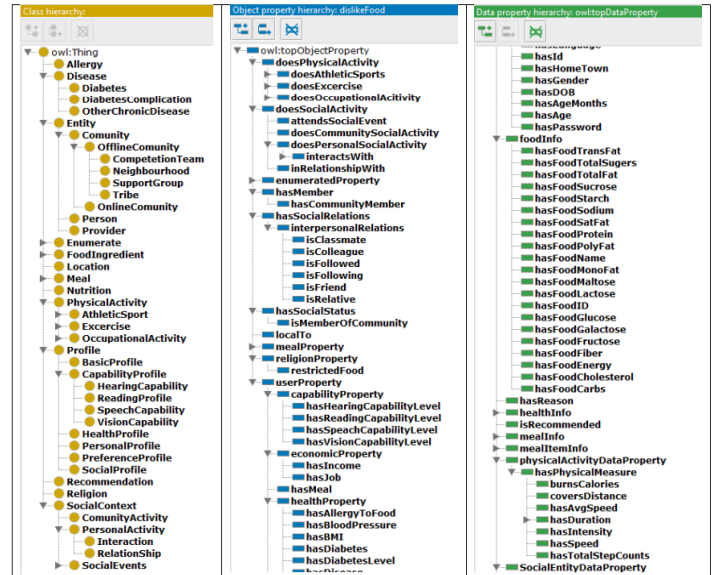


Figure 5: Partial view of ontology NAOnto

## 4. Implementation

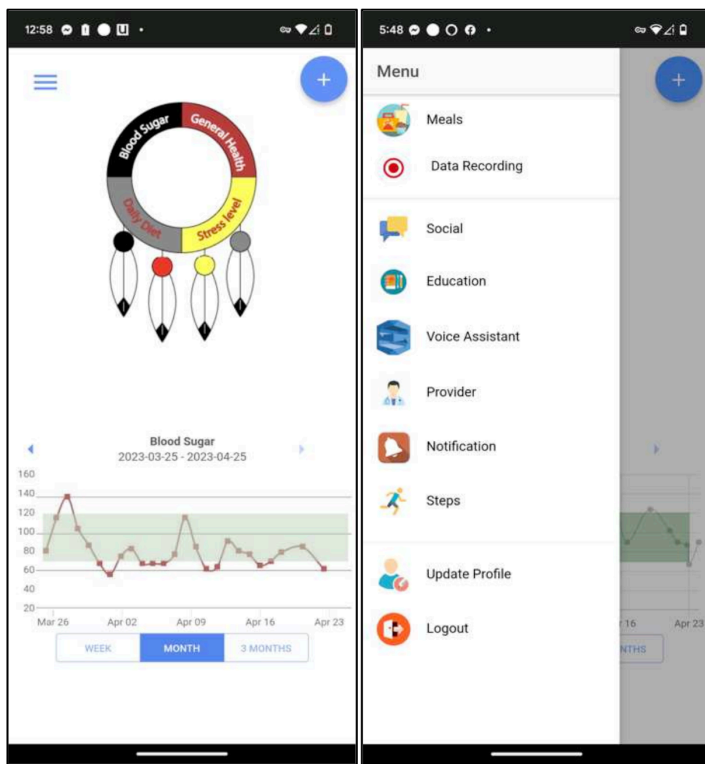
The development and deployment of our diabetes self-management mobile app leveraged the capabilities of the Ionic framework and Javalin server, optimizing it for both functionality and security.

Ionic was selected for its exceptional cross-platform capabilities, allowing us to create a single codebase that functions seamlessly on both iOS and Android platforms. This choice significantly streamlined the development process, ensuring a consistent user experience across different devices.

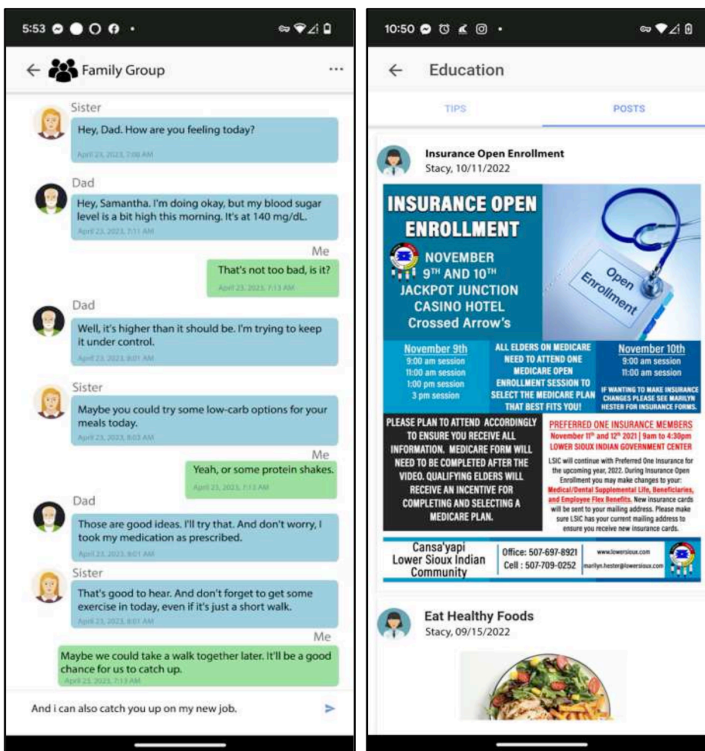
Javalin, a lightweight Java and Kotlin web framework, was employed for handling HTTP requests between the app and our backend server. It efficiently routes these requests to the appropriate handlers, ensuring swift and accurate data processing.

To safeguard user data, all communications between the mobile app and the Javalin server were encrypted using HTTPS protocols. Additionally, data stored on the SQL server was secured using RSA encryption, reflecting our commitment to data security and user privacy.

The app features several unique interfaces, each designed with specific user needs in mind. In figure 6(a) shows the integration of the Lower Sioux tribal logo into our diabetes management system offers a culturally sensitive and intuitive interface. The health



(a) (b)



(c) (d)

Figure 6: Screenshots of the mobile app.

wheel, divided into sections representing blood sugar levels, general health, stress levels, and diet, allows users to input and track their data efficiently. In figure 6(b) menu encompasses a broad range of features including meal recommendations, health data recording, social networking, educational resources, and

direct connectivity with healthcare providers. It's a holistic state of tools designed to facilitate diabetes self-management. Family Group Communication Feature is shown in figure 6(c) Acknowledging the importance of family in NA culture, this feature enables family members to support each other in diabetes management, promoting a collaborative care approach. Culturally feature enables family members to support each other in diabetes management, promoting a collaborative care approach. Culturally Tailored Education and Events is shown in figure 6(d). This aspect of the app keeps users informed about cultural events and educational resources, fostering community engagement and participation.

Each app function underwent rigorous testing to ensure optimal performance and usability. We conducted multiple use case studies to assess the feasibility and effectiveness of features such as meal planning and restaurant menu recommendations. Specific app functions like personalized meal planning and restaurant menu screening were evaluated through user studies. These studies provided insights into the usability and impact of these features on the self-management of diabetes among NA users.

While individual functions have been evaluated, a comprehensive assessment by the tribal community is pending. We plan to engage with the Lower Sioux community to gather feedback and assess the overall effectiveness of the app. This evaluation will offer vital insights into the app's acceptability, practicality, and its impact on health outcomes in the NA community.

To ensure sustained engagement with our mobile app, we have developed a comprehensive strategy focusing on personalization, gamification, community involvement, and continuous improvement. Personalization is key; the app learns and adapts to user interactions, offering tailored recommendations to maintain relevance. Gamification techniques, such as challenges and rewards, are introduced to engage users and encourage consistent app usage. As part of our cyclical Community-Based Participatory Research (CBPR) process, depicted in Figure 1, we commit to regular content updates, providing fresh information and resources that invite users to keep engaging with the app. Community features like forums and support groups are incorporated to cultivate a supportive network, enhancing user motivation through shared experiences and successes. Additionally, we have established a feedback mechanism that enables users to contribute to the app's evolution, ensuring it continuously aligns with their preferences and needs. This strategic approach aims to integrate the app into users' daily routines, making it a valuable, long-term asset in diabetes self-management.

In summary, our diabetes self-management app combines technological innovation with cultural sensitivity. It's designed not just as a tool for health management, but as a culturally attuned companion for Native Americans. Our goal is to empower users through a blend of modern technology and respect for traditional values and practices, thereby facilitating effective diabetes management in the NA community.

## 5. Conclusion

In the concluding section of our paper, we emphasize the successful design and development of a mobile application that is

specifically tailored to the needs and cultural nuances of Native Americans (NAs) with diabetes. Our approach, characterized by a deeply collaborative and iterative design process, incorporated valuable insights from the Lower Sioux Indian Community. This collaboration was instrumental in shaping an app that is not only technologically sound but also deeply respectful and reflective of the unique cultural traditions, local resources, and socio-economic realities of the NA population.

The app offers bespoke dietary recommendations, aligning with the traditional dietary habits and available local food resources of NAs. This personalized approach ensures that the guidance is not only health-conscious but also culturally appropriate. Acknowledging the integral role of family in NA communities, the app incorporates features that enable and encourage family involvement in diabetes management, promoting a collective approach to health and wellness. The app provides users with information on local resources, including traditional foods and community events, facilitating a connection between users and their local environment. By embedding cultural elements into the app's interface and functionality, we ensure that the technology resonates with the users on a cultural level, enhancing engagement and relevance.

Moving forward we aim to conduct extensive user studies within the Lower Sioux community. These studies will involve a larger sample size to ensure a more diverse and representative feedback. A key focus of our upcoming studies will be to evaluate the app's effectiveness in improving health outcomes among NAs with diabetes. This will involve tracking various health indicators and user engagement metrics over time. The feedback obtained from these comprehensive studies will be pivotal in further refining the app. Our goal is to continuously evolve the app's features and usability to better align with the needs and preferences of NA users. While our initial focus has been on the Lower Sioux community, we envisage expanding the reach of the app to other NA communities, adapting it to different tribal traditions and needs.

In summary, our paper not only details the creation of a culturally tailored mobile app for NAs with diabetes but also sets a roadmap for its future development and enhancement. By continuing to engage with the community and rigorously assessing the app's impact, we aim to create a dynamic tool that not only supports diabetes management but also fosters a deeper connection between health technology and cultural identity. This endeavor reflects our commitment to developing solutions that are not only technologically advanced but also culturally empathetic and responsive to the unique needs of diverse populations.

### Conflict of Interest

The authors declare no conflict of interest.

### Acknowledgment

The authors express their gratitude to Dr. Donald Warne, Darin Prescott, Stacy Hammer and the Lower Sioux Indian community for their invaluable support, collaboration, valuable suggestions, and feedback. Their contribution has been instrumental in the development and improvement of our app, and we are thankful for their ongoing partnership and commitment to improving healthcare outcomes for Native American communities affected by diabetes. This work was supported by the National Science Foundation (NSF) with award numbers: 1722913 and 2218046.

### References

- [1] National Diabetes Statistics Report | Diabetes | CDC, Jan. 2024.
- [2] Disparities | Fact Sheets, Jan. 2024.
- [3] A.H. Ariel-Donges, E.L. Gordon, B.N. Dixon, A.J. Eastman, V. Bauman, K.M. Ross, M.G. Perri, "Rural/urban disparities in access to the National Diabetes Prevention Program," *Translational Behavioral Medicine*, **10**(6), 1554–1558, 2020, doi:10.1093/TBM/IBZ098.
- [4] M. Sarche, P. Spicer, "Poverty and health disparities for American Indian and Alaska native children: Current knowledge and future prospects," *Annals of the New York Academy of Sciences*, **1136**, 126–136, 2008, doi:10.1196/ANNALS.1425.017.
- [5] R. Struthers, F.S. Hodge, B. Geishirt-Cantrell, L. De Cora, "Participant Experiences of Talking Circles on Type 2 Diabetes in Two Northern Plains American Indian Tribes," *Diabetes Care*, **26**(8), 1094–1115, 2003, doi:10.1177/1049732303256357.
- [6] C.O. Airhienbuwa, C.L. Ford, J.I. Iwelunmor, "Why Culture Matters in Health Interventions," *Health Affairs*, **32**(1), 78–84, 2013, doi:10.1177/1090198113487199.
- [7] M.W. Kreuter, S.N. Lukwago, D.C. Bucholtz, E.M. Clark, V. Sanders-Thompson, "Achieving Cultural Appropriateness in Health Promotion Programs: Targeted and Tailored Approaches," *Health Affairs*, **22**(2), 133–146, 2003, doi:10.1177/1090198102251021.
- [8] M. Barrera, F.G. Castro, L.A. Strycker, D.J. Toobert, "Cultural adaptations of behavioral health interventions: a progress report," *Journal of Consulting and Clinical Psychology*, **81**(2), 196–205, 2013, doi:10.1037/A0027085.
- [9] A.L. Asad, T. Kay, "Toward a multidimensional understanding of culture for health interventions," *Social Science & Medicine*, **144**, 79–87, 2015, doi:10.1016/J.SOCSCIMED.2015.09.013.
- [10] A. Quevedo Rodríguez, A.M. Wägner, "Mobile phone applications for diabetes management: A systematic review," *Endocrinología, Diabetes y Nutrición*, **66**(5), 330–337, 2019, doi:10.1016/J.ENDINU.2018.11.005.
- [11] P.P. Brzan, E. Rotman, M. Pajnkihar, P. Klanjsek, "Mobile Applications for Control and Self Management of Diabetes: A Systematic Review," *Journal of Medical Systems*, **40**(9), 1–10, 2016, doi:10.1007/S10916-016-0564-8/METRICS.
- [12] S. Sneha, S. Thalla, I. Rischie, H. Shahriar, "Health Internet Technology for Chronic Conditions: Review of Diabetes Management Apps," *Journal of Health Management and Practice*, **20**(2), 17431, 2024, doi:10.2196/17431.
- [13] M.M. Kebede, C.R. Pischke, "Popular diabetes apps and the impact of diabetes app use on self-care behaviour: A survey among the digital community of persons with diabetes on social media," *Frontiers in Endocrinology*, **10**(MAR), 135, 2019, doi:10.3389/FENDO.2019.00135/BIBTEX.
- [14] F. Petersen, "Impact of Culture on the Adoption of Diabetes Self-Management Applications: Cape Flats, South Africa," *ArXiv.Org*, 2021.
- [15] A. Millan-Ferro, A.E. Caballero, "Cultural approaches to diabetes self-management programs for the Latino community," *Current Diabetes Reports*, **7**(5), 391–397, 2007, doi:10.1007/S11892-007-0064-9/METRICS.
- [16] E. Chacko, "Culture and therapy: complementary strategies for the treatment of type-2 diabetes in an urban setting in Kerala, India," *Social Science & Medicine*, **56**(5), 1087–1098, 2003, doi:10.1016/S0277-9536(02)00105-3.
- [17] C.M. Goody, L. Drago, "Using Cultural Competence Constructs to Understand Food Practices and Provide Diabetes Care and Education," *Diabetes Spectrum*, **22**(1), 43–47, 2009, doi:10.2337/DIASPECT.22.1.43.
- [18] R. Hendawi, S. Alian, J. Li, "A Smart Mobile App to Simplify Medical Documents and Improve Health Literacy: System Design and Feasibility Validation," *Journal of Health Management and Practice*, **20**(2), 17431, 2024, doi:10.2196/17431.
- [19] B. Maharjan, J. Li, J. Kong, C. Tao, "Alexa, What Should I Eat?: A Personalized Virtual Nutrition Coach for Native American Diabetes Patients Using Amazon's Smart Speaker Technology," *2019 IEEE International Conference on E-Health Networking, Application and Services, HealthCom 2019*, 2019, doi:10.1109/HEALTHCOM46333.2019.9009613.
- [20] M.S.A.T. Zadeh, J. Li, S. Alian, "Personalized Meal Planning for Diabetic Patients Using a Multi-Criteria Decision-Making Approach," *2019 IEEE International Conference on E-Health Networking, Application and*

- Services, HealthCom 2019, 2019,  
doi:10.1109/HEALTHCOM46333.2019.9009593.
- [21] W. Ul Hasan, K. Tuz Zaman, M.S.A.T. Zadeh, J. Li, "Eat This, Not That!  
- a Personalised Restaurant Menu Decoder That Helps You Pick the Right  
Food," 2022 IEEE International Conference on E-Health Networking,  
Application and Services, HealthCom 2022, 43–48, 2022,  
doi:10.1109/HEALTHCOM54947.2022.9982770.
- [22] V. Pandey, J. Li, S. Alian, "Evaluation and Evolution of NAOnto - An  
Ontology for Personalized Diabetes Management for Native Americans,"  
2021 7th International Conference on Computer and Communications,  
ICCC 2021, 1635–1641, 2021, doi:10.1109/ICCC54389.2021.9674339.
- [23] J. Li, S. Alian, "Design and development of a biocultural ontology for  
personalized diabetes self-management of American Indians," 2018 IEEE  
20th International Conference on E-Health Networking, Applications and  
Services, Healthcom 2018, 2018,  
doi:10.1109/HEALTHCOM.2018.8531107.

**Copyright:** This article is an open-access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).



# Smart Agent-Based Direct Load Control of Air Conditioner Populations in Demand Side Management

Pegah Yazdkhasti\*, Julian Luciano Cárdenas–Barrera, Chris Diduch

Department of Electrical & Computer Engineering, Smart Grid Research Center, University of New Brunswick, Fredericton, Canada

## ARTICLE INFO

Article history:

Received: 10 November, 2023

Revised: 12 January, 2024

Accepted: 12 January, 2024

Online: 06 February, 2024

Keywords:

Demand-Side Management

Smart Grid

System Identification

Direct Load Control

Load Forecast

Rebound Effect

## ABSTRACT

The integration of fluctuating renewable resources such as wind and solar into existing power systems poses challenges to grid reliability and the seamless incorporation of these resources. To address the inherent variability in renewable generation, direct load control emerges as a promising method for demand-side management. Thermostatically controlled appliances, like air conditioners, hold a significant role in this approach. However, effective direct load control necessitates accurate load magnitude estimation and the potential for load shifting. In this paper, we introduce a smart-agent architecture that employs a mathematical model to forecast aggregated power consumption behavior, even when changes are introduced by the controller. To assess system performance, a numerical simulator was developed, demonstrating the system's adaptability to changes, its self-retraining capability, and its continuous improvement in predicting aggregated power consumption.

## 1. Introduction

This paper builds upon the research initially introduced at ICCAD-2023 [1], where the authors presented a smart agent to make a mathematical model for a population of air conditioners and forecast their power consumption and control it. This paper expands the mathematical formulation for modeling and provides justification for the linear approximation. Additionally, it improves the forecasting model by introducing a new formulation for reserve capacity to reshape the aggregated power demands.

The utilization of clean resources, such as wind and solar power, for electricity generation has experienced substantial growth in the past decade [2]. However, the integration of renewable resources poses challenges to existing power grids, primarily due to the rapid fluctuations in generation, diminishing the reliability and quality of electric power systems [3]. While energy storage devices like batteries can mitigate the swift fluctuations of wind and solar power [4], the cost of deploying renewable energy increases with the use of energy storage. Direct Load Control (DLC) emerges as a cost-effective demand-side management (DSM) method to address the intermittency of generation [5]. Recent studies demonstrate that adjusting set points or operation times of thermostatically controlled loads (TCL), such as electric water heaters or air conditioners, effectively controls aggregated power consumption [6, 7, 8, 9].

The system operator plays a crucial role in balancing generation and demand sides, and an accurate forecast of the power consumption is essential for this task. In a DLC program, the system operator needs to understand the controller's capacity to increase or decrease the load, providing feasible instructions.. This research focuses on developing a methodology for fast and accurate load controllability forecasting.

In recent years, significant efforts have been dedicated to developing mathematical models and control strategies aimed at effectively managing the aggregate demand of air conditioners. One approach involves designing a control strategy to adjust the thermostat set point, thereby reducing peak demand and alleviating the impact of renewable energy variability [10, 11]. Alternatively, centralized or distributed/decentralized control strategies have been explored, focusing on directly managing the on/off states of AC compressors [11, 12, 13, 14, 15]. However, these methods fall short in providing practical means to estimate essential model parameters like thermal capacitance/resistance, or to dynamically adapt the model to changing conditions. Another challenge lies in the absence of information regarding the capacity to increase or decrease the total load within a future time window.

In response to these challenges, several studies have shifted focus towards forecasting the controllable load. Some utilize neural networks to predict air conditioning power consumption

\*Corresponding Author: Department of Electrical & Computer Engineering, Smart Grid Research Center, University of New Brunswick, Fredericton, Canada, Email: [pegah.ykh@unb.ca](mailto:pegah.ykh@unb.ca)

based on temperature, humidity, and historical power consumption [16, 17, 18]. Others employ regression-based algorithms coupled with Kalman filters to forecast air-conditioning load using historical data [19]. Least square support vector machines have also been proposed for short-term load forecasts [20]. However, these methods may face impracticalities when integrated with a control system that alters load behavior. Furthermore, none of these methods offer insights into how much the load can be changed due to a control action or formulate the control capacity.

Introducing a dynamic process and a state space model, our proposed methodology provides an efficient and accurate forecast of the controllable load for the system operator. By constructing a model adaptable to changes in thermal parameters and load behavior, our method surmounts limitations associated with traditional approaches. Furthermore, it equips the system operator with the ability to estimate the controllable capacity of the load, a critical factor in load balancing tasks.

The effectiveness of our methodology was rigorously evaluated through simulations, comparing its performance with existing methods. The results demonstrate that our proposed method surpasses existing approaches in terms of both accuracy and speed. This underscores its potential to offer a more efficient and reliable solution for managing the demand side in response to the integration of renewable resources.

In conclusion, this paper introduces a novel methodology for forecasting the controllable load of air conditioners, capable of adapting to changes in thermal parameters and load behavior. The proposed methodology not only delivers an accurate and efficient load forecast but also empowers the system operator to estimate the controllable capacity of the load. With promising results in simulations, our method holds the potential for practical implementation, promising improved demand side management in response to the integration of renewable resources.

The remainder of this paper is structured as follows: Section 2 provides an overview of the system architecture, while Section 3 offers a concise introduction to the mathematical model of the system. The learning algorithm is detailed in Section 4, and Section 5 outlines the process of generating forecasts. The capacity to control the load is formulated in Section 6. The paper concludes with the presentation of simulation results in Section 7.

## 2. System Architecture

The objective of this study is to devise a smart agent for the efficient management of power consumption in thermostatically controlled appliances. This system operates within a hierarchical structure, collaborating with a system operator or a virtual power plant (VPP). Its role is to receive a target aggregated power consumption, interact with a load aggregator, and send control signals to adjust the load. The goal is to minimize the discrepancy between the measured and desired power consumption while maintaining customers' comfort. Figure 1 provides an overview of the system, which encompasses a load aggregator with three distinct modules:

1. Learning Model Module: This module is tasked with estimating a mathematical model that articulates the power consumption of the aggregated load based on input parameters.

2. Forecasting Module: Leveraging the estimated mathematical model and input forecasts, this module generates predictions for the baseline load and reserve capacity. Input forecasts encompass variables like weather conditions, occupancy patterns, and other factors influencing power consumption. The baseline load forecast represents the expected power consumption without any control action, while the reserve capacity forecast indicates the potential power reduction or increase through controller actions. These forecasts guide the controller in determining appropriate control signals to align the load with the desired power consumption.
3. Responsible for adhering to the desired aggregated load set by the system operator, this module controls the load. Further details on a simplified implementation of such a controller are presented in [21].

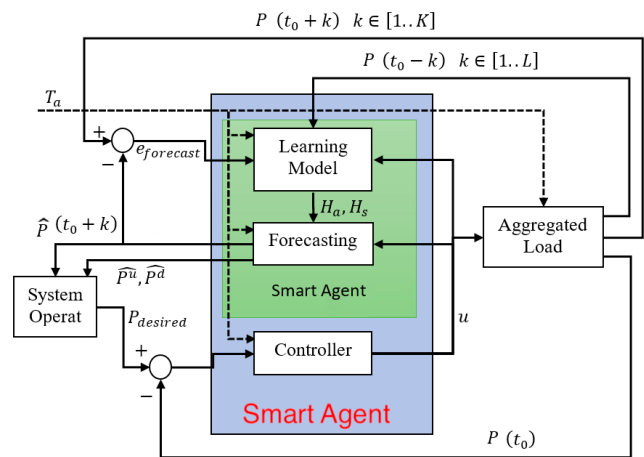


Figure 1: A block diagram for direct load control system [1]

In Figure 1, the power consumption of the load population (Aggregated Load) is input into the load aggregator at the current time denoted by  $t_0$ . Historical data from prior measurements ( $P(t_0 - k)$  for times preceding  $t_0$ ) is utilized by the 'Learning Model' to establish a mathematical model for the system ( $H_a$  and  $H_s$ ). Subsequently, the 'Forecasting' module utilizes this model to generate forecasts for the baseline load and control capacity. In the future, upon obtaining actual measurements of the aggregated load, the forecast error can be calculated as the difference between forecasted and actual power ( $e_{forecast}$ ). This error serves as a trigger mechanism, initiating the training process.

## 3. Modeling the Aggregated Load

The controllable load in this research are a population of air conditioners. A Mathematical model can provide a powerful tool to analyze the behavior of such system. The following sections will present a commonly used model for single AC and a population of ACs.

### 3.1. Individual Load

In this context, we denote  $m(k)$  as the on/off state of the air conditioner at time  $k$ , and  $T$  as the room temperature. The dynamics of room temperature changes are governed by Equation (1), a widely adopted formulation in the literature [22, 23, 12, 21]:

$$\dot{T}(t) = -\frac{1}{CR}(T(t) - T_a(t) + m(t)RP_{nom}\eta) \quad (1)$$

Here,  $R$  denotes the thermal resistance of the room,  $C$  is the thermal capacitance of the room,  $T_a$  represents the ambient temperature,  $P_{nom}$  corresponds to the nominal power rating of the air conditioner, and  $\eta$  stands for the coefficient of performance. The air conditioner switches on ( $m(t) = 1$ ) when the room temperature surpasses the higher threshold setpoint, aiming to reduce the room temperature until it reaches the lower threshold, at which point the air conditioner turns off ( $m(t) = 0$ ).

### 3.2. Aggregated Load

The power consumption of an air conditioner (AC) can be conceptualized as a system with two inputs: the thermostat setpoint (as the controllable input) and ambient temperature (as the disturbance signal). Other parameters, such as thermal resistance and capacitance, may be considered as external hidden parameters of the system.

Simulation results demonstrate that the aggregated power demand reaches a steady-state value when the inputs remain constant, as depicted in Figure 2. In general, the steady-state response of the aggregated power demand to the inputs is nonlinear. However, this surface can be approximated by a plane described by:

$$P_{ss}(T_a, T_s) = P_{ss0} + \alpha T_a + \beta T_s, \quad T_a < 34^\circ\text{C}, T_s \leq T_a \quad (2)$$

Figure 2 illustrates the simulated steady-state aggregated power demand vs. the value obtained from (3). As expected, around the nonlinear zone of high ambient temperature, the plane exhibits larger differences than the simulated power demand. However, for the middle sections of the surface, it provides a good approximation for the steady-state aggregated power demand, with the error range over the middle part of the surface being less than 5%.

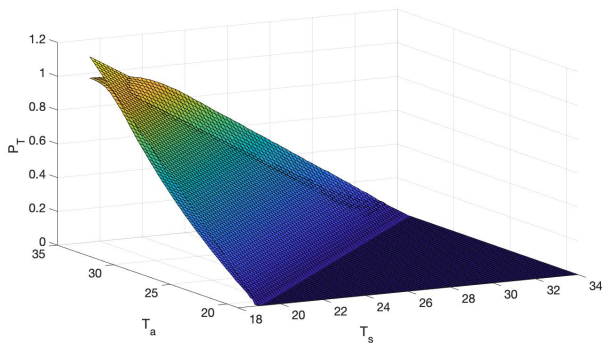


Figure 2: Comparison of the Approximated Plane Surface given by (2) with Simulated Aggregated Power Demand

Through numerical calculations, the coefficients for the best-fitting plane have been determined as:

$$P_{ss}(T_a, T_s) = 0.075T_a - 0.075T_s \quad (3)$$

With this observation, it becomes plausible to consider modeling this system as a Linear Time-Invariant (LTI) system, with the equilibrium point representing the DC gain of the model. Assuming the system is given by:

$$\delta P(s) = H_a(s)\delta T_a(s) + H_s(s)\delta T_s(s) \quad (4)$$

Here,  $\delta P$  denotes changes in the aggregated demand from the equilibrium point due to variations in ambient temperature and thermostat setpoints ( $\delta T_a, \delta T_s$ ).  $H_a$  and  $H_s$  represent the LTI transfer functions for  $T_a$  and  $T_s$  respectively. The steady-state output is then given by:

$$\delta P_{ss} = H_a(0)\delta T_a + H_s(0)\delta T_s \quad (5)$$

This equation mirrors (2), where  $H_a(0) = \alpha$  and  $H_s(0) = \beta$ .

Now, by approximating the aggregated demand with a LTI system, the dynamic response of the system can be captured by[24]:

$$P(k) = y_a(k) + y_s(k) + P_0 \quad (6)$$

Here,  $P_0$  denotes the aggregated demand at the equilibrium point (or  $P_{ss}$ ), while  $y_a$  and  $y_s$  represent the change in aggregated power consumption with respect to the equilibrium point, caused by variations in the ambient temperature and thermostat set point, respectively. These LTI system can be represented in state space model as:

$$H_a : \begin{cases} x_a(k+1) = A_a x_a(k) + B_a u_a(k), \\ y_a(k) = C_a x_a(k) \end{cases} \quad (7)$$

and

$$H_s : \begin{cases} x_s(k+1) = A_s x_s(k) + B_s u_s(k), \\ y_s(k) = C_s x_s(k) \end{cases} \quad (8)$$

where the matrices  $A_a, B_a, C_a, A_s, B_s$  and  $C_s$  should be identified through the learning process.

## 4. Learning Model

The purpose of this module is to formulate an approximation model for the population of air conditioners, eliminating the need for knowledge about any of the physical parameters introduced in Section 3.

The details of the model learning process is presented in [24]. As a short summary, the learning procedure initiates when the forecast error surpasses a predefined threshold level. At this juncture,  $T_a$  and  $T_s$  are determined, and the controller is temporarily disabled to prevent any artificial fluctuation in the aggregated power consumption.

The learning process encompasses two steps:

1. Determining  $H_a$  while holding  $T_s$  constant. In this scenario, changes in the output power are solely governed by variations in  $T_a$ .
2. Determining  $H_s$  when a small change is applied to  $T_s$ , and the effect of changes in  $T_a$  can be estimated and removed using  $H_a$ . The formulations for these steps are elaborated in detail in [24].

## 5. Load Forecast

To generate a load forecast, one can utilize the forecasted values of the model inputs in equations (7)-(8) to obtain a forecast for the output. However, this requires measurements of the states, which are not physical parameters that can be directly measured. As a result, a state observer is required to estimate these states.

### 5.1. State Estimation

The overall architecture of the state observer for this system is shown in Figure 3. In the figure, the changes in the output power are denoted by  $y$ , while the estimated impact of  $u_s$  and  $u_a$  are denoted by  $\hat{y}_s$  and  $\hat{y}_a$ , respectively. The difference between the estimated changes and the actual change in the aggregated power is then fed to the state model through observer gains  $H_s$  and  $H_a$  to improve state estimation.

$$\begin{cases} \hat{x}_a(k+1) = A_a \hat{x}_a(k) + B_a u_a(k) + H_a C_a (x_a(k) - \hat{x}_a(k)) \\ \hat{y}_a(k) = C_a \hat{x}_a(k) \end{cases} \quad (9)$$

and similarly for the  $x_s$  state:

$$\begin{cases} \hat{x}_s(k+1) = A_s \hat{x}_s(k) + B_s u_s(k) + H_s C_s (x_s(k) - \hat{x}_s(k)) \\ \hat{y}_s(k) = C_s \hat{x}_s(k) \end{cases} \quad (10)$$

where  $\hat{x}_a$  and  $\hat{x}_s$  are the estimated states.

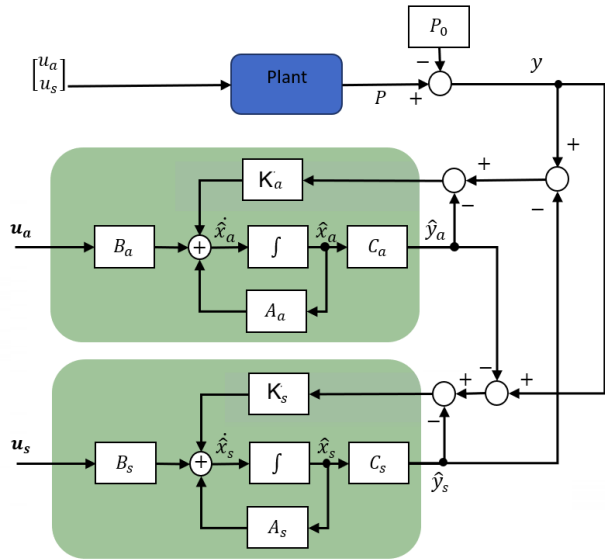


Figure 3: The state observers for each input [1].

These equations can be rewritten in a matrix format as:

$$\begin{cases} \hat{x}(k+1) = \mathcal{A} \hat{x}(k) + \mathcal{B} \begin{bmatrix} u_a \\ u_s \end{bmatrix} + \begin{bmatrix} K_a \\ K_s \end{bmatrix} (y(k) - C \hat{x}(k)) \\ \hat{y}(k) = C \hat{x}(k) \end{cases}$$

where

$$\mathcal{A} = \begin{bmatrix} A_a & 0 \\ 0 & A_s \end{bmatrix}, \quad \mathcal{B} = \begin{bmatrix} B_a & 0 \\ 0 & B_s \end{bmatrix} \quad (11)$$

and

$$C = [C_a \quad C_s] \quad (12)$$

By considering the state error:  $\tilde{x}(k) = x(k) - \hat{x}(k)$ , we will have:

$$\tilde{x}(k+1) = (\mathcal{A} - KC) \tilde{x}(k) \quad (13)$$

And  $K$  can easily be determined so that the dynamics of the state error would be faster than the  $y$  system through functions such as pole placement technique in MATLAB.

### 5.2. Forecasting Baseline Load

When forecasting the baseline load, we operate under the assumption that the controller does not influence the load. Let's consider a forecasting horizon of  $L$  samples, assuming that the forecast is generated at time index  $k_0$ . Within this interval, we have:

$$\hat{u}_a(k) = \hat{T}_a(k) - \hat{T}_{a0} \quad k \in [k_0 + 1, k_0 + L] \quad (14)$$

$$\hat{u}_s(k) = 0 \quad k \in [k_0 + 1, k_0 + L] \quad (15)$$

Here,  $\hat{T}_a(k)$  is the forecasted ambient temperature at time  $k$ . Then the forecast of the baseline aggregated power will be:

$$\hat{P}_b(k) = P_0 + \hat{y}_a(k) \quad k \in [k_0 + 1, k_0 + L] \quad (16)$$

where  $P_0$  is the power consumption at equilibrium point and

$$\begin{cases} \hat{x}_a(k+1) = A_a \hat{x}_a(k) + B_a \hat{u}_a(k) \\ \hat{y}_a(k) = C_a \hat{x}_a(k) \end{cases} \quad (17)$$

where the initial state  $\hat{x}_a(k_0)$  can be obtained from the state observer that was presented earlier.

Equation (16) operates under the assumption that no previous control actions have occurred, implying no fluctuations in the aggregated power resulting from prior thermostat changes. Nevertheless, past control actions may continue to influence future power consumption. Given that the forecasting module receives information about each control signal to the thermostat setpoints, it can consistently revise the state space variables of the system. Consequently, the forthcoming changes in  $\hat{y}_s$  can be determined by:

$$\begin{cases} \hat{x}_s(k+1) = A_s \hat{x}_s(k) \\ \hat{y}_s(k) = C_s \hat{x}_s(k) \end{cases} \quad (18)$$

and the baseline load will be:

$$\hat{P}_b(k) = \hat{P}_0 + \hat{y}_a(k) + \hat{y}_s(k) \quad k \in [k_0 + 1, k_0 + L] \quad (19)$$

The payback effect, also known as rebound, is a well-established phenomenon in DSM. It characterizes the transient surge in aggregated power that occurs once control over the aggregated load is relinquished, allowing the power to revert to its baseline. This transient, initiated when the system shifts from controlled to uncontrolled operation, has the potential to give rise to unexpected secondary demand peaks.



The controller, tasked with executing demand-side management measures, holds the key to influencing the rebound effect by sustaining control over the devices even beyond the system operator's requested interval. However, this strategy may not always be practical or desirable, particularly if associated with high costs or energy consumption. Hence, in this section, we assume that the controller will exclusively operate during the specified interval, refraining from additional adjustments thereafter.

To equip the system operator with valuable insights into the repercussions of control actions, we introduce a method for estimating the payback of any desired demand even before initiating control. This approach aids the system operator in evaluating the acceptability of control actions and preemptively avoiding an escalation in future peak demands. By furnishing a dependable estimate of the payback effect, this method empowers the system operator to optimize the control strategy and mitigate potential risks to power system operation.

### 5.3. Approximating Control Signal

Let us consider a scenario where the system operator or VPP issues a dispatch, denoted by  $\Delta P_d(k)$ , for a control interval of  $C$  samples. In this scenario, the desired load can be considered as a change to the baseline load, which is given by:

$$\Delta P_d(k) = P_d(k) - \hat{P}_b(k) \quad (20)$$

Here,  $\Delta P_d(k)$  represents the amount of power that needs to be added to or reduced from the baseline load. To effect this change, the thermostat set points,  $\hat{u}_s$ , need to be adjusted so that they create appropriate  $\hat{y}_s$  values.

Now, the controller needs to generate a  $\hat{u}_s(k)$  that results in the output  $\hat{y}_s(k)$ , matching  $\Delta P_d(k)$ . This assumes that the linear model remains valid when  $y_s$  changes by  $\Delta P_d$ .

$$\begin{cases} \hat{x}_s(k+1) = A_s \hat{x}_s(k) + B_s \hat{u}_s(k) \\ \hat{y}_s(k) = \Delta P_d(k) = C_s \hat{x}_s(k) \end{cases} \quad (21)$$

This equation can be rewritten as:

$$\Delta P_d(k+1) = C_s [A_s \hat{x}_s(k) + B_s \hat{u}_s(k)] \quad (22)$$

And the control signal,  $\hat{u}_s(k)$  will be:

$$C_s B_s \hat{u}_s(k) = \Delta P_d(k+1) - C_s A_s \hat{x}_s(k) \quad (23)$$

If the order of the system model is  $n$ , then the matrices  $C_s$  and  $B_s$  will have dimensions  $1 \times n$  and  $n \times 1$ , respectively. The product  $C_s B_s$  will then result in a scalar value, and:

$$\hat{u}_s(k) = (C_s B_s)^{-1} (\Delta P_d(k+1) - C_s A_s \hat{x}_s(k)) \quad (24)$$

subject to:

$$-\alpha \leq u_s(k) \leq \alpha \quad (25)$$

for customers' comfort (will be described more in detail in Section 6). Then, the thermostat set points will be:

$$\hat{T}_s(k) = T_{s0} + \hat{u}_s(k) \quad (26)$$

for  $k$  in the control interval. It should be noted that initial state variable,  $\hat{x}_s(k_0)$ , is estimated using the state observer described in Section 5.1.

Figure 4 illustrates an example of an arbitrary desired aggregated load, represented by the top curve. The baseline load forecast,  $\hat{P}_b$ , is also shown in this figure, where the forecast interval is 6 : 00 – 10 : 00 am. The desired aggregated load for the interval of 7 : 00 – 8 : 30 am (control interval) is denoted by  $P_d$ . Equations (24)-(26) were utilized to determine how the thermostat set points should be adjusted in order to achieve the desired load. The resulting adjustments are illustrated by the bottom curve of Figure 4.

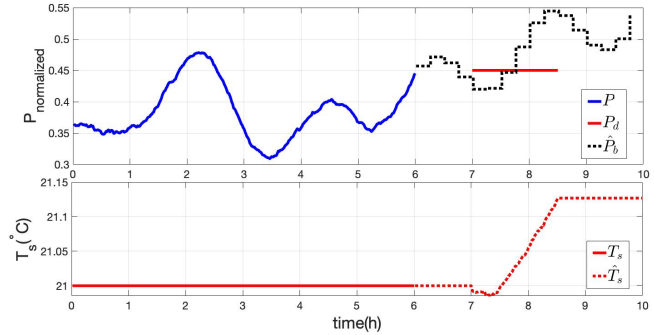


Figure 4: Example of desired load vs. calculated thermostat set point adjustments to follow the load.

### 5.4. Forecasting Rebound

Once the system operator changes the shape of the power demand over an interval, the profile of the demand will also change in the subsequent intervals, which is referred to as the payback or rebound effect of the control action. If an optimal controller adjusts the thermostat set points to follow the desired load demand and stops controlling it after that interval, the forecasting module can predict the shape of the demand after the control interval even before it starts. This feature can be beneficial for the system operator to estimate the shape of the load after a proposed desired aggregated load and determine if the consequences of the control actions are acceptable (e.g., not producing a larger peak in the future). However, it should be noted that various control strategies may affect this forecast.

Let us assume that the system operator issues a desired aggregated power demand, denoted by  $P_d(k)$ . This desired load can be considered as a change to the baseline load:

$$P_d(k) = \hat{P}_b(k) + \Delta P_d(k), \quad (27)$$

where  $\Delta P_d(k)$  is the amount of power that should be added to or reduced from the baseline load. This change should be achieved by adjusting the thermostat set points ( $\hat{u}_s$ ) to create an appropriate  $\hat{y}_s$ . Thus, the controller should create a  $\hat{u}_s(k)$  to match  $\hat{y}_s(k)$  with  $\Delta P_d(k)$ , so an estimation of the control signal will be:

$$\hat{u}_s(k) = (C_s B_s)^{-1} (\Delta P_d(k+1) - C_s A_s \hat{x}_s(k)) \quad (28)$$

Using (28), an estimation of the thermostat control signal can be calculated with the given desired aggregated load. With this

estimated  $u_s$  for the control interval and the forecast of the ambient temperature, a new forecast can be generated to approximate the payback of a control action.

$$\begin{cases} \hat{x}_s(k+1) = A_s \hat{x}_s(k) + B_s \hat{u}_s(k) & \text{for } k \text{ in control interval} \\ \hat{x}_s(k+1) = A_s \hat{x}_s(k) & \text{otherwise} \\ \hat{y}_s(k) = C_s \hat{x}_s(k) \end{cases} \quad (29)$$

The aggregated power demand will be:

$$\hat{P}_{\text{payback}}(k) = \hat{P}_0 + \hat{y}_d(k) + \hat{y}_s(k) \quad (30)$$

## 6. Feasible Desired Demand

One of the key parameters in DLC is ensuring that the controller does not interfere with customers' comfort. For air conditioners, the controller can adjust the thermostat set point to change the aggregated power consumption to follow a desired load. However, if the controller sets the thermostat to a very high or very low value, it may cause discomfort for the end-users and prompt them to opt-out of the DLC program. Therefore, the controller cannot change the thermostat set point to any arbitrary value. The set point must be changed within a certain range to ensure that customers remain comfortable.

This means that the system operator must send a desired demand that can be followed without disturbing the customers. Such a demand is called a feasible desired demand. To ensure that only feasible desired demands are sent, a method will be presented in the next subsections that defines more precisely the customers' comfort and the criteria for feasible desired demands.

### 6.1. Customers' Comfort

One method of ensuring customers' comfort in DLC is to limit the range of change for the thermostat set point. The preferred thermostat set point for a group of loads can be denoted by  $T_s^0$ . The controller can then adjust the  $T_s$  within the range of  $[T_s^0 - \alpha, T_s^0 + \alpha]$ , where  $\alpha$  determines the comfort zone. Once the  $T_s$  reaches these limits, the controller cannot adjust it beyond this range, thus losing control over the aggregated power. This signifies the point when there is no more reserve capacity available.

It is crucial to note that this method of limiting the range of  $T_s$  affects the controller's ability to adjust the aggregated power consumption to follow a desired load. Infeasible desired loads may arise due to the limitations imposed by the comfort zone. Thus, the system operator should only send feasible desired loads that can be followed without disturbing the customers' comfort. In the following subsections, the concept of customers' comfort and feasible desired loads will be defined more precisely, and a method for informing the system operator about the criteria for feasible desired loads will be presented.

### 6.2. Formal Formulation

A feasible desired dispatch is a power profile that satisfies three conditions. First, it should be within the range, i.e.  $0 \leq P_d(k) \leq P_{max}$ .

Second, the controller should be able to produce appropriate changes to the thermostat set points, so that the aggregated demand of the loads would follow the desired load. Third, the changes on the thermostat set points should not exceed the customers' comfort level, as described previously in Section 6.1.

This means that the desired load should be within the range of feasible set point changes. If the desired load is outside of this range, the controller will not be able to follow it without causing discomfort to customers, and therefore, it is not a feasible desired demand.

Assuming that the aggregated power can be modeled with a second-order system, the transfer function representation is given by:

$$\frac{Y_s(z)}{U_s(z)} = \frac{b_1 z^{-1} + b_2 z^{-2}}{1 + a_1 z^{-1} + a_2 z^{-2}} \quad (31)$$

Here,  $Y(z)$  and  $U(z)$  represent the z-transform of the  $y_s(k)$  and  $u_s(k)$  signals, respectively. This equation in time domain will be:

$$b_1 u_s(k-1) + b_2 u_s(k-2) = y_s(k) + a_1 y_s(k-1) + a_2 y_s(k-2) \quad (32)$$

In this equation,  $y_s(k)$  represents the required changes that should be compensated by changing the thermostat set points, and  $u_s(k)$  is the control signal which is the change in thermostat set point, issued by the controller to follow  $y_s(k)$ . It is assumed that the initial conditions are zero, i.e.:

$$u_s(k) = 0 \quad \text{for } k < 0 \quad (33)$$

and

$$\Delta P_d(k) = 0 \quad \text{for } k \leq 0 \quad (34)$$

If we require  $y_s(k)$  to track  $\Delta P_d(k)$  over the control interval, then it can be shown that the solution for  $u_s(k)$  is:

$$u_s(k) = \rho \Delta P_d(k+1) + \sigma \Delta P_d(k) + \tau \Delta P_d(k-1) + \sum_{i=1}^{k-2} ((-b_2/b_1)^{k-1-i} \tau \Delta P_d(i)) \quad (35)$$

where

$$\begin{cases} \rho = \frac{1}{b_1} \\ \sigma = \frac{a_1 b_1 - b_2}{b_1^2} \\ \tau = \frac{a_2 b_1^2 - a_1 b_1 b_2 + b_2^2}{b_1^3} \end{cases} \quad (36)$$

As discussed earlier, a given desired demand is feasible if, for all time steps  $k$ , it satisfies the following two conditions:

$$T_s^0 - \alpha \leq u_s(k) \leq T_s^0 + \alpha \quad (37)$$

or:

$$T_s^0 - \alpha \leq \rho \Delta P_d(k+1) + \sigma \Delta P_d(k) + \tau \Delta P_d(k-1) + \sum_{i=1}^{k-2} ((-b_2/b_1)^{k-1-i} \tau \Delta P_d(i)) \leq T_s^0 + \alpha \quad (38)$$

To ensure the feasibility of a desired demand profile, it is essential to verify that it satisfies Equation (38) for all  $k$  samples of the profile. However, in order to obtain a general form for feasibility, it may be useful to study the feasibility of specific class of desired demand profiles.

An example class is the trapezoidal shape, where  $y_s$  starts from 0 with a constant slope, reaches a maximum value of  $P_d^{max}$  and remains constant for  $k_c$  samples, then goes back down to 0 with a constant slope.

One approach to studying feasibility is to determine the length of time that  $u_s(k)$  can satisfy (37) when  $\Delta P_d$  is a ramp. This evaluation can be repeated for different values of the ramp rate.

Figure 5 displays two ramp-shaped desired demand profiles and the corresponding control signal  $u_s$ . The  $u_s$  curve for each profile is computed using Equation (35) until the customers' comfort level is reached at  $T_s \pm \alpha$ , where  $\alpha = 0.5^\circ C$ . The plot illustrates that higher ramp slopes require more changes in  $T_s$ , causing the system to reach its limit faster. Consequently, the feasibility of profiles with steeper slopes is lower than those with slower rates of change. Figure 10 compares the feasibility intervals of ramp-shaped desired demand profiles with various slopes.

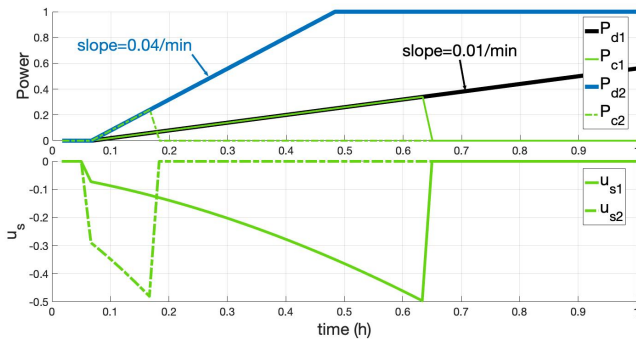


Figure 5: Feasibility intervals for two ramp shape desired demand profiles

It should also be mentioned that the ramp-rate is given for normalized power, which means that the actual rate for power will be obtained by multiplying this rate with the number of individual loads  $N$  and nominal power  $P_{nom}$ . For instance, a rate of  $0.02/min$  for a group of 1,000 air conditioners with a nominal rating of  $2kW$  would correspond to a ramp rate of  $40kW/min$ .

Feasibility studies can also be conducted for the class of trapezoidal-shaped desired demand profiles with varying slope and  $P_d^{max}$ . Figure 6 provides an example of such a profile, along with the calculated  $T_s$  needed to track the desired demand when  $\alpha$  is assumed to be  $0.5^\circ C$ . As the slope and  $P_d^{max}$  values change, so will the corresponding feasibility interval. Figure 12 illustrates how the feasibility interval is affected by these parameters.

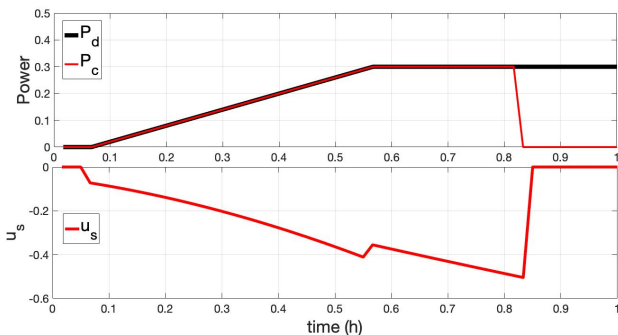


Figure 6: Feasibility interval for a trapezoidal-shaped desired demand

## 7. Simulation Results

To assess the effectiveness of the proposed method, a numerical simulator was developed in MATLAB to simulate a population of  $N = 10,000$  individual air conditioners. As explained in Section 3, we assume that all ACs have the same nominal power, and their thermal capacities follow a log-normal distribution. The simulation employs values listed in Table 1, which have also been used in several other studies.

Table 1: Simulation Parameters [13]

Parameter	Value	Description
$P_{nom}$	2kW	Nominal power rating of the AC
$\eta$	3	Coefficient of performance
$R$	$2^\circ C/kW$	Thermal resistance
$\mu_C$	$3.6kWh/^\circ C$	Mean value of the thermal capacitance
$\sigma_{rel}$	0.2	Standard deviation of log-normal distribution for $C$

### 7.1. System Modeling

For the given population of ACs in this simulation, the mathematical system that was identified is formulated as:

$$A_a = \begin{bmatrix} 1.975 & -.977 \\ 1 & 0 \end{bmatrix}, B_a = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C_a = \begin{bmatrix} 0.0027 & -.0025 \end{bmatrix}$$

$$A_s = \begin{bmatrix} 1.898 & -.903 \\ 1 & 0 \end{bmatrix}, B_s = \begin{bmatrix} 1 \\ 0 \end{bmatrix}, C_s = \begin{bmatrix} -.1381, .1376 \end{bmatrix}$$

### 7.2. Baseline Forecast

Figure 7 shows an example of a forecast profile for the next four hours, generated by the forecasting module at 6:00am on July 3, 2020, using the ambient temperature profile in Barbados. The dashed curve represents the forecasted normalized aggregated power calculated using Equation (16), while the solid curve represents the actual normalized aggregated power calculated using a numerical simulator. The figure illustrates that the forecast module is capable of predicting future power consumption with high accuracy, demonstrating the effectiveness of the mathematical model.

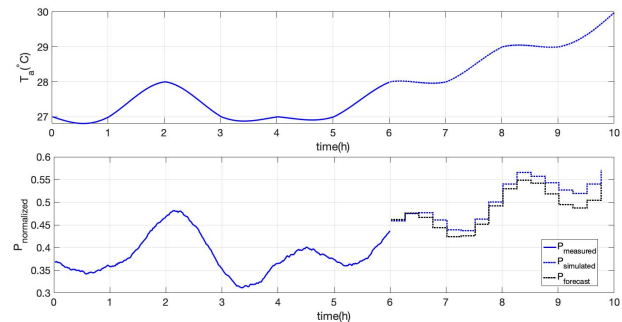


Figure 7: A forecast of the aggregated power consumption, vs. simulated aggregated power consumption in 15 minutes intervals for 4 hours.

Comparison metrics such as the mean absolute percentage error (MAPE), mean absolute error (MAE), and root mean square error (RMSE) are commonly used by researchers to assess the accuracy of forecast models. The proposed method's performance is evaluated using these metrics and compared with other methods in Tables ???. The results show that the proposed method has slightly improved the performance of the baseline forecast. However, the main strength of this method is its ability to predict the load in the presence of control actions, as discussed in the next section. It should also be noted that the MAE and RMSE are calculated based on the normalized power, which is the total power demand divided by the number of loads and the nominal power of each device.

### 7.3. Forecasting Rebound Effect

Figure 8 illustrates the impact of a desired aggregated load on the original baseline forecast ( $\hat{P}_b$ ) and the resulting baseline with payback ( $\hat{P}_{payback}$ ). The blue curve represents the original forecast for the baseline load, while the green curve depicts the desired aggregated load. The orange curve shows the resulting forecast after applying the desired load, which includes the payback effect. The payback effect causes the load profile to be higher after the control interval, compared to the profile if no control action was taken. This figure demonstrates how the desired aggregated load can be used to adjust the power consumption profile over a specified interval.

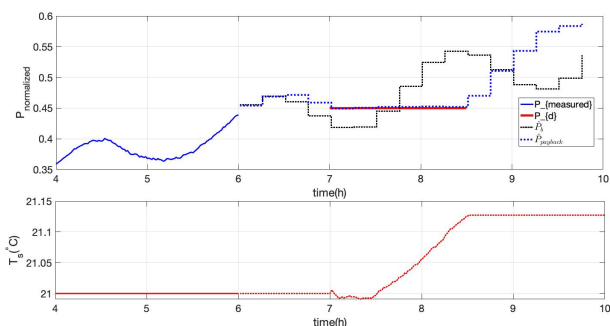


Figure 8: Baseline forecast ( $\hat{P}_b$ ) and forecast with payback effect ( $\hat{P}_{payback}$ ).

### 7.4. Forecasting Control Capacity

Figure 9 illustrates an example of using a high-rate slope as a desired demand (the red curve). Applying the mathematical model, the required temperature set point is depicted in green in this figure. Applying these values for the thermostat set points on a simulated population of ACs will result in the aggregated power demand that is shown by the solid blue curve. If we apply a controller to follow this desired demand, the thermostat set point and the aggregated power look like the dotted blue curves.

In Figure 10, the feasibility intervals of desired demand profiles with different slopes are compared. The graph indicates that as the ramp speed increases, the time during which the load can be controlled to align with the desired power consumption decreases. This phenomenon is attributed to the necessity of preserving customers' comfort as discussed earlier.

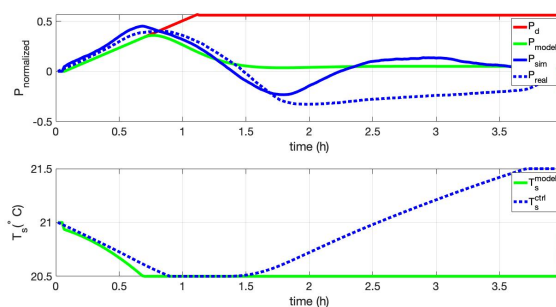


Figure 9: Feasibility of a high-speed ramp desired demand obtained from the model and simulation

Figure 11 illustrates an example of using a trapezoidal-shape as a desired demand (the red curve). Applying the mathematical model, the required temperature set point is depicted in green in this figure. Applying these values for the thermostat set points on a simulated population of ACs will result in the aggregated power demand that is shown by the solid blue curve. If we apply a controller to follow this desired demand, the thermostat set point and the aggregated power look like the dotted blue curves.

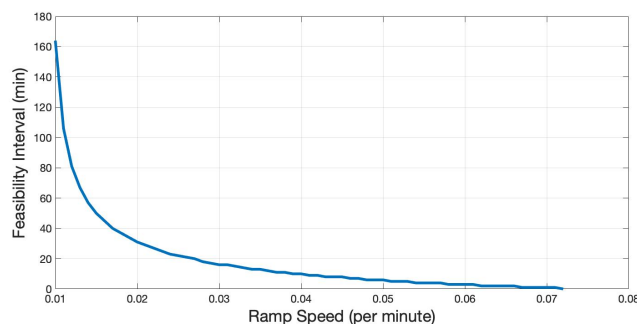


Figure 10: Feasibility intervals for ramp shape desired demand profiles with varying slopes

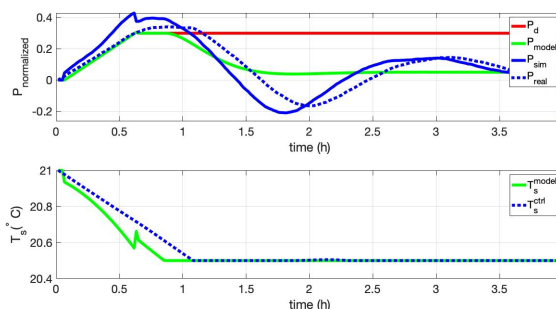


Figure 11: Feasibility of trapezoidal desired demand obtained from the model and simulation

An important aspect of the forecast module is to communicate this information to the system operator. With this knowledge, the operator can make informed decisions on how to generate the desired demand, taking into account the capacity of the controller to track the load.



Figure 12 visually demonstrates the impact on the feasibility interval when altering the slope and  $P_d^{max}$  values of the trapezoidal shape. Once more, the observation is evident: a higher slope (faster rate) and a larger maximum value lead to reduced controllability intervals in order to maintain customers' comfort.

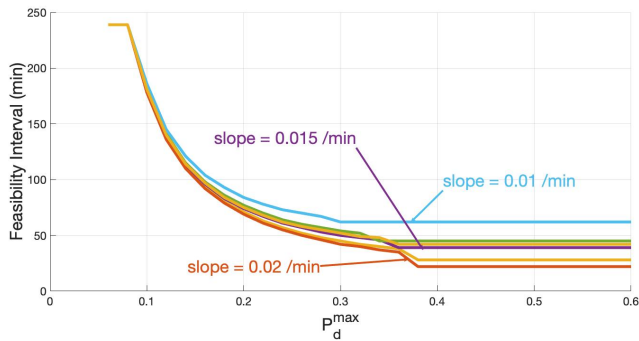


Figure 12: Feasibility intervals for trapezoidal shaped desired demand profiles with varying slopes and  $P_d^{max}$

## Conclusions

Thermostatically controlled devices have the potential to be utilized as controllable loads for direct load control, presenting a practical solution to address rapid fluctuations in power generation. The effective management of such loads requires precise predictions of future power consumption and the ability to adjust capacity relative to the baseline.

This study introduces an intelligent agent specifically designed to learn the behavior of a population of air conditioners. It provides a mathematical model for forecasting parameters such as baseline load, controllable load, and the impact of past and present control actions on the forecast. Moreover, the agent assesses potential alterations in the load. The proposed learning mechanism offers two significant advantages: 1) it eliminates the need for knowledge about physical parameters like thermal capacitance or resistance, and 2) it demonstrates adaptability to changing parameters over time. This makes it a practical and versatile plug-and-play load aggregator suitable for deployment in various locations and environments.

## References

- [1] P. Yazdkhasti, C. P. Diduch, "Smart Agent-Based Direct Load Control of Air Conditioner Populations in Demand Side Management," in 2023 International Conference on Control, Automation and Diagnosis (ICCAD), 1–7, 2023, doi:10.1109/ICCAD57653.2023.10152358.
- [2] M. Fan, K. Sun, D. Lane, W. Gu, Z. Li, F. Zhang, "A Novel Generation Rescheduling Algorithm to Improve Power System Reliability with High Renewable Energy Penetration," IEEE Transactions on Power Systems, **33**(3), 3349–3357, 2018, doi:10.1109/TPWRS.2018.2810642.
- [3] A. Halder, X. Geng, P. R. Kumar, L. Xie, "Architecture and algorithms for privacy preserving thermal inertial load management by a load serving entity," in 2017 IEEE Power Energy Society General Meeting, 1–1, 2017, doi:10.1109/PESGM.2017.8274310.
- [4] Y. Cheng, "Methods for mitigating the effects of intermittent energy production of photovoltaic sources," in 2011 International Conference on Power Engineering, Energy and Electrical Drives, 1–6, 2011, doi:10.1109/PowerEng.2011.6036564.

- [5] L. Hucheng, W. Xinwei, Y. Yubo, S. Wei, G. Yanfeng, "Simulation and analysis of electric water heater load regulation model based on direct load control," in 2017 IEEE Conference on Energy Internet and Energy System Integration (EI2), 1–5, 2017, doi:10.1109/EI2.2017.8245591.
- [6] M. Shaad, C. P. Diduch, M. E. Kaye, L. Chang, "A basic load following control strategy in a direct load control program," in 2015 IEEE Electrical Power and Energy Conference (EPEC), 339–343, 2015, doi:10.1109/EPEC.2015.7379973.
- [7] T. Mega, S. Kitagami, S. Kawawaki, N. Kushiro, "Experimental Evaluation of a Fast Demand Response System for Small/Medium-Scale Office Buildings," in 2017 31st International Conference on Advanced Information Networking and Applications Workshops (WAINA), 291–295, 2017, doi:10.1109/WAINA.2017.84.
- [8] M. Shad, A. Momeni, R. Errouissi, C. P. Diduch, M. E. Kaye, L. Chang, "Identification and Estimation for Electric Water Heaters in Direct Load Control Programs," IEEE Transactions on Smart Grid, **8**(2), 947–955, 2017, doi:10.1109/TSG.2015.2492950.
- [9] P. Yazdkhasti, C. P. Diduch, "A Mathematical Model for the Aggregated Power Consumptions of Air Conditioners," in 2018 the 6th international conference on smart energy grid engineering, 1 – 8, 2018.
- [10] S. Kundu, N. Sinitsyn, S. Backhaus, I. Hiskens, Modeling and control of thermostatically controlled loads, Power Systems Computation Conference, 2011.
- [11] W. Zhang, J. Lian, C. Y. Chang, K. Kalsi, "Aggregated Modeling and Control of Air Conditioning Loads for Demand Response," IEEE Transactions on Power Systems, **28**(4), 4655–4664, 2013, doi:10.1109/TPWRS.2013.2266121.
- [12] J. Mathieu, S. Koch, D. Callaway, "State estimation and control of electric loads to manage real-time energy imbalance," IEEE Transactions on Power Systems, **28**(1), 430–440, 2013, doi:10.1109/TPWRS.2012.2204074.
- [13] N. Mahdavi, J. H. Braslavsky, M. M. Seron, S. R. West, "Model Predictive Control of Distributed Air-Conditioning Loads to Compensate Fluctuations in Solar Power," IEEE Transactions on Smart Grid, **8**(6), 3055–3065, 2017, doi:10.1109/TSG.2017.2717447.
- [14] S. H. Tindemans, V. Trovato, G. Strbac, "Decentralized Control of Thermostatic Loads for Flexible Demand Response," IEEE Transactions on Control Systems Technology, **23**(5), 1685–1700, 2015, doi:10.1109/TCST.2014.2381163.
- [15] K. Meng, D. Wang, Z. Y. Dong, X. Gao, Y. Zheng, K. P. Wong, "Distributed control of thermostatically controlled loads in distribution network with high penetration of solar PV," CSEE Journal of Power and Energy Systems, **3**(1), 53–62, 2017, doi:10.17775/CSEEJPES.2017.0008.
- [16] C.-R. Chen, S.-C. Shih, S.-C. Hu, "Short-term Electricity Forecasting of Air-conditioners of Hospital Using Artificial Neural Networks," in 2005 IEEE/PES Transmission Distribution Conference Exposition: Asia and Pacific, 1–5, 2005, doi:10.1109/TDC.2005.1547136.
- [17] D.-X. Niu, Q. Wanq, J.-C. Li, "Short term load forecasting model using support vector machine based on artificial neural network," in 2005 International Conference on Machine Learning and Cybernetics, volume 7, 4260–4265 Vol. 7, 2005, doi:10.1109/ICMLC.2005.1527685.
- [18] S. Singh, S. Hussain, M. A. Bazaz, "Short term load forecasting using artificial neural network," in 2017 Fourth International Conference on Image Information Processing (ICIIP), 1–5, 2017, doi:10.1109/ICIIP.2017.8313703.
- [19] C. Perfumo, J. H. Braslavsky, J. K. Ward, "Model-Based Estimation of Energy Savings in Load Control Events for Thermostatically Controlled Loads," IEEE Transactions on Smart Grid, **5**(3), 1410–1420, 2014, doi:10.1109/TSG.2014.2298840.
- [20] R. Gao, X. Liu, "Short-term load forecasting based on least square support vector machine combined with fuzzy control," in Proceedings of the 10th World Congress on Intelligent Control and Automation, 1048–1051, 2012, doi:10.1109/WCICA.2012.6358034.

- [21] P. Yazdkhasti, C. P. Diduch, "A PID controller for Direct Load Control of Thermostatically Controlled Appliances," in 2019 IEEE 58th Conference on Decision and Control (CDC), 1913–1918, 2019.
- [22] N. Mahdavi, J. H. Braslavsky, C. Perfumo, "Mapping the Effect of Ambient Temperature on the Power Demand of Populations of Air Conditioners," *IEEE Transactions on Smart Grid*, **9**(3), 1540–1550, 2018, doi: 10.1109/TSG.2016.2592522.
- [23] C. Perfumo, E. Kofman, J. Braslavsky, J. Ward, "Load Management: Model-Based Control of Aggregate Power for Populations of Thermostatically Controlled Loads," *Energy Conversion and Management*, **55**, 36–48, 2012.
- [24] P. Yazdkhasti, C. P. Diduch, "A Model-Based Short-term Load Forecast Methodology for Aggregated Power Consumption of Thermostatically Controlled Appliances in DSM," in 2021 IEEE Energy Conversion Congress and Exposition (ECCE), 727–732, 2021, doi:10.1109/ECCE47101.2021.9595608.
- [25] L. Von Krannichfeldt, Y. Wang, G. Hug, "Online Ensemble Learning for Load Forecasting," *IEEE Transactions on Power Systems*, **36**(1), 545–548, 2021, doi:10.1109/TPWRS.2020.3036230.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license ( <https://creativecommons.org/licenses/by-sa/4.0/> ).

# Comparing Kalman Filter and Diffuse Kalman Filter on a GPS Signal with Noise

Maximo Giovanni Tandazo Espinoza\*

Universidad Politécnica Salesiana, Computer Science, Guayaquil, Ecuador

## ARTICLE INFO

### Article history:

Received: 21 November, 2023

Revised: 21 January, 2024

Accepted: 21 January, 2024

Online: 21 February, 2024

### Keywords:

Kalman Filter

Fuzzy Logic

Noise

Measurement

Filtered

## ABSTRACT

The navigation control of an autonomous vehicle can be determined by the coordinates of a GPS (Global Positioning System) positioning system, angular velocity, and acceleration with an INS (Inertial Navigation System). However, the errors associated with these devices do not allow it to be the only measurement system used in an autonomous vehicle. The need arises to implement tools that determine the system's state reliably at any instant and perform the necessary control actions to fulfill the trajectory optimally, considering the system's internal model. Therefore, applying a Diffuse Kalman filter is vital, allowing information integration from GPS and other devices. This work was divided into three essential parts such as the Kalman filter, the fuzzy control, and the simulation of a GPS sensor signal, taking into account that, in this last part, a comparison is made with the behavior of a Diffuse Kalman filter. In general, due to the comparisons of the position estimations in GPS measurements, it is evident that the DKF achieves more efficient reliability values since the position estimation error is reduced.

## 1. Introduction

The navigation control of an autonomous vehicle can be determined by the coordinates of a GPS (Global Positioning System), angular velocity, and acceleration with an INS (Inertial Navigation System). However, the errors associated with these devices do not allow it to be the only measurement system used in an autonomous vehicle. The need arises then to implement tools that will enable the system to reliably determine the state of the system at any instant and perform the necessary control actions to fulfill the trajectory optimally, considering the system's internal model. Therefore, applying a Diffuse Kalman filter (DKF) is critical at this stage since it allows information integration from GPS and other devices. This work was divided into 3 essential parts: the Kalman filter (KF), the fuzzy control, and the simulation of a GPS sensor signal, considering that in this last part, a comparison is made with the behavior of a Diffuse Kalman filter. With the position estimation in GPS measurements, high-reliability values are achieved. These variations depend very much on how the position measurements are taken to model the noise to which it is exposed. One of the advantages of the Kalman filter is that it avoids acquiring very accurate sensors since, if sensors with more accurate clocks and more advanced position measurement techniques were considered, the costs would rise considerably.

The Kalman filter originates in the paper "A New Approach

to Linear Filtering and Prediction Problems" published by Rudolf Emil Kalman in 1960, where he describes a recursive solution to the problem of linear filtering of discrete data [1]. Nowadays, it is a widely used method to optimally estimate the states of a dynamic system in real-time from the indirect noisy measurements that are taken from it [2]. These real-time estimates of the system state are valuable when operating in the open loop; however, if closed-loop operation is considered, they can be used to control and keep the vehicle in the desired direction.

Filtering is desirable in engineering and embedded systems situations since a good filtering algorithm can eliminate noise from the various eventualities to which the process to be monitored is exposed. There are likely fluctuations or disturbances caused by the environment or the existing sensor(s) characteristics. The Kalman filter is a method for estimating the variables of various processes. The Kalman filter estimates the states of a linear system in mathematical terms. The filter is an algorithm based on the state space model of a plan to assess the future shape and output by optimally filtering the output signal, and depending on the delay of the incoming samples, it can serve as an estimator or a filter but in both cases, it can eliminate noise. The filter applied to a model formulated in the state space allows a unified treatment of various aspects, such as the estimation of model parameters, the prediction of values, or the analysis of the system dynamics. This paper presents a review of the principal associated ideas. Among its applications are demographic

\*Corresponding Author: Maximo Giovanni Tandazo Espinoza, Universidad Politécnica Salesiana, Computer Science, Guayaquil, Ecuador, Email: [mtandazo@ups.edu.ec](mailto:mtandazo@ups.edu.ec)

estimation, signal processing, navigation systems, predicting the behavior of economic variables, and image processing, among others. Due to its wide field of action, it is essential to know how it works to have the basic tools that solve several practical problems simply and optimally [3].

## 2. Materials and Methods

### 2.1. The Kalman Filter

Rudolf Emil Kalman's 1960 work "A New Approach to Linear Filtering and Prediction Problems" contains the original description of the Kalman filter. It outlines a recursive solution to the linear filtering of discrete data problems. Recursive least squares estimate is at the heart of state-space models, which encompass the work he did [1].

The Kalman filter is an ideal state estimation method for dynamic systems with stochastic disturbances. More specifically, the Kalman filter provides a linear recursive algorithm that estimates the states of a dynamic system with discrete noise in real-time, ideally and with minimal variance error. Radar, ballistic missile trajectory calculation, satellite navigation, and video and laser tracking systems are just a few of the many industrial, civil, and military uses for it. Real-time applications of the Kalman filter have been developed in response to the swift advancement of faster computers [4]. The Kalman filter is also seen as an effective and versatile procedure to combine the output of sensors with noise to estimate the states of a system with uncertain dynamics.

In essence, this filter is a set of mathematical formulas that construct an optimal predictor-corrector type estimator, meaning that given certain conditions, it minimizes the estimated covariance error [5]. A variety of time series models can be handled with the state-space representation, which is a helpful notation for estimating stochastic models where mistakes in the system's measurement are assumed. Specific applications include the depiction of other models that need maximum likelihood approximation and the modeling of time-varying parameters and unobservable components. The filter is a mathematical process that uses a mechanism for correction and prediction. This algorithm predicts the next state based on its previous estimation by introducing a correction term proportionate to the prediction error and minimizing it statistically [5, 6, 7, 8].

The Kalman filter is the most popular linear estimation method for estimating the states and parameters of linear and nonlinear dynamic systems. Using a mathematical model of the plant and a measurement model of the sensor systems [9], the Kalman filter can anticipate the future values of any dynamic system. In attempting to collect the most accurate estimates of the system state, inaccurate data still frequently appears. According to [10], the user's driving style can alter the battery's state of charge (SOC), which is influenced by several factors, including temperature, internal resistance, capacity, and charge and discharge rates, among others. The Kalman filter is used to remove erroneous data due to these causes. The process consists of calculating the new state and its uncertainties.

Within the state-space notation, the derivation of the Kalman filter rests on the assumption of normality of the initial state vector and system perturbations in such a way that it is possible to calculate the likelihood function on the prediction error with which the

estimation of the unknown parameters of the system is carried out (see figure 1).

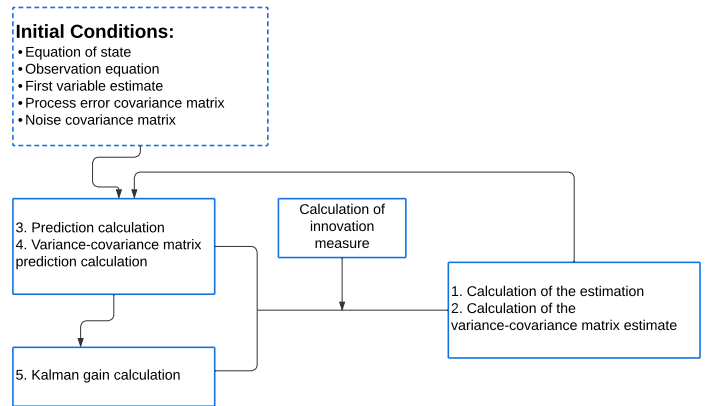


Figure 1: Kalman Algorithm Summary Diagram

### 2.2. Process estimation

The filter aims to solve the general problem of estimating the state  $X \in R^n$  of a discrete-time controlled process, which is governed by the linear equation of the stochastic differential equation as follows:

$$X_k = Ax_{k-1} + Bu_k + w_{k-1} \quad (1)$$

With a measurement  $z \in R^m$  that is:

$$Z_k = Hx_k + v_k \quad (2)$$

The random variables  $w_k$  and  $v_k$  represent the process and measurement error respectively. They are assumed to be independent of each other, white noise, and with normal probability distribution:

$$P(w) \approx N(0, Q) \quad (3)$$

$$P(v) \approx N(0, R) \quad (4)$$

Although they can be considered constant, the process noise covariance  $Q$  and the measurement noise covariance  $R$  are matrices that could vary with each time step or measurement in practice. Matrix  $A$  of size  $n \times n$  relates the state at the previous time step  $k-1$  to that at the current time step  $k$ . Matrix  $B$  of size  $n \times l$  relates the optional control input  $u \in R^l$  to the state  $x$ . Matrix  $H$  of size  $m \times n$  relates the state to the measurement  $z_k$ . In practice, these matrices may change each time, but they are usually assumed to be constant.

The KF estimates the process mentioned above using a feedback control. In other words, it evaluates the process at a particular moment and then uses the observed data to get feedback. Two sets of equations can be employed to obtain the filter: one updates the time or prediction equations, and the other updates the observed data or update equations. Those of the first group are responsible for the state projection at time  $k$ , taking as reference the state at time  $k-1$ , and for the intermediate update of the state covariance matrix. The second group of equations is responsible for the feedback, i.e., they incorporate new information into the previous estimate, resulting in an improved state estimate [11].



Forecasting equations can alternatively be defined as equations that update in real-time, while correction equations incorporate new information. A forecast-correction method for various issues can be used to characterize the final estimation procedure. To correct the projection using the new measurement and forecast the new state and its uncertainty, the Kalman filter employs a projection correction mechanism. This cycle is shown in Figure 2.

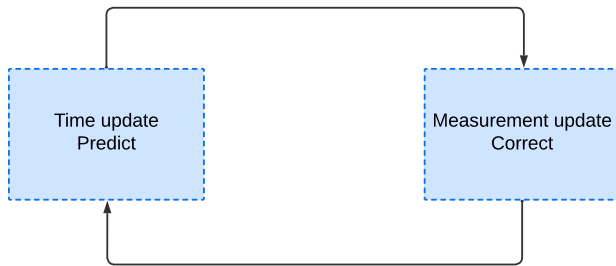


Figure 2: Kalman Filter Cycle

The first step consists of generating a forecast of the future state in time, taking into account all the information available at that time, and in the second step, an improved prediction of the state is generated so that the error is statistically minimized. The specific equations for forecasting and status correction are given below.

$$\hat{x}^- = A\hat{x}_{k-1} + Bu_k \quad (5)$$

$$P_x^- = AP_{k-1}A^T + Q \quad (6)$$

The equations in Table 2 forecast the state estimates and the forward covariance from k-1 to k. Matrix A relates the state at the previous time k-1 to the form at the current time k; this matrix could change for different points in time (k). Q represents the covariance of the random disturbance of the process trying to estimate the state.

$$K_k = P_k^- H^T (H P_k^- H^T + R)^{-1} \quad (7)$$

$$\hat{x}_k = \hat{x}_k^- + K_k (Z_k - H\hat{x}_k^-) \quad (8)$$

$$P_k = (I - K_k H) P_k^- \quad (9)$$

The first task during the correction of the state projection is calculating the Kalman gain,  $K_k$ , using equation (7). This weighting factor or growth is selected to minimize the error covariance of the new state estimate. The next step is to measure the process to obtain  $Z_k$  and then generate a new state estimate incorporating the new observation as in equation (8). The final step is to get a new assessment of the error covariance using equation (9).

After each pair of updates, both time and measurement, the process is repeated, taking the new state and error covariance estimates as a starting point. This recursive nature is one of the striking features of the Kalman filter [4, 7]. Figure 3 gives a complete picture of the operation of the filter, combining Figure 2 with equations (5), (6), (7), (8) and (9). The KF uses the least squares method to recursively generate an estimator of the state at time k, which is

linear, unbiased, and of minimum variance. The filter is in line with the Gauss-Markov theorem, giving the KF enormous power to solve a wide range of problems in statistical inference.

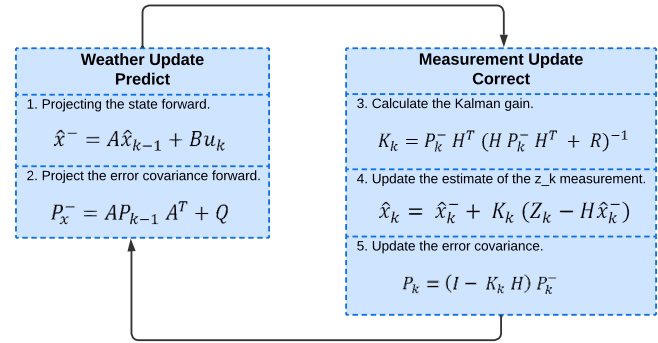


Figure 3: Kalman Filter Overview

It avoids the influence of possible structural changes in the estimation. Selective estimation starts with an initial sample and updates the estimates by successively incorporating new observations until all the information is covered. This means that the most recent assessment of the factors is influenced by the long history of the series, which, in the case of structural changes, could distort it. Sequential estimations can correct this bias, although doing so will increase the standard error. As a result, the KF employs the complete series history, much like recursive methods do. It does, however, have the benefit of removing the chance of estimation bias in structural modifications by aiming to estimate a stochastic route for the coefficients instead of a deterministic one.

The filter stands out because it can predict the state of a model in the past, present, and future, even when the modeled system's specific characteristics are unclear. One of the main distinguishing features of the Kalman approach is the dynamic modeling of a system. A linear transition from one period to the next characterizes linear dynamic models, which include most of the models frequently used in time series analysis.

One of the filter's drawbacks is that to initiate the recursive algorithm, initial conditions of the mean and variance of the state vector must be met as starting conditions. There is still disagreement about what these starting circumstances should be. For example, in a Bayesian approach, this filter requires a priori values of the initial coefficients and their respective variances to be specified. One way may be to obtain this information by estimating a model similar to the desired one but with fixed coefficients for a sample subperiod. On the other hand, it is necessary to specify the variances for which minimal and proportional conflicts are suggested about those obtained for the initial coefficients [12].

The development of the Kalman filter, as found in the original paper, assumes a broad knowledge of probability theory, specifically with the issue of Gaussian conditionality in random variables, which may cause a limitation for its study and application. When developed for autoregressive models, the results are conditioned to the past information of the variable in question. In this sense, time series forecasting represents the strength or inertia currently in the system and is efficient only in the short term.

### 2.3. Fuzzy Logic

Fuzzy logic (FL) is an extension of Boolean logic by Lotfi Zadeh in 1965 based on the mathematical theory of fuzzy sets, which is a generalization of classical set theory [13], has been developed basically in different disciplinary practices, especially in those related to industrial process control, the computer sector and numerous applications in economics [14, 15].

Fuzzy logic can be defined as:

- Mathematics that generalizes two-valued (0,1) logic for reasoning under uncertainty.
- Theories and technologies that employ fuzzy sets, which are set with boundaries based on a membership degree [16].

The first goal of fuzzy logic is to alleviate difficulties in developing and analyzing complex systems involving conventional mathematical tools [17]. It is also motivated by the assumption that human reasoning does not always have well-established boundaries. Fuzzy logic can be used to model and control complex and nonlinear systems or systems that need to be better defined for conventional modeling and control techniques. Fuzzy logic is a technology for developing intelligent control and information systems, as it offers a practical way to design nonlinear control systems. It achieves nonlinearity through a linear approximation of the system elements. The basic building block for fuzzy control systems is the if-then rule set, which performs a functional mapping [18].

Fuzzy logic techniques are based on four basic concepts:

- Fuzzy sets are those having smooth boundaries. Ambiguity.
- Linguistic variables are those whose values fit into a fuzzy set and can be qualitative and quantitative.
- Distribution of Possibilities: Assigning a fuzzy set imposes a linguistic variable's value limitation.
- A fuzzy if-then rule is a knowledge representation scheme for a functional mapping or a logical formula that generalizes an implication into two logical values.

Fuzzy Sets allow the elimination of fixed and exact constraints by using the membership of a set through its degree of membership. A group's membership degree is expressed by a number between zero and one, where zero means altogether outside the scene, one means entirely in the set, and a number between zero and one means partially inside the group. In this way, a gradual and smooth transition can be described from outside the location to inside the set. So, a fuzzy set is defined by a function that maps objects in a domain concerning their membership value in the group. It is important to remember that a fuzzy set is always expressed in a context, even if the context is not explicit [19].

Linguistic variables allow their value to be described qualitatively (linguistic term) and quantitatively (corresponding to a membership value). The linguistic term is used to express concepts and knowledge in human communication, while the membership function is helpful for input data processing. A linguistic variable is a composition of a symbolic and numeric variable. For situations with a very definite boundary between the possible and the impossible,

fuzzy logic offers an alternative, in which it generalizes the distinction between the possible and the impossible through a degree of possibility.

Fuzzy if-then rules have been applied to many disciplines, such as control systems, decision-making, pattern recognition, and system modeling. Conceptually, these rules generalize a logical inference called Modus Ponens, in which the inferred conclusions are modified by a degree of membership in which the antecedent is satisfied. Mathematically, it can be seen as an interpolation scheme because it allows the fusion of multiple fuzzy rules when all their conditions are comfortable to a certain degree [16].

This theory is based on sets of fuzzy or fuzzy numbers, which denote, in essence, groups of elements belonging with varying intensities or degrees to a specific category. It allows multiple levels between the extreme values of each interval, even with the opportunity to establish references of resemblance between the limits and their internal nuances [15].

The above contrasts with the ideal world posed by Classical Logic (CL), which is based on the membership or not of the elements to each category. Consequently, LD has a more remarkable resemblance to the reality of social phenomena, where expressions are used whose boundaries are not clearly defined, as in the case of the terms familiar, quickly, approximate, old, novice, warm, experienced, fleeting, firm, submissive, authoritarian, etc., making it possible to classify an object or phenomenon into several conceptual categories at the same time, depending on the scale used by the person making the judgment.

In this representation (Figure 4), the concept that is qualified in a fuzzy way is the linguistic variable, while the different values it takes or is associated with constitute the linguistic values. In addition, each linguistic value is, in turn, another fuzzy set, and the range of values that the linguistic variable can take is known as the Discourse Universe, Universal Set, or simply Domain (U).

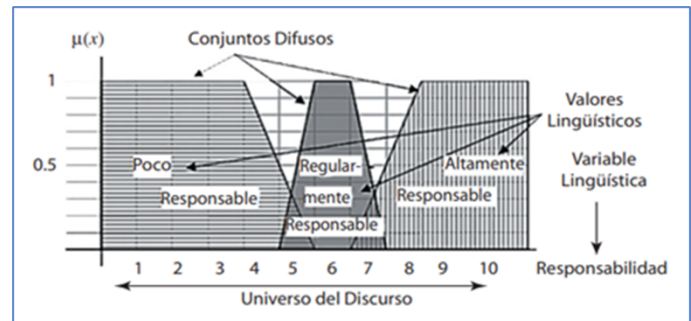


Figure 4: Belonging Functions

The Belonging Function ( $\mu$ ) assigns to each element of  $U$  a degree of membership or belonging to the fuzzy set, which is always in the interval  $[0, 1]$ ; if it takes the value one (1), it means that it fully complies with the condition or characteristic of the fuzzy set, while a value of zero (0) would be equivalent to saying that the element does not meet that condition.

The difference between FL and CL is that the former can establish degrees of membership to an element of the set, which implies the validity of partial membership (valuations between 0 and 1). At

the same time, in the latter, this is impossible since CL only allows total membership (value 1) or exclusion (value 0) to each category [20].

### 3. Results

#### 3.1. Diffuse Kalman Filter

Based on the discrete Kalman filter presented in the previous subsection, it is used to design a fuzzy control to manage the covariance matrix as it directly affects the filter performance. It has been determined that the main equations that model a Kalman Filter are given in table 1:

Table 1: Main Equations that model the Kalman Filter

Item	Equation	Description
1	$P_x^- = AP_{K-1}A^T + Q$	A priori value of the covariance of the estimated error.
2	$K_K = P_K^- H^T (HP_K^- H^T + R)^{-1}$	Calculation of the correction gain.
3	$P_K = (I - K_K H)P_K^-$	A posteriori value of the covariance of the estimated error.

It can be seen in Figure 5 that between the variables  $P_K^-$  (a priori value of the estimated error covariance) and  $P_K$  (a posteriori value of the estimated error covariance), there is a difference that serves as input to the fuzzy control, as proposed in [21]. The other input value to the system is  $P_K$ , which serves to know the covariance of the error of the state vector after correction and to monitor how it approaches zero.

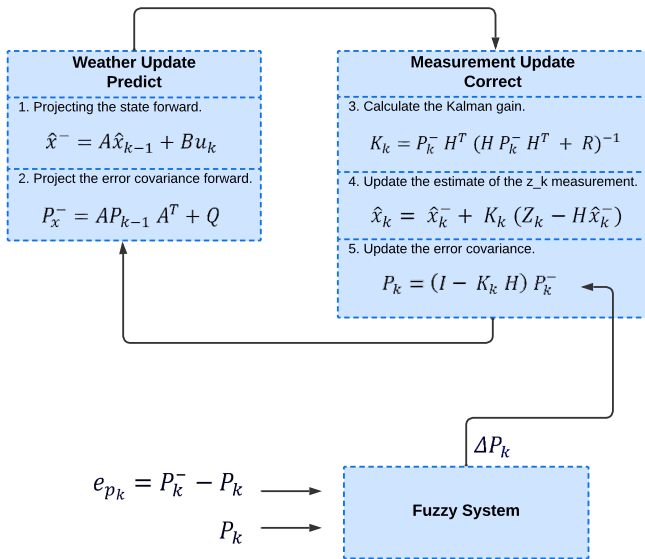


Figure 5: Operation of the Kalman Filter with Fuzzy control

The output of the fuzzy system is the value of the decrement/increment that is made to that same value of  $P_K$ . This forces

the system to approach a covariance error equal to zero faster, depending on the covariance's position and the error between the two covariance values (a priori and a posteriori). A general diagram of the filter operation is shown in Figure 5.

The fuzzy system rules were represented (Table 2) because the input  $P_k$  has the values of [Z, SP, LP] (Zero, Small Positive, and Large Positive), the input  $e_{P_k}$  has the values of [LN, SN, Z] (Large Negative, Small Negative and Zero), and the output  $\Delta P_k$  can take the values of [LN, SN, Z] (Large Negative, Small Negative and Zero).

Table 2: Values in the fuzzy rules.

Item	Inputs	Values	Description
1	$P_k$	[Z, SP, LP]	(Zero, Small Positive and Large Positive)
2	$e_{P_k}$	[LN, SN, Z]	(Large Negative, Small Negative and Zero)
3	$\Delta P_k$	[LN, SN, Z]	(Large Negative, Small Negative and Zero)

Figure 6 shows the rule matrix for the Diffuse Kalman Filter and figure 7 shows the diffuse fuzzy control with two inputs  $P_k, e_{P_k}$  and an output  $\Delta P_k$ .

		$e_{P_k}$		
		LN	SN	Z
$P_k$	Z	Z R1	Z R2	Z R3
	SP	Z R4	Z R5	SN R6
	LP	Z R7	SN R8	LN R9

Figure 6: Rule matrix for the Diffuse Kalman filter

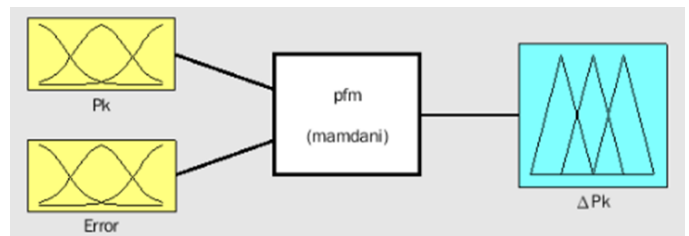


Figure 7: Diffuse Control

The following table 3 describes the rules to be used in fuzzy control:

Table 3: Values in the fuzzy rules.

Item	Rules
1	If ( $P_K$ is Z) and ( $e_{P_k}$ is LN) then ( $\Delta P_k$ is Z)
2	If ( $P_K$ is Z) and ( $e_{P_k}$ is SN) then ( $\Delta P_k$ is Z)
3	If ( $P_K$ is Z) and ( $e_{P_k}$ is Z) then ( $\Delta P_k$ is Z)
4	If ( $P_K$ is SP) and ( $e_{P_k}$ is LN) then ( $\Delta P_k$ is Z)
5	If ( $P_K$ is SP) and ( $e_{P_k}$ is SN) then ( $\Delta P_k$ is Z)
6	If ( $P_K$ is SP) and ( $e_{P_k}$ is Z) then ( $\Delta P_k$ is SN)
7	If ( $P_K$ is LP) and ( $e_{P_k}$ is LN) then ( $\Delta P_k$ is Z)
8	If ( $P_K$ is LP) and ( $e_{P_k}$ is SN) then ( $\Delta P_k$ is SN)
9	If ( $P_K$ is LP) and ( $e_{P_k}$ is Z) then ( $\Delta P_k$ is LN)

The rules are interpreted as follows: for the first rule, if  $P_k$  is close to zero and the value of  $e_{P_k}$  is large negative, meaning that the difference between  $P_k^- - P_k$  is large, it can be deduced that  $P_k$  is a value close to zero and is approaching quickly, so no value needs to be added or subtracted.

For the second rule if  $P_k$  is close to zero and the value of  $e_{P_k}$  is small negative, that is, the difference between  $P_k^- - P_k$  is small, it can be deduced that  $P_k$  is a value close to zero and is fast approaching, so no value needs to be added or subtracted.

For the third rule, if  $P_k$  is close to zero and the value of  $e_{P_k}$  is zero, meaning that the difference between  $P_k^- - P_k$  is zero, it can be deduced that  $P_k$  is a heat close to zero and is approaching fast, so no value needs to be added or subtracted.

For the fourth rule, if  $P_k$  is a small positive value and the value of  $e_{P_k}$  is a large negative, that means the difference between  $P_k^- - P_k$  is large, it can be deduced that  $P_k$  is a value close to zero, so no value needs to be added or subtracted from it.

For the fifth rule, if  $P_k$  is a small positive value and the value of  $e_{P_k}$  is a small negative, meaning that the difference between  $P_k^- - P_k$  is small, it can be deduced that  $P_k$  is a value close to zero, so no value needs to be added or subtracted.

For the sixth rule, if  $P_k$  is a small positive value and the value of  $e_{P_k}$  is zero, meaning that the difference between  $P_k^- - P_k$  is zero, it can be deduced that  $P_k$  is a value that is not close to zero, so a small value needs to be subtracted from  $P_k$  to make it try to approach zero faster.

For the seventh rule, if  $P_k$  is a large positive value and the value of  $e_{P_k}$  is a large negative, meaning that the difference between  $P_k^- - P_k$  is large, it can be deduced that  $P_k$  is a value close to zero, so no value needs to be added or subtracted from it.

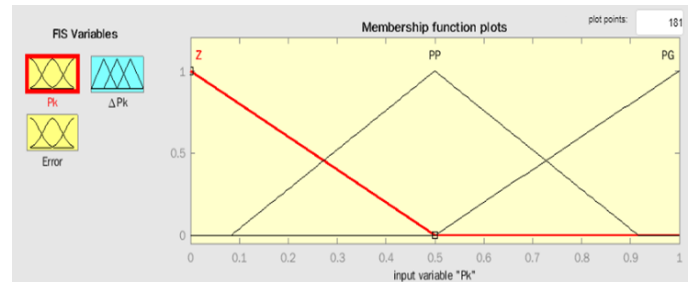
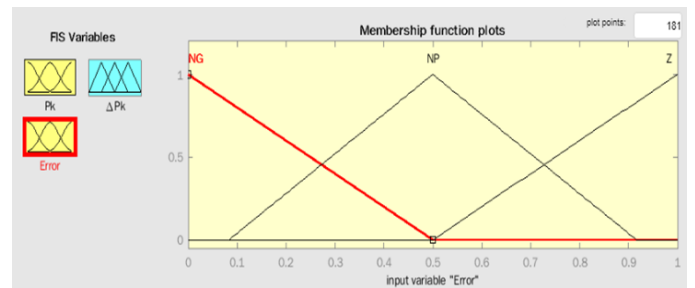
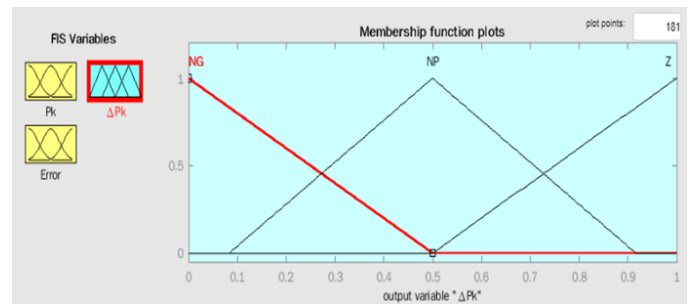
For the eighth rule, if  $P_k$  is a large positive value and the value of  $e_{P_k}$  is a small negative, meaning that the difference between  $P_k^- - P_k$  is small, it can be deduced that  $P_k$  is a value that is not close to zero, so no small value needs to be subtracted from it.

For the ninth rule, if  $P_k$  is a large positive value and the value of  $e_{P_k}$  is zero, this means that the difference between  $P_k^- - P_k$  is large, it can be deduced that  $P_k$  does not try to approach zero, so it is not necessary to subtract a large value from  $P_k$  to make it try to approach zero faster.

Figures 8, 9, and 10 show the membership functions used by the fuzzy control. Based on the above rules, it was decided to add a fuzzy part to the existing simulation, so the MATLAB Fuzzy Logic toolbox was used, with the following features:

- Fuzification : Singleton uncertainty.

- T-Norm : Minimal.
- Implication : Mamdani
- Defuzzification : Centroid


Figure 8:  $P_K$  Membership Function

Figure 9:  $e_{P_K}$  Membership Function

Figure 10:  $\Delta P_K$  Membership Function

### 3.2. Simulation

This section analyzes, compares, and interprets the simulations of the KF and the DKF with the characteristic values of a GPS sensor.

#### 3.2.1. Latitude

Figure 11 shows the comparison between the two simulations, showing that the error in the latitude processing of the DKF approaches a value closer to zero than the conventional KF.

Figure 12 shows the comparison of the prediction of the Kalman Filter versus the Diffuse Kalman filter on the latitude signal concerning time, and it can be seen that the FKF prediction fits better concerning the original signal with noise.



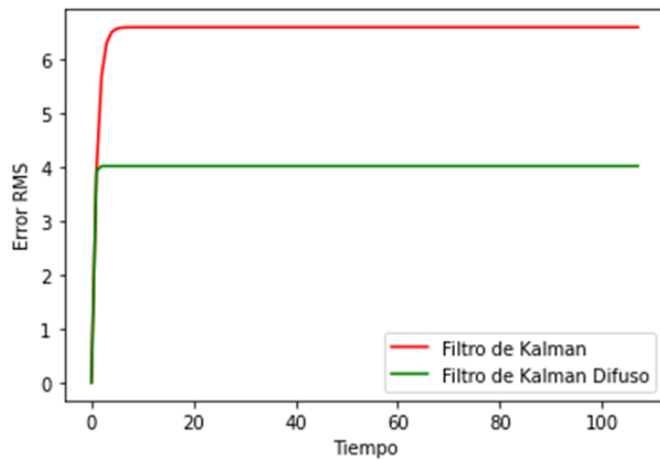


Figure 11: Comparison of Latitude Error (KF vs FKF)

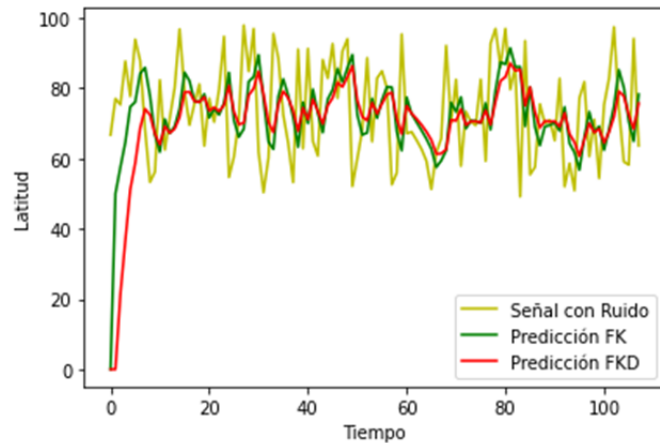


Figure 12: Comparison of Latitude Prediction (KF vs FKF)

Figure 13 shows the comparison of the Kalman Filter versus the Diffuse Kalman filter update on the latitude signal concerning time, and it can be seen that the FKF update fits better concerning the original signal with noise.

### 3.2.2. Longitude

Figure 14 shows the comparison between the two simulations of the conventional Kalman filter and the Diffuse Kalman filter, where it is noted that the longitude error in the Diffuse Kalman filter approaches a value closer to zero than the conventional Kalman filter.

Figure 15 shows the comparison of the prediction of the Kalman Filter versus the Diffuse Kalman filter on the longitude signal concerning time, and it can be seen that the FKF prediction fits better concerning the original signal with noise.

Figure 16 shows the comparison of the Kalman Filter versus Diffuse Kalman filter update on the longitude signal concerning time and it can be seen that the FKF update fits better concerning the original signal with noise.

### 3.2.3. Altitude

Figure 17 shows the comparison between the two simulations of the conventional Kalman filter and the Diffuse Kalman filter, where it is noted that the altitude error in the Diffuse Kalman filter approaches a value closer to zero than the conventional Kalman filter.

Figure 18 shows the comparison of the prediction of the Kalman Filter versus the Diffuse Kalman filter on the altitude signal concerning time and it can be seen that the FKF prediction fits better concerning the original signal with noise.

Figure 19 shows the comparison of the Kalman Filter versus Diffuse Kalman filter update on the altitude signal concerning time and it can be seen that the FKF update fits better concerning the original signal with noise.

## 4. Discussion

KF is a widely used filtering technique for predicting the state of a dynamic system considering noisy measurements. However, traditional KF is designed for linear and Gaussian systems, which limits its applicability to complex and nonlinear systems. The DKF is an extension of the KF that allows estimation in nonlinear and non-Gaussian systems by combining the Kalman Filter with fuzzy logic [22].

Fuzzy logic helps to represent uncertainty and imprecision in complex systems. DKF uses fuzzy sets instead of probability distributions to represent system states and measurements. This helps to handle noisy measurements better and capture uncertainty in nonlinear systems.

The advantage of the DKF is that it can handle complex, nonlinear systems, making it applicable in a wide range of areas, such as robotics, navigation, economics, and biomedicine. In addition, DKF provides more accurate and robust estimation than other nonlinear methods, such as particle filters. However, DKF also has some limitations. The main one is computational complexity, as the computation of fuzzy logic and updating estimates may require more computational resources than the traditional Kalman filter. In addition, proper selection of the fuzzy sets and DKF parameter settings can be challenging and require specialized knowledge.

DKF can handle noisy measurements and non-Gaussian error distributions by combining fuzzy logic with the Kalman filter. This provides a more accurate and robust estimation of the system state, which is crucial in applications where accuracy is essential, such as real-time navigation or complex process monitoring.

DKF can also handle smooth transitions between various models or regimes, making it more adaptable and flexible. This is particularly crucial in dynamic environments where systems face disturbances, condition changes, or unforeseen events.

Using the KF and DKF to treat the GPS signal proved an excellent option for reducing sensor error. However, there are much more efficient techniques, such as the particle filter, that effectively treat non-Gaussian noise and nonlinear models, especially in applications where the computational cost is cheap and the sampling rate is moderate, so this line of research can obtain essential contributions. Still, it must be addressed in the present work since it is not part of the stated objectives.

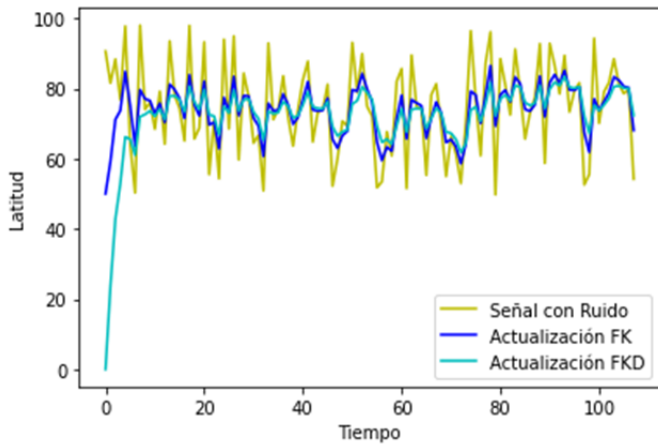


Figure 13: Latitude Updating Comparison (KF vs FKF)

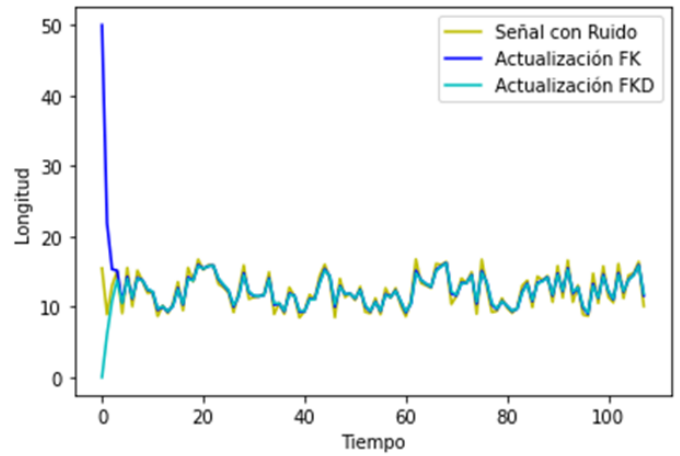


Figure 16: Comparison of longitude Updating (KF vs FKF)

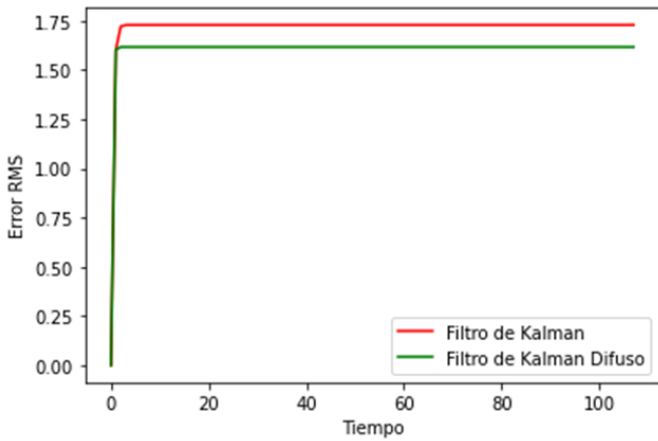


Figure 14: Comparison of longitude Error (KF vs FKF)

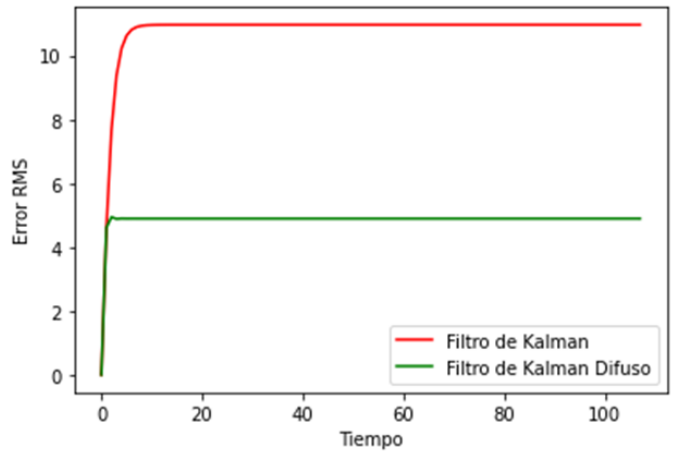


Figure 17: Comparison of Altitude Error (KF vs FKF)

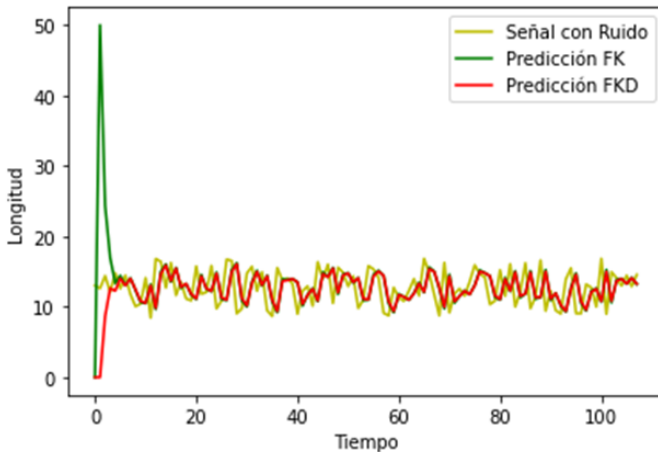


Figure 15: Comparison of longitude Prediction (KF vs FKF)

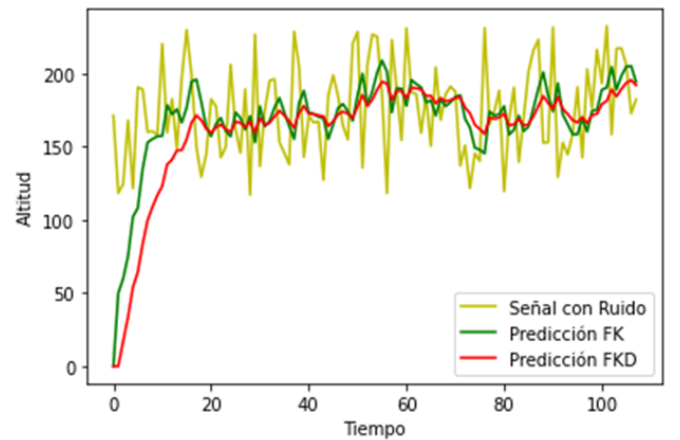


Figure 18: Comparison of Altitude Prediction (KF vs FKF)

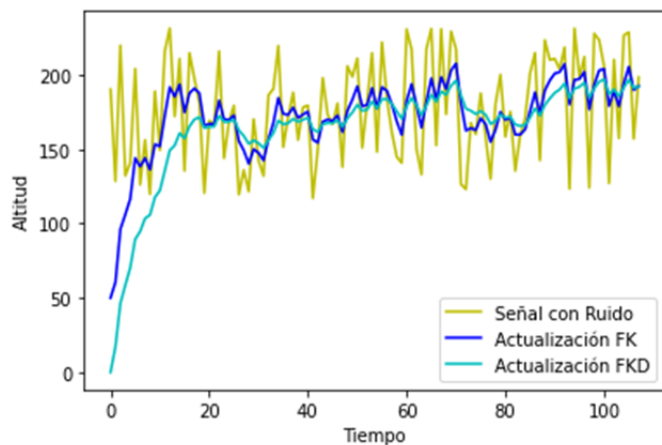


Figure 19: Comparison of Altitude Updating (KF vs FKF)

## 5. Conclusions

This work proposes using the fuzzy Kalman filter with a GPS sensor to maximize the accuracy in the prediction of position and velocity in an autonomous vehicle. This work was divided into 3 essential parts: the Kalman filter, the fuzzy control, and the simulation of the signal from a GPS sensor, taking into account that in this last part, a comparison is made in the behavior between the discrete Kalman filter and the Fuzzy Kalman filter.

In general, high-reliability values are achieved with position estimation in GPS measurements. The variations depend very much on how the position measurements are taken to model the noise to which it is exposed. One of the advantages of the Kalman filter is that it avoids the purchase of costly and accurate sensors since if sensors with more accurate clocks and more advanced position measurement techniques were considered, the costs would rise considerably.

Finally, the benefit of using techniques related to Fuzzy Logic in the design of fuzzy control is compelling, as it reduces the position estimation error. The fuzzy Kalman filter is worth studying because of its ability to handle complex and nonlinear systems, improve the robustness and accuracy of estimates, incorporate expert information, and adjust to model changes. Because of these qualities, the fuzzy Kalman filter is a valuable tool in many estimation problems. A comparison and analysis of the Fuzzy Kalman Filter, the Extended Kalman Filter, and the Particle Filter is suggested for future work.

## References

- [1] R. E. Kalman, "A New Approach to Linear Filtering and Prediction Problems." ASME. J. Basic Eng., 82(1): 35–45, 1960, doi:10.1115/1.3662552.
- [2] E. A. Wan, R. van der Merwe, "The unscented Kalman filter for nonlinear estimation," Proceedings of the IEEE 2000 Adaptive Systems for Signal Processing, Communications, and Control Symposium (Cat. No.00EX373), 153–158, 2000.
- [3] D.-J. Jwo, S.-Y. Lai, "Navigation Integration Using the Fuzzy Strong Tracking Unscented Kalman Filter," The Journal of Navigation, 62(2), 303–322, 2009, copyright - Copyright © The Royal Institute of Navigation 2009; Última actualización - 2023-11-28.

- [4] C. K. Chui, G. Chen, "Kalman Filtering with Real-Time Applications," Springer Berlin, Heidelberg, XV, 191, 2009, doi:10.1007/978-3-662-02508-6.
- [5] J. T. Jalles, "Structural Time Series Models and the Kalman Filter: A Concise Review," SSRN, 541, 153–158, 2009, doi:10.2139/ssrn.1496864.
- [6] G. Bishop, G. Welch, et al., "An introduction to the kalman filter," Proc of SIGGRAPH, Course, 8(27599-23175), 41, 2001.
- [7] R. J. Meinhold, N. D. Singpurwalla, "Understanding the kalman filter," Am. Stat., vol. 37(No. 2), 123–127, 1983.
- [8] H. W. Sorenson, "Least-squares estimation: from Gauss to Kalman," IEEE Spectrum, 7(7), 63–68, 1970, doi:10.1109/MSPEC.1970.5213471.
- [9] B. Yildirim, Sigmarho Kalman Filter Implementation and Analysis, Ph.D. thesis, California State University, 2020, copyright - Database copyright ProQuest LLC; ProQuest does not claim copyright in the individual underlying works; Última actualización - 2023-06-21.
- [10] A. Khalid, R. K. Syed Abdul, N. U. Ain, M. Awais, A. S. Majid, J. E. M. Carreño, J. C. Vasquez, J. M. Guerrero, B. Khan, "Comparison of Kalman Filters for State Estimation Based on Computational Complexity of Li-Ion Cells," Energies, 16(6), 2710, 2023.
- [11] M. Grewal, A. Andrews, "Kalman filtering: theory and practice using MATLAB," New York: John Wiley and Sons, 14, 2001, doi:10.1002/9780470377819.
- [12] R. L. Thomas Doan, C. Sims, "Forecasting and conditional projection using realistic prior distributions," Econometric Reviews, 3(1), 1–100, 1984, doi:10.1080/07474938408800053.
- [13] J. A. Goguen, "L. A. Zadeh. Fuzzy sets. Information and control, vol. 8 (1965), pp. 338–353. - L. A. Zadeh. Similarity relations and fuzzy orderings. Information sciences, vol. 3 (1971), pp. 177–200." The Journal of Symbolic Logic, 38(4), 656–657, 1973, doi:10.2307/2272014.
- [14] T. J. Ross, "Fuzzy Logic with Engineering Applications," Wiley, 3, 1–100, 2010, doi:10.1002/9781119994374.
- [15] I. P. B. Leon, "Logica difusa para principiantes. Teoria y practica," Publicaciones UCAB, 3, 2007.
- [16] J. Yen, R. Langari, "Fuzzy logic: Intelligence, control, and information," Prentice Hall, Upper Saddle River NJ, 1999.
- [17] C. Ni, X. Ma, "An integrated long-short term memory algorithm for predicting polar westerlies wave height," Ocean Engineering, 215, 107715, 2020, doi:https://doi.org/10.1016/j.oceaneng.2020.107715.
- [18] C. Lu, "Concrete materials compressive strength using soft computing techniques," Multiscale and Multidisciplinary Modeling, Experiments and Design, 2023, doi:10.1007/s41939-023-00276-4.
- [19] R. Srivastava, M. P. S. Bhatia, "Quantifying modified opinion strength: A fuzzy inference system for Sentiment Analysis," in 2013 International Conference on Advances in Computing, Communications and Informatics (ICACCI), 1512–1519, 2013, doi:10.1109/ICACCI.2013.6637404.
- [20] M. G. Voskoglou, "Fuzzy Logic as a Tool for Assessing Students' Knowledge and Skills," Education Sciences, 3(2), 208–221, 2013, doi:10.3390/educsci3020208.
- [21] R. Woo, E.-J. Yang, D.-W. Seo, "A Fuzzy-Innovation-Based Adaptive Kalman Filter for Enhanced Vehicle Positioning in Dense Urban Environments," Sensors, 19(5), 2019, doi:10.3390/s19051142.
- [22] B. Parrell, V. Ramanarayanan, S. Nagarajan, J. Houde, "The FACTS model of speech motor control: fusing state estimation and task-based control," bioRxiv, 2019, doi:10.1101/543728.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

# A Novel Metric for Evaluating the Stability of XAI Explanations

Falko Gawantka<sup>\*1</sup>, Franz Just<sup>1</sup>, Marina Savelyeva<sup>1</sup>, Markus Wappler<sup>2</sup>, Jörg Lässig<sup>1,2</sup>

<sup>1</sup>University of Applied Sciences Zittau/Görlitz, Faculty of Electrical Engineering and Computer Science, Görlitz, 02826, Germany

<sup>2</sup>Fraunhofer IOSB, Advanced System Technology (AST), Görlitz, 02826, Germany

## ARTICLE INFO

Article history:

Received: 15 November, 2023

Revised: 16 January, 2024

Accepted: 17 January, 2024

Online: 21 February, 2024

Keywords:

eXplainable AI

Evaluation

Stability of explanations

## ABSTRACT

Automated systems are increasingly exerting influence on our lives, evident in scenarios like AI-driven candidate screening for jobs or loan applications. These scenarios often rely on eXplainable Artificial Intelligence (XAI) algorithms to meet legal requirements and provide understandable insights into critical processes. However, a significant challenge arises when some XAI methods lack determinism, resulting in the generation of different explanations for identical inputs (i.e., the same data instances and prediction model). The question of explanation stability becomes paramount in such cases. In this study, we introduce two intuitive methods for assessing the stability of XAI algorithms. A taxonomy was developed to categorize the evaluation criteria and the ideas were expanded to create an objective metric to classify the XAI algorithms based on their explanation stability.

## 1 Introduction

As artificial intelligence (AI) becomes increasingly integrated into people's lives, a notable trend can be observed: automation of processes and domains that significantly impact the well-being and future prospects of individuals. For example, the initial screening of job applicants can be considered. Given the substantial human resource expenditure involved in this process, a compelling case can be made for its automation. Candidate scores can be calculated using artificial intelligence, as demonstrated in previous research [1]. Refer to Figure 1 for more details. Nevertheless, the decision-making process utilized by the underlying machine learning model presents a challenge. The accumulation of applicant information often results in a lack of transparency, a concern shared by HR managers, companies, applicants, and regulators alike.

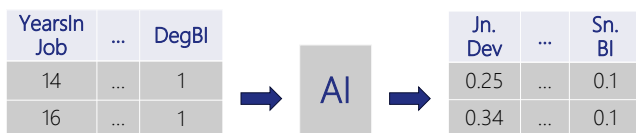


Figure 1: The job application pipeline is depicted here. On the left, the input values predicted by the ML model are displayed, while on the right, the corresponding results are presented. Transparency issues arise when the information density is reduced to a single numeric value. This lack of transparency makes it neither intuitive nor comprehensible for the stakeholders involved. The question of: „What impact does feature  $x$  had on the prediction?“ arise.

<sup>\*</sup> Corresponding Author: University of Applied Sciences Zittau/Görlitz, Faculty of Electrical Engineering and Computer Science, Görlitz, 02826, Germany, Email: falko.gawantka@hszg.de

<sup>1</sup>This domains could harm people or the future of a person, for instance here the HR use case represents an automated decision that has an impact on a person.

One viable approach to address this concern is the integration of explanation algorithms, often referred to as XAI algorithms. This strategy enables visualization of model decisions in an understandable way. Additionally, both the transparency and the validity of the predictive model can be improved, ensuring that the needs and expectations of the relevant stakeholders are met.

However, it is essential to recognize that not all explanatory algorithms can be used without restrictions in high-stake areas <sup>1</sup>. This journal paper extends the work originally presented in IEEE ICCI\*CC'22 [1]. The stability of the explanations was investigated through different XAI approaches and ranked according to their results. This was achieved by quantifying the feature importance values (FIVs). In this work, the state of the art for the evaluation of XAI explanations is presented, and the taxonomies found are expanded or modified in terms of the aspect: *stability*. The corresponding research questions are as follows.

- RQ1: What does the *stability* of an explanation mean?
- RQ2: How can the explanation of an XAI algorithm be measured in terms of *stability*?

By that, the following hypothesis can be stated: „High *stability* of an explanation is observed when the explanation undergoes minimal changes in response to minor variations in unimportant features of the data instance.“



For example, in the case of an HR applicant selection process, a change in the input features (i.e. 14 vs 16 *YearsInJob*) should not produce a substantially different explanation.

The paper is organized as follows. In section 2, metrics and evaluation techniques for XAI algorithms are presented, and the foundations of XAI and representative methods are discussed in detail. In section 3, an approach to quantify explanation stability is introduced. The details and limitations of previous approaches serve as a basis for a new scoring metric. The results are shown in section 4. Potential future research directions are outlined in the discussion in section 5.

## 2 Related Work and Background

For a deeper understanding and examination of XAI algorithms, an overview of evaluation taxonomies, methodologies, and the underlying calculation models is provided, forming the theoretical foundations. In addition to that, some background information are provided to enhance the understanding of XAI. This work aims to be considered as best practices in the selection process of XAI algorithms for real-world use cases, with the objective of uncovering a comprehensive understanding of the model with high *stability* to the respective explanation of XAI.

### 2.1 Related Work on XAI evaluation

To explore the stability of an explanation result, this subsection first defines the relevant properties/criteria. It then provides an overview of the measuring techniques and presents the mathematical foundations.

#### 2.1.1 Notions by different Taxonomies

In the work of [2], a comprehensive taxonomy is introduced, a distinction is made between *the user aspect*, *the explanation aspect* and *the model aspect*. The focus of this work lies on *the explanation aspect*. In this aspect there are two further sub-criteria defined by [2], which have to do with stability: *Identity* and *Separability*. *Identity* is defined by: “identical instances should have identical explanations” and *Separability* is defined by: “non-identical instances should not have identical explanations” [2]. To describe these two aspects, one can think of the ends of a scale.

A paper cited by [2] was that of [3] that defines *Stability* as: It represents how similar are the explanations for similar instances. While consistency compares explanations between different models, stability compares explanations between similar instances for a fixed model. High stability means that slight variations in the feature values of an instance do not substantially change the explanation, unless these slight variations also strongly change the prediction.

Nonetheless, a lack of stability can also be created by non-deterministic components of the explanation method, such as a data sampling step. Regardless of that, high stability is always desirable. This definition brings another dimension into play. The author in [2] speaks only of “identical instances”, and [3] also mentioned the predictor by constructing a context around the instance and saying that there is also a predictor and an instance.

Following the definitions in [2], the author in [4] describes in the work the so-called *Co-12 properties*. The *Consistency* property coincides with the *Identity* criterion from [2]. *Consistency* is defined as: “Identical inputs should have identical explanations” [4]. The term inputs can be defined by the data instance as well as the ML model. In this context, the model’s perspective is integrated, with the rest remaining consistent with the definition in [2]. And there is also a second aspect in [4], which can be equated with *Stability* in [3]. The term *Continuity* introduced by [4] is defined as follows: “Similar inputs should have similar explanations” [4]. The use of similar inputs could mean both the model view and the data view.

Besides the *Stability* notion by [3] and the term *Continuity* by [4] there is another work [5], where the author measured the *Stability* of XAI algorithms. The underlying assumption was that: If the inputs are almost identical, then it is expected that only minor changes appear in the explanation. This can be used as a scoring metric for the stability of an XAI algorithm. The term *almost identical* refers hereby to *similar* and therefore matches the understanding of *Stability* by [3] and the notion *Continuity* from the work of [4]. The author in [5] gives no explicit definition, but: “similar instances and an identical model should have similar explanations”, could be suggested as a common understanding for the property of *Stability*. Another property or criterion investigated in [1] was the *Reproducibility/Stability of the explanation*. This refers to an aspect in which an identical model predicts an identical data instance several times (i.e., 10 times, 100 times, and 1000 times). In the context of the reproducibility of explanation stability, *Reliability* would be a more appropriate term than the one chosen by the author.

The entire notions of the taxonomies are grouped and summarized on the *Stability Scale* and could be seen in Figure 2.

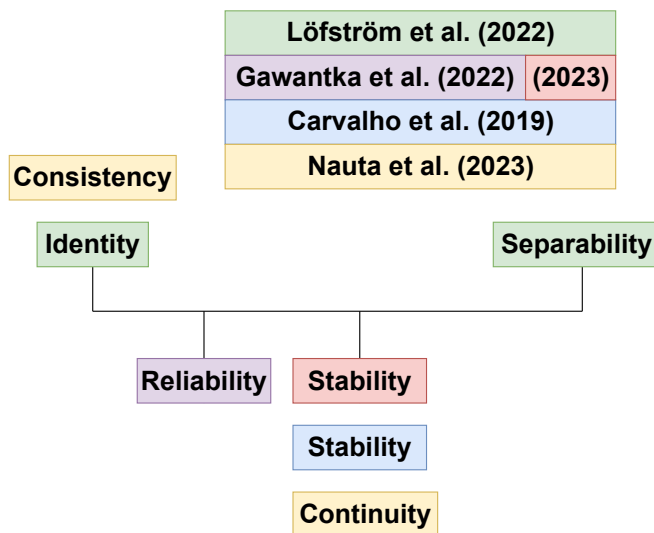


Figure 2: Stability Scale: The scale visualize the different terms and group them on the scale to show what terms could be used interchangeably.

The focus of this work lies in unifying different schools and terms, proposing concrete measuring techniques, and establishing a strong foundation from the mathematical perspective of XAI evaluation metrics.

### 2.1.2 Quantifying XAI through Feature Importance Values

A formal introduction into machine learning is initiated, drawing inspiration from a more detailed section in [6]. A comprehensive introduction to artificial intelligence is available in the textbook [7]. The notion of explanation method is taken and adapted from [8].

Feature importance values are used to quantify the explanations. These are numerical scores that help to evaluate the relative significance of individual features of the model's predictions, thus offering transparency and interpretability. It is important to note that the choice of the feature importance method matters and that different methods may lead to varying results. Therefore, careful consideration is needed to select the most appropriate method for a particular application. Adopting the notions presented in subsection 2.1.1 to the feature importance values will provide evaluation metrics to support the selection of data-driven methods.

In the context of supervised learning, XAI and the feature importance values will be considered. This is based on two fundamental components: the model and the data. Given an input domain  $X$  and an output domain  $Y$ , supervised learning aims to predict an output  $y \in Y$  for given  $x \in X$ . In the example of classifying job applicants, the input domain  $X$  reflects the data structure of the applicant profile, while  $Y$  encodes the job profiles. The prediction mechanism is described through a mathematical model that can be deterministic or probabilistic.

A *deterministic model* is a function  $f : X \rightarrow Y$  that predicts exactly for a given input  $x \in X$  an output  $f(x) \in Y$ . This can be a classification prediction, in case  $Y$  is a discrete set, or this can be a regression prediction, if  $Y$  is a continuous set. For example  $f$  assigns a job profile  $y \in Y$  to each applicant for a job  $x \in X$ .

A *probabilistic model* is a probability distribution  $p : X \times Y \rightarrow [0, 1]$ . It can be recognized as a full distribution, giving the probability of assignment of  $x \in X$  and  $y \in Y$  by  $p(x, y)$ , or it can be recognized as a conditional distribution, providing the probability of output  $y \in Y$  for chosen  $x \in X$  by  $p(y|x)$ . In the latter case for example,  $p$  would encode the probability that an applicant's profile  $x \in X$  will be classified as job profile  $y \in Y$ .

In either deterministic model or probabilistic model, the underlying function is usually parameterized. Training the model means optimizing the parameters to achieve good predictions, based on a data set  $\{(x^1, y^1), \dots, (x^m, y^m)\} \subseteq X \times Y$ .

The input data are commonly structured data  $X = X_1 \times \dots \times X_n$  with the notion of a *feature* being applied context-dependent to any primitive domain  $X_k$  or to any data entry  $x_l \in X_k$  or instance  $x \in X$ . The output  $y \in Y$  is called the label.

In the remainder of the paper, all models are considered to be deterministic. The set of all (deterministic) models is denoted by  $\mathcal{F} = X \rightarrow Y$ .

Having a (trained) model  $f$ , an explanation method  $\Phi : \mathcal{F} \times (X = (X_1, \dots, X_n)) \rightarrow \mathbb{R}^n$  attributes an importance score  $\Phi(f, x)_i$  to each feature  $x_i$  describing its impact on prediction  $f(x)$ . The range of feature importance values is method-dependent. Common choices are  $(-\infty, \infty)$  for SHAP,  $(0, 1)$  for Gini importance,  $(0, \infty)$  for Gain,  $(-1, 1)$  for Correlation Coefficients.

Using this notion, *Consistency* means that  $\Phi(f, x)$  evaluates to the same value, regardless of how often it is calculated. Obviously, this is always fulfilled, making the term irrelevant in the context of

feature importance values. But still, a similar metric might be useful, as certain XAI methods, e.g. surrogate explainer methods, train a model and try to approximate the original ML model, e.g. LIME is such a method. In fact, each LIME instance  $\Phi(f, x)$  might behave differently depending on the training data points used. Demanding that explanations at least do not deviate much would be a natural extension of *Consistency*.

Thus, in [1], another evaluation criterion known as *Reliability* is mentioned here. For every call of  $\Phi$  an element is added to the result set  $\{\Phi^{(r)}(f, \cdot)\}_r$ . One can expect that a statistical quantity that measures the deviation, such as the length of confidence intervals, will decrease with increasing  $r$ . However, its rate of decrease provides a measure for *Reliability*.

The definition of *Reliability* is shown in the following equation:

$$\{\Phi^{(r)}(f, \cdot)\}_r = \left\{ \begin{pmatrix} fiv_{i=1} \\ \vdots \\ fiv_{i=s} \end{pmatrix}_{r=1}, \dots, \begin{pmatrix} fiv_{i=1} \\ \vdots \\ fiv_{i=s} \end{pmatrix}_{r=c} \right\} \quad (1)$$

In Equation 1 the result of a single call from  $\Phi$  is shown as vector. For every call  $r$  an explanation vector ( $\vec{F}$ ) is calculated, so  $fiv_i \in \vec{F}$ . The result is a set of vectors that represents a matrix.

Based on this aspect, the mean over the 1st feature importance value from  $r$ -repetitions could be defined by:

$$\mu_{F_{(i=1)}} = \frac{1}{c} \sum_{r=1}^c fiv_{(i=1)r} \quad (2)$$

Therefore the variance could be calculated by the mean:

$$\sigma_{F_{(i=1)}}^2 = \frac{1}{c} \sum_{r=1}^c (fiv_{(i=1)r} - \mu_{F_{(i=1)}})^2 \quad (3)$$

The standard deviation for a single feature importance value is then calculated from:

$$\sigma_{F_{(i=1)}} = \sqrt{\sigma_{F_{(i=1)}}^2} \quad (4)$$

Finally, the deviation can establish the foundation for the confidence interval of a single feature importance value.

$$CI_{F_{(i=1)}} = \mu_{F_{(i=1)}} \pm (z = 1.96) \cdot \frac{\sigma_{F_{(i=1)}}}{\sqrt{c}} \quad (5)$$

In the previous equations, an approach was presented to calculate the confidence interval of the single feature importance value of the feature importance vector  $\vec{F}$ . When this is done over the entire feature importance vector and then the mean is built on top of it, this matches the metric from [1]. This calculation process is denoted by  $R(x)$ .

The intuitive definition of *Stability* involves small perturbations of an instance that result in only minor changes to the explanations under the same model. To be numerically tangible this definition has to be measurable, hence the notion of correlated small changes needs to be quantified. In [9] the explanation methods SHAP and LIME are evaluated in terms of Lipschitz continuity, using the best Lipschitz constant as a stability measure. This approach is generalized while keeping it local (depending on the chosen data point). Furthermore, it will be demonstrated that the different stability measures previously used fall under this umbrella.

Lipschitz continuity can be defined in a general sense for metric spaces as follows: A function  $f : X \rightarrow Y$  between metric spaces  $(X, d_X)$  and  $(Y, d_Y)$  is called *Lipschitz continuous* if there is a real constant  $L \geq 0$  (referred to as *Lipschitz constant*) such that for all  $x, \tilde{x} \in X$ ,

$$d_Y(f(x), f(\tilde{x})) \leq L d_X(x, \tilde{x}). \quad (6)$$

This is, the distance of the images of  $f$  is bounded by the distance of the respective arguments up to a constant factor. In case of local explanation methods, e.g. as SHAP and LIME are, an instance-independent constant is not attainable. Therefore, we will base the definition on a data point-dependent Lipschitz factor within a neighborhood of the data point.

Let  $X$  be an input domain, equipped with the metric  $d_X$ . Let  $Y$  be an output domain equipped with the metric  $d_Y$ . Let  $f \in \mathcal{F} = X \rightarrow Y$  be a model, and  $\Phi : \mathcal{F} \times X \rightarrow X$  be an explanation method. An explanation  $\Phi(f, x)$  can be termed  $\varepsilon$ -stable if there is a local real constant  $L(x) \geq 0$  such that

$$d_Y(\Phi(f, x), \Phi(f, \tilde{x})) \leq L(x) d_X(x, \tilde{x}) \quad (7)$$

for all  $\tilde{x} \in N_\varepsilon(x) = \{\tilde{x} \in X : d_X(x, \tilde{x}) \leq \varepsilon\}$ . The Lipschitz factor  $L(x)$  can be seen as a stability measure. The lower  $L(x)$ , the smaller the deviation, the higher the stability. Finding small Lipschitz factors can be done using the Monte Carlo method. Examples from the literature will be provided that specify the notion of stability.

**Example 2.1.** In [9] the metrics are induced by norms, resulting in  $N_\varepsilon(x) = \{\tilde{x} \in X : \|x - \tilde{x}\| \leq \varepsilon\}$ . The best Lipschitz constant is given by

$$L(x) = \max_{\tilde{x} \in N_\varepsilon(x) \setminus \{x\}} \frac{\|\Phi(f, x) - \Phi(f, \tilde{x})\|}{\|x - \tilde{x}\|}. \quad (8)$$

An approximation can be done by sampling with  $\{x_1, \dots, x_n\} \subset N_\varepsilon(x)$ .

**Example 2.2.** In [10] an evaluation metric for XAI methods called Max-Sensitivity was introduced. The concept is taken up by [8]. To derive the term, the inequality is revisited (7) and relaxed to

$$d_Y(\Phi(f, x), \Phi(f, \tilde{x})) \leq \varepsilon L(x) = L'(x). \quad (9)$$

Now, again using norms as distance metrics,  $L'(x)$  can be obtained by

$$L'(x) = \max_{\tilde{x} : \|x - \tilde{x}\|_\infty \leq \varepsilon} \|\Phi(f, x) - \Phi(f, \tilde{x})\|_2. \quad (10)$$

This is in fact the Max-Sensitivity metric. Again, it can be approximated by sampling.

**Example 2.3.** The work [5] considers a similarity measure by first generating data points by perturbing a feature around 1% and then evaluating the ratio of original feature importance value and perturbed feature importance value for that feature. Furthermore, the feature importance values are normalized so that they form a probability distribution.

Modifying this approach by allowing perturbations of at least  $\varepsilon$ , which bounds the values around one percent, allows to extend this measure to

$$S_i(x) = \max_{\tilde{x} \in N_\varepsilon(x)} \left| 1 - \frac{\Phi(f, x)_i}{\Phi(f, \tilde{x})_i} \right|. \quad (11)$$

Therefore,  $S_i(x)$  is the best similarity to the feature  $i$  that all  $\tilde{x} \in N_\varepsilon(x)$  fulfill. As a combined score for all features could be

$$S(x) = \max_{\tilde{x} \in N_\varepsilon(x)} \frac{\|\Phi(f, x) - \Phi(f, \tilde{x})\|_\infty}{\|\Phi(f, \tilde{x})\|_\infty}. \quad (12)$$

On one hand, this is similar to Lipschitz-based stability criteria; on the other hand, it effectively demonstrates that the criteria proposed in [5] normalize the feature importance values.

## 2.2 Background of XAI Taxonomies

At this point, eXplainable Artificial Intelligence is introduced as a field of study within the domain of artificial intelligence, along with an explanation of its associated focuses and terminology.

In [11, p. 3511] the author defines XAI as „the study of explainability and transparency for socio-technical systems, including AI.“ The 2019 XAI Taxonomy by [12] introduces important terms in explainable artificial intelligence. Later, in [13] used this taxonomy as a decision-making tool to choose the appropriate explanatory algorithms for the IBM AIX 360 tool. The tree structure of the taxonomy of [13] was transformed into a tabular format and can be seen in Table 1. The focus of this work is to provide an overview of explanatory models, which is highlighted in the table. The *data* leaf is not shown in the table and the interactive path is actually empty. The explainability of the *model* has a *local* and a *global path*; that is, there is a distinction between explaining the data sample and explaining the predictive model. Local interpretations of a single data instance can be done *ante-hoc*, i.e., the predictor is so comprehensible that a closer look serves as an explanation, for example, when the ML model is a simple decision tree. The other path, called XAI *post-hoc* methods, needs at least one execution of the predictor to produce information about the decision process of the ML model. The *global* model explanation is divided into two parts. The *direct* methods have the notion of direct interpretable and could be seen as *ante-hoc*. *Post-hoc* refers, in contrast to that, to algorithms that can create an explanation.

Due to rapid development in the field of XAI, the results of Table 1 are supplemented by other essential terms. The simple XAI taxonomy of [14], as well as the more sophisticated categories of [15], served as the basis for the following taxonomy, which has three sub-categories. The first sub-category is shown in Table 2.

Table 1: This IBM XAI Taxonomy is derived from the decision tree presented in [13], and serves as a guide to find a suitable XAI method. This overview focuses specifically on the model explanation aspect and does not present specific algorithms.

model				
local			global	
ante-hoc	post-hoc		direct	post-hoc
	samples	features		surrogate visualize

In Table 2, the distinction between XAI algorithms is made based on the creation of the explanation. For example, *Local Perturbation* methods try to modify the input and find important features as well as their corresponding feature importance values [14, 15]. The methods in *Leveraging Structure* use internal information from the ML model, such as gradients, like *Backprop* in the paper by



[14]. *Meta Explanations* abbreviated with *Meta-Expl.* exploit explanations from different explanatory methods, and *Architecture Modification* abbreviated with *Arch. Mod.* algorithms alter the predictor with the goal of finding simpler representations [15]. *Examples Methods* derive explanations from input data and use the output as an interpretation for similar inputs.

Table 2: This table presents the functional part by the XAI Taxonomy according to [14] and [15]. The methods are classified based on their functional interaction with the predictor model.

<i>functional approach</i>				
Local Perturbation (Perturbation)	Leveraging Structure	Meta-Expl.	Arch. Mod. (Backprob)	Examples

Another classification uses the output of the explanation approaches, which is shown in Table 3. The first sub-category is the weighting of feature importance values that provides information on which features have an influence and how great this influence is. *Surrogate Models* try to approximate more complex models by focusing only locally or by generally using *ante hoc* explanatory models. In addition, example-based methods are also used here. [15]

Table 3: The scope of this part from the introduced XAI Taxonomy according to [14] and [15] is focused on the results that different explanation models provide. As it can be seen there are *Feature Importance* as well as *Example* explanations and *Surrogate Models* that can explain black box models.

<i>result approach</i>		
Feature Importance	Surrogate Models	Examples

The conceptual approach, shown in Table 4, is the last category introduced by [15] and covers all aspects that were introduced by [14] in 2020. Further breakdowns are made into the *usage*, the *scope* of the explanation, and the *other dimensions*. The *usage* includes procedures that can be explained from the outside or XAI algorithms that provide a *model agnostic (model-agn.)* or *model specific (model-spec.)* explanation after the predictor model. The *scope* describes the range of what is explained, e.g., the complete/global ML model or individual data instances. *Other dimensions* include outputs, among others [15]. IBM researchers suggest the following approaches for a feature-based explanation: CEM [16], LIME [17], and SHAP [18]. To understand the decision underlying the model, it is not enough to explain the sample (data instance), but it is fundamental to explain the impact of each feature on the decision. Based on the explanations for each feature, it may also be possible to identify bias. Considering the subject of *evaluation*, it is necessary to generate nearly equal outputs, falling into the sub-category of *Feature Importances*. In the proposed use case shown in section 1, and given the axiom that explanations are an advantage of decision support systems (DSS), the easiest way to implement this is to use *local post-hoc* Methods. A summary of suitable methods and their limitations is presented.

Table 4: This view of the XAI taxonomy from [14] and [15] highlight the explanatory capability of XAI methods. The main questions addressed here are: what is being explained and whether the explanation can be obtained from the XAI method before the machine learning model makes predictions. Additional information related to the XAI approach and the problem domain is also provided.

<i>conceptual approach</i>					
<i>stage (usage)</i>		<i>scope</i>		<i>other dim.</i>	
ante-hoc	<i>post-hoc</i>	glo.	loc.	res.	issues
	model-agn.				
	model-spec./ (intrinsic)				

### 2.3 Background of XAI Algorithms

In this sub-section are reviewed three XAI methods such as LIME (Local Interpretable Model-Agnostic Explanations by [17]), SHAP (SHapley Additive exPlanations by [18]) and CIU (Contextual Importance and Utility by [19]). To describe these algorithms, two definitions are needed, namely deterministic and nondeterministic algorithm.

An algorithm whose behavior depends entirely on the input data is called *deterministic*. Thus, processing the same input data always leads to the same result. A *nondeterministic* algorithm is an algorithm that specifies several ways to process the same input data - without any specification of which option will be chosen, which can lead to either the same or different output [20].

#### 2.3.1 LIME

The idea behind LIME is to consider the local model as a black box model. The mode of operation of LIME is based on perturbing an original data point as input into the black box model and using the resulting predictions to train an interpretable surrogate model, which locally approximates the predictions of the black box model. The explanation provided by LIME is defined by:

$$\xi(x) = \operatorname{argmin}_{g \in G} \mathcal{L}(f, g, \pi_x) + \Omega(g) \tag{13}$$

In Equation 13,  $\xi$  is the explanation of instance  $x$ , which is obtained through an optimization task. The function  $g$  is an interpretable local model, and  $G$  is a class of potentially interpretable models. The function  $f$  is the original predictor and  $\pi_x$  defines the radius of the neighborhood around instance  $x$ .  $\mathcal{L}$  is the loss function that measures the accuracy of the prediction of the instance  $x$  with respect to the interpretable model  $g$  and the original prediction of  $f$  in the area of  $\pi_x$  around the original prediction.  $\Omega$  is a complexity measure of  $g$  and serves as a penalty function.

LIME calculates feature importance values that show the contribution of each feature for and against a prediction in a certain class. LIME values are numerical, where a negative numerical value indicates that this feature is not in favor of the prediction. On the other hand, a positive numerical value shows that this feature has a positive influence on the prediction.

The advantages of LIME are the simple explanations diagrams and the ability to process various types of input [21]. However, LIME has some disadvantages: it provides only local explanations and does not show this to the end user. The interpretation of the



importance plots is often hard to understand by nonexperts. The algorithm has consistency issues and can be attacked by adversarial examples [21].

The consistency issues with LIME are not only related to minimal changes in features but also to the fact that such changes can lead to completely different explanations. In previous experiments with LIME, it has been shown that by iterating with LIME over the same data instance 10, 100 and 1000 times, the inconsistency is also present [1]. These results and [22] show that the LIME algorithm is nondeterministic.

### 2.3.2 SHAP

The idea of the SHAP algorithm has its origins in the game theory. Calculate the extent to which a coalition (set of features) contributes or does not contribute to a particular classification based on the so-called Shapley values. The implementation used was the implementation by [18], known as *SHap Additive exPlanations*. The following definition describes the generation of explanations by the algorithm [18, 23]:

$$g(z') = \phi_0 + \sum_{i=1}^M \phi_i z'_i \quad (14)$$

The function  $g$  describes the explanatory model and  $z'$  describes the data instance to be interpreted. The variable,  $z'$  may have only a subset of all features. The explanation is generated by a linear model, where  $\phi_i \in \mathbb{R}$  and  $z'_i$  are either zero or one, to represent the presence or absence of a value from the feature set  $z'$  [18].

The author in [18] presents various explainer models in the framework, and the model used in [1] was the kernel explainer. The computational model approximates Shapley values with perturbations of  $z'$ . Thus, the complexity problem of computing Shapley values could be solved. SHAP also generates feature importance values such as LIME. A higher FIV of SHAP means that it contributes more to a prediction than a smaller one. A negative feature value argues against a prediction in a class and a positive value argues for a particular prediction.

SHAP also has some advantages: first, there is the sophisticated mathematical model for greater consistency and accuracy [17, 21, 23]. Besides that, the better intuition about feature weights, which is more related to humans, is another reason mentioned in [21]. Even though SHAP comes in hand with several advantages, the algorithm is non-deterministic [24], the approximation time of the Kernel SHAP is an issue [21, 23] and SHAP is also vulnerable to adversarial attacks [21].

### 2.3.3 CIU

The third model-agnostic algorithm is based on Decision Theory, more specifically, based on the subdomain of Multiple Criteria Decision Making (MCDM). In contrast to LIME and SHAP, this approach distinguishes between the measured importance and the utility of an attribute. Based on the relevance of the features, the focus lies on contextual importance ( $CI$ ). This is described in [19, 25] as follows:

$$CI_j(\vec{C}, \{i\}) = \frac{C_{max_j}(\vec{C}, \{i\}) - C_{min_j}(\vec{C}, \{i\})}{absmax_j - absmin_j} \quad (15)$$

The explanation model CIU calculates with the function  $CI_j$  the importance of feature  $i$  in the feature vector  $\vec{C}$  for an output label (value)  $j$ . The function  $C_{max_j}$  determines the maximum output of the prediction  $j$  for a certain feature  $i$ . The  $C_{min_j}$  calculation follows a similar approach. The functions  $absmax_j$  and  $absmin_j$  determine the highest and lowest prediction from the given data set. More details are provided in [19, 25, 26].

The CIU values are in the interval between 0 and 1, where a higher numerical value represents a greater influence on the prediction and vice versa. A value of 0 means that the attribute does not influence the prediction. In contrast to the other methods, in which negative values argue against a classification, this is not the case here (i.e., values are between 0 and 1).

The advantages of the CIU algorithm are the different working calculation model and the absence in the questions of consistency (such as in LIME), as well as the calculation time as in SHAP [1]. Thus, this algorithm is suitable as a control method in the group of model-agnostic methods like in [9]. In [1] it is shown that CIU has the best stability during different runs with the same instances to explain, i.e. in terms of repeatability it is a good benchmark for SHAP and LIME. The problem with the value interval of 0 to 1 was solved in [9] by introducing a threshold value as a decision boundary. If feature values exceed this limit, the value has a positive influence and vica versa. Despite the advantages of the CIU algorithm, it is also non-deterministic, due to its randomized data sampling for numeric values, which the HR use case currently is [25].

## 3 The Idea

According to [2] there are two key criteria when you want to quantify the stability of explanations. The first aspect is the *Identity* criteria, i.e., that you should have identical explanations for identical instances. The second criterion is the *Separability*, which means that when the data instances are not identical, the explanation should not be identical.

The initial proposal of this work is to understand that these two criteria are at different ends of the same scale. The next step is to expand this scale by the criteria that are also presented in [subsubsection 2.1.1](#). A representation of the so-called *Stability Scale* is shown in [Figure 3](#). The following sub-sections explain the development of the scale in greater detail.

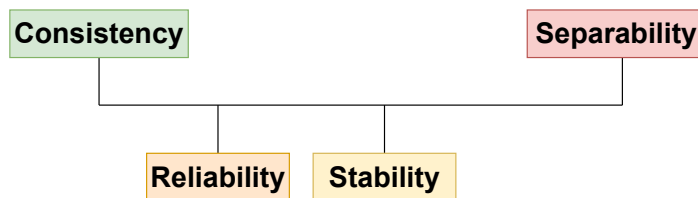


Figure 3: Stability Scale and their aspects (sub-criteria, properties)

### 3.1 Unification of Notions

For the standardization of notions (terms) the left end of the scale in [Figure 3](#), serves as a fix point, from that other properties (terms), or criteria are explained. In [Figure 2](#) there are two notions shown at

this fix point: *Consistency* and *Identity*. According to the definition of *Identity* in [2] it is mainly focused on the instances and this could lead to the idea that only the data instances are mentioned. In contrast, the definition of *Consistency* by [4] includes the data instances and the predictor is considered. That means, the scope is broader and not only data-centered. Due to this, the notion *Consistency* is preferred, because of its wider scope.

Since the term *Reliability*, which comes next to *Consistency*, is not considered in the other works, it is not necessary to unify this notion with another.

*Stability* comes next to *Reliability* and in Figure 2 it is shown that there are three notions. The definitions of the notions are very similar, according to the majority usage of the term *Stability*, it could be used as a common notion.

The right end of the scale, which is denoted with *Separability*, is only mentioned by [2] and there is also no need for a standardization.

It could be summarized that the initial idea, denoted in this work as *Explanation-Stability*, could be seen as an aspect of an explanation produced by an XAI algorithm. When it comes to details, it could be stated that this aspect of stability has sub-criteria and the initial notion is not sufficient, that means the *Stability Scale* represents all kinds of aspects related to the proposed notion *Explanation-Stability*.

### 3.2 The Notions of this work in Detail

In the following, it is explained why the notions (sub-criteria) are ordered as seen in Figure 3. As a foundation, the arrangement of the sub-criteria of *Explanation Stability* in Table 5 was employed.

Table 5: Comparison of sub-criteria from the Explanation Stability according to [1, 2, 3, 4, 5]. The abbreviations used in the table are: id. as *identical*, sim as *similar*. The aspects of the comparison are the input, the output (that is, the explanation), and the constraints. The yellow-highlighted terms indicate the differences from the previous row, specifically in comparison to the stricter criterion.

Notion	Inputs		Explanation	Constraints
	model	data		
<i>Consistency</i>	id.	id.	id.	determinism
<i>Reliability</i>	id.	id.	highly sim.	–
<i>Stability</i>	id.	sim.	sim.	–
<i>Separability</i>	id.	non-id.	non-id.	different instances

The sub-criteria *Consistency* is the most strict criterion because of the common definitions of [2] and [4]: “identical inputs should result in identical outputs”. Despite this, [3] mentioned in their definition of *Stability* that: “Nonetheless, a lack of stability can also be created by non-deterministic components of the explanation method, such as a data sampling step. Regardless of that, high stability is always desirable”. In this work, where *Stability* is examined in detail, the constraint of determinism is assigned to the *Consistency* criteria as the strongest requirement for stability.

A weaker requirement in terms of stability is *Reliability*. In Table 5, the second row summarizes this criterion. The inputs are the same as in the *Consistency* row, with the explanation being only *highly similar* compared to the first criterion. Therefore, this stability aspect could be fulfilled by nondeterministic XAI algorithms. To

ensure uniform inputs and solely vary the output – specifically, the explanation property – the stability aspect here is weaker compared to *Consistency*.

Another stability property denoted by *Stability* is not as strong as the *Reliability* criterion. The only identical input is the model (i.e., the predictor), while the data instances are only similar, representing slight variations of an original data point. Consequently, the explanations are only *similar*, not *identical* or *highly similar*. This implies that the stability aspect, denoted as *Stability*, has less stringent requirements than both *Consistency* and *Reliability*. Consequently, it can be concluded that *Stability* is the least demanding stability criterion represented in Figure 3.

The *Separability* property, situated at the opposite end of the scale, asserts that two fundamentally different instances should not produce the same explanation on the same ML model. This high instability contrasts with or complements *Consistency*, measuring an opposite aspect on the *Stability Scale*. The summarized information is visualized in Figure 4.

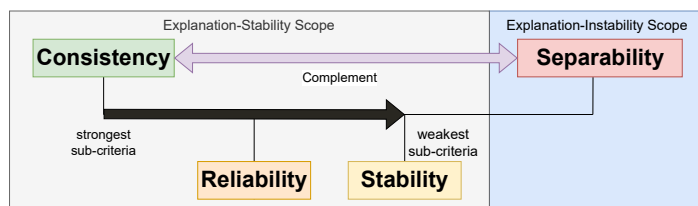


Figure 4: The proposed *Stability Scale* is shown here with additional context, starting from the left with the strongest criterion, up to the right with the complement of stability. The relations between the stability aspects: *Consistency*, *Reliability* and *Stability* are presented in this overview.

### 3.3 Mapping of Stability Criteria and Measures

At this point, it can be stated that there is an overarching concept for the stability of an explanation produced by an XAI approach. This term was introduced as *Explanation-Stability* and has three sub-criteria of stability. In order to quantify the *Explanation-Stability* objectively and to take all aspects into account, one should measure all sub-criteria of stability.

However, the properties of the XAI algorithms should be taken into account (subsection 2.3). In the proposed use case of [1], all XAI procedures are post-hoc approaches that generate local explanations. The explanation is of the type where each predicted feature  $f(x_i)$  is associated with a feature importance value  $\Phi(f, x)_i$ . Therefore, for every explanation method  $\Phi$ , the stability has been measured.

In case that all XAI approaches lack in terms of determinism, this stability criterion is very hard to fulfill by the used approaches. Therefore, more realistic stability measures such as *Reliability* and *Stability* are focused when measuring nondeterministic XAI algorithms. An overview of the measurement approaches assigned to the XAI methods used is shown in Table 6.

After assigning concrete stability measures, it is now denoted as *Explanation-Stability-Metric* due to the use of single stability-metric approaches.

Table 6: Assigning the proposed stability measures to non-deterministic, post-hoc explanation methods is discussed in detail in section 2.

XAI Algorithm	Explanation-Stability		
	Consistency	Reliability	Stability
LIME (2.3.1)	violate determinism	R(x)	S(x)
SHAP (2.3.2)			
CIU (2.3.3)			

### 3.4 Problems of Reliability and Stability

Even though the proposed approaches of the reliability measurement, presented in [1], as well as the stability analysis in [5] deliver highly valuable information about the quality of a certain XAI algorithm, the information are limited to one sub-criteria of the explainable models. The confidence interval method in [1] focuses on approaching the problem of *Reliability*, by conducting a test with the same data samples/models and with the expectation of highly similar data explanations. However, this approach lacks information regarding the stability of an XAI model.

To solve that, the paper [5] implements a method to evaluate the stability of an XAI model, by conducting changes on the input data. In contrast to [1], however, information regarding the reliability are missing. In addition, outliers are treated as regular explanations, which negatively effects the overall stability performance of the algorithm. An example for that can be seen in Figure 5, where several XAI algorithms are compared based on the similarity of the explanations in the feature "Income". By that, it can be seen that for instance LIME suffer from a high data dispersion because of several outliers. Due to a 1% change in one of the input features, the explanation similarity is expected to be in a range of 99% to 101%, where 100% refers to an identical explanation. However, algorithms like LIME seem to be prone to outliers. To avoid an overall biased score, those outliers needs to be handled accordingly.

Besides those problems, both papers do not provide a reference point, specifically an optimality criterion that serves as a baseline for gauging the algorithm's performance at its optimum. Instead, the only available approach is to compare algorithm A with another XAI method to determine whether algorithm A outperforms or underperforms in relation to algorithm B.

It can be concluded that stability measures are highly valuable when, combined because they address both stability aspects. Additionally, the use of an optimality criterion for the reliability metric is particularly interesting when evaluating the stability of an individual XAI algorithm. When considering the stability metric, it is important to take outliers into account.

### 3.5 Improvements of Reliability and Stability

To solve the problems mentioned in subsection 3.4, the current article presents a novel metric to evaluate current XAI solutions by combining the strengths of several XAI criteria. In particular, based on the approaches presented in equations  $R(x)$  and Equation 12, the metric returns a vector of the corresponding scores to evaluate the XAI models. Since multiple aspects are considered, the metric provides an objective overall perspective on the performance of an XAI algorithm. In addition to that, the result vector enables

comparability to other XAI models to find the best solution for the corresponding use-case.

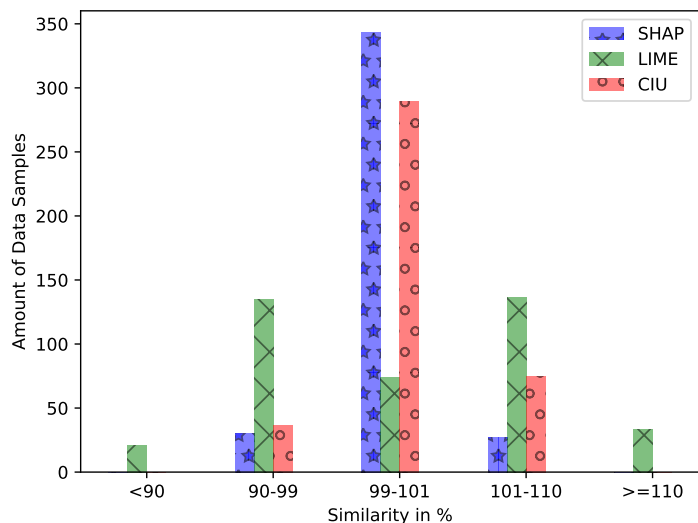


Figure 5: Algorithm Comparison about the similarity of the explanations for the feature „Income“ - representation of the feature importance similarities for the data pairs. The similarity measures the percentage match between the explanation of the original data sample as well as the one with a feature change of around 1%. The results of SHAP, LIME and CIU are compared.

## 4 Results

The motivation for the novel approach to assess the stability aspects of XAI explanations was derived from the need to unify various notions. A comprehensive foundation was established through an extensive review of survey papers and articles on stability, including references such as [1]–[5]. This research aided in the identification and consolidation of the most relevant concepts and terms, forming the basis for the proposed *Stability Scale*, as depicted in Figure 2.

After gathering all key ideas and notions, a suitable process for unifying terms is detailed in subsection 3.1. The proposed and consistent scale is presented in Figure 3. The suggested notion of the XAI explanation stability is *Explanation-Stability*, it is important to emphasize that this is an overarching concept. For a more detailed view, the *Explanation Stability* is defined by three sub-criteria. These are visualized on a scale and ordered by their strictness according to Table 5. In fact, by defining the table and identifying these sub-criteria, research question RQ1 is fully answered.

Considering the real-world scenario presented in [1], it can be concluded that the *Explanation-Stability* construct of nondeterministic XAI approaches deteriorates in only two of the three criteria, namely *Reliability* and *Stability*. Furthermore, the outstanding idea of this work is to propose a concrete measurement approach for every stability aspect (sub-criterion). The sub-criterion of *Reliability* is measured by  $R(x)$  and *Stability* is measured by  $S(x)$ , this concept is denoted as *Explanation-Stability-Metric*. Measures are described in depth in subsection 2.1.2 and improvements are identified in subsection 3.4.

This level of explicitness is unprecedented; unlike [2, 3], which provided only evaluation criteria and definitions. These works and



considerations were pivotal when addressing the overarching question: “How can we evaluate XAI in general?” When the focus narrows to evaluating specific aspects, such as the actual explanation, and the scope is tightly defined, sources are rarely limited. Developing a new standard for evaluating XAI methods entails constructing a robust foundation from an initial and intuitive idea. This process involves being highly specific and thereby bridging existing gaps.

In addressing research question RQ2, the newly proposed approach, known as *Explanation-Stability-Metric*, has the potential to provide at least a partial answer to nondeterministic XAI approaches. With additional effort, expanding the metric to incorporate the stability aspect of *Consistency*, it could also be applied to deterministic XAI methods.

The key idea is to provide a holistic measure of *Explanation-Stability*; therefore, it is necessary to furnish information about stability. Upon the development of an initial outcome from the execution of this metric, a stability vector would provide all stakeholders with sufficient information to better trust, build, and comprehend AI models. Furthermore, this work should contribute to making autonomous systems applicable in high-risk domains.

## 5 Discussion

Future work and assumptions are considered to enhance the new stability assessment approach outlined in this study. Firstly, the question arises: “How useful is the addition of the stability metric for *Consistency* concerning nondeterministic XAI methods?”. This can be achieved by incorporating a distance measure that supports multidimensional data, such as the Euclidean distance. This enhancement provides a more objective overview of stability, which is valuable even when these methods produce less accurate scores than deterministic methods.

In the proposed stability metric, a complementary aspect to *Consistency*, referred to as *Separability*, was identified. Its addition could prove beneficial not only for measuring stability aspects but also for assessing *instability*, as mentioned in subsection 3.2. This broadens the scope of stability and introduces a related aspect into the metric. However, the inclusion of these two additional criteria requires further exploration in future work.

Considering the scope of stability, measuring all four sub-criteria could prove valuable, potentially making this approach applicable to deterministic XAI algorithms.

When considering the representation of the results, a second significant question arises: how to combine the measures of all sub-criteria into a single score. Several reasons justify this additional processing step. The fundamental concept of a metric is to distill relevant information into a real-valued single score. This approach offers a significant advantage, preventing information overload for those seeking insights.

As AI systems become more prevalent in our daily lives, ensuring information about stability is accessible to non-experts is essential. A single-score approach facilitates this by offering a concise and easily understandable representation.

However, aggregating the information also introduces a few disadvantages as it may lead to the loss of detailed information

when reducing the sub-criteria scores of *Explanation-Stability* into a single score. When the resulting scores from each sub-criterion are condensed into a single value, questions such as “Could *Stability sub-criterion A* potentially compensate for the lower performance of *Stability sub-criterion B*, and to what extent does this influence exist?” arise.

Other important considerations include “How strong is the correlation between these stability aspects?”, “Is there a dominant stability criterion?” and “Are there domain-dependent sub-criteria that should be prioritized, such as in the medical field or other sensitive use cases?” These are the research questions that will need to be explored as this metric approach is implemented in real-world use cases.

## 6 Conclusion

In conclusion, the development of autonomous systems must adhere to specific quality criteria. Given that these systems frequently employ artificial intelligence for critical tasks, assessing the stability of complex machine learning models is imperative for comprehension, creation, and improvement. This work aims to lay the groundwork for future endeavors in verifying the stability of such models and their applicability in real-world scenarios.

**Conflict of Interest** The authors declare no conflict of interest.

## References

- [1] F. Gawantka, A. Schulz, J. Lässig, F. Just, “SkillDB - An Evaluation on the stability of XAI algorithms for a HR decision support system and the legal context,” in 2022 IEEE 21st International Conference on Cognitive Informatics & Cognitive Computing (ICCI\*CC), 183–190, 2022, doi:10.1109/ICCICC57084.2022.10101657.
- [2] H. Löfström, K. Hammar, U. Johansson, “A Meta Survey of Quality Evaluation Criteria in Explanation Methods,” in J. De Weerd, A. Polyvyanyy, editors, *Intelligent Information Systems*, 55–63, Springer International Publishing, Cham, 2022.
- [3] D. V. Carvalho, E. M. Pereira, J. S. Cardoso, “Machine learning interpretability: A survey on methods and metrics,” *Electronics*, **8**(8), 832, 2019.
- [4] M. Nauta, J. Trienes, S. Pathak, E. Nguyen, M. Peters, Y. Schmitt, J. Schlötterer, M. van Keulen, C. Seifert, “From anecdotal evidence to quantitative evaluation methods: A systematic review on evaluating explainable ai,” *ACM Computing Surveys*, **55**(13s), 1–42, 2023.
- [5] F. Gawantka, F. Just, M. Ullrich, M. Savelyeva, J. Lässig, “Evaluation of XAI Methods in a FinTech Context,” in *International Workshop on Artificial Intelligence and Pattern Recognition*, 143–154, Springer, 2023.
- [6] M. Schuld, F. Petruccione, *Machine Learning with Quantum Computers*, Quantum Science and Technology, Springer International Publishing, 2021.
- [7] S. Russell, P. Norvig, *Artificial Intelligence: A Modern Approach* (4th Edition), Pearson, 2020.
- [8] I. Kakogeorgiou, K. Karantzas, “Evaluating explainable artificial intelligence methods for multi-label deep learning classification tasks in remote sensing,” *International Journal of Applied Earth Observation and Geoinformation*, **103**, 102520, 2021, doi:https://doi.org/10.1016/j.jag.2021.102520.
- [9] S. Hariharan, R. Rejimol Robinson, R. R. Prasad, C. Thomas, N. Balakrishnan, “XAI for intrusion detection system: comparing explanations based on global and local scope,” *Journal of Computer Virology and Hacking Techniques*, 1–23, 2022.



- [10] C.-K. Yeh, C.-Y. Hsieh, A. Suggala, D. I. Inouye, P. K. Ravikumar, "On the (In)fidelity and Sensitivity of Explanations," in H. Wallach, H. Larochelle, A. Beygelzimer, F. d'Alché-Buc, E. Fox, R. Garnett, editors, *Advances in Neural Information Processing Systems*, volume 32, Curran Associates, Inc., 2019.
- [11] D. Minh, H. X. Wang, Y. F. Li, T. N. Nguyen, "Explainable artificial intelligence: a comprehensive review," *Artificial Intelligence Review*, 1–66, 2021.
- [12] V. Arya, R. K. E. Bellamy, P. Chen, A. Dhurandhar, M. Hind, S. C. Hoffman, S. Houde, Q. V. Liao, R. Luss, A. Mojsilovic, S. Mourad, P. Pedemonte, R. Raghavendra, J. T. Richards, P. Sattigeri, K. Shanmugam, M. Singh, K. R. Varshney, D. Wei, Y. Zhang, "One Explanation Does Not Fit All: A Toolkit and Taxonomy of AI Explainability Techniques," *CoRR*, **abs/1909.03012**, 2019.
- [13] Q. V. Liao, M. Singh, Y. Zhang, R. Bellamy, "Introduction to explainable AI," in *Extended Abstracts of the 2021 CHI Conference on Human Factors in Computing Systems*, 1–3, 2021.
- [14] A. Das, P. Rad, "Opportunities and Challenges in Explainable Artificial Intelligence (XAI): A Survey," 2020, doi:10.48550/ARXIV.2006.11371.
- [15] T. Speith, "A review of taxonomies of explainable artificial intelligence (XAI) methods," in *2022 ACM Conference on Fairness, Accountability, and Transparency*, 2239–2250, 2022.
- [16] A. Dhurandhar, P.-Y. Chen, R. Luss, C.-C. Tu, P. Ting, K. Shanmugam, P. Das, "Explanations based on the missing: Towards contrastive explanations with pertinent negatives," *Advances in neural information processing systems*, **31**, 2018.
- [17] M. T. Ribeiro, S. Singh, C. Guestrin, "'Why Should I Trust You?': Explaining the Predictions of Any Classifier," in *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD '16*, 1135–1144, Association for Computing Machinery, New York, NY, USA, 2016, doi:10.1145/2939672.2939778.
- [18] S. M. Lundberg, S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," in *Proceedings of the 31st International Conference on Neural Information Processing Systems, NIPS'17*, 4768–4777, Curran Associates Inc., Red Hook, NY, USA, 2017.
- [19] K. Främling, "Decision theory meets explainable AI," in *Explainable, Transparent Autonomous Agents and Multi-Agent Systems: Second International Workshop, EXTRAAMAS 2020, Auckland, New Zealand, May 9–13, 2020, Revised Selected Papers 2*, 57–74, Springer, 2020.
- [20] R. W. Floyd, "Nondeterministic algorithms," *Journal of the ACM (JACM)*, **14**(4), 636–644, 1967.
- [21] M. Knap, "Model-Agnostic XAI Models: Benefits, Limitations and Research Directions," 2022.
- [22] A. Holzinger, A. Saranti, C. Molnar, P. Biecek, W. Samek, "Explainable AI methods-a brief overview," in *International Workshop on Extending Explainable AI Beyond Deep Models and Classifiers*, 13–38, Springer, 2020.
- [23] C. Molnar, "Interpretable machine learning," <https://christophm.github.io/interpretable-ml-book/>, 2022, accessed on 04 16, 2023.
- [24] R. Heese, S. Mücke, M. Jakobs, T. Gerlach, N. Piatkowski, "Shapley Values with Uncertain Value Functions," in *International Symposium on Intelligent Data Analysis*, 156–168, Springer, 2023.
- [25] K. Främling, "Contextual Importance and Utility: A Theoretical Foundation," in G. Long, X. Yu, S. Wang, editors, *AI 2021: Advances in Artificial Intelligence*, 117–128, Springer International Publishing, Cham, 2022.
- [26] K. Främling, "Explainable AI without Interpretable Model," *CoRR*, **abs/2009.13996**, 2020.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Investigating Heart Rate Variability Index Classification in *Macaca fascicularis* and Humans: Exploring Applications for Personal Identification and Anonymization Studies

Daisuke Hirahara<sup>1</sup>, Itaru Kaneko<sup>\*2</sup>, Junji Nishino<sup>3</sup>, Junichiro Hayano<sup>4</sup>, Oscar Martinez Mozos<sup>5</sup>, Emi Yuda<sup>2</sup>

<sup>1</sup>Department of AI Research Lab, Harada Academy, 2-54-4 Higashitaniyama, Kagoshima, Kagoshima 891-0113, Japan

<sup>2</sup>Tohoku University, GSIS, Sendai, 980-8579, Japan

<sup>3</sup>The University of Electro-Communication, Tokyo, 182-8585, Japan

<sup>4</sup>Nagoya City University, Nagoya, 467-8601, Japan

<sup>5</sup>Örebro University, Institutionen för Naturvetenskap och Teknik, Örebro, 701 82, Sweden

### ARTICLE INFO

Article history:

Received: 09 November, 2023

Revised: 25 January, 2024

Accepted: 26 January, 2024

Online: 22 February, 2024

Keywords:

*Macaca fascicularis*

Privacy risks

Heart rate patterns

### ABSTRACT

*In this paper, we determine the feasibility of differentiating between the heart rate patterns of *Macaca fascicularis* and human infants by comparing pertinent hyperparameters. This verification process was undertaken to ascertain the suitability of *Macaca fascicularis* heart rate data as a testbed for evaluating heart rate parameter privacy safeguarding methodologies. The biological characteristics of *Macaca fascicularis* bear significant resemblance to those of humans, which consequently renders them useful subjects in medical experiments alongside other laboratory animals. The process of capturing heartbeat data from *Macaca fascicularis* is notably akin to the methodologies used to record human cardiac activity. In other hand, the recent years have witnessed the construction of extensive heart rate databases, thus raising important considerations surrounding privacy in their usage. Heartbeat recordings, indeed, can provide a wealth of diverse information, necessitating careful handling to maintain data privacy. Specifically, a Holter monitor, a type of electrocardiogram device, can record cardiac electrical activity for over 24 hours. The statistical indices derived from these recordings prove useful for various types of analysis, and simultaneously hold information relating to individual behaviors and health conditions. The extent to which individuals can be identified within such expansive databases is a topic warranting exploration; however, few individuals have granted consent for their data to be used for such research purposes. Given this scenario, since the protection of personal data is not a requisite for *Macaca fascicularis*, the proposition of employing *Macaca fascicularis* data to investigate the potential for individual identification appears to be a plausible approach. The experiment verified the similarity of cynomolgus monkey heart rate data to human heart rate data. The results are similar, suggesting that it is appropriate to use cynomolgus monkey heart rate data for personality identification experiments.*

## 1. Introduction

*Macaca fascicularis* exhibit a significant similarity to humans in terms of biological characteristics, making them suitable subjects in medical experiments alongside other laboratory

animals [1] and [2]. Heartbeat data from *Macaca fascicularis* can also be recorded using methodologies similar to those employed for humans [3]. Concurrently, a large heart rate variability database has been developed in recent years [4], [5], [6] and [7]. Personal identity can be an issue when using large heart rate databases. Therefore, research is needed to verify individual identifiability, which is our overall goal.

\*Corresponding Author: Itaru Kaneko, Tohoku University, GSIS, Sendai, 980-8579, Japan, [kyhsubmit@it-aru.com](mailto:kyhsubmit@it-aru.com)

Heart rate recordings are rich sources of data. Notably, wearable Holter electrocardiogram devices have the capacity to record cardiac electrical activity over a span exceeding 24 hours. The statistical indices derived from these recordings serve multiple analytical purposes and can also aid in predicting individual behaviors and health conditions.

Analysis of these databases revealed that various personal attributes can be extracted from heart rate [8], [9] and [10]. It is therefore essential to consider the protection of personal data contained in heart rate variability. Therefore, it is important to ascertain the identifiability of heartbeats and explore methods to enhance their anonymization.

Information pertaining to such health conditions, particularly cardiac diseases, is of a private nature and ought to be safeguarded from public disclosure.

Hence, when disseminating such expansive databases for research purposes, it is imperative to apply processing techniques ensuring the heartbeat data contained therein cannot be attributed to a specific individual. Various methodologies have been proposed to hinder individual identification, including anonymization strategies.

There is no absolute method to confirm that personal identification has been entirely thwarted. The effectiveness of identification prevention measures is typically evaluated empirically. Public tests, wherein attempts are made to identify specific individuals using test data, will be carried out, with the resultant outcomes undergoing evaluation. This necessitates the availability of appropriate test data.

Nevertheless, employing human data for such public testing initiatives is untenable due to inherent privacy implications.

The crux of this matter lies in the fact that the data that needs to remain unidentified pertains to personal health characteristics—specifically diseases—that the individual would prefer to keep private. Should the test results lead to the identification of both the disease and the individual, this would pose serious ethical concerns. Contributing data to such tests would bring about substantial risks and minimal benefits to the subjects. Consequently, procuring consent for such use is virtually unfeasible.

Consequently, we propose the use of Macaca fascicularis data for identification trials. Given that there is no requirement for the protection of personal data pertaining to Macaca fascicularis, there exist no ethical issues associated with identifying individual entities or symptoms within this data set.

Nonetheless, if the heart rate data of Macaca fascicularis can be statistically distinguished from that of humans, it would be improper to incorporate the data of Macaca fascicularis into human data and investigate the feasibility of individual identification. This is due to the fact that if we can discern the data of Macaca fascicularis, it would inevitably lead to the recognition of data specific to an individual entity.

Hence, it is crucial to establish that distinguishing between Macaca fascicularis and humans is sufficiently challenging. The primary aim of this study is to undertake this verification.

In essence, within our research purpose, the data of Macaca fascicularis and humans do not have to be entirely indistinguishable. If we select human data that bears similarity to the data of Macaca fascicularis, our objective is to ensure that distinguishing between Macaca fascicularis and humans proves challenging. Should this validation be successful, we can draw the conclusion that the Macaca fascicularis data, which represents the ultimate objective of this study, can be effectively utilized.

## **2. Proposed method**

### *2.1. Data preparation*

Thirty seven neonatal human samples and Macaca fascicularis ECGs, a single case of ECG data obtained from a Macaca fascicularis were utilized for this study.

Thirty seven neonatal Human heartbeat samples were obtained from ALLSTAR (Allostatic State Mapping by Ambulatory ECG Repository) database, those are all available data in the database. ALLSTAR is database based in Japan (<https://allstar.jpn.org/>). The data contained within the ALLSTAR database was gathered utilizing a Holter electrocardiograph, a product of Suzuken Co., Ltd., Japan.

The Macaca fascicularis used in this study was male and it is the only data we could use. The heart rate data for these specimens were collected using a two-lead electrocardiograph provided by Tokyo University of Agriculture and Technology, Japan.

### *2.2. Analysis Method*

A prevalent technique for the detection of R-waves in both the Macaca fascicularis and human neonatal data involves the extraction of baseline fluctuation components, followed by the isolation of feature points through waveform detection processing. For this study, we opted for a device that is capable of automatically calculating RRI.

Utilizing the heart rate variability index as a high-dimensional feature, we sought to determine whether it could be visualized using three representative methods of classification visualization.

Visualization using PCA, t-SNE, UMAP were executed.

PCA (Principal Component Analysis):

PCA stands for Principal Component Analysis and is a classical technique used for dimensionality reduction and information extraction of data. It transforms the data into a low-dimensional space, mainly by finding new axes (principal components) that maximize the variance of the data.

Since PCA extracts features in the range of linear transformations, it is specifically useful in this study to see the relationship between features and original variables in linear transformations.

t-SNE (t-Distributed Stochastic Neighbor Embedding) is a nonlinear dimensionality reduction technique. High-dimensional data can be mapped into a low-dimensional space while maintaining data similarity. This is particularly suitable for visualizing cluster structures. Neighborhoods in high-dimensional space are also mapped to neighborhoods in low-dimensional space.

This transformation not only provides visualization, but also gives insight into how each piece of data is categorized. In this study, it is useful to visualize nonlinear structures.

UMAP (Uniform Manifold Approximation and Projection) is a newer technique than t-SNE. It performs non-linear dimensionality reduction like t-SNE, but makes it possible to capture the overall structure while preserving the local structure of high-dimensional data. It is used for visualization and clustering of high-dimensional data.

It reflects the local and global structures of high-dimensional data in a well-balanced manner. In addition, it is resistant to noise and provides stable and good results.

In this experiment, we expected that PCA could confirm the relationship between the feature value and the original space in the linear transformation relationship, that visualization of the cluster structure by t-SNE and that UMAP would provide a well-balanced view of local and global structures.

Commonly utilized features of heart rate variability were computed and employed for multi-dimensional analysis. These features encompass HR, MNN, SDNN, SDNN, rMSSD, TPW, ULF, VLF, LF, HF, VHF, , and LFHF. These features are widely employed in various heart rate variability analyses[7][8].

### 3. 3D visualization

#### 3.1. 3D visualization using PCA

Figure 1 presents a three-dimensional visualization created through the principal component analysis (PCA) method. In this visualization, the principal components are derived through PCA, with the three most significant components being transformed into a three-dimensional format. The data points corresponding to Macaca fascicularis are circled in red.

Since PCA is a technique used for converting high-dimensional data into lower dimensions, it amalgamates dimensions in a high-dimensional linear space and extracts the dimension contributing most significantly to the data's variability. Consequently, dimensions that do not illustrate data differences are minimized, allowing for a reduction in dimensions while preserving information vital for data identification to the maximum possible extent.

In the figure 1, Macaca fascicularis is situated peripherally within the overall distribution. Nonetheless, numerous human data points exist in proximity to those of Macaca fascicularis, including closely neighboring human samples, thereby rendering the differentiation between the two challenging.

#### 3.2. 3D visualization using t-SNE

Figure 2 presents the data visualization in three dimensions via the t-SNE method.

Since t-SNE (t-Distributed Stochastic Neighbor Embedding) is a dimensionality reduction technique that is designed to maintain data differences while reducing the dimensionality of high-dimensional data to low dimensionality, a key feature of t-SNE is its inclusion of non-linear transformations, which enhances its performance in data clustering and the preservation of similarity.

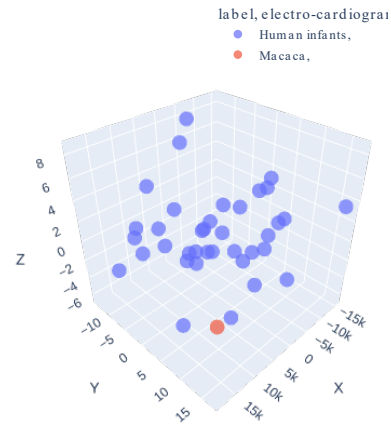


Figure 1: 3D visualization using PCA. 3-dimensional representation is obtained by selecting the top three eigenvectors associated with the largest eigenvalues of the covariance matrix. These eigenvectors capture the most significant variance in the original data, ensuring that the essential information is retained while reducing computational complexity and facilitating data interpretation. The selected 3-dimensional subspace preserves the main patterns and relationships present in the original data, enabling efficient data analysis and visualization. Macaca fascicularis is circled in red. It can be seen that there are human samples in the vicinity and it is difficult to distinguish them. Local structure in the cluster is not clear in case of PCA analysis.

In this analysis, t-SNE compresses from a high dimension to a 3-dimensional space.

This technique performs non-linear transformations to preserve the similarity of data. In other words, data points from the same class are located in close proximity to each other even in a three-dimensional space, while data points from different classes are more distantly spaced. Hence, some degree of clustering can be discerned in the figure, making it easier to understand the relationships between different groups of data in the original high-dimensional space.

Given that t-SNE reduces dimensions while preserving similarities between data points in the high-dimensional space, the figure tends to reflect some relative distances and similarities between the data in the original high-dimensional space.

Moreover, patterns and similarities within the data can be discerned more clearly with t-SNE than with PCA. This is attributed to t-SNE's ability to emphasize local data structures. Data points that were closely situated in high-dimensional space should also tend to be close together in 3D space, allowing local features to be more clearly discerned in the figure.

However, it should be noted that dimensionality reduction by t-SNE does not necessarily preserve the global structure. Therefore, it should be considered that the global arrangement in the 3D space may not necessarily reflect the global arrangement in the higher dimensions.

#### 3.3. 3D visualization using UMAP

Figure 3 shows a three-dimensional visualization by UMAP.

UMAP embeds data in a three-dimensional space while preserving data similarity, resulting in a distribution that reflects the features and clusters of the original high-dimensional data.



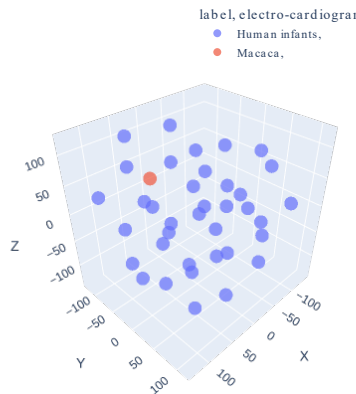


Figure 2: Visualization using t-SNE. Macaca fascicularis is circled in red. It can be seen that there are three or four clusters but local structure of the clusters are punctuated. It can be seen that there are human samples in all around it and it is difficult to distinguish from them.

Similar to t-SNA, data belonging to the same class are plotted close together in 3D space and data of different classes are plotted far apart. As a result, clusters of data are clearly visible in the figure, making it easier to understand the relationships between groups of data.

Also, like t-SNA, UMAP emphasizes local structures, so data that were close together are arranged close together in 3D space. Data that are plotted close together on the figure are likely to be nearby even in a high-dimensional space, making it easier to read the cluster structure.

Unlike t-SNA, UMAP also reflects global data placement to some extent. The structure between clusters can also be read in the figure, but in the case of UMAP, it is thought that this reflects the structure in a high-dimensional space.

This figure seems to be useful for understanding the characteristics of high-dimensional data by performing dimensionality reduction and visualization of data, but it is not true that there is no information loss due to dimensionality reduction.

#### 4. 2D visualizations

Figures 4 to 6 provide two-dimensional (2D) visualizations, achieved through further dimensionality reduction from the previous three-dimensional (3D) representations. While 2D inherently contains fewer dimensions than 3D, this can actually be advantageous when visualizing data. When 3D visualizations are projected onto a 2D plane (such as a computer screen or printed page), some information can be lost due to the lack of depth perception. This could potentially obscure important features or relationships in the data. Therefore, if the 3D visualization isn't explored with a 3D viewer or software that allows for rotation and inspection from different angles, it can sometimes be more insightful to create a 2D visualization from the outset. This can ensure that the most significant data relationships are immediately visible, and are not hidden by the 3D to 2D transformation at the.

##### 4.1. 2D visualization using PCA

Figure 4 presents a two-dimensional (2D) visualization using Principal Component Analysis (PCA). Here, the dimension that contributes the least to the variability of the data has been removed, effectively reducing the data from three dimensions down to two.

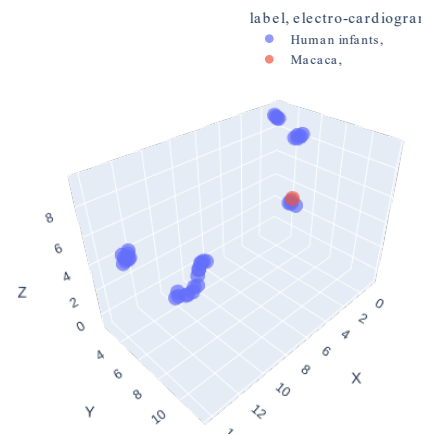


Figure 3: 3D visualization using UMAP. There are three or four clusters and the local structure of those are also visible. Macaca fascicularis is circled in red. It can be seen that it is in one of the clusters and there are human samples in the all-around it and it is difficult to distinguish from them.

The reduction process in PCA is designed to retain the features that account for the most variability in the data. Consequently, this 2D visualization is better suited to display the overall, or global, structure of the data. While PCA is very effective at illustrating these global trends and differences, it's worth noting that it may not accurately portray the more nuanced, local structures within the data set. These may become obscured or lost during the dimensionality reduction process. Thus, while PCA is an invaluable tool for examining broad trends in a data set, it may not fully capture the complexity of the data if there are intricate local patterns or clusters.

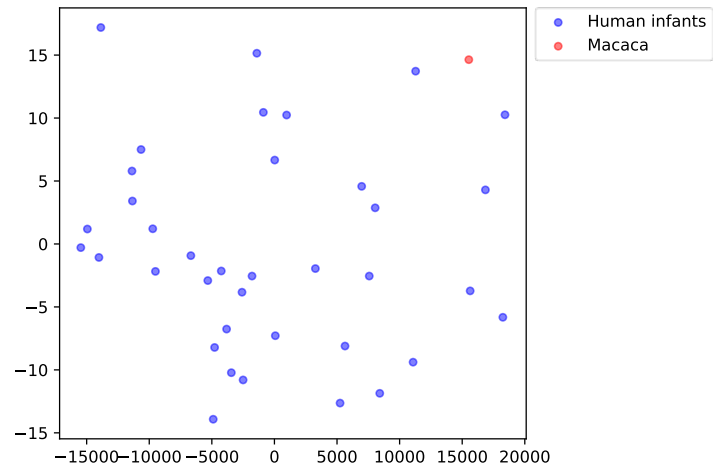


Figure 4: 2D visualization using PCA. Acquire the maximum 2D after PCA and convert it to 2D. Macaca fascicularis is circled in red. It can be seen that there are human samples in the vicinity and it is difficult to distinguish them. Local structure in the cluster is not clear.

##### 4.2. 2D visualization using t-SNE

Figure 5 represents the data's compression into two dimensions using t-SNE (t-Distributed Stochastic Neighbor Embedding). Like in the case of Figure 4, this 2D visualization method emphasizes the preservation of local structures and similarities between data points. Therefore, it provides a more detailed view of the internal structure within each cluster, often producing a more intuitive representation than PCA.

This 2D visualization helps reveal four discernible clusters, with data from *Macaca fascicularis* appearing within one of these clusters. However, as previously mentioned, t-SNE may not necessarily preserve the global structure of the data. When comparing this t-SNE visualization with the PCA representation, this limitation becomes evident.

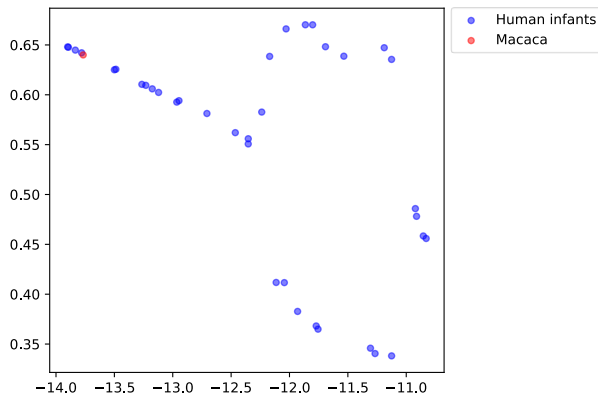


Figure 5: 2D visualization using t-SNE. It can be seen that there are three or four clusters but local structure of the clusters are punctuated. *Macaca fascicularis* is circled in red. It is difficult to distinguish *Macaca fascicularis*.

It's important to note that while t-SNE generally requires higher computational resources compared to PCA, in this case, due to the relatively small size of the dataset, this computational cost is not a significant concern.

#### 4.3. 2D visualization using UMAP

Figure 6 illustrates a two-dimensional visualization using the UMAP (Uniform Manifold Approximation and Projection) method. UMAP is particularly adept at preserving both local and global structures in data, leading to robust and informative visualizations.

In this figure, UMAP outperforms t-SNE by capturing intricate local cluster structures in a more comprehensible manner. While t-SNE may cause some clusters to collapse, UMAP is able to differentiate within-cluster data points and maintain clearer demarcations.

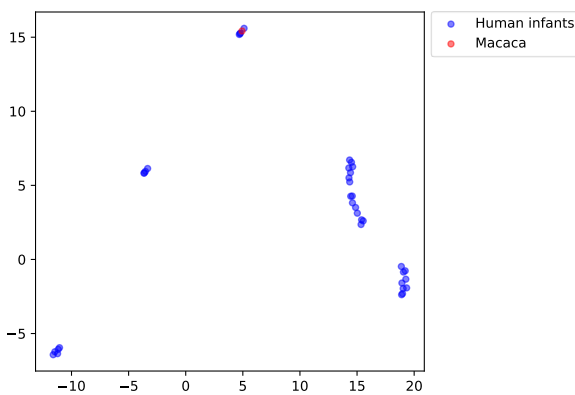


Figure 6: 2D visualization using UMAP. It can be seen that there are three or four clusters and the local structure of those are also visible. *Macaca fascicularis* is circled in red. It can be seen that it is in one of the clusters and there are human samples in the all-around it and it is difficult to distinguish from them.

Moreover, data points that were dispersed across separate clusters in the t-SNE visualization appear more integrated in this [www.astesj.com](http://www.astesj.com)

UMAP visualization. Consequently, it's easier to discern three or four distinct clusters. The data points for *Macaca fascicularis* are situated within one of these clusters, showcasing a fairly even distribution.

Since, UMAP is less sensitive to variations in parameters, allowing the global structure of the visualization to better represent the actual distribution of data in the high-dimensional space, global structure is best visualized in this figure.

#### 5. PCA after UMAP

Figure 7 shows a 2D visualization of UMAP (Uniform Manifold Approximation and Projection) followed by PCA. UMAP is particularly good at preserving both local and global structure in your data, resulting in robust and informative visualizations. PCA can derive the principal factors by linear transformation of the obtained dimensions. By combining these, you can see the structure visualized by UMAP on the space where the factors of the overall structure are orthogonalized in this figure.

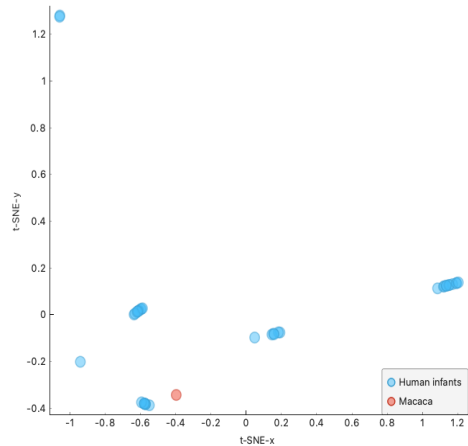


Figure 7: 2D visualization using UMAP (Uniform Manifold Approximation and Projection) followed by PCA. UMAP is preserving local and global structure in data, resulting in robust and informative visualizations. PCA can derive the principal factors by linear transformation of the obtained dimensions.

#### 6. Discussion

Throughout the conducted analyses, our findings consistently validate that the *Macaca fascicularis* dataset exhibits a distribution pattern that closely aligns with that of the human data, posing difficulties in effectively distinguishing the *Macaca fascicularis* data. Furthermore, the cynomolgus dataset demonstrates integration within this structure, upon the cluster structure observed in the human data.

Therefore, the findings of this study indicate that straightforward visual features and dimensionality reduction techniques may not be sufficient to differentiate between the heart rate variability indices of *Macaca fascicularis* and humans when their heart rates are similar. This implies that when the mechanical pulsation mechanism (the mechanical model) of the heart is consistent, the heart rate variability index exhibits analogous characteristics, making it challenging to distinguish between *Macaca fascicularis* and humans using conventional methods.

These results provide an intriguing insight into the similar physiological patterns shared between certain primates and humans when observed through the specific lens of heart rate

variability indices. Furthermore, they also underscore the need for more nuanced or advanced methods for distinguishing between such closely aligned data sets.

This study might have significant implications for a range of applications where such differentiation could be crucial. For instance, in medical research involving primate models, or in the development of biometric identification systems that rely on heart rate variability as a distinguishing factor. Further research is needed to identify and develop more sophisticated techniques that can make this crucial differentiation more accurately and efficiently.

A clear limitation of this study is the scant amount of data for both *Macaca fascicularis* and humans. However, the ultimate aim of this research is to utilize *Macaca fascicularis* data to verify the individuality of human heartbeats. For this purpose, it is not strictly necessary to establish whether there is discrimination between humans and *Macaca fascicularis*; it is sufficient to demonstrate that the level of discrimination between humans and *Macaca fascicularis* is considerably low.

If *Macaca fascicularis* possess traits that humans do not, it's assumed that it would be particularly easy to identify individuals in human heartbeat data. Therefore, it could actually be beneficial for the experiments on individual identification and anonymization. This opens up opportunities for future research and experimentation with larger datasets, more sophisticated analytical techniques, and alternative data sources.

To bolster the reliability of these results, future research will need to gather a larger sample size and consider an approach that incorporates a variety of feature values. Techniques such as data co-mingling and masking are commonly employed to safeguard personally identifiable information, and it's imperative that datasets remain devoid of any personally identifiable information or personal identifiers.

Additionally, integrating different types of machine learning and statistical techniques may also improve the ability to differentiate between the two species. Involving more complex techniques, such as deep learning or support vector machines, could also be considered to help refine the model. Overall, this research provides an interesting starting point, but more extensive work is required to validate these results and to explore the potential implications further.

To bolster the reliability of these findings, our future work will encompass the collection of additional data samples and exploration of a methodology that integrates a variety of feature values. It is crucial to highlight that in the handling of such data, methods such as data commingling and masking are often employed to safeguard personally identifiable information. This ensures that the datasets remain void of personal identifiers and any information that can be traced back to individuals.

In summary, the objective of this experiment was that cynomolgus monkeys are highly similar to humans, and a highly reliable answer was obtained through a multifaceted investigation. In the future, as we acquire more data, we will be able to use these results to advance research toward the broader goal of identifying individual heartbeats.

## 7. Conclusion

This research substantiated the claim that the cardiac data of *Macaca fascicularis* are substantially analogous to human heart rhythms and thus can be aptly employed to validate various method intended to prevent personal identification of cardiac data. This research provided highly reliable answer, derived through a multifaceted investigation. The deployment of *Macaca fascicularis* cardiac rhythm data as a test set for such validation is expected to enable research on more secure processing of cardiac rhythm data. And once the privacy preservation technique for heart rhythm data is possible to be provided, large and efficient use of large-scale shared heart rhythm data repositories will be possible. Clear limitation of this study is the small number of *Macaca fascicularis* and we need to address this limitation in future research. Even current result is possible to support such result, still it is possible to improve precision and reliability of result by extending data size and analysis method.

## References

- [1] B. Koo, D. Lee1, P. Kang, K. Jeong, S. Lee, K. Kim, Y. Lee, J. Huh, Y. Kim, S. Park, Y.B. Jin, S. Kim, J. Kim, Y. Son and S. Lee, "Reference values of hematological and biochemical parameters in young-adult cynomolgus monkey (*Macaca fascicularis*) and rhesus monkey (*Macaca mulatta*) anesthetized with ketamine hydrochloride," PMID: 32257895 PMID: PMC7081622, 2019.
- [2] Y. Uno, S. Uehara, H. Yamazaki, "Genetic polymorphisms of drug-metabolizing cytochrome P450 enzymes in cynomolgus and rhesus monkeys and common marmosets in preclinical studies for humans," PMID: 29277691, DOI: 10.1186/s42826-019-0006-0.
- [3] Y. Noguchi, I. Tawara, K. Kondo, H. Nigi, T. Tanaka, "Electrocardiographic studies in the Japanese monkey (*Macaca fuscata*) with special reference to the effect of anesthesia with barbiturates," *Primates* 10, 273-283, 1969, DOI: 10.1007/BF01730348.
- [4] S. Aziz, S. Ahmed, M. Alouini, "ECG-based machine-learning algorithms for heartbeat classification," *Sci Rep* 11, 18738 (2021), DOI: 10.1038/s41598-021-97118-5.
- [5] A. L. Goldberger, L. A. N. Amaral, L. Glass, J. M. Hausdorff, P. C. Ivanov, R. G. Mark, J. E. Mietus, G. B. Moody, C. Peng, H. E. Stanley, "PhysioBank, PhysioToolkit, and PhysioNet: Components of a New Research Resource for Complex Physiologic Signals," *Circulation* 101(23):e215-e220, 2000 (June 13), doi: 10.1161/01.CIR.101.23.e215.
- [6] E. Yuda, Y. Furukawa, Y.a Yoshida, J. Hayano & ALLSTAR Research Group, "Association Between Regional Difference in Heart Rate Variability and Inter-prefecture Ranking of Healthy Life Expectancy: ALLSTAR Big Data Project in Japan," *Lecture Notes of the Institute for Computer Sciences, Social Informatics and Telecommunications Engineering* book series (LNICST, volume 194), 2017.
- [7] E. Yuda, M. Kisohara, Y. Yoshida & J. Hayano, "Constituent factors of heart rate variability ALLSTAR big data analysis," *Wireless Networks*, 28(3):1287-1292, 2022, DOI: 10.1007/s11276-018-01898-0.
- [8] T. Yukishita, K. Lee, S. Kim, Y. Yumoto, A. Kobayashi, T. Shirasawa, "Age and Sex-Dependent Alterations in Heart Rate Variability Profiling the Characteristics of Men and Women in Their 30s," *ANTI-AGING MEDICINE*. 2010; 7: 94-99, 2010.
- [9] R. K. Tripathy, A. Acharya, S. K. Choudhary, "Gender Classification from ECG Signal Analysis using Least Square Support Vector Machine," *American Journal of Signal Processing*, 2(5): 145-149, 2012 doi: 10.5923/j.ajsp.20120205.08.
- [10] I. Kaneko, J. Hayano, E. Yuda, "How can gender be identified from heart rate data? Evaluation using ALLSTAR heart rate variability big data analysis," *BMC Research Notes* 16, Article number: 5, 2023, doi:10.1186/s13104-022-06270-2.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Development of a GSM-RC Automated Device for Measuring Mobile Communication Signal Strength and Meteorological Parameters

Giwa Abdulgafar Babatunde<sup>1</sup>, Ewetumo Theophilus<sup>1</sup>, Ojo Joseph. Sunday<sup>1\*</sup>, Adedayo Kayode David<sup>1</sup>, Owolabi Gbenga Ayodele<sup>2</sup>

<sup>1</sup>Department of Physics, Federal University of Technology Akure, 2340001, Nigeria

<sup>2</sup>Engineering and Maintenance Unit, NNPC Retail Ltd, Abuja, Nigeria

### ARTICLE INFO

Article history:

Received: 13 December, 2023

Revised: 27 January, 2024

Accepted: 28 January, 2024

Online: 21 February, 2024

Keywords:

GSM-RC Device

GSM received signal strength

Meteorological parameters

Field test

### ABSTRACT

The automated Global System for Mobile Communication Signal Strength and Radio Climatological (GSM-RC) measuring device is an integration of different electronic sensors in a box for an in-situ measuring system. The sensor, data logging, and communication subsystems are integrated for transmitting information on meteorological parameters (MPs) and GSM signal strength level (SSL). The goal is to develop a device that could simultaneously measure MP and SSL of GSM communication systems in any location of interest. This is to reduce significant errors due to a lack of synchronization among multiple devices. To accomplish this objective, we designed an atmospheric sensing system with GSM SSL, temperature, relative humidity, and pressure sensors integrated as a GSM-RC unit. An Arduino microcontroller unit was used to wirelessly transmit the data collected by various sensors in each subsystem and stored on a micro-SD card. A statistical analysis of the SSL between the GSM-RC and the Samsung Galaxy A10s mobile reveals a correlation of roughly 0.99. The ANOVA analysis of variance demonstrates no noticeable distinction between the SSL from developed and conventional devices. The P-value is about 0.93, with  $\alpha$ -value of 0.05. The MPs were validated with a standard Vintage Pro weather station, and the data were statistically correlated with accuracies close to unity. A field test was carried out with the device to measure the SSL through the GSM and the selected MP in Akure from January to December 2022. The findings indicate a weak and poor correlation between temperature and signal strength, while relative humidity and pressure have a positive and weak correlations with the signal strength. This implies that an increase in SSL leads to a slight decrease in temperature while the relative humidity and pressure increase slightly. Other than being affordable in terms of production and deployment, the device has also solved the problem of labour-intensiveness arising from bulkiness.

### 1. Introduction

Over the years, when considering the impact of meteorological parameters (rain, temperature, pressure, relative humidity, etc.) on the received signal strength (RSS), scientists usually carried out research based on modeling and in-situ measurement of the meteorological parameter using separate devices for measurements [1]. Among the devices that can measure some of these parameters, there are both fixed and mobile types. Examples of such devices are the automatic weather station, Vintage Pro and Vantage Vue, and divers of rain gauges, among others. However, studies show that most of these equipments are expensive and bulky, and the results have always led to significant errors due to a lack of synchronization [2-3]. RSS is also measured with a signal strength

meter, a signal strength application installed on mobile phones, SATlink, and a field strength meter. In order to give room for synchronization, there is a need to develop a device that will address the aforementioned drawbacks in accounting for the exact and concurrent measurements of the GSM signal strengths and meteorological parameters.

The act of measuring involves expressing the relationship between two quantities that are directly accessible to the terminal element of measuring equipment. Electrical signals can be obtained directly when a variation of an intermediate quantity causes some modification of the intense qualities of the sensitive device, or indirectly when a physical quantity, such as speed, heat, etc., generates an electrical quantity, such as in thermo-electricity-conductivity, permeability, and permittivity [4]. Electronic methods, whose benefits are numerous and unquestionable, have

\*Corresponding Author: Ojo Joseph. Sunday, Department of Physics Federal University of Technology Akure, Nigeria, & ojojs\_74@futa.edu.ng

[www.astesj.com](http://www.astesj.com)

<https://dx.doi.org/10.25046/aj090115>



made measuring electrical signals simple. It is necessary to determine the quantities of electrical values using the proper measuring device. These tools enhance or add to human sense-making, perception, calculation, and evaluation abilities. The transducer, whose main purpose is to convert some physical variables to electrical variables, is one of the key components in this regard [4]. The current communication systems face challenges from the quest for ubiquitous connectivity due to some physical restrictions. Also, customers demanded their service providers to deliver good quality services. Therefore, systems that can be swiftly implemented and offer bandwidth-efficient communications must be produced by manufacturers. Instrumentation and measurements serve a crucial and essential part in achieving this goal [5] Rigid testing is carried out throughout the early stage of equipment development to evaluate system functioning and performance as well as to guarantee system interoperability. Additionally, the complexity of communication signals is putting additional stress on design teams, which already have a lot of work to do. The developer must not only conduct conformance testing but also swiftly determine the underlying reasons for potential technical issues using measurement results.

To properly emphasize the importance of measurement, it is helpful to refer to a common and well-known wireless communication system that uses radio frequency –RF [5]. The signal strength metre, also referred to as an electromagnetic radiation tester, is a special receiver designed to gauge the electromagnetic wave (EM wave) strength emitted by any transmitter operating within the same frequency bands and to show the output on a regulated scale, liquid crystal display (LCD) or light-emitting diode (LED) monitored [6]. The amount of carrier wave modulation and antenna directivity, the distance between the points of reception and transmission, the attenuation the signal experiences as it travels between the two points, the strength of interference at the receiving end, the fading caused by direct and indirect signals, and the quality of the signal are all factors needed for the ability of a signal strength meter to receive a satisfactory signal from a given broadcasting station at any given point [4].

Given the significance and complexity of electromagnetic waves, such as the frequency-modulated signal (FM signal), there is a need for a direct measuring device to identify their presence and strength. This would enable designers and technicians to evaluate transmitter performance and predict their projected range. Making a field test is one method of obtaining this conclusion, although this sometimes requires moving across different distances with a radio receiver and comparing attenuation levels. According to technical literature, a communication system is a setup made up of physical structures, various types of equipment (transmitters, receivers, and repeaters), and different add-ons or enhancements (encryption, security solutions, and interoperability or networking), intending to disseminate information following user needs. Each component must have a single goal, work together technically, follow similar practices, react to controls, and generally operate in harmony [7].

The majority of wireless localization systems employ either timing information or features of the received signal to estimate the distance to the positioning beacons, with RSSI being the most often used signal-related aspect. However, RSSI data are not always available from the hardware, or they are supplied with very low

granularity, limiting the positioning system's accuracy. Other signal properties could be employed to estimate the distance between the mobile unit and the positioning beacons to solve this problem [8]. Simulation techniques, empirical models, deterministic models, theoretical models, and stochastic models have all been used to study GSM signal strength and some weather parameters. However, these existing methods have flaws in accounting for exact and concurrent measurements of weather parameters and propagated GSM signal strength with the same real-time synchronizations. As a result of the need to address the aforementioned flaw, a GSM signal strength and radio climatological measuring device was developed to evaluate the quality of specific signal strength and instantaneous climatological parameters.

The loss in signal strength of a GSM signal should not be overlooked because it has a significant impact on the quality of the call received. Because some of the available devices could not account for both real-time and concurrent measurements of the signal strength under meteorological conditions, individual measurements of the parameters were used instead. To solve the aforementioned shortcomings, a portable, easy-to-repair, low-cost, and easy-to-deploy GSM signal strength and radio climatological measurement device is the goal of this work. The developed device will replace the time-consuming and inconvenient methods of measuring GSM signal strength and radio-climatological parameters individually.

## 2. Method and Materials

The goal of this study is to develop a cheap, miniaturized, and efficient device that simultaneously measures atmospheric parameters and received signal strength (RSS). In this section, an outline of the necessary hardware and software used in developing the GSM-RC automated device is highlighted. The hardware section analyzes in detail the mode of selection of various sensors, while the software section highlights the Arduino integrated development environment (IDE) and other software tools used during the design. Furthermore, the data collection procedures and field tests using the developed device are also presented. Figure 1 presents a block diagram of the design and operation of the automated GSM-RC device

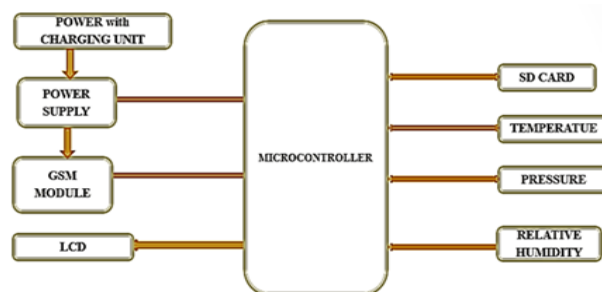


Figure 1: Systematic block diagram representation of the methodology adopted

### 2.1. GSM-RC major subsystem

The automated GSM-RC is composed of three main sections: the communication subsystem, the data logging subsystem, and the sensors subsystem. The sections cover the various components of each subsystem. The discussion is focused on how the components

used in this study function and operate. The GSM-RC assembly process is shown in Figure 2.

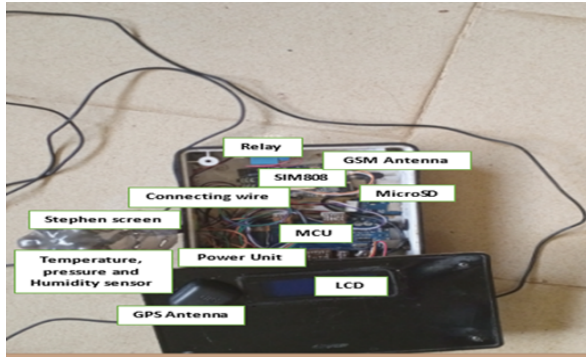


Figure 2: Assembly stage of the GSM-RC

### 2.2. Sensor subsystem

This system is made up of locally available electronic sensors that access physical features such as temperature, humidity, pressure, and GSM signal strength. The BME280 is a temperature sensor categorized as an integrated silicon temperature sensor. It is based on piezo-resistive technology, which has a semiconducting material (usually silicon) that changes the resistance when a mechanical force like atmospheric pressure is applied and can be integrated into the system over the inter-integrated circuit (I2C) interface. It has a  $\pm 2^\circ\text{C}$  accurate analog output temperature sensor with an operating voltage range of 3 V to 5 V and a wide temperature range of  $-40^\circ\text{C}$  to  $+85^\circ\text{C}$ . Also known as a pressure sensor, it acts as a transducer, generating an output signal as a function of pressure. It has a measuring range between 300 hPa and 1100 hPa. The HTU31 is a humidity sensor with an analog-output integrated circuit. It uses a laser-trimmed, thermoset polymer capacitive sensing element with on-chip integrated signal conditioning.

### 2.3. Data logging and Communication Subsystem

While the communication system assures reliable data transmission to the storage device, the data logging system enables accurate data collection, organization, transmission, and processing. This subsystem integrates a number of components to make the device conform to a normal mobile station and complete the required task. There is a SIM808 module that is integrated with a high-performance GSM/GPRS engine, a GPS engine, and a Bluetooth engine. SIM808 is a quad-band GSM/GPRS module that works on frequencies 850 MHz (GSM), 900 MHz (EGSM), 1800 MHz (DCS), and 1900 MHz (PCS), thereby obtaining the Received Signal Strength of the GSM-RC with respect to the network's frequency. In case the communication links stop, a microSD card module is provided as a storage device. The measured data can be kept on the micro-SD card for later analysis because radio signals can be easily interfered with during undesirable weather conditions, rendering transmission null [9]. At such times, there may be trouble with transmission to the ground station. All the components in the major subsystems' output data are in quantitative form and need to be analyzed. This is accomplished by using an Arduino ATmega 2560 microcontroller to wirelessly transmit the data collected by various sensors in each subsystem. During the SIM808's mission, this microcontroller creates the communication link between the SIM808 and the base transceiver station. Real-time data processing

and archiving are made feasible by the microcontroller that receives the data. Individually and with the aid of the Arduino IDE software, these modules and the circuits that go with them were created and tested. Following the implementation of the circuits, inferences might be made using the stored data.

### 2.4. Power sizing

The most accurate way for determining how much power is utilized by each component in various subsystems is power sizing. Making plans for the type of battery to utilise is made easier with this knowledge [10]. The formula  $P = IV$ , where P is the electric power in Watts, V is the voltage differential in Volts, and I is the current in mA, was used to extract data on power consumption for each component from multiple data sheets. The data obtained in Table 1 guides us in choosing the right battery to be used by the automated GSM-RC device subsystems. Another factor is the weight of the device, in that the battery chosen should be light.

Table 1: Voltage, current and power ratings of various components of the GSM-RC

Components	Voltage (V)	Current (mA)	Power (mW)
Arduino	5.0	20.0	100.00
GSM Module	12.0	44.0	528.00
Relay	5.0	20.0	100.00
HTU31	5.0	1.0	5.00
BME280	3.3	0.6	1.98
SD Card Shield	3.3	5.0	16.50
Total Power		90.6	751.48

### 2.5. Software design and programming

The Arduino IDE is a piece of open-source software used to code the Arduino ATmega board. Using the C and C++ programming languages, this software makes it simple to develop instructions to programme the various sensors. An onboard computer, or microcontroller, such as an Arduino board, takes input and transforms the input signal into a readable output. The primary logical decisions related to the instructions given by the Arduino IDE software are made by this microcontroller. The board has been coded to read and interpret the values of the sensor outputs. The processed data from these sensors will be displayed on the liquid crystal display as digital values.

### 2.6. Testing

Boiling water in a kettle on the stove produced a humid environment. Because of the abundant steam that is released into the air, the humidity has increased. As soon as water hits the boiling point, it starts to emit steam and re-evaporate back into the atmosphere. The sensor was placed in a humid environment to monitor its measurement, and the humidity level increased with time. The temperature value recorded by the temperature sensor drops to that of room temperature when it is moved away from the heat source. The temperature rises as it gets closer to the heat source. The microcontroller and SIM808 modules' LEDs light red when the power is turned on. When the sensing signal LED starts flashing, the module detector is pressed and held in place before being released, and launching a search for data requires more than 3 minutes. Thus, all exhibit appropriate behavior in their responses.

### 3. Results and Discussion

In this section, a number of data sets that were collected from the in-situ measurements of the automated GSM-RC are illustrated graphically and in Table format.

#### 3.1. Data from field test

The device was deployed to the field to ascertain the status of the real-time measurement. Table 2 presents a sample of a typical measurement on the field. Although not shown in Table 2, the measured result shows that the device can successfully measure the GSM signal strength up to -165 dBm and the selected radio climatological parameters [temperature (°C), humidity (%), and pressure (mbar)]. The device was also capable of measuring the number of base stations and both satellites in view and in operation. The logging time was rapid, with a response time of 30 seconds or less. The temperature sensor in use measures temperature with a resolution of 0.25 °C between -40 °C and 85 °C, and its maximum humidity range is about 99%.

#### 3.2. Validating the device measurement with an existing equipment

The data obtained was validated with the existing equipment. A Vantage Vue Integrated Sensor Suite (ISS) has been installed at the Physics Department, Federal University of Technology Akure, Nigeria. The equipment measures, among others, meteorological parameters such as temperature, pressure, and relative humidity. They were measured with the DAVIS Vantage Vue Integrated Sensor Suite (ISS) weather station at 1- minute integration time and co-located with the developed device. The data were respectively stored in the data loggers of both the developed and the DAVIS Vantage Vue Integrated Sensor Suite (ISS) weather station devices, from where they were downloaded and extracted. For performance purposes, the sensor temperature, relative humidity, and pressure of the developed device were enclosed in a prototype Stephen screen against precipitation and direct heat radiation from outside sources while still allowing air to circulate freely around them. This is done to validate the accuracy of the sensors. The responses obtained are presented in Figures 3–5 for temperature, relative humidity, and pressure, respectively. The characteristic equations obtained for each of the parameters were shown in equations (1) – (3), respectively:

For Temperature (BME280),

$$T_p = 0.7624T_m + 17.0991 \quad (1)$$

For HTU31,

$$RH_p = 0.6433RH_m + 16.489 \quad (2)$$

For Pressure,

$$P_p = 0.5275P_m + 498.38 \quad (3)$$

where  $T_p, RH_p$  and  $P_p$  are the predicted temperature, relative humidity and pressure, while  $T_m, RH_m$  and  $P_m$  are the measured temperature, relative humidity and pressure. A correlation analysis

was conducted, resulting in a correlation coefficient of 0.8324 for relative humidity, 0.9827 for temperature, and 0.9519 for pressure between the developed and standard devices. In addition, the signal strength of the GSM/GPRS device was correlated with a Galaxy A10s. Table 3 shows the received signal strength measurements taken at the Postgraduate Physics Laboratory, the Federal University of Technology, Akure, Nigeria at 10-minute intervals. The correlation coefficient between the devices is close to unity (0.99), indicating a strong agreement. The signal strength received between the two devices show no noticeable distinction based on the ANOVA test as presented in Table 4. The p-value of 0.93 is higher than the significance level ( $\alpha$ ) of 0.05, suggesting no statistically significant variance.

#### 3.3. Statistical Validation

To prove the validity of the analytical method, statistical analysis of the data collected during the validation was also presented. The computation of the correlation ( $R^2$ ), root mean square error (RMSE), mean absolute error (MAE), and standard deviation (S.D) are the main variables utilized for the interpretation of the results of analytical technique validation. Correlation ( $R^2$ ) was estimated using:

$$R^2 = 1 - \frac{RSS}{TSS} \quad (4)$$

where  $R^2$  is the coefficient of determination, RSS is the sum of squares residual and TSS is the total sum of squares.

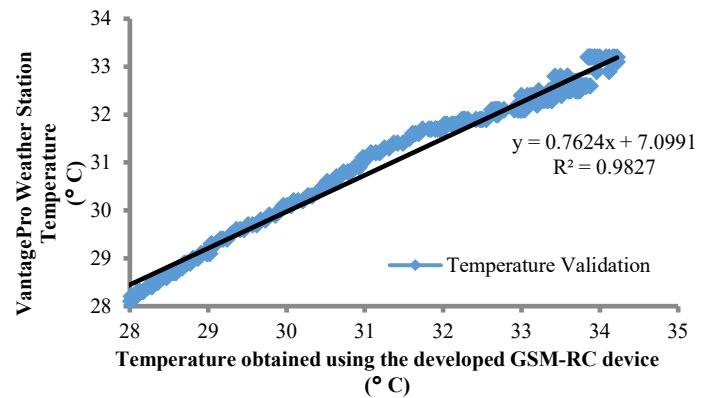


Figure 3: The validation of the temperature for the GSM-RC Device with Vantage Pro Weather Station

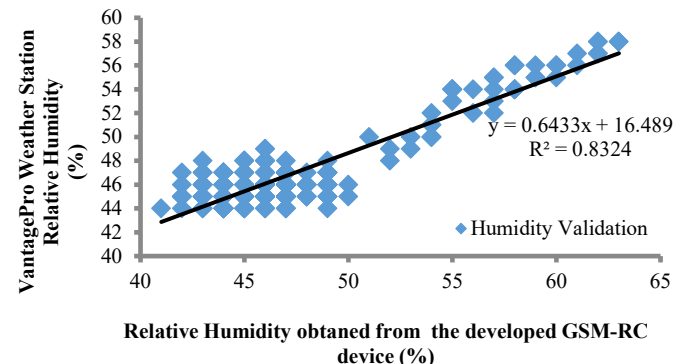


Figure 4: The validation of the relative humidity of the GSM-RC Device with Vantage Pro Weather Station

Table 2: Data obtained from the developed measuring device during the performance evaluation

Time	Signal Quality	Signal Strength (dBm)	Sat. View	Sat. Used	Lat. (°)	Long. (°)	Temp. (°C)	RH (%)	Pressure (mbar)	Heat Index (°C)	Radio Refractivity (N-units)
23:26:30	11	-92	8	7	7.2989	5.1343	27	53	968.11	27.60	328.52
23:26:56	11	-92	9	8	7.2989	5.1343	27	53	968.08	27.60	328.51
23:27:24	11	-92	8	7	7.2989	5.1343	27	53	968.14	27.60	328.53
23:27:53	11	-92	8	7	7.2989	5.1343	27	53	968.07	27.60	328.51
23:28:19	11	-92	8	8	7.2989	5.1343	27	53	967.98	27.60	328.49
23:28:45	10	-94	7	7	7.2989	5.1343	27	53	968.08	27.60	328.51
23:29:13	10	-94	8	7	7.2989	5.1343	27	53	968.02	27.60	328.50
23:29:43	10	-94	9	6	7.2989	5.1343	27	52	968.08	27.54	327.04
23:30:12	10	-94	9	7	7.2989	5.1343	27	53	968.07	27.60	328.51
23:30:41	10	-94	9	8	7.2989	5.1343	27	53	968.04	27.60	328.50
23:31:09	11	-92	8	8	7.2989	5.1343	27	53	968.06	27.60	328.51
23:31:35	10	-94	9	8	7.2989	5.1343	27	53	968.00	27.60	328.49
23:32:02	11	-92	10	9	7.2989	5.1343	27	53	968.08	27.60	328.51
23:32:29	11	-92	9	8	7.2989	5.1343	27	53	968.06	27.60	328.51
23:32:55	10	-94	9	8	7.2989	5.1343	27	53	967.93	27.60	328.47
23:33:23	10	-94	7	6	7.2989	5.1343	27	53	967.95	27.60	328.48
23:33:51	11	-92	8	7	7.2989	5.1343	27	53	967.94	27.60	328.48
23:34:19	10	-94	9	8	7.2989	5.1343	27	53	967.92	27.60	328.47
23:34:48	10	-94	9	8	7.2989	5.1343	27	53	967.95	27.60	328.48
23:35:15	10	-94	9	8	7.2989	5.1343	27	53	967.88	27.60	328.46
23:35:44	11	-92	8	8	7.2989	5.1343	27	53	967.92	27.60	328.47
23:36:11	11	-92	9	8	7.2989	5.1343	27	53	967.94	27.60	328.47
23:36:39	11	-92	8	7	7.2989	5.1343	27	53	967.93	27.60	328.47
23:37:06	11	-92	9	8	7.2989	5.1343	27	53	967.97	27.60	328.48
23:37:34	10	-94	9	8	7.2989	5.1343	27	53	967.89	27.60	328.46
23:38:03	11	-92	8	7	7.2989	5.1343	27	53	967.98	27.60	328.49
23:38:32	9	-96	9	8	7.2989	5.1343	27	53	967.87	27.60	328.46
23:39:00	11	-92	8	7	7.2989	5.1343	27	53	967.88	27.60	328.46



Table 3: Comparison of the received signal strength between the GSM-RC and Samsung Galaxy A10s.

Time interval (minutes)	GSM-RC RSS (dBm)	Samsung Galaxy A10s RSS (dBm)
0	-80	-79
10	-80	-79
20	-70	-71
30	-72	-71
40	-56	-56
50	-76	-75
60	-76	-76
70	-74	-74
80	-66	-67
90	-65	-66
100	-77	-78
110	-73	-74
120	-61	-63
	correlation	0.992291

Table 4: The ANOVA results of the signal strength received between GSM-RC and Galaxy A10s

ANOVA: Single Factor						
OUTPUT						
Groups	Number	Addition	Mean	Variance		
GSM-RC device	13	-926	-71.2308	54.02564		
Galaxy A10s	13	-929	-71.4615	46.9359		
ANOVA trend	SS	diff	MS	F	P-value	F crit
Between Groups	0.346154	1	0.346154	0.006857	0.934691	4.259677
Amidst Groups	1211.538	24	50.48077			
Summation	1211.885	25				

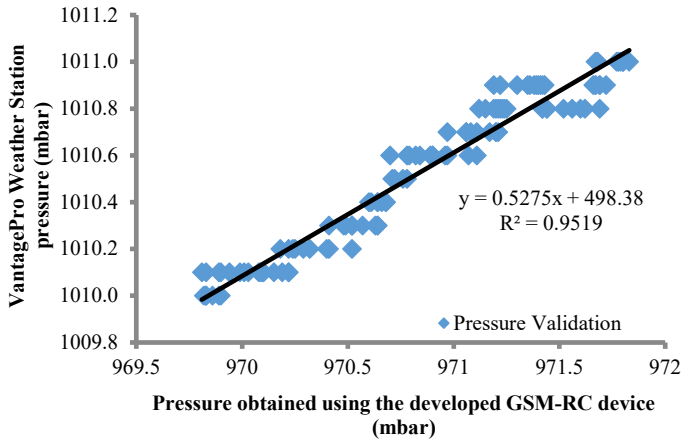


Figure 5: The validation of the pressure for the GSM-RC Device with Vantage ProWeather Station

The Root Mean Square Error (RMSE) was estimated using:

$$RMSE = \sqrt{\frac{\sum(x_i - \hat{x}_i)^2}{N}} \tag{5}$$

where RMSE is the root mean square error,  $x_i$  is the actual observation of time series,  $\hat{x}_i$  is the estimated time series and N is the number of non-missing data points. The Mean Absolute Error (MAE) was also calculated using:

$$MAE = \frac{\sum|y_i - x_i|}{n} \tag{6}$$

where MAE is the mean absolute error,  $y_i$  is the prediction,  $x_i$  is the true value and n is the total number of data points.

The standard Deviation (S.D) was calculated using:

$$S = \sqrt{\frac{\sum(x_i - \bar{x})^2}{n-1}} \tag{7}$$

where  $x_i$  is the actual observation,  $\bar{x}$  is the mean and n is the total number of observations. A statistical software programme was used to carry out these calculations. Table 5 shows the result obtained from the statistical validation of the radio climatological parameters between the VantagePro weather station and the developed device.

Table 5: Statistical validation of measured data for Radio climatological parameters between the developed GSM-RC device and the Standardize Vantage Pro

Parameters	$R^2$	RMSE	MAE	S.D
Temperature (T)	0.983	0.181	2.499	3.144
Relative Humidity (RH)	0.832	0.350	5.391	9.994
Pressure (P)	0.957	1.917	4.917	1.554

### 3.4. Seasonal analysis of the measurement obtained from the Field

This section discusses the seasonal analysis of the data obtained using the developed device. Each of the weekly hourly

variables was averaged and presented for comparison over the months of the year. GSM signal strength, temperature, and relative humidity were accessed and further evaluated to establish the relationship between them. April, July, and January were chosen to represent the commencement, intense, and dry seasons, respectively. For the MTN network during the commencement of the rainy season, intense rainy season, and dry season, the GSM signal strength was correlated with the temperature, pressure, and relative humidity

### 3.5. Influence of Atmospheric parameter on RSS during the commencement of rainy season (April)

The correlation among signal strength, temperature, humidity, and pressure on a typical day in April is shown in Figures 6 (a-c). Data gathered daily was used to analyze the information for 24 consecutive hours. Since the correlation follows the same pattern, only the reading taken on April 1, 2022, is presented. Figure 6a shows that while the relative humidity profile is high at night, the values begin to decline in the middle of the day. In Figure 6b, during the day, temperatures typically oscillate, which can cause changes in atmospheric pressure. As observed in the morning hours (Figure 6b), when the temperatures are cooler, atmospheric pressure is higher than during the hot afternoon hours when the air is less dense. Similarly, at night, when temperatures cool again, atmospheric pressure tends to increase. In the early hours, as shown in Figures 6a and 6c, the humidity is high but the pressure profile is low. This could be influenced by the fact that water molecules are higher than oxygen and nitrogen molecules, which are the primary components of the atmosphere. As the amount of water vapor in the air (humidity) increases, the air becomes less dense, which may lead to a decrease in atmospheric pressure observed during those hours [11].

It was also observed that there are empirical relationships between the GSM signal strength, temperature, relative humidity, and pressure. The empirical correlations between the RSS, temperature, relative humidity, and pressure are presented in Table 6. The developed linear relationships are:

$$RSS_T = m_1T + C_1, \tag{8}$$

for signal strength in dBm and temperature in °C;

$$RSS_{RH} = m_2RH + C_2 \tag{9}$$

for signal strength and relative humidity in %; and

$$RSS_P = m_3P + C_3 \tag{10}$$

for signal strength and pressure in mbar

Generally, the results presented in Table 6 show very weak correlations between signal strength, temperature, relative humidity, and pressure. It also shows a varying positive and negative correlation for each of the examined days in April. Previous researchers have also examined the variance or correlation between signal strength and meteorological parameters [3], [12-13]. The degree of linear dependence of GSM

RSS on temperature, relative humidity, and pressure has also been quantified using the Pearson correlation coefficient to provide a better understanding.

On signal strength, temperature has a small but significant impact except for the third and fifth days, which are positive, with all the relationships showing a linear, negative trend. In terms of impact

size, there are various variances as well (regression coefficient). As presented in Table 6, there is a low and negative correlation between GSM RSS and temperature, while the correlation coefficient varies. This could be due to the characteristics of the measurement site, as reported by [14], [15], and [16]. Additionally, on the selected days, the GSM RSS was correlated with the temperature.

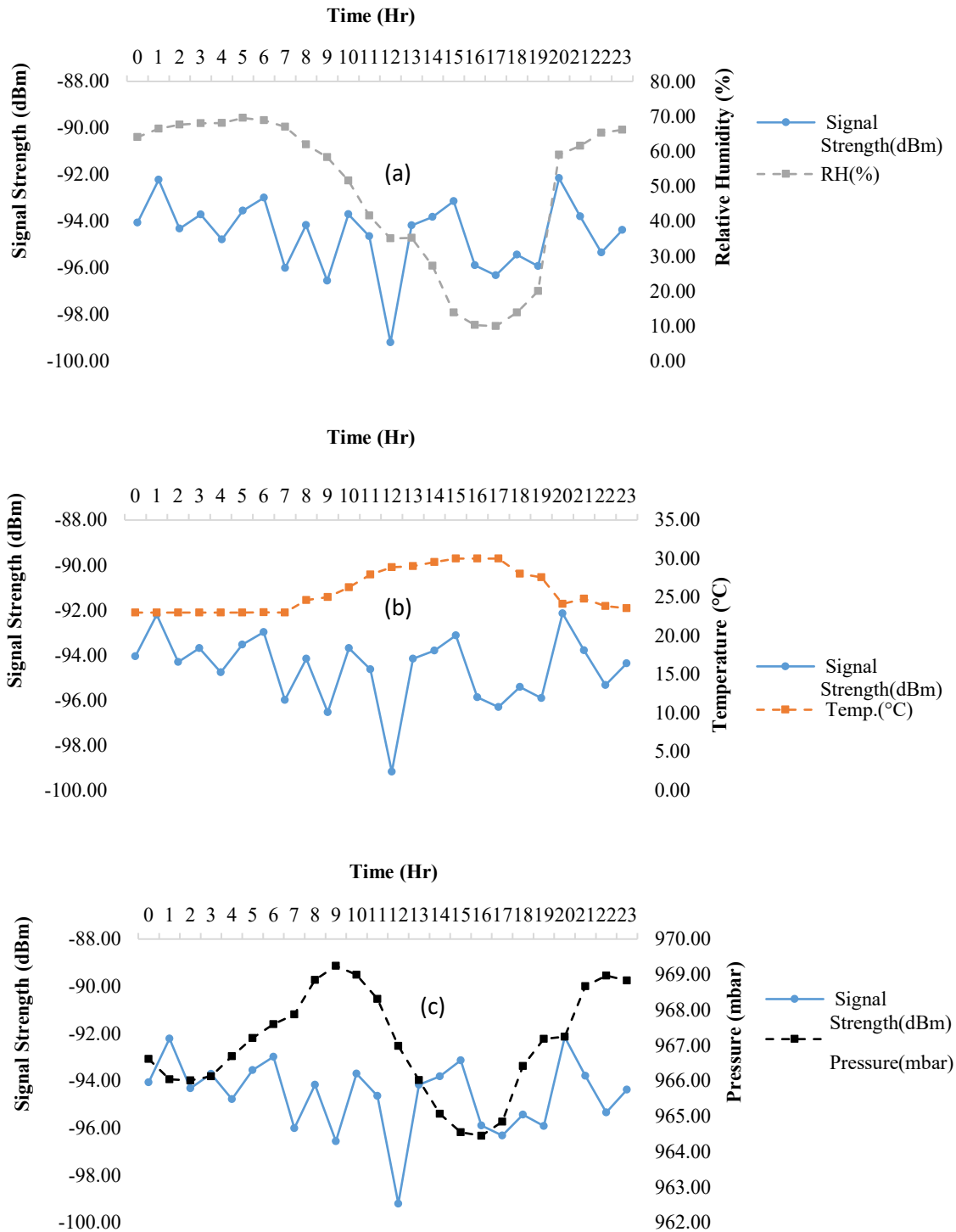


Figure 6: Mean hourly variation of GSM signal strength with (a) relative humidity (b) temperature and (c) pressure on a typical day (01-04-2022)

Table 6: Empirical correlation between GSM RSS and the selected atmospheric parameters

Days	T, RH and Pressure regressed on RSS	Slope ( $m_1, m_2$ and $m_3$ )	Intercept ( $c_1, c_2$ and $c_3$ )	correlation, coefficient, r
01-04-22	T	-0.2120	-37.829	-0.3774
	RH	0.0260	-95.859	0.3661
	Pressure	-0.0537	-42.7110	-0.0510
02-04-22	T	-0.0319	-90.594	-0.0646
	RH	0.0076	-91.761	0.1268
	Pressure	0.2515	-335.000	0.2818
03-04-22	T	0.4298	-102.201	0.2803
	RH	-0.0741	-87.930	-0.4089
	Pressure	-0.6151	503.459	-0.3353
04-04-22	T	-0.2091	-89.853	-0.2052
	RH	0.0223	-95.895	0.1724
	Pressure	0.5051	-583.801	0.3544
05-04-22	T	0.1562	-96.942	0.2115
	RH	-0.0207	-92.111	-0.2664
	Pressure	-0.0740	-21.4267	-0.0593
06-04-22	T	-0.0870	-90.189	-0.1138
	RH	0.0085	-92.768	0.1049
	Pressure	-0.4817	-379.316	-0.3722
07-04-22	T	-0.4309	-84.147	-0.4159
	RH	0.0543	-97.575	0.4912
	Pressure	0.2245	-312.466	0.1504

The third and fifth days, which have positive and low network correlation values, are an exception. For the selected days in April 2022, there is a weak but positive correlation between relative humidity and GSM RSS. This indicates that there may be additional parameters influencing it as it varies and may probably be caused by the humid nature of the season as reported in the work of [14] and [15]. GSM RSS and relative humidity exhibit a very weak negative correlation on the third and fifth days. Pressure when correlated with GSM RSS shows a very low correlation for

the first, fifth and seventh days while the second, third, fourth and sixth days show a moderate correlation.

The peak (maximum) and dip (minimum) trend of the GSM RSS data are also shown in Table 7 together with the temperature, relative humidity, and pressure. Table 7 shows that the signal strength fluctuates for the network, having its maximum and minimum values between -87.14 (03-04-22) and -102.50 (03-04-22). Similar to this, analogue temperature and relative humidity variations were converted to digital values to obtain their



corresponding varying values. For temperature, the peak and least values were 31.99 °C (02-04-2022) and 22.70 °C (06-04-2022). Humidity has peak values of 74.80% (07-04-2022) and a least value of 9.00 (02-04-2022 and 03-04-2022) while pressure has a peak value of 971.29 mbar (02-04-2022) and a dip value of 962.49 mbar (03-04-2022). In particular, 07-04-2022 has the least value for temperature, and the highest peak value for humidity, and the signal strength fluctuation is significant. This might be due to the transition from the dry to the wet seasons, when the atmosphere gradually becomes moister as the temperature drops, the humidity value increases. This might prompt scattering of the radio signal and affects the signal quality.

### 3.6. Influence of Influence of Atmospheric parameter on RSS during Intense rainy season (July)

The correlation of signal strength with temperature, relative humidity, and pressure on a typical day during July is shown in Figures 7 (a–c). The analysis is done for 24 consecutive hours on intense rainy days. Since the correlation follows the same pattern, only the reading taken on July 1, 2022, is presented. The result presented in Figure 6a shows that the relative humidity profile is strong and essentially consistent during the chosen days. When it is foggy, daytime recorded data show that the humidity profile is often high because of the low temperature. The average temperature and humidity in cloudy climates are 27.73 °C and 84.04%, respectively, as presented in Figures 7a and 7b. This could be due to the decrease in temperature during the rainy season as there is an increase in cloud cover. As more clouds form in the atmosphere, they can reflect a greater amount of sunlight back into space, which can lead to a decrease in temperature at the Earth's surface. Additionally, as rain falls from the clouds, it can cool the air, leading to further decreases in temperature. At the same time, higher values of humidity were observed. This is due to the increase in moisture in the air, which is caused by several factors such as increased evaporation from bodies of water, increased transpiration from plants, and the influx of moist air masses from other regions. The implication is that during this condition, the nature of the radio signal will be weak due to the high humidity.

It was also observed that the GSM signal strength, temperature, relative humidity, and pressure all exhibit empirical connections, as presented in Table 8. The RSS, temperature (T), relative humidity (RH), and pressure (P) all showed a linear relationship in the result. The developed linear relationships are in agreement with equations (8)–(10).

For the selected days in July, Table 8 generally reveals very weak correlations between signal strength, temperature, relative humidity, and pressure. It also exhibits a periodic positive and negative correlation for each of the selected days. Based on the Pearson correlation coefficient, the linear dependence of the GSM RSS with temperature, relative humidity, and pressure was evaluated

The correlation between pressure and the RSS slightly varies with an increase in signal strength. As presented in Table 8, the first and second days show a negative trend, while the fifth, sixth, and seventh days show a linear but weak trend. The temperature values have a negligible but noticeable impact, as seen on the third

day. Further analysis shows that except for the first day, which displays a weak negative correlation, other days show a weak positive trend. This is in agreement with the work of Christian and Michael [17], also in July. The implication is that the trend may be influenced by other factors due to the nature of the period [14–15]. Similarly, the sixth day in July has a positive link with the other days. However, on the second, third, and fourth days of the selected days, there exists a weakly negative correlation.

The temperature, relative humidity, and pressure are all displayed in Table 9, along with the peak and dip of the GSM RSS data. Table 9 shows that the signal strength fluctuates for the network, with a peak value at -89.11 (04-07-22) and the least value at -96.46 (02-07-22).

Similar to this, analog temperature and relative humidity variations were converted to digital values to obtain their corresponding varying values. For temperature, it has its peak value at 29.96 °C (07-04--2022) and its least value, at 24.20 °C (07-01--2022), while relative humidity has its peak value at 89.40 % (01-07--2022) and the least at 69.65 (02-07--2022 and 05-07-2022).

For pressure, the peak was observed at 973.37 mbar (06-07-2022) and the least value at 968.70 mbar (04-07-2022). The highest peak value of the GSM RSS, the least temperature value, the peak relative humidity value, and the lowest pressure value were all observed on July 4, 2022, in particular. This could be due to the consistent increase in atmospheric moisture as the rainy season intensifies and the temperature decreases. Another reason could be as a result of relative humidity, which is constantly higher during the rainy season than actual rainfall.

### 3.7. Influence of atmospheric parameters on RSS during the dry season (January)

The correlation of signal strength with temperature, relative humidity, and pressure on a typical day during January is also shown in Figures 8 (a–c). The information in Figures 8 (a–c) was analyzed using data that was collected for 24 consecutive hours daily. Figures 8a and 8b show that the relative humidity profile is at its minimum and the temperature profile is at its maximum on the chosen days. This could be related to how the earth responds to solar insolation, which is a primary driving force behind the weather conditions that were being experienced when the Harmattan season was at its peak. Since the moisture content is lower during this time, light will move at a faster rate [18]. The Harmattan season is connected to the northeasterly wind that crosses the Sahara Desert while there is a dry climate [19]. Another factor that can cause an increase in temperature during the dry season is the absence of rain or cloud cover. As there is little to no precipitation, the sun's energy is able to penetrate the atmosphere more directly, which can lead to a greater amount of solar radiation being absorbed at the Earth's surface. This could result in higher temperatures, particularly in regions with a high incidence of sunlight. At the same time, the dry season is characterized by a decrease in humidity. This is because there is little to no moisture in the air, and there is little to no precipitation to introduce moisture into the atmosphere. Additionally, any moisture that is present may be quickly evaporated due to the high temperatures and low relative humidity.

Table 7: Summary of variation of GSM RSS and Radio climatological parameters during the Commencement of rainy season (April).

Day	Signal Strength (dBm)		Temperature (°C)		Humidity (%)		Pressure (mbar)	
	Highest Peak	Dip	Highest Peak	Dip	Highest Peak	Dip	Highest Peak	Dip
01-04-22	-92.15	-99.20	30.00	23.00	69.54	10.03	969.24	964.45
02-04-22	-89.45	-94.27	31.99	23.00	67.00	9.00	971.29	965.69
03-04-22	-87.14	-102.50	31.46	24.33	59.57	9.00	969.61	962.49
04-04-22	-91.08	-99.35	28.71	23.00	61.77	11.10	970.78	965.57
05-04-22	-90.82	-97.00	29.43	23.00	67.66	11.03	970.25	966.22
06-04-22	-90.62	-97.25	29.43	22.70	67.05	11.24	970.33	966.08
07-04-22	-91.44	-100.37	30.79	22.99	74.80	9.85	970.73	964.63

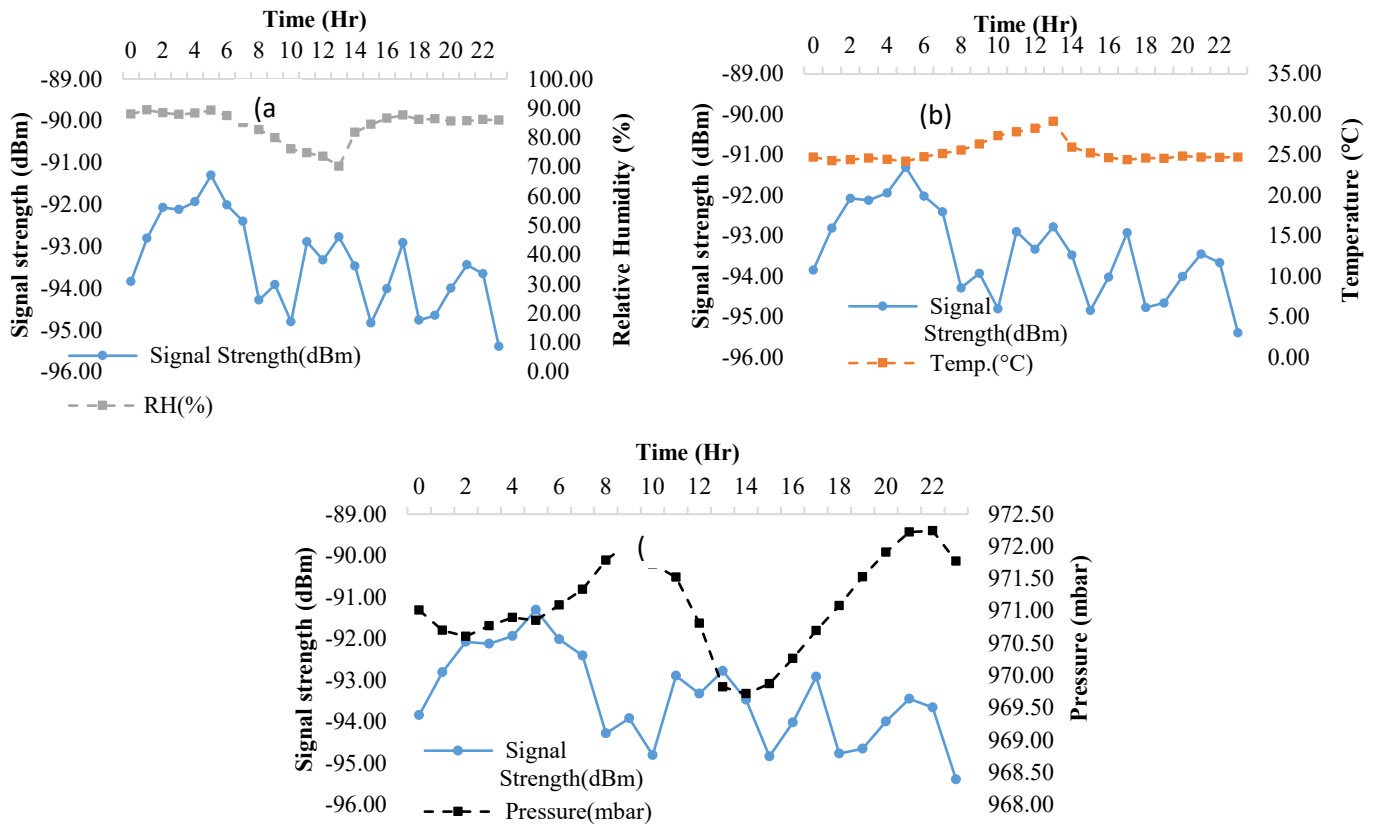


Figure 7: Mean hourly variation of GSM signal strength with (a) relative humidity (b) temperature (c) pressure on a typical day (01-07-2022)

Table 8: Empirical correlation between GSM RSS and the selected atmospheric parameters

Days	T, RH and Pressure regressed on RSS	Slope ( $m_1, m_2$ and $m_3$ )	Intercept ( $c_1, c_2$ and $c_3$ )	correlation, coefficient, $r$
01-07-22	T	-0.0730	-91.550	-0.0924
	RH	0.0371	-96.526	0.1809
	Pressure	0.6923	290.609	-0.2662

02-07-22	T	0.1984	-99.374	0.3053
	RH	-0.0702	-88.399	-0.3668
	Pressure	-0.2381	137.209	-0.2726
03-07-22	T	0.9808	-118.947	0.4130
	RH	-0.0353	-89.278	-0.0614
	Pressure	-0.3343	232.636	-0.2069
04-07-22	T	0.0398	-91.254	0.0734
	RH	-0.0529	-86.017	-0.3016
	Pressure	-0.1069	13.584	-0.1565
05-07-22	T	-0.1861	-87.137	-0.1277
	RH	0.0973	-99.985	0.1566
	Pressure	0.0669	-157.115	0.0350
06-07-22	T	-0.7069	-73.671	-0.4410
	RH	0.2318	-111.378	0.4203
	Pressure	-0.0413	-52.352	-0.0303
07-07-22	T	0.0092	-92.408	0.0138
	RH	0.0074	-92.763	0.0310
	Pressure	0.0319	-123.132	0.0446

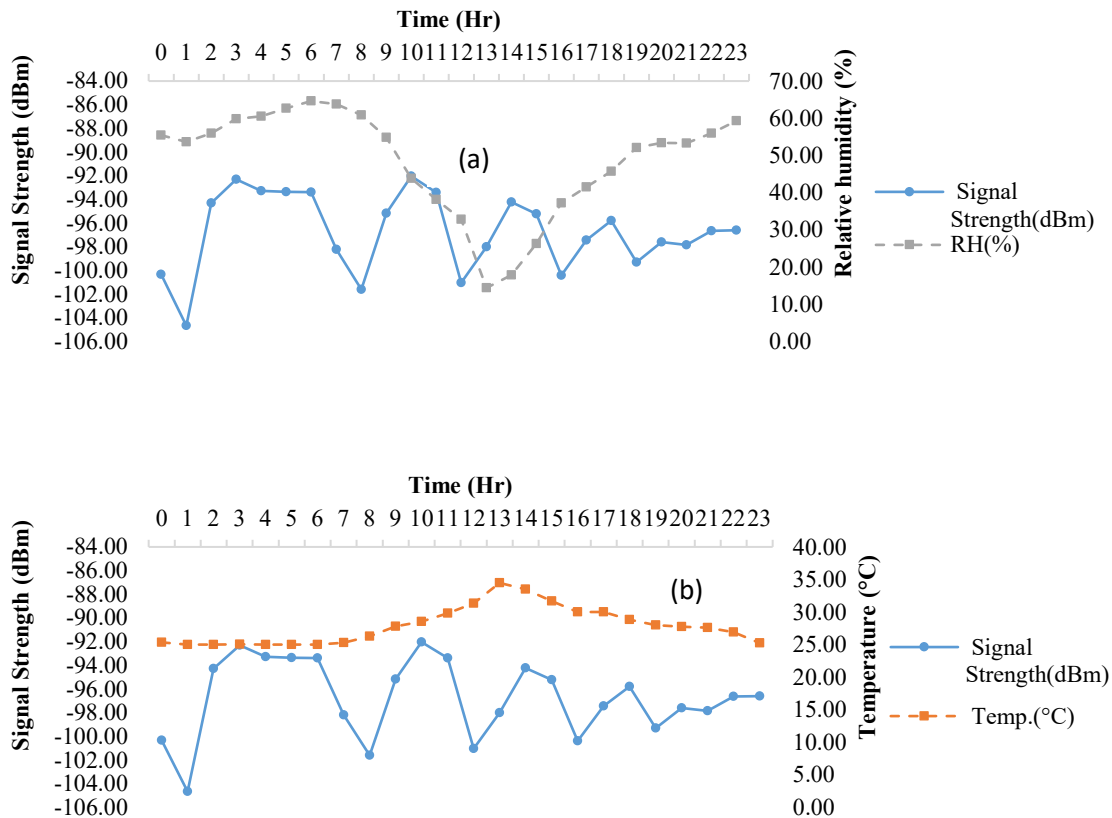
Table 9: Summary of variation of GSM RSS and Radio climatological parameters during the intense rainy season (July).

Day	Signal Strength (dBm)		Temperature (°C)		Humidity (%)		Pressure (mbar)	
	Highest Peak	Dip	Highest Peak	Dip		Highest Peak	Dip	Highest Peak
01-07-22	-91.31	-95.40	29.17	24.20	01-07-22	-91.31	-95.40	29.17
02-07-22	-92.33	-96.46	29.62	24.50	02-07-22	-92.33	-96.46	29.62
03-07-22	-89.56	-95.74	28.86	26.33	03-07-22	-89.56	-95.74	28.86
04-07-22	-89.11	-91.74	29.96	26.45	04-07-22	-89.11	-91.74	29.96
05-07-22	-89.43	-95.67	28.81	25.60	05-07-22	-89.43	-95.67	28.81
06-07-22	-90.00	-96.19	28.93	25.33	06-07-22	-90.00	-96.19	28.93
07-07-22	-91.14	-94.31	28.47	24.87	07-07-22	-91.14	-94.31	28.47

Also in Figure 8a, the variation of GSM signal strength did not correlate with the low relative humidity readings that were recorded on this day. This could be due to the Harmattan dust haze and other pollutants in the atmosphere during this month. Moreover, attenuation by scattering is greater for relatively large wavelengths of GSM signal strength compared to dust particle size, which might disturb the atmosphere and radio wave propagation. This is consistent with the findings by [19] and earlier researchers who also observed radio signal obstruction during Harmattan and an absence of radio line-of-sight [20-22]. Since the correlation follows the same pattern, only the reading taken on January 1, 2022, is fully discussed. For the selected days in January, Table 10 generally reveals moderate correlations between signal strength, temperature, relative humidity, and pressure. It also exhibits a periodic positive and negative correlation for each of the selected days. Based on the Pearson correlation coefficient, the linear dependence of the GSM RSS with temperature, relative humidity, and pressure was evaluated. As the RSS increases, the correlation between pressure and RSS declines with slight changes. As shown in Table 10, the relationship between pressure and RSS was positively linear and moderate on the second and seventh days, except for the other days, which were weak. Apart from the first day (onset), temperature and humidity show a clear link with RSS throughout the days. Possibly, this was influenced by how the planet reacts to solar insolation, which is a major cause of the weather that was present. Further analysis shows that except for the third day, when

the humidity had a positive but relatively moderate correlation with RSS for all the selected days, there was a negative but moderate correlation between temperature and RSS across the selected days. Nonetheless, different correlation coefficient values can exist, as reported by [14], [16], and [23].

Table 11 shows that the signal strength fluctuates for the network, with the peak value and least value between -52.00 dBm (06-01-22 and 07-01-2022) and -104.67 dBm (01-01-22), respectively. Similar to this, analog temperature and relative humidity variations were converted to digital values to obtain their corresponding varying values. For temperature, its peak value was observed at 36.16 °C (06-01-22 and 07-01-2022), and the least value was 22.11 °C (04-01-2022). The highest peak value for relative humidity was noticed at 87.82% (06-01-22 and 07-01-2022), and the least value was 07.24 (05-01-2022). For pressure, its peak value was 973.05 mbar (01-01-2022) and its dip was 965.94 mbar (05-01-2022). The study discovered that on the selected days in January, this season is characterized by high temperatures (36.16°C) and low relative humidity values (7.24%). This could be as a result of the Harmattan period, during which the amount of atmospheric moisture decreases and the temperature of the environment rises. Also, these could be linked to the way the planet reacts to the amount of solar radiation that strikes the earth's surface as electromagnetic energy, which is a significant factor in the weather conditions observed. Since the moisture content is lower during this time, light will move at a faster rate. This could explain the peak value experienced for signal strength (-52.00 dBm) during the season.





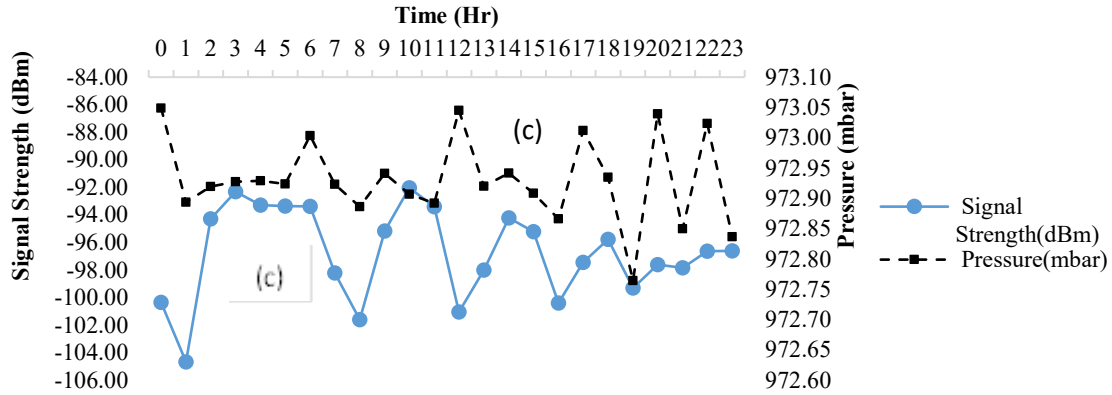


Figure 8: Mean hourly variation of GSM signal strength with (a) relative humidity (b) temperature (c) pressure on a typical day (01-01-2022)

Table 10. Empirical correlation between GSM RSS and the selected atmospheric parameters

Days	T, RH and Pressure regressed on RSS	Slope ( $m_1, m_2$ and $m_3$ )	Intercept ( $c_1, c_2$ and $c_3$ )	correlation, coefficient, r
01-01-22	T	-0.0415	-95.6028	-0.0361
	RH	0.0090	-97.1964	0.0391
	Pressure	1.4132	-1471.80	0.0301
02-01-22	T	-2.6253	-10.1192	-0.4670
	RH	0.1975	-87.5551	0.4336
	Pressure	3.8263	-3787.00	0.5734
03-01-22	T	0.7956	-98.0658	0.3890
	RH	-0.0811	-74.1352	-0.4094
	Pressure	-0.6970	597.890	-0.2190
04-01-22	T	-0.8650	-70.6787	-0.3681
	RH	0.1118	-98.4487	0.3480
	Pressure	-0.4120	305.883	-0.0660
05-01-22	T	-0.3410	-86.8400	-0.4880
	RH	0.0544	-98.4200	0.4960
	Pressure	0.2199	-309.200	0.1258
06-01-22	T	-0.5695	-40.6099	-0.6886
	RH	0.1389	-66.7231	0.6630
	Pressure	-0.0110	-46.8500	-0.0040
07-01-22	T	-0.2548	-48.2391	-0.3328
	RH	0.0744	-60.7671	0.3863
	Pressure	1.7714	-17772.9	0.5807

Table 11: Summary of variation of GSM RSS and Radio climatological parameters during the intense rainy season (July).

Day	Signal Strength (dBm)		Temperature (°C)		Humidity (%)		Pressure (mbar)	
	Highest Peak	Dip	Highest Peak	Dip	Highest Peak	Dip	Highest Peak	Dip
01-01-22	-92.04	-104.67	34.47	25.00	64.63	14.40	973.05	972.76
02-01-22	-66.37	-104.50	29.00	24.43	57.85	10.73	970.73	966.66
03-01-22	-69.43	-82.51	30.00	23.50	62.37	09.00	971.24	967.41
04-01-22	-71.66	-100.45	32.05	22.11	77.45	08.95	972.04	968.39
05-01-22	-88.61	-99.79	33.99	23.00	68.77	07.24	971.75	965.94
06-01-22	-52.00	-62.55	36.16	24.69	87.82	41.48	971.67	967.64
07-01-22	-52.00	-62.40	36.16	24.63	87.82	41.48	971.16	967.64

#### 4. Conclusion

This work has developed a handy and locally-based device that can simultaneously measure the received signal strength and atmospheric parameters, rather than the cumbersome method of measuring them individually. The global system for mobile communication received signal strength and radio climatological measuring devices has been successfully developed. A statistical analysis of the received level of signal between the GSM-RC and the Samsung Galaxy A10s mobile reveals a correlation of roughly 0.99. The results of an ANOVA analysis of variance demonstrate that there is not a noticeable distinction between the signal strength values from the developed and conventional devices. The P (value) is determined to be 0.93 and higher  $\alpha$  (value) of 0.05. The values obtained were in good agreement with those obtained using a standard device. The developed instrument, which was validated against the standard device, measured relative humidity and temperature, with their accuracy close to unity. A field test using this device shows some correlations between the radio-climatological parameters for some selected seasons and the GSM RSS. The findings indicate a weak and poor correlation between temperature and GSM signal strength, while relative humidity and pressure are positive but weak correlation. RSS has the highest value during the dry season (January), when the atmosphere is mostly perturbed, while temperature variation is low in July and high in January. Pressure and humidity attained their peak values during the intense rainy season (July) and their lowest values during the dry season (January). The implication of this is that the pressure and susceptibility of the radio signal to losses increase with increasing humidity at the study location. The developed device has been found to be cheap, handy, and user-friendly.

#### Conflict of Interest

All authors declare no conflict of interest in this paper.

#### References

- [1] J. S Ojo, A. Ayegba, A.T Adediji "Impact of Atmospheric Parameters and Noise Temperature on Digital Terrestrial Television Signal Strength over Karshi Area, Abuja, North-Central, Nigeria", IOP Conference Series: Earth and Environmental Science, 2021, doi: 10.1088/1755-1315/665/1/012048
- [2] J. J. Popoola, I.O Megbowon, V.S Adeloje V. S. (2009). Performance Evaluation and Improvement on Quality of Service of Global System for Mobile Communications in Nigeria. Journal of Information Technology Impact (JITI), **118**, 52–60, 2009.
- [3] U. U Abraham., U.O Okpo., E.O Elijah. "Instantaneous GSM signal strength with weather and environmental factors". Amer J Eng Res: **4**, 104-115, 2015.
- [4] A. M. Zungeru, M. S. Ahmed, G. M Doko, D. Abdul-Azeez. "Modelling and Development of A Frequency Modulated Field Strength Indicator" International Journal of Emerging Trends in Engineering and Development, **1(4)**, 192-205, 2014
- [5] L. Angrisani, D. Petri, M. Yeery, "Instrumentation and Measurement in Communication Systems". IEEE Instrumentation and Measurement Magazine, **18(2)**, 4–10. 2015, DOI: 10.1109/MIM.2015.7066676
- [6] B.L. Therajar, A.K Theraja, A testbook of electrical technology. S. Chand & Company Ltd & Ram Nagar, New Delhi, 2005.
- [7] M. H. Weik., Communications Standard Dictionary, 2nd Edition, New York, NY, USA: Van Nostrand Reinhold Company Inc , 1983.
- [8] M. Beigl, C. Henning, T. R. Roth-Berghofer, A. Kofod-Petersen, K. R. Coventry, H. R. Schmidtke "Modeling and Using Context". Proceedings 7th International and Interdisciplinary Conference, CONTEXT 2011, Karlsruhe, Germany, September 26-30, 2011..
- [9] G. F. Sa., Lessons Learned from Smart Machines-CTIC-UNI Program for Developing for Capacity Building in Peruvian Undergraduates and Perspectives in The Near Future. Smart Machines-CTIC-UNI Program, 2021.
- [10] D. C Marubin, S. S Yi, "Development of Appropriate Power Distribution Design for Can-Sized satellite CanSAT" Journal of Advanced Industrial Technology and Application, **2(2)**, 1-10, 2021.
- [11] L. Ali, I. Alam, S. A. A. Shah, M. Yaqoob, "Various Meteorological Parameters Effect on GSM Radio Signal Propagation for a Moderate Area", International Conference on Frontiers of Information Technology. 115-120, 2017, DOI: 10.1109/FIT.2017.00028
- [12] B. G Ayatunji, B. Musa, H. Mai-Unguwa, L. A Sunmonu, A. S Adewumi, A. Sa'ad L Kado A, "Atmospheric humidity and its implications on UHF signal over Gusau, North West, Nigeria", Int J Sc Eng Res: **9**, 599-611, 2018..
- [13] J. Amajama 2016. "Impact of atmospheric temperature on UHF radio signal". Int J Eng Res Gen Sc: **4**, 619-622
- [14] j. Luomal, I. Hakala,.. "Effects of Temperature and Humidity on Radio Signal Strength in Outdoor Wireless Sensor Networks". Computer Science and Information Systems, **5**, 1247-1255, 2015. doi: https://doi.org/10.15439/2015F241
- [15] K.E Ukhurebor, O. J, O. I Umukoro "Influence of meteorological variables on UHF radio signal: recent findings for EBS, Benin City, South-South, Nigeria". 2nd Int Conf Sc Sust Dev, IOP Conf. Series: Earth and Enviro Sc:

pp. 179, 2018.

- [16] S. Sabu, S. Renimol, D. Abhiram, B. Premlet, "A study on the effect of temperature on cellular signal strength quality". 2017 International Conference on Nextgen Electronic Technologies: Silicon to Software (ICNETS2), 38-4, 2017. .
- [17] S. A Christian U. O Michael. "Design and Construction of Weather Instrument and Its Use in Measurements to Determine the Effects of Some Weather Parameters on GSM Signal Strength". *Advances in Applied Sciences*. **6 (4)**, 142-154, 2021.
- [18] O. L Ojo. "Geo-spatial distribution of radio refractivity and the influence of fade depth on microwave propagation signals over Nigeria". *International Journal of Physical Science (Academic Journal)* , **18(3)**, 72-83 , July 2023. doi.org/10.5897/IJPS2023.5036.
- [19] D. Dajab, "Perspectives on the effects of harmattan on radio frequency waves". *Journal of Applied Sciences Research*, **2(11)**, 1014-1018, 2006.
- [20] M. U. Onuu, A. Adeosin "Investigation of propagation characteristics of UHF waves in Akwa Ibom State, Nigeria". *Indian J Rad Space Phy*: **37**, 197-203, 2008
- [21] A Eyo, O. E., Menkiti, A. I., Udo, S. O.. "Microwave signal attenuation in Harmattan weather along Calabar-Akamkpa line-of-sight link". *Turk J Phys*: **27**, 153-160, 2003..
- [22] A. S Abdullahi. J. S Ojo, A. T Adediji, "Statistical Analysis of Atmospheric Parameters, Noise Temperature and Digital Terrestrial Television Signal Strength Over JOS Metropolis, Plateau State, Nigeria". July 2022 *Heliyon* **8(3)**:e09988
- [23] L. Sa'du, N. C Maduka, A. Kado, I. Abdulmuninu, M. Bello. "Assessment of mobile network signal strength for GSM networks in Gusau, Zamfara State, Nigeria". *Inter J Innov Sc, Eng Tech*: **6**, 10-20, 2022.

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).

## Mathematical Model of Wind Turbine Simulator Based Five-Phase Permanent Magnet Synchronous Generator with Nonlinear Loads and Harmonic Analysis

Peerawat Meesuk, Vijit Kinnares\*

School of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL), Bangkok, 10520, Thailand

### ARTICLE INFO

Article history:

Received: 16 November, 2023

Revised: 28 January, 2024

Accepted: 29 January, 2024

Online: 21 February, 2024

Keywords:

Wind Turbine Simulator  
Permanent Magnet Synchronous  
Harmonic Effect

### ABSTRACT

This paper presents mathematical model of a wind turbine simulator based five-phase permanent magnet generator supplying nonlinear load. The mathematical model of wind turbine characteristics together with available tool blocks of the five-phase permanent generator and semiconductor devices of an AC-DC converter formed as a nonlinear load is implemented on MATLAB /Simulink to investigate the harmonic effect on performance of the generator. The detailed descriptions of the proposed model are fully given. The harmonic analysis is also provided. The validity of the proposed model is verified by simulation using MATLAB /Simulink in terms of dynamic responses of rotor speed, torque and power quality of the generator. It is found that the nonlinear load significantly affects the electromagnetic torque ripple and the distortions of both voltage and current of the generator. Moreover, the proposed system offers higher nonlinear load voltage and faster response compared to a conventional three-phase permanent magnet synchronous generator system. The electromagnetic torque ripple is reduced by 88% and the total harmonic distortions of the phase voltage and the stator current are more or less 7% and 60% which exceed the limits of the harmonic standards.

### 1. Introduction

Wind energy is one of the abundant renewable resources that is getting much attention around the world since it is clean, naturally occurring and environmentally friendly energy without burning fossil. In a wind turbine power generation system, the kinetic energy generated by the wind power is converted into mechanical torque at the shaft of the wind turbine which is coupled either directly or indirectly via gear box to the generator. Then the generator converts mechanical energy harvested from the wind energy by the turbine into electrical energy through magnetic field. However, the use of a wind turbine power generation system is necessary to study many components such as potential of locally generated wind speeds, electricity standards and regulations, system performance, expense and payback on investments, etc. In order to obtain the best performance of the system in terms of maximum efficiency achievement and cost-effectiveness, several research works and development of the wind turbine power generation systems have been paid attention [1,2]. Although an induction generator (IG) has various advantages over a major counterpart permanent magnet synchronous generator (PMSG), particularly for low maintenance, high ruggedness, and low cost, it

gives lower efficiency and draws reactive power resulting in poor power factor for both self-excited and grid connected applications. Performance of a three-phase self-excited induction generator (SEIG) operating as a single-phase generator supplying nonlinear loads for standalone applications was investigated in [1]. However, the electromagnetic torque ripple due to harmonic effect has not been reported yet. Normally the SEIG needs a capacitor bank for reactive compensation for both voltage buildup process and voltage regulation but it significantly affects a variation of the generator frequency which is a drawback of the SEIG. Harmonic analysis and experimental tests of a three-phase SEIG with nonlinear loads were proposed in [3]. The obtained results confirm that the harmonics associated with the nonlinear loads affect the voltage and current waveforms. A PMSG is more attractive for small scale applications due to high efficiency, high power density, better voltage regulation for standalone applications [4]. Control of grid connection of an axial flux permanent magnet generator (AFPMG) (i.e. another type of PMSG) using a single-phase inverter with multifunctionality for real power transfer and reactive and harmonic compensations with various types of nonlinear loads was implemented in both simulation and hardware by [4,5]. However, the wind turbine simulator using motor drives as a prime mover has not been proposed. An example of a wind turbine

\*Corresponding Author: Vijit Kinnares, School of Engineering, King Mongkut's Institute of Technology Ladkrabang (KMITL), [kkwijit@kmitl.ac.th](mailto:kkwijit@kmitl.ac.th)



simulator emulating actual wind turbine characteristics implemented by low cost hardware can be found in [6] which is useful in a laboratory test for wind energy conversion research. However, mathematical model of the simulator implemented on simulation was not focused. Unlike standalone applications, for grid connected applications, a large variation of the frequency of voltage and current of the generator is not necessary which is the major advantage of a wind turbine power generation system since an ac-dc rectifier converts alternating current into a direct current form. Generally, a five-phase permanent magnet synchronous machine is introduced in motoring operation rather than as a generator for electrical propulsion due to low torque pulsations, good fault tolerance, high power density and low loss [6-8]. Although the smaller ripple of the dc output voltage in the five-phase PMSG system associated with a rectifier than that in the three-phase PMSG system was presented in [9], harmonic effect on electromagnetic torque pulsation and harmonic analysis were not reported. There are a few publications related to performance evaluation of a five-phase PMSG with nonlinear load. Moreover so far the harmonic effect on the performance of the five-phase PMSG associated with wind energy conversion has not been investigated yet. With the shortcomings of the previous works mentioned earlier, therefore this work contributes to mathematical model setup applied in simulation using MATLAB/Simulink computer program of the wind simulator based five-phase PMSG with nonlinear loads and harmonic analysis for investigating harmonic effect on the generator performance.

The paper is organized as follows. A wind power generation system is introduced briefly followed by mathematical model of wind turbine characteristics in Section 2. Section 3 describes mathematical model of the five-phase PMSG presented in a rotating d-q-x-y frame modeling followed by an available tool block in the MATLAB/Simulink. Section 4 deals with the nonlinear load model and harmonic analysis. Section 5 describes the proposed Simulink model and simulation results in terms of dynamic response and steady state conditions for various performances of the generator like terminal voltage and current waveforms, torque, speed, and so on together with discussion. Finally, conclusion and suggestion are given.

## 2. Wind Power Generation System

The electrical power generation system for grid connected applications used in the simulation is referred to Figure 1. It consists of a wind turbine simulator, a five-phase permanent magnet synchronous generator and a nonlinear load using a diode-based rectifier and a smoothing filter capacitor. Generally, an inverter circuit is used to convert DC input into AC output for grid connection in order to transfer the electrical energy produced from the wind energy to the grid. In the simulation, the model of the inverter is not taken into account. The equivalent resistor is used for representation of the transferred power and the converter power loss.

In steady state, the mechanical power generated at the shaft of the wind turbine can be obtained as [6]

$$P_{WT} = 0.5\rho\pi R^2 v_{wind}^3 C_p(\lambda, \beta) \quad (1)$$

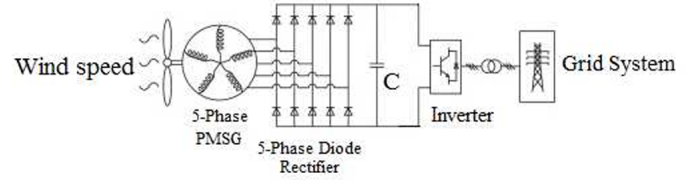


Figure 1: Overview of an electric power generation system using wind energy conversion system for grid connected applications.

where

$P_{WT}$  is the mechanical power of the wind turbine ,

$\rho = 1.25 \text{ kg/m}^3$  which is the air density ,

$R$  is the turbine rotor radius,

$v_{wind}$  is the wind velocity,

$C_p$  is the power coefficient,

$\beta$  is the pitch angle, and

$\lambda$  is the tip speed ratio.

The power coefficient is given by:

$$C_p(\lambda, \beta) = C_1 \left( \frac{C_2}{\lambda_i} - C_3\beta - C_4 e^{-\frac{C_5}{\lambda_i}} \right) + C_6\lambda \quad (2)$$

and

$$\frac{1}{\lambda_i} = \frac{1}{\lambda + 0.08\beta} - \frac{0.035}{\beta^3 + 1} \quad (3)$$

The coefficients  $C_1$  to  $C_6$  are as follows.  $C_1$  is 0.5176 ,  $C_2$  is 116 ,  $C_3$  is 0.4 ,  $C_4$  is 5 ,  $C_5$  is 21 and  $C_6$  is 0.0068.

The tip speed ratio is given by:

$$\lambda = \frac{\pi n R}{30 v_{wind}} \quad (4)$$

The shaft torque of the wind turbine is expressed as

$$T_{WT} = \frac{30 P_{WT}}{\pi n} \quad (5)$$

where  $T_{WT}$  is the mechanical torque of the wind turbine and  $n$  is the rotor speed of the wind turbine. Schematic diagram of the proposed system for the simulation on a computer program is shown in Figure 2 in accordance with Figure 1. The power transferred to the grid and the converter power loss are lumped together represented by resistance  $R_{eq}$ . Figure 3 shows the simulation model implemented on MATLAB/Simulink simulation program as a wind turbine simulator. The mathematical expressions in (1)-(5) are used to generate wind turbine characteristics providing the mechanical torque applied to the PMSG tool block model. Detailed parameter values can be found in [6]. The wind turbine characteristics are plotted by using the

wind turbine simulator as shown in Figures 4 and 5 for input mechanical torque and mechanical power against the rotor speed, respectively. As can be seen, below maximum values, the mechanical torque is proportional to the squared rotor speed whilst the mechanical power is proportional to the cube of the rotor speed at a given wind speed.

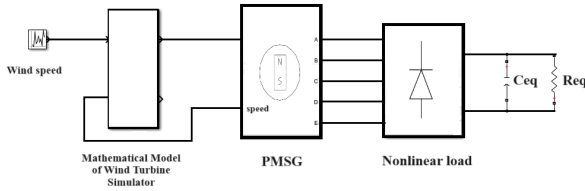


Figure 2: Schematic Diagram of wind turbine simulator based five-phase permanent magnet synchronous generator with nonlinear load.

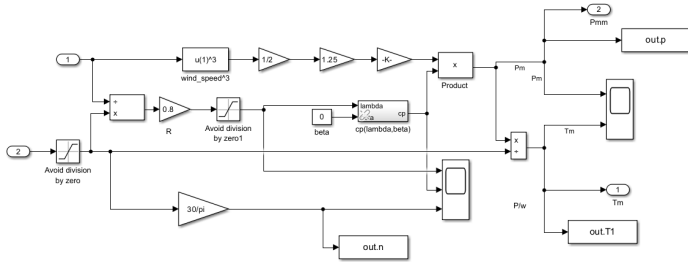


Figure 3: Mathematical model of wind simulator using MATLAB/SIMULINK

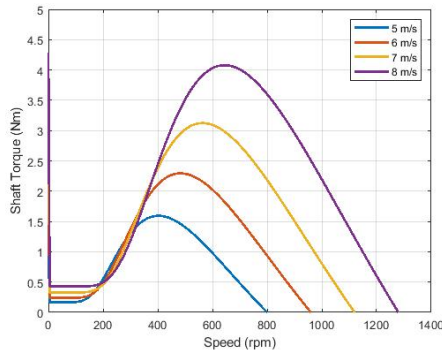


Figure 4: Applied mechanical shaft torque versus rotor speed with variation of wind speed

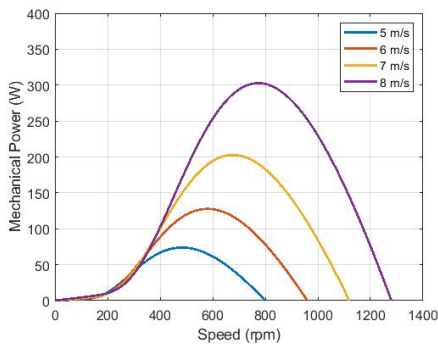


Figure 5: Applied mechanical power against rotor speed with variation of wind speed.

The results from Figures 4 and 5 are in accordance with [6]. Theses confirm the validity of the proposed wind turbine simulator in a mathematical model form.

### 3. Five-Phase Permanent Magnet Synchronous Generator

The stator winding for each phase is displaced by 72 electrical degrees. The 5-phase PMSG can be modeled in a rotating d-q-x-y frame for analysis of dynamic response and steady state conditions [6]. More details can be also found in [7], [8] and [10]. The mathematical equations can be expressed as follows.

$$v_d = R_s i_d + L_d \frac{di_d}{dt} - \omega L_q i_q \quad (6)$$

$$v_q = R_s i_q + L_q \frac{di_q}{dt} + \omega L_d i_d + \omega \phi_f \quad (7)$$

$$v_x = R_s i_x + L_{ls} \frac{di_x}{dt} \quad (8)$$

$$v_y = R_s i_y + L_{ls} \frac{di_y}{dt} \quad (9)$$

$$v_o = R_s i_o + L_{ls} \frac{di_o}{dt} \quad (10)$$

The relationship between the mechanical torque ( $T_l$ ) and the electromechanical torque ( $T_{em}$ ) can be given as

$$J \frac{d\Omega}{dt} + f_r \Omega = T_{em} - T_l \quad (11)$$

where  $v_d, v_q, v_x, v_y, i_d, i_q, i_x, i_y$  are the stator voltages and stator currents in the d-q-x-y, respectively.  $J$  is the moment of inertia,  $f_r$  is the viscous coefficient, and  $\Omega$  is the rotor speed and  $T_l$  is the load torque.

The electromagnetic torque can be expressed as

$$T_{em} = \frac{5}{2} p (\phi_f i_q - (L_d - L_q) i_d i_q) \quad (12)$$

where  $p$  is the pole pair,  $L_d$  is the direct axis inductance,  $L_q$  is the quadrature axis inductance,  $i_q$  is the quadrature axis current,  $i_d$  is the direct axis current,  $L_{ls}$  is the leakage inductance and  $\phi_f$  is the flux generated by permanent magnet in the rotor. These equations are used for calculation in the machine model. The available tool block for the PMSG model in MATLAB/SIMULINK is illustrated in Figure 6.  $T_m$  is the input mechanical torque which the positive sign is for the motoring operation and negative sign represents the generator operation. “m” is the output for measuring various variables and machine performance like stator currents, flux linkages, and mechanical outputs such as speed, electromagnetic torque, etc. The rated values of the used model of the 5-phase PMSG are output torque of 8 Nm and the rotor speed of 2000 RPM.

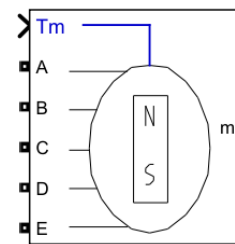


Figure 6: Tool block model of a five-phase permanent magnet synchronous generator available in MATLAB/ Simulink.

### 4. Nonlinear Load model and Harmonic Current Analysis

In this work, the nonlinear load is connected in parallel with the generator output consisting of a five-phase full bridge rectifier

including 10 diodes and a smoothing filter capacitor connected in parallel at the output of the rectifier. The per phase simplified model of the PMSG with nonlinear load is shown in Figure 7 which various harmonic currents produced by the nonlinear load are represented by harmonic current sources. The time domain of harmonic currents can be expressed as

$$i_s(t) = \sum_{n=1}^{\infty} I_n \cos(n\omega t - \theta_n) \quad (13)$$

where  $I_n$  is the amplitude of the source current at harmonic order  $n$  and  $\theta_n$  is the phase difference angle of the current with respect to the source voltage. The nonlinear load model in [11] with Matlab/Simulink is shown in Figure 8 where  $R_{eq}$  represents the converter loss and the active power drawn by the grid system in accordance with Figure 1. The values of the smoothing filter capacitor and  $R_{eq}$  are  $5000 \mu F$  and  $10 \Omega$  respectively. The diode model consists of a RC snubber circuit including  $R_s$  of  $250 \Omega$ , connected in series with  $C$  of  $250 nF$ , forward voltage drops of  $0.8 V$  and conduction resistance ( $R_{on}$ ) of  $0.001 \Omega$ . In order to analyze the propagation of a harmonic source in the PMSG, the simplified equivalent circuit can be depicted as shown in Figure 9 which is treated in the same manner as the circuit reported in [12] for a three-phase SEIG. The transfer function of the harmonic current propagation can be expressed as

$$T(s) = \frac{I_{mh}}{I_h} \quad (14)$$

By using KCL (Kirchoff's current law), the harmonic current at order  $h$  injected into the machine yields

$$I_{mh} = \frac{R_L}{R_L + R_S + sL_S} I_h \quad (15)$$

In this study, the parameters of the PMSG are as follows.  $R_S$  is  $0.0485 \Omega$ ,  $L_S$  is  $8.5 mH$  and  $R_L$  is  $10 \Omega$ . Therefore, the transfer function of the harmonic current propagation is

$$T(s) = \frac{I_{mh}}{I_h} = \frac{R_L}{R_L + R_S + sL_S} = \frac{10}{0.0085s + 10.0485} \quad (16)$$

From (16) the frequency response can be illustrated in magnitude and phase forms by using the Bode Plot as shown in Figure 10. According to these results the cut-off frequency is approximately  $100 kHz$ . It implies that the harmonic frequency below this value can affect the PMSG. Note that the harmonic currents can give the effect of a decrease in efficiency, heating and reducing the lifetime of the machine [12].

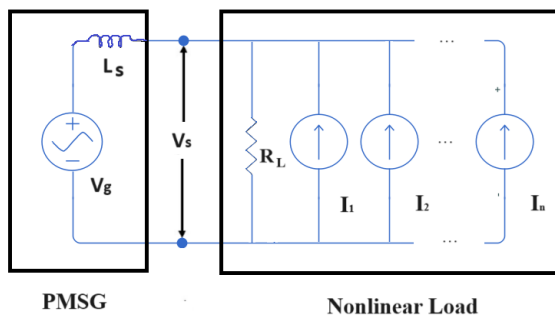


Figure 7: Simplified model of PMSG considering nonlinear load.

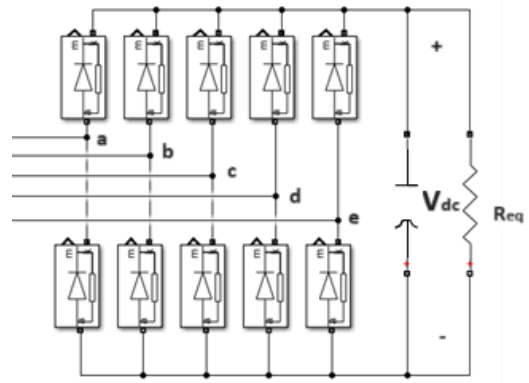


Figure 8: Model of the nonlinear load using MATLAB/ Simulink.

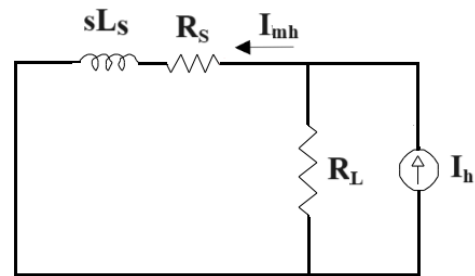


Figure 9: Simplified equivalent circuit for considering harmonic current propagation on PMSG.

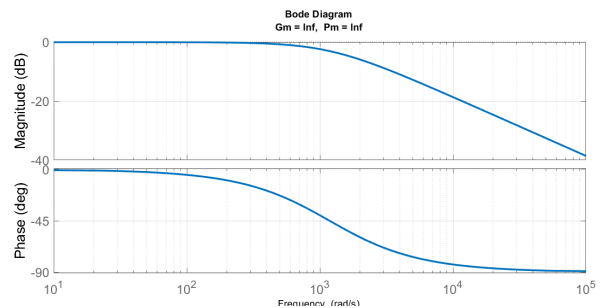


Figure 10: Frequency response for harmonic current propagation.

## 5. Simulation Results and Discussion

The simulation model of the wind turbine simulator based five- phase PMSG supplying nonlinear load is implemented by using MATLAB/SIMULINK as shown in Figure 11. It consists of a five-phase PMSG tool block and subsystems of the wind turbine simulator using the proposed mathematical model, and the nonlinear load. Wind speed is required for the command. The wind turbine simulator provides the input mechanical torque for the PMSG. The rotor speed of the PMSG is used to feedback to the wind simulator. Various waveforms and signals are measured by measurement units.

Figure 12 illustrates comparative waveforms between the conventional three-phase PMSG and the proposed five-phase PMSG of the machine terminal voltage and nonlinear load voltage (i.e. dc link voltage or output of the rectifier, Vdc) at the wind speed of  $6 m/s$  during startup and steady state conditions.

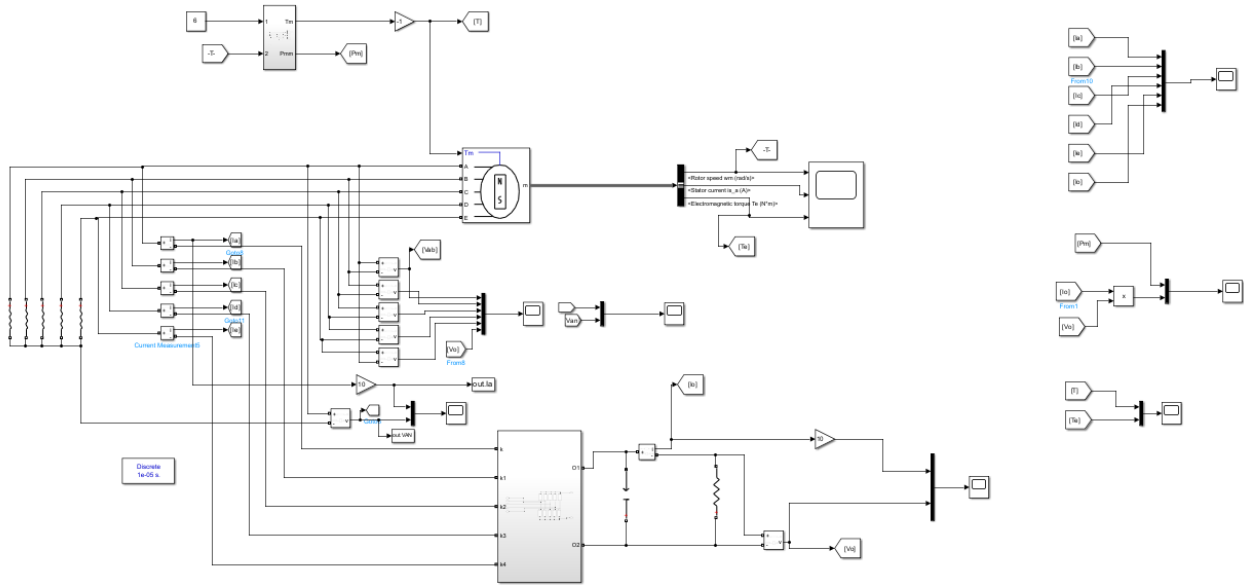


Figure 11: Simulation model of wind turbine simulator based five-phase PMSG.

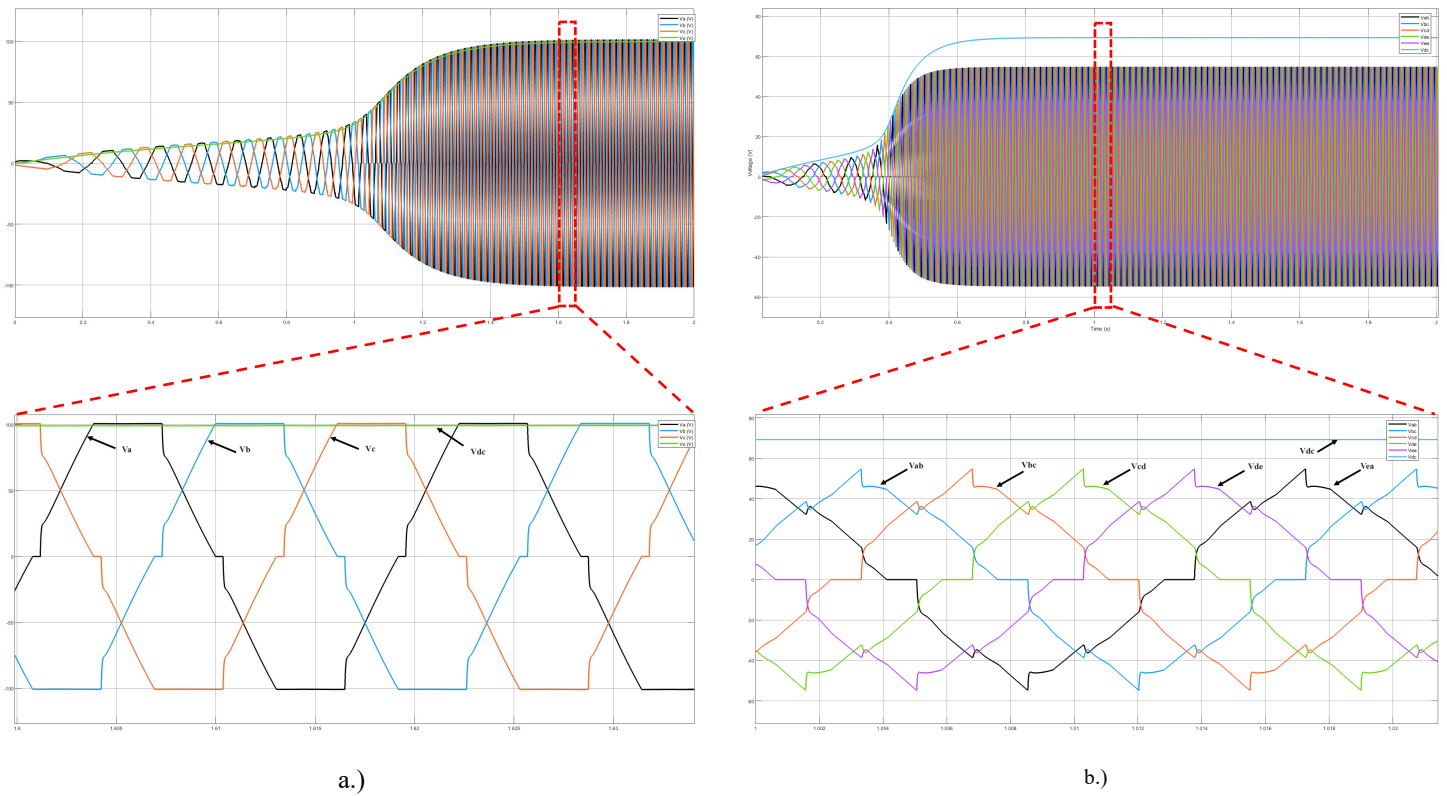


Figure 12: Waveforms of output voltage of rectifier and terminal voltage during startup and steady state conditions and corresponding enlarged waveforms a.) conventional three-phase PMSG b.) proposed five-phase PMSG

Clearly, with the same input mechanical torque for both cases the voltage and its frequency are gradually increased and then abruptly grown up until steady state. The dc output voltage of the rectifier and dynamic response for the proposed PMSG are higher and faster response than those for the conventional three-phase PMSG, respectively. The enlarged waveforms show the distortion due to

the propagation of the harmonic currents associated with the nonlinear load. The dc voltage of the rectifier output is much higher than the peak value of the terminal voltage which is the advantage of the proposed five-phase PMSG compared to the conventional three-phase PMSG. Moreover, the dc output contains insignificant ripple. Apparently, enlarged waveforms of the



terminal voltages are distorted from sinusoid due to the propagation of the harmonic currents associated with the nonlinear load for both cases. The five-phase PMSG with nonlinear load offers less distortion of voltage waveforms than those for the three-phase PMSG. Figure 13 illustrates comparative waveforms between the conventional three-phase PMSG and the proposed five-phase PMSG of the corresponding stator currents and dc output current waveform during start up and steady state conditions. The dc output current of the rectifier and dynamic response for the proposed PMSG are higher and faster response than those for the conventional three-phase PMSG, respectively. Obviously, the inrush currents are present during a suddenly increase in terminal voltage for both cases. Apparently, the enlarged current waveforms are distorted due to the nonlinear load.

Figures 14 and 15 show dynamic response during startup of the input mechanical torque and the electromechanical torque of the PMSG for linear and nonlinear loads with the same condition for a conventional three-phase PMSG and the proposed five-phase PMSG, respectively. Noticeably, the electromagnetic torque contains higher ripple for the nonlinear load compared to the linear load due to the large value of the smoothing filter capacitor and harmonic effect. When comparing the ripple of the electromagnetic torque and dynamic response for the nonlinear load between the conventional three-phase PMSG and the proposed five-phase PMSG, obviously, the proposed five-phase PMSG has lower ripple and faster response than the three-phase PMSG. During the growth of the terminal voltage, the transient torque occurs for both linear and nonlinear loads. The peak values of the mechanic torque for both cases are almost equal. The significant difference between the input mechanical torque and the electromagnetic torque during transient response represents the large acceleration torque required for the linear load resulting in faster response (i.e. see equation (11)). At the steady state, the absolute input mechanical torque is little higher than the absolute electromagnetic torque due to friction and windage components.

Figures 16 and 17 illustrate the enlarged waveforms during the steady state of the input mechanical torque and the electromagnetic torque for linear and nonlinear loads. Clearly, the electromagnetic torque has higher ripple for the nonlinear load than that for the linear load. It is found that the electromagnetic torque ripple is about 9.34% which is in good agreement with [9]. Note that the electromagnetic torque ripple for the five-phase PMSG is lower than that for the conventional three-phase PMSG reported in [9]. The peak-to-peak electromagnetic torque ripple of the proposed five-phase PMSG is reduced about 88.18%. compared to that of the conventional three-phase PMSG. In order to confirm the higher ripple of the electromagnetic torque for the conventional three-phase PMSG with nonlinear load, Figure 18 illustrates comparative harmonic spectra of the electromagnetic torque between the three-phase PMSG and the proposed PMSG by using FFT analysis tool block. Obviously, the electromagnetic torque of the conventional three-phase PMSG contains large low order harmonic components (i.e. the sixth harmonics) resulting in high pulsation.

Figure 19 shows the input mechanical power and the electrical power of the nonlinear load. During transient response in startup duration the input power is drastically higher than the output power since the input power is drawn with large amount for acceleration of the generator. Eventually, the input power is higher than the output power with a fixed value. The difference of both powers is present due to the machine losses including core loss, copper loss,

and friction and windage loss, semiconductor losses including conduction and switching losses, and snubber loss in the model. As a consequence, it implies that the efficiency of the system can be determined which is not included in here. Comparison between linear and nonlinear loads for phase voltage and current waveforms are shown in Figure 20.

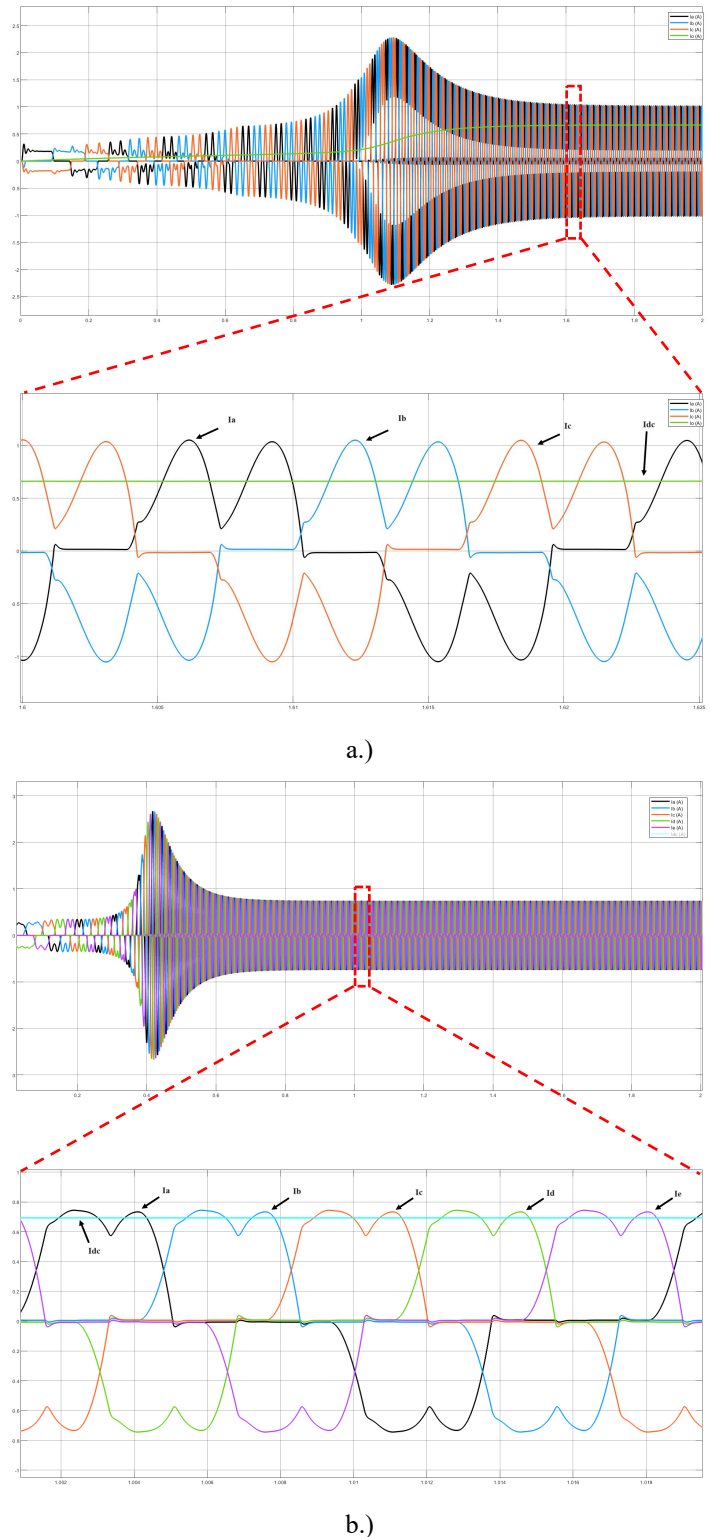


Figure 13: Waveforms of stator currents during start up and steady conditions and enlarged waveforms  
a.) conventional three-phase PMSG b.) proposed five-phase PMSG

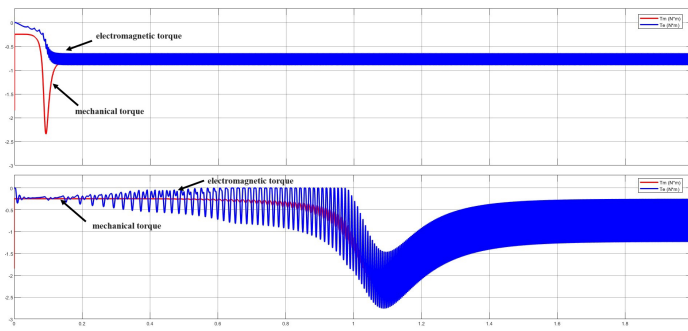


Figure 14: Input mechanical torque and electromagnetic torque of conventional three-phase PMSG a.) linear load b.) nonlinear load

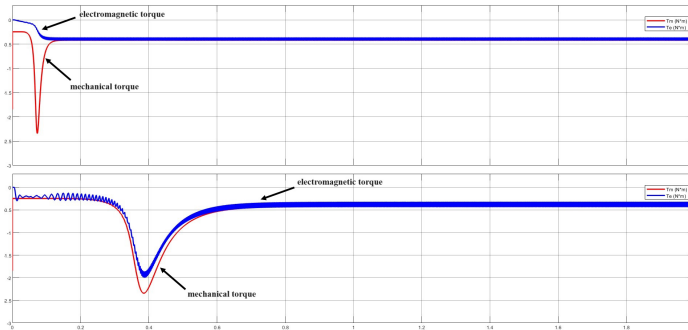


Figure 15: Input mechanical torque and electromagnetic torque of the proposed five-phase PMSG a.) linear load b.) nonlinear load

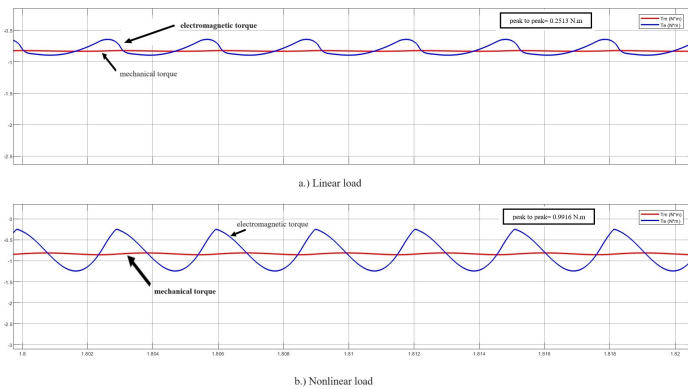


Figure 16: Instantaneous input mechanical torque (red signal) and instantaneous electromagnetic torque (blue signal) under steady state of conventional three-phase PMSG.

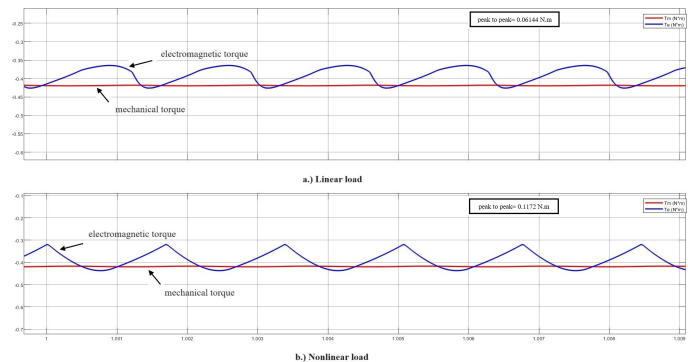
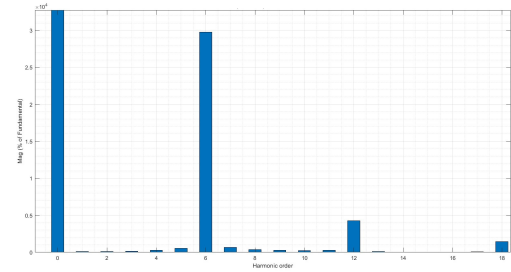
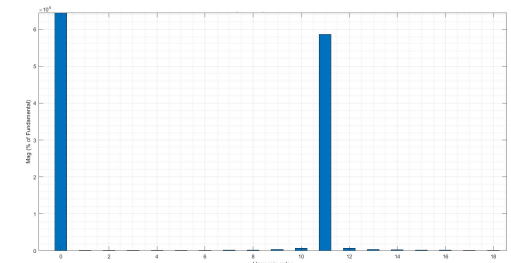


Figure 17: Instantaneous input mechanical torque (red signal) and instantaneous electromagnetic torque (blue signal) under steady state of the proposed five-phase PMSG.

Apparently, unlike for linear load without distortion, both terminal voltage and stator current waveforms of the generator for the nonlinear load are significantly distorted from sinusoidal waveforms. This is the fact that the harmonic currents associated with the nonlinear load are injected to the machine resulting in the distortion of sinusoidal waveforms.



a.)



b.)

Figure 18: Electromagnetic torque spectra for nonlinear load a.) conventional three-phase PMSG b.) proposed five-phase PMSG

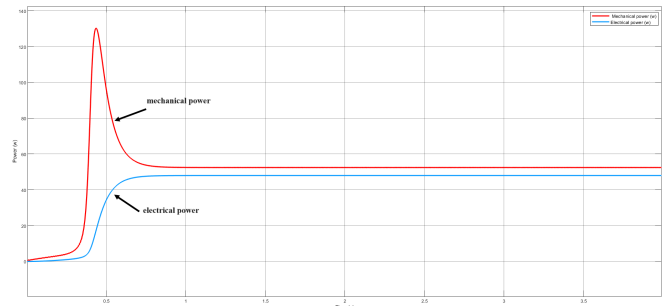
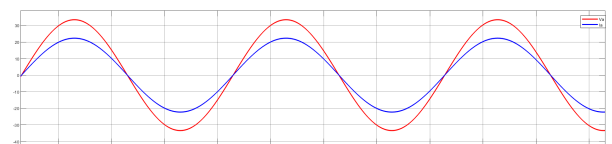
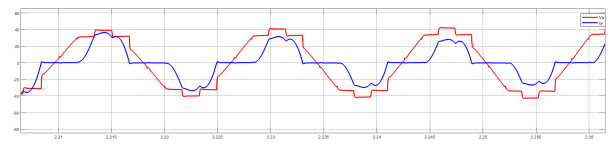


Figure 19: Transient response and steady state conditions during startup of input mechanical power and the output electrical power of the nonlinear load.



a.) Linear load

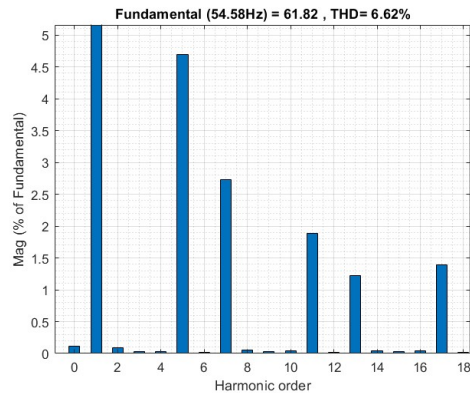


b.) Nonlinear load

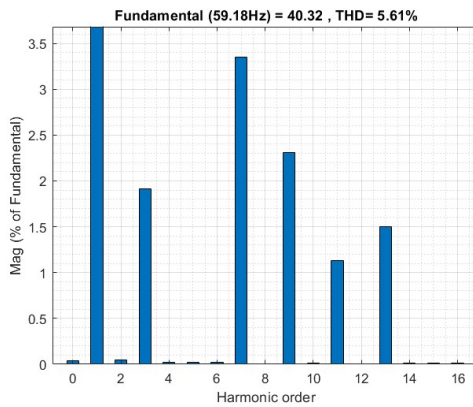
Figure 20: Terminal voltage and stator current waveforms for nonlinear load and linear load.

Figures 21 and 22 illustrate comparative harmonic spectra of both terminal voltages and stator currents between the three-phase PMSG and the proposed PMSG for the nonlinear load by using FFT analysis tool block. For the proposed PMSG case, the total harmonic distortion of the phase voltage and stator current are more or less 7 % and 60 %, respectively. These high figures are not within acceptable levels, particularly for the current case [13]. The solution for this problem could be introduced by either passive or active power filters. Note that these values depend on load level conditions. Quite clearly, unlike the conventional PMSG, the magnitudes of the third harmonics for both current and voltage are dominant. Unlike the conventional three-phase PMSG, some multiples of three of harmonics are still present which are different from a three-phase system in which the multiples of three of harmonic orders are absent. As a consequence, the harmonic distortion for the conventional three-phase PMSG is better than that for the proposed PMSG. The harmonic orders are in agreement with those reported in [14-17].

five-phase PMSG requires a smaller size of the smoothing capacitor.



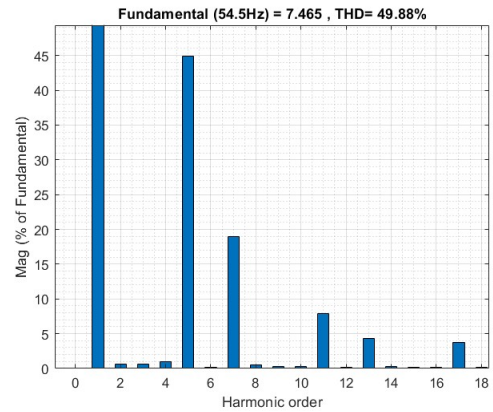
a.)



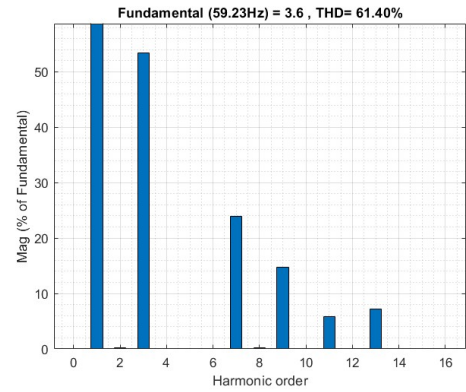
b.)

Figure 21: Terminal voltage harmonic spectra for nonlinear load a.) conventional three-phase PMSG b.) proposed five-phase PMSG

Figure 23 show a comparison of the ripple of the dc output voltage for the nonlinear load between the conventional three-phase PMSG and the proposed five-phase PMSG. Obviously, unlike the conventional three-phase PMSG, the ripple exists twice frequency of the ac terminal voltage of the machine. As a result, the proposed five-phase PMSG has lower ripple of dc output voltage lower than three-phase PMSG. It implies that the proposed

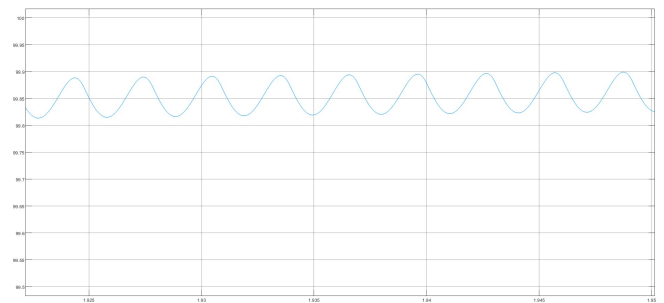


a.)

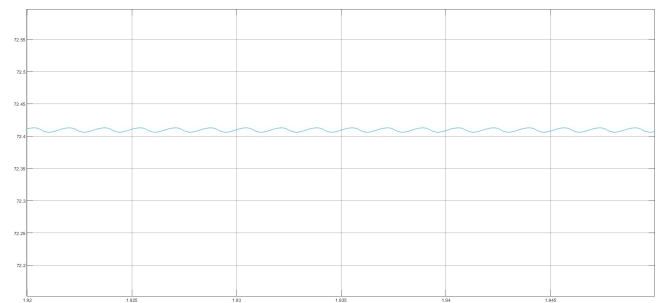


b.)

Figure 22: Stator current harmonic spectra for nonlinear load a.) conventional three-phase PMSG b.) proposed five-phase PMSG



a.)



b.)

Figure 23: Ripple voltage of the DC output voltage with nonlinear load a.) conventional three-phase PMSG b.) proposed five-phase PMSG

Figure 24 displays transient responses of the rotor speed, the stator phase current and the electromagnetic torque during a step change in wind speed from 5 to 8 m/s at  $t=2$  seconds. At the beginning of the startup operation, the rotor speed, the electromagnetic torque (i.e. negative sign), and the stator current and its frequency are gradually increased. Then at about  $t=0.5$  seconds, they are rapidly grown up until they reach to the steady state. When the wind speed is changed in a step, the high peak electromagnetic torque and high inrush stator current occur. Obviously after transient response the electromagnetic torque and stator current fall down to steady state values which are higher than the previous ones before disturbance. The rotor speed is increased slowly which looks like the response of the first order of a derivative equation in accordance with (11).

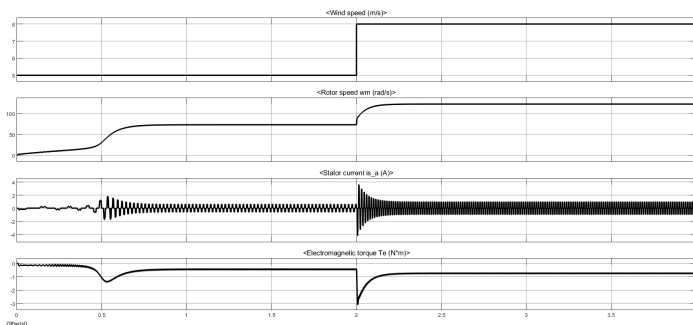


Figure 24: Step change in wind speed from 5 m/s to 8 m/s

In order to investigate the performance of the proposed system during wind speed changes according to the realistic conditions, wind speeds with random input ranging between 5 to 8 m/s, have been applied. The dynamic responses of the rotor speed, the electromagnetic torque and stator current with such random wind speeds are illustrated in Figure 25. The rotor speed response is slower than the wind speed command. The final value of the rotor speed after disturbance is promotional to the wind speed. During a step change in either an increase or a decrease in the wind speed, the high inrush current and high peak torque occur particularly during startup.

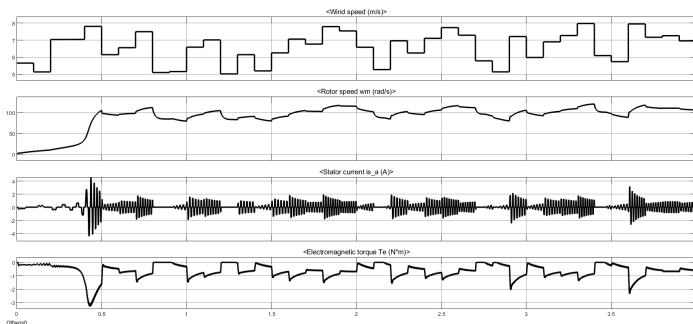


Figure 25: Dynamic response for random input wind speed for the nonlinear load.

## 6. Conclusion

This paper has proposed mathematical model of a wind turbine simulator based five-phase permanent synchronous generator supplying nonlinear load using MATLAB/Simulink. The related mathematical equations of the machine model are also described. The harmonic analysis for propagation and the harmonic effect on the PMSG have been given. The descriptions of Simulink model

of the proposed system are fully given. The investigation of the performance and power quality of the PMSG with the nonlinear load under various conditions has been conducted. It is found that the harmonics associated with the nonlinear load adversely affects the performance and power quality of the PMSG. The nonlinear load causes degradation of the performances of the proposed PMSG in terms of higher electromagnetic torque ripple and slower response when compared to the linear load case. The proposed five-phase PMSG offers some advantages over the conventional three-phase PMSG in terms of lower electromagnetic torque, higher output dc voltage value and lower ripple of the dc output voltage. However, its disadvantage is higher harmonic distortions of both terminal voltage and stator current which are over limits of the standards. The suggestion is a requirement of harmonic solutions such as either passive or active power filters to mitigation of such harmonics.

## References

- [1] P. Nakorn, P. Machot, V. Kinnarees, and C. Manop, "Study of Three phase Self-excited Induction Generator Operating as Single-phase Induction Generator Supplying Non-linear Load," In 2021 18th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTICON), 806-809, 2021, DOI: 10.1109/ECTI-CON51831.2021.9454845.
- [2] C. Xue, W. Song, and X. Feng, "Finite control-set model predictive current control of five - phase permanent - magnet synchronous machine based on virtual voltage vectors," IET Electric Power Applications, **11**(5), 836-846,2017, doi.org/10.1049/iet-epa.2016.0529.
- [3] T. Kamel, D. Abdelkader, and B. Said, "Vector control of five-phase permanent magnet synchronous motor drive," in 2015 4th International Conference on Electrical Engineering (ICEE) , 1-4 , 2015, DOI: 10.1109/INTEE.2015.7416853 .
- [4] P. Wannakarn, and V. Kinnarees, "Single-phase grid connected axial flux permanent magnet generator system with reactive power compensation functionality," in 2012 10th International Power & Energy Conference (IPEC) ,338-341,2012, DOI: 10.1109/ASSCC.2012.6523289.
- [5] P. Wannakarn, and V. Kinnarees, "Single-phase grid connected axial flux permanent magnet generator system with harmonic mitigation functionality for various types of nonlinear loads," International Review of Electrical Engineering, **13**(2), 157-164,2018, DOI: 10.1109/ASSCC.2012.6523289.
- [6] N. Thodsaporn, and V. Kinnarees, "Wind turbine simulator equipped with real-time monitoring and user-friendly parameter setup controlled by C2000 microcontroller," In 17th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON), 713-716, 2020, DOI: 10.1109/ECTI-CON49241.2020.9158236.
- [7] H. C. Chen, C. H. Hsu, and D. K. Chang, "Position sensorless control for five-phase permanent-magnet synchronous motors," in 2014 IEEE/ASME International Conference on Advanced Intelligent Mechatronics 794-799, 2014, DOI: 10.1109/AIM.2014.6878176.
- [8] F. Yu, X. Zhang, M. Qiao, and C. Du, "The direct torque control of multiphase permanent magnet synchronous motor based on low harmonic space vector PWM," in 2008 IEEE International Conference on Industrial Technology, 1-5, 2008, DOI: 10.1109/ICIT.2008.4608494.
- [9] N. E. A. M. Hassanain, Comparison between three-and five-phase permanent magnet generators connected to a diode bridge rectifier, Ph. D Thesis, University of Strathclyde Department of Electronics and Electrical Engineering, 2009.
- [10] S. Rhaili, A. Abbou, S. Marhraoui, N. El Hichami, , and A. V. Hemeyine, "Robustness investigation of Vector Control of Five-phase PMSG based Variable-Speed Wind Turbine under faulty condition." In 2018 Renewable Energies, Power Systems & Green Inclusive Economy (REPS-GIE) , 1-6, 2018, DOI: 10.1109/REPSGIE.2018.8488809.
- [11] P. Meesuk, and V. Kinnarees, "Mathematical Model of Wind Turbine Simulator Based Five Phase Permanent Magnet Synchronous Generator Supplying Non-Linear Loads," in 2023 9th International Conference on Engineering, Applied Sciences, and Technology (ICEAST), 168-171, 2023,



DOI: 10.1109/ICEAST58324.2023.10157898.

- [12] R.Nazir, "Analysis of Harmonic Currents Propagation on the SelfExcited Induction Generator with Nonlinear Loads," *Journal of Electrical Engineering and Technology*, **9**(6), 1935-1943, 2014, <http://dx.doi.org/10.5370/JEET.2014.9.6.1935>.
- [13] K. Wirtayasa, C. Y. Hsiao, N. C. Yang, "High-Performance Five-Phase Axial Flux Permanent Magnet Generator for Small-Scale Vertical Axis Wind Turbine," *Applied sciences*, 2020, **12**(7), 3632, [doi.org/10.3390/app12073632](https://doi.org/10.3390/app12073632).
- [14] M. McCarty, T. Taufik, A. Pratama, and M. Anwari, "Harmonic analysis of input current of single-phase controlled bridge rectifier," in *2009 IEEE Symposium on Industrial Electronics & Applications*, **1**, 520-524, 2009, DOI: 10.1109/ISIEA.2009.5356404.
- [15] R. R. Kumar, S. K. Singh, R. K. Srivastava, A. S. S. Vardhan, R. M. Elavarasan, R. K. Saket, E. Hossain, "Modeling of airgap fluxes and performance analysis of five phase permanent magnet synchronous generator for wind power application," *IEEE Access*, **8**, 195472-195486, 2014.
- [16] F. Bu, H. Liu, W. Huang, H. Xu, K. Shi, "Optimal third harmonic injection-based control for a five phase dual stator winding induction generator DC generating system," *IEEE Transactions on Industrial Electronics*, 2018, **65**(11), 9124-9134.
- [17] R. R. Kumar, A. Kumari, P. Devi, and C. Chetri, "Design and Performance Characteristics of Dual-Rotor Magnetically Decoupled Stator Five-Phase Permanent Magnet Synchronous Generator for Wind Power Applications.," 2021, In *Advances in Electrical and Computer Technologies: Select Proceedings of ICAECT 2020*, 1261-1274, 2020, DOI: 10.1109/ACCESS.2020.3034268.
- [18] E. Ebrahimzadeh, F. Blaabjerg, X. Wang, and C. L. Bak, "Harmonic stability and resonance analysis in large PMSG-based wind power plants," *IEEE Transactions on Sustainable Energy*, **9**(1), 12-23, 2017, DOI: 10.1109/TSTE.2017.2712098.
- [19] H. Abusannuga, and M. Özkaymak, "Performance of Vertical Axis Wind Turbine Type of Slant Straight Blades," *Advances in Science, Technology and Engineering Systems Journal*, **6**(4), 292-297, 2021, DOI:10.25046/aj060432.
- [20] M. H. Rashid, *Power electronics: circuits, devices, and applications.*, Pearson Education India, 2009.
- [21] Ned, M., Ullmann, F., & Robbins, *Power electronics.*, Pearson Education India, 2003.
- [22] Simulink reference, from [https://www.mathworks.com/help/pdf\\_doc/simulink/simulink\\_ref.pdf](https://www.mathworks.com/help/pdf_doc/simulink/simulink_ref.pdf)

**Copyright:** This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY-SA) license (<https://creativecommons.org/licenses/by-sa/4.0/>).