



# INTERNATIONAL RESEARCH JOURNAL ON ADVANCED SCIENCE HUB

e-ISSN : 2582 - 4376  
Open Access

## RSP SCIENCE HUB

(The Hub of Research Ideas)

Available online at www.rspsciencehub.com

Special Issue of First International Conference on Innovations in Engineering Sciences (ICIES 2020)

### Analysis of Student Academic Performance through Expert systems

Kandula Neha<sup>1</sup>, Dr. S Jahangeer Sidiq<sup>2</sup>

<sup>1</sup>Assistant Professor, Artificial Intelligence, Lovely Professional University (Punjab)

<sup>2</sup>Assistant Professor, Computer Science and Engineering, Lovely Professional University (Punjab)

<sup>1</sup>neha09kandula@gmail.com, <sup>2</sup>jahangir.25453@lpu.co.in

#### Abstract

*Predicting the performance of students is one of the most important topics required for learning contexts such as colleges and universities, as it helps to design successful mechanisms that boost tutorial outcomes and prevent dropouts among various items. These are benefited by automating the many processes involved in the activities of usual students which handle huge volumes of information collected from package tools for technology-enhanced learning. Thus, the careful analysis and interpretation of these information would provide us with valuable data regarding the data of the students and therefore the relationship between them and hence the tutorial tasks. This data is the supply which feeds promising algorithms and methods able to estimate the success of the students. During this analysis, virtually many papers were analysed to show radically different trendy techniques widely applied to predict the success of students, along with the goals they need to achieve in this area. These computing-related techniques and approaches are mainly machine learning techniques, deep learning techniques, Artificial Neural Networks & Neural Networks Convolution, etc. This paper demonstrates the analysis and their comparisons of various methods used to forecast Student Academic success.*

**Keywords:** Deep learning, Machine Learning, Prediction, Academic Performance, Analysis

#### 1. Introduction

Educational facilities area unit running in Associate in Nursing more and more competitive and complicated environment and facing a burden in dealing with world and national economic, these facilities try to establish modifications such as the rising one to improve the magnitude relationship of scholars in specific disciplines and to ensure that the quality of learning programmes area units each Area of students unit the first stakeholders within the academic establishments and their success play a notable role in the social and financial growth of an extremely nation by delivering innovative alumni. There's a fundamental interest in pedantic institutions staying up and providing massive student datasets for useful simple leadership. Furthermore, the use of net innovation has become an integral part of

this educational era in many universities, improving the essential exchange of information about students, teachers and their contact with learning and educational frameworks. Advanced education plays an significant part in a society's progress and growth. It is an environment that offers lots of membership information, such as students, teachers, offices and training programmes.

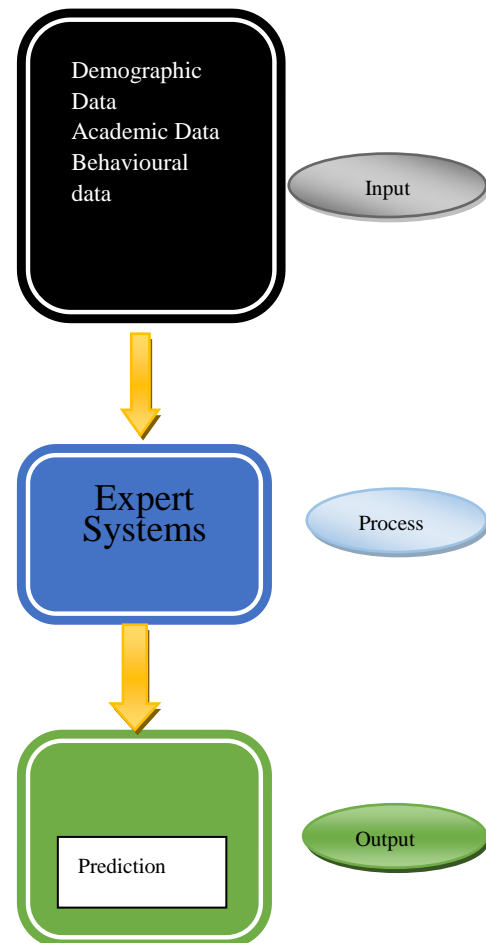
Scholars' success can be a major concern for different stakeholders as well as researchers, superiors and organisations. Therefore students should strive flat out for glorious marks, so they can rise to educational partners' standards. Student's success can be an basic concern of academic institutions, for the glorious educational accomplishment of a particular university will cause that university's level to increase. The

success of the student may not be inheritable by operation of the co-curriculum and measurement of earnings. Most of the researchers, however, stated that graduation is the student's live performance. Data science stands out as a discipline to replace. It is seen as an associate in the incorporation of traditional disciplines such as data processing, analytics, distributed systems and databases into nursing. It was important to merge current methods to view abundantly usable information in useful data for society, individuals and organisations. Information science tools are widely used to identify procedures, evaluate bottlenecks, verify enforcement and even suggest changes.

Data science depends on the knowledge from account data, observations, and prophetic models. This attempt involves curation and cleaning up at one end, and disseminating information at the other. After the successful creation and copy of systems and software packages the unit area capable of efficiently store, retrieve and efficiently store information on methods, focus has now shifted to predictive and correlative analytics. The fields of machine learning and information science have been robustly illustrated in Hell, this demand for a new approach, one that has hitherto seen some example in the syllabus of education and in the literature of scientific discipline. To employ information science in education, students would obtain data on their success in terms of their colleagues' importance or development. On the contrary, directors and decision-makers will take advantage of the information science to try to improve the environment of tutoring.

There is typically a positive thing that should be able to anticipate the actions of potential students in order to improve instructional style and set up programmes on the programme provided to the scholars for educational support and steering. This can be anywhere data processing is involved (DM comes in). DM techniques analyze datasets and extract info to rework it into graspable structures for later use. Machine Learning (ML), cooperative Filtering (CF), Recommender Systems (RS) and Artificial Neural Networks (ANN), Convolution Neural Networks (CNN) are the most machine techniques that method this info to predict students' performance, their grades. Nowadays, there is a significant amount of analyses and

studies that follow on the lines of predicting the actions of the students, among various related topics of interest within the educational space. In reality, several papers are printed in journals and conference on this subject. Hence the main objective of this analysis is to provide an exhaustive overview of the various theoretical techniques and algorithms applied to the present



**Fig 1 : Framework**

**Methodologies Used :** There are few Methodology techniques used for the prediction of students performance and few of them are listed below.

**Education Data Mining:** Data mining, additionally popularly referred to as data Discovery in information, refers to extracting or "mining" data from massive amounts of information. data processing techniques area unit accustomed care for massive volumes of information to get hidden patterns and relationships useful in deciding. whereas data processing and data discovery in information area

unit of times treated as synonyms, data {processing} is truly a part of the data discovery process. Various techniques and few kind of algorithms like Classification, , Regression, Neural Networks, Rules of association, Trees, Genetic, Clustering , Techniques in Nearest Neighbour etc., area unit used for data discovery from databases. These techniques and have different ways for processing of data and temporary mention in understanding in own ways.

**A. Classification :** Classification is that it uses the most efficiently applied data processing method, a series of pre-classified examples for the creation of a model that will classify the database population to huge. This time approach which is called tree or classification algorithms based on neural networks. The method of classification of knowledge applies to learning and classification. In Learning the unit of coaching information field defined by the law of classification. Consider the unit of information field used to estimate the consistency of classification rules in classification. If the precision is sufficient the rules would apply to the latest Tuples of information. The classifier-training rule uses these pre-classified examples to see the set of parameters needed for correct discrimination. The rule these parameters uses and then encodes into a model known as a classifier.

Classification approach can be used for successful suggests that nevertheless it becomes costly to classify teams or groups of objects so bunch would be used as a pre-processing method for collection and classification of attributes set. [19]

**C. Regression technique** For postulation, will be customised. Multivariate analysis will be used to model the relation between one or more variables depending on freelance and a few variables. We already appear to toll-known attributes in the processing of freelance variables, and response variables square measure what we would like to expect. Unfortunately, many real-world issues do not seem to be just a guess. Hence, several complicated techniques ( e.g., supply regression, call trees, or neural nets) may also be required to forecast future values. Typically similar varieties of the model can be used for each regression and classification. . for instance, the CART (Classification and Regression Trees) decision tree rule are accustomed build every classification tree (to classify categorical response variables) and

regression trees (to forecast continuous response variables). Neural networks may also turn out every classification and regression models. [19]

**D. Association rule:** Association and correlation is typically to look out frequent item set findings among huge data sets. this kind of finding helps businesses to create sure selections, like catalog vogue, cross-selling, and shopper wanting behaviour analysis. Association Rule algorithms have to be compelled to be able to generate rules with confidence values however one. however the quantity of gettable Association Rules for a given dataset is generally really giant and a high proportion of the principles unit usually of little or no (if any) worth. [19]

**E. Neural networks :** Neural network may be a collection of connected input/output units and each affiliation options a weight gift with it. The network learns in the coaching half by changing weights so that the correct class Tuple labels can be predicted. Neural networks have the exceptional capacity to derive that indicates difficult or inaccurate data, and may well be familiar with extract patterns and spot trends that unit of measurement is too complex to be identified by either humans or different computer techniques. This measurement unit is like minded for inputs and outputs that are continuously measured. Neural networks unit of measurement best at distinctive patterns or data trends and as intended for analysis or declaration desires [19]

**Machine learning:** Machine learning techniques such as fully connected feed forward Artificial Neural Network, Naïve Bayes, Decision Tree, have been used to build the machine learning model.

### Artificial Neural Networks

Artificial Neural Network represents a group of input unites and output unites. They are linked by weighted connexions to every other. The ANN learns by excessively adjusting the weights of the links, so it can predict the right target mark for a few instances of the input file. Backpropagation algorithmic method is one of the noted learning algorithms used to train the ANN. ANN has some advantages, such as its high resistance to yelling data sets and its strong success in classifying patterns which have not been trained so it is used in things until there is some knowledge about the relationship between the category mark and the

options inside the dataset.

The ANNs have many worldwide applications such as image and written recognition, voice recognition, laboratory medicine, and pathology. There are several kinds of ANNs that would help their design and style to be listed. One example may be a fully connected multilayer feed forward ANN during which the network has an input layer Associate in Nursing, one or more hidden layers, and also the output layer. [17]

In this analysis a three-layer fully linked feed forward ANN was used. The network is made up of input layer Associate in Nursing, 2 hidden layers, and also the output layer. The main layer consists of twenty input units, neurons, while the primary hidden layer includes six hidden units. The second layer concealed has 3 hidden units. The fourth layer is the output layer which only has 1 unit capacity. The linear measure Rectifier was used so the activation of the hidden units is done.

**Naive Bayes :** Naive Bayes classification model is called because the theorem network's simplest variant is. This model assumes that given the target attribute status each function attribute is freelance from the opposite attributes. . instance  $x$  within the dataset contains values for the attributes  $a_1, a_2, \dots, a_i$ . The target perform  $f(x)$  equals any price from predefined finite set  $V=(v_1, v_2, \dots, v_j)$ . Naïve mathematician model uses the subsequent equation.[17]

**Decision Tree :** A model for a decision tree represents a tree structure which is just like a flow diagram. -- internal node during this structure represents a check on the attribute of the dataset, while each limb represents the check result. Moreover, each leaf node represents a target function code, and the higher initial node within the tree also represents the base node. Decision trees can be either binary or non-binary. Call trees are common strategies for classifying victims that they do not want prior knowledge on the subject domain or sophisticated classification criteria structure. Moreover, they can actually be resurrected to the laws of classification and that they can be interpreted clearly. This / tree classification technique has been used in many real word applications such as monetary research, medicine, biology, development and uranology.

**Data Mining Techniques:** As several machine learning algorithms square measure won't to train

the scholar prophetic models therefore after analyzing several of them for comparison functions we have a tendency to have elite supplying regression and support vector machine using sequent stripped improvement 2 wide famous classifiers for prediction.

**a) Logistics Regression:** Regression of logistics is widely used anywhere categorical goal variable is used and the most famous constant quantity approach is taken into account. Generally, providing square measure linear models of regression models used for predicting some categorical dependent worth primarily based on some freelance predictor meaning. It can also solve the problems of binary (two classes) classification as multinomial (two or additional classes) values. Logistic regression model operates on logic perform design to find the linear combination or prevalence possibilities of any targeted group worth supporting predictor values (continuous or discrete).

**b) Support Vector Machine(SVM):** Additionally named SVM support vector machine is one of the most common machine learning algorithms used for classification, prediction, and regression. Since these algorithms work best with analysing large datasets with many predictor attributes, SVM algorithms have a nice use in text recognition or categorization, and bioinformatics. The SVM model is regarded as a discriminative maximum-margin model and operates on the principle of classifying information point into 2 or additional categories by locating the associate degree optimistic decision boundary which is assumed to be N-dimensional hyper plane and should be at a massive distance from the datum of each category.

**Deep Learning :** Deep learning classifiers are given to the data set collected from educational environments. Data is pre-processed and check for missing values. Classifiers are applied on the data set to build the models. Models are tested with test data to predict the students' performance and the best models yielding high accuracy are considered. The Methodologies Proposed in this research work has the important phases such as Artificial Neural Network, Supervised and Unsupervised Learning, Deep Neural Networks and Convolution Neural Networks.

**Artificial Neural Network:** Artificial neural network approaches used to predict academic



success. This is also intended to identify aspirants to enter university profession at multiple levels in line with the probability of reaching a success standard. This study identifies shortcomings for students during the academic career from various viewpoints in courses, and offers ways to prevent them. Many educational writers concentrated on learning varieties and their recommendations.

They depend on the fact that students have various temperament types that appear to have different learning varieties that influence students' performance in each subject. Through this study

we can build an artificial neural network to predict academic success that has supported the type of student training. Neural networks are efficient and widespread implementation in enormous variety of information mining applications that exceed classifiers according to the study. This main objective of this analysis is to examine whether square neural networks measure a correct classifier to predict LMS (learning management system) scholars performing in academic data processing context.

**Table.1 Comparative Research work of Data Mining and outcomes :**

SI No	Title	Techniques used	Outcome
1	Edifice an Educational Framework using Educational Data Mining and Visual Analytics[11]	Classification and Association Rule mining	Analysis of the students' performance.
2	Analysis of Girls Vocational High School Students Academic Failure Causes with Data Mining Techniques[12]	Clustering Technique	Analysis of failure causes of high school students
3	Educational Data Mining Classifier For Semester One Performance to Improve Engineering Students Achievement [13]	Classification Technique and Neuro-Fuzzy classifier	Assessment of students at early in the beginning of the course
4	Students' Employability Prediction Model through Data Mining [14]	Classification Bayesian methods, Multilayer Perceptions and Sequential Minimal Optimization (SMO), Ensemble Methods and Decision Trees	Predict the employability of the PG students
5	Mining Educational Data to Analyze Students' Performance [15]	Classification decision tree method	Predict the enrolment of students in a particular course for higher education system in the university and to identify the dropouts.
6	Students Academic Failure Prediction using Data Mining [16]	Classification and Feature selection Algorithm	Designed a model to predict student failure at the end of the semester.

**Convolution Neural Network** : CNN is commonly used for prediction functions, as CNN provides several settings that need to be outlined and manipulated to achieve maximum precision in performance. These variables have a strong effect on the CNN model's efficiency (i.e., efficiency of

activation, variety of layers, variety of vegetative cells in each layer, and hyper-parameter). If the data set is obtained, the model will be coached with a series of tests and various models will be used to test the predicted model.

Authors follow the variety of epoch and number of layers as key variables for manipulating until maximum precision is achieved. As a result of the

experiments, the accuracy measures that have been obtained will be contrasted with few works in the literature that have discussed the success prediction of the behaviour of the students. End Results of Few Methodologies and their accuracy

**Table2 - Methods Accuracy**

Method	Accuracy
Multi Class Classifier	99.81
SVM	93.9
Naïve Bayes	97.45
Random Forest	99.3
Decision Tree	98.9
CNN	100

**Conclusion:** This survey paper gives the Prediction of students performance and how it has become an huge trend these days in research work. In order to predict their performances many algorithms and techniques have been designed and have been used to get the accurate values. These predictions are been done based on many parameters of the individual students performances like behaviour of students in their home and classroom , their extra circular activities, classroom activities, behaviours in classroom, participation in quizzes and group discussions etc. All these parameters makes an huge impact on their academic performances. A Survey has been done on many Deep Learning based algorithms such as Artificial Neural Networks, Machine Learning, Support Vector Machines , Convolution Neural Network etc to draw conclusions based on their accuracy and to select the best deep learning algorithm to predict the students performance.

**Reference :**

- [1].Sri Lakshmi, A V Krishna Prasad "Research Methodologies for Student Performance Evaluation Using Educational Analytics Tools and Approaches", International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8, Issue-1S4, June 2019
- [2].J. Han, M. Kamber, and J. Pei, Data Mining: Concepts and Techniques, 3rd ed. Morgan Kaufmann publications, 2012.
- [3].Daud, Ali, et al., "Predicting student performance using advanced learning analytics," Proceedings of the 26th

international conference on world wide web companion. International World Wide Web Conferences Steering Committee, 2017.

- [4].Mansur, Andi Besse Firdausiah, Norazah Yusof, and Ahmad Hoirul Basori, "Personalized Learning Model based on Deep Learning Algorithm for Student Behaviour Analytic," Procedia Computer Science vol. 163, pp. 125-133, 2019.
- [5].Patil, Akhilesh P., Karthik Ganesan, and Anita Kanavalli, "Effective deep learning model to predict student grade point averages," IEEE International Conference on Computational Intelligence and Computing Research (ICCIC), pp. 1-6, 2017.
- [6].Waheed, Hajra, et al., "Predicting academic performance of students from VLE big data using deep learning models," Computers in Human Behavior, vol. 104, pp. 106189, 2020.
- [7].Students' Academic Performance Dataset, [Online], available <https://www.kaggle.com/aljarah/xAPI-Edu-Data>, Access on August 2019.
- [8].Predicting Students' Academic Performance Through Supervised Machine Learning, 978-1-7281-6899-9/20//\$31.00 ©2020 IEEE
- [9].R.Arora, "Comparative analysis of classification algorithms on different datasets using WEKA," International Journal of Computer Applications, vol. 54(13),2012.
- [10]. Arsad, PM, Buniyamin, N and Manan, J-IA. 2013. A neural network students' performance prediction model (NNSPPM). 2013 IEEE International Conference on Smart Instrumentation, Measurement and Applications (ICSIMA), (November): 1–5.
- [11].Dr. S. Anupam Kumar (2016) "Edifice an Educational Framework using Educational Data Mining and Visual Analytics" I.J. Education and Management Engineering, 2, 24-30 .
- [12].Deperlioglu, Omer & Sibel Birtil, Fatma. (2016). Analysis of Girls Vocational High

School Students' Academic Failure Causes with Data Mining Techniques. The Anthropologist.

- [13]. Usamah Bin Mat, Norlida Buniyamin and Pauziah Mohd Arshad (2016) Engineering & Technical Education Research Group (Enter), Faculty of Electrical Engineering, Universiti Teknologi MARA, Middle-East Journal of Scientific Research 24 (2): 338-346, 2016.
- [14]. Mishra, Tripti & Kumar, Dharminder & Gupta, Sangeeta. (2016). Students' employability prediction model through data mining. 11. 2275-2282.
- [15]. Brijesh Kumar Baradwaj, Saurabh Pal (2011) International Journal of Advanced Computer Science and Applications, (IJACSA) Vol. 2, No. 6
- [16]. Jantawan, Bangsuk & Tsai, Cheng-Fa. (2013). The Application of Data Mining to Build Classification Model for Predicting Graduate Employment.
- [17]. Hussein Altabrawee, Osama Abdul Jaleel Ali, Samir Qaisar Ajmi, "Predicting Students' Performance Using Machine Learning Techniques", Journal of University of Babylon, Pure and Applied Sciences, Vol.(27), No.(1): 2019
- [18]. Student Academic Performance Prediction using Supervised Learning Techniques <https://doi.org/10.3991/ijet.v14i14.10310>
- [19]. Mining Educational Data to Analyze Students' Performance, (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 2, No. 6, 2011.
- [20]. A Comparative Analysis of Techniques for Predicting Student Performance, Proceedings of the 9th International Conference on Educational Data Mining