Automated Detection of Depression and Anxiety Using Lexical and Phonestheme Features in Spanish Texts

Elizabeth Martinez[1], Juan Cuadrado[1], Juan Carlos Martinez-Santos[1] and Edwin Puertas[1,*,†]

[1]*Universidad Tecnologica de Bolivar, School of Engineering, Cartagena de Indias 130010, Colombia.*

Abstract

Depression and anxiety are common mental disorders characterized by persistent sadness, lack of interest, and heightened worry or fear. Detecting these disorders is essential for timely intervention and support. This paper presents our approach to detecting depression, anxiety, or neither using text data, focusing on a multiclass classification task. We utilize a dataset of labeled messages from Telegram groups related to mental health. Our methodology involves data pre-processing, lexical feature extraction, phonestheme embeddings, VAD (Valence Arousal Dominance) analysis, and emotion detection. During the training phase, we obtained promising results through extensive experimentation and model refinement. However, we encountered challenges in the MentalRiskES evaluation, particularly regarding latency and real-time detection speed. This research contributes to the field of automated mental health assessment and provides insights into the potential of text analysis techniques for identifying depression and anxiety. We acknowledge the need for further improvement and remain dedicated to advancing our methodology to support better individuals affected by these mental disorders.

Keywords

Mental Risk, Depression, Anxiety, Lexical Features, Phonesthemes Embedding, VAD, Emotion Detection

1. Introduction

Depression and anxiety are prevalent mental health disorders that affect millions of individuals globally, posing significant challenges for timely detection and intervention [1]. Traditional methods, including clinical assessments and self-reporting, are often limited by subjectivity and resource constraints [2, 3]. As a result, there is a growing interest in developing automated methods that utilize natural language processing (NLP) and machine learning (ML) techniques [4].

Substantial research efforts have focused on detecting depression and anxiety by analyzing a variety of linguistic and contextual features, sentiment analysis, and different ML algorithms [5, 6, 7]. Despite these efforts, accurately identifying these mental health disorders from textual data remains challenging due to the subtle and nuanced nature of language expressions

IberLEF 2024, September 2024, Valladolid, Spain

*Corresponding author.

eayala@utb.edu.co (E. Martinez); jflechas@utb.edu.co (J. Cuadrado); jcmartinezs@utb.edu.co (J. C. Martinez-Santos); epuerta@utb.edu.co (E. Puertas)

0000-0001-6592-347X (E. Martinez); 0000-0002-8226-1372 (J. Cuadrado); 0000-0003-2755-0718 (J. C. Martinez-Santos); 0000-0002-0758-1851 (E. Puertas)

associated with mental health [8]. Moreover, the limited availability of annotated datasets for training and evaluation further complicates the development of effective detection models [9].

This paper introduces our methodology for detecting depression and anxiety using the IberLEF 2024 Mental Risk dataset [10], which includes labeled messages from Telegram groups centered on mental health issues [11, 12]. Our approach involves a comprehensive pipeline that begins with data pre-processing, such as the removal of stop words and punctuation, followed by the extraction of lexical features to identify linguistic patterns [13]. To further enhance our model, we incorporate phonestheme embeddings, Valence Arousal Dominance (VAD) analysis [14], and emotion detection [15], utilizing phonetic information to capture contextual representations [16].

During the training phase, our model yielded promising results, achieving approximately 81% accuracy, macro F1, and macro Fp scores in multiclass classification using RandomForest [17]. The model also performed competitively when integrating lexical features and phonesthemes [18]. However, the evaluation conducted by IberLEF and the MentalRiskES organization [11] revealed a notable decline in performance, highlighting the necessity for further refinement and enhancement of our methodology.

2. Related Work

The automatic detection of depression and anxiety using machine learning and natural language processing has gained significant attention in recent years. Researchers have explored various approaches to identify mental health signals in text data, leveraging linguistic and contextual features. This section provides an overview of significant studies in this domain.

De Choudhury et al. [19] pioneered using Twitter data to predict depression among users, highlighting social media's potential as a data source for mental health prediction.

Burdisso et al. [20] proposed a text classification framework for the early detection of depression in social media streams. Their approach, based on supervised learning techniques, outperformed traditional models in both computational efficiency and interpretability.

Chiong et al. [21] examined the effects of text pre-processing and feature extraction on depression detection using machine learning classifiers, demonstrating effective detection even without explicit depression-related keywords.

Amanat et al. [22] developed a model using recurrent neural networks (RNN) and long short-term memory (LSTM) networks for detecting depression in textual data. Their model achieved a high precision of 99.0

Babu and Kanaga [23] reviewed sentiment analysis for detecting depression in social media texts, emphasizing the role of multiclass classification techniques and deep learning algorithms in improving detection accuracy.

Mustafa et al. [24] utilized sentiment analysis and psychological attributes to classify depression levels in social media users, underscoring the importance of feature selection and combination to enhance classifier performance.

These studies collectively underscore the potential of automated methods to improve the speed and reach of mental health diagnostics, providing timely intervention and support.

3. Data

For the 2024 edition of the MentalRisk competition, hosted by IberLEF, a newly tailored dataset has been introduced to detect depression and anxiety. This dataset comprises labeled messages sourced from public Telegram groups popular among users discussing various mental health-related topics. These groups, primarily Spanish-speaking, cover a broad range of subjects directly tied to mental disorders.

The organizers of MentalRisk IberLEF 2024 carried out the extraction and anonymization of conversations from public Telegram groups. Telegram, a cross-platform, encrypted, and cloud-based messaging service, allows users to send and receive messages individually or in groups. Using the Prolific service, the organizers connected annotators to a specialized labeling tool to manually review and label each user's chat history. Ten annotators reviewed each user's history.

The probability of a disorder is established by dividing the number of annotators that identified evidence of the disorder by the total number of annotators (ten).

The dataset for Task 1 was divided into three subsets, each associated with different mental disorders. Specifically, the corpus, comprising 884 users, was split into trial, training, and testing sets of 20, 464, and 400 users, respectively. On average, each user contributed approximately 50 messages, providing a robust dataset for developing and evaluating detection models.

4. Architecture

This section provides a detailed overview of the predictive model designed to address Task 1 of the Mental Risk Challenge in the IberLEF 2024 competition. Task 1 focuses on multiclass classification to determine whether users, based on their textual messages, exhibit signs of depression, anxiety, or neither.

To effectively tackle this task, our model employs a structured approach comprising several key stages. Initially, the process involves reading and pre-processing the data, where multiple messages from the same user are concatenated into a single string to maintain the continuity of information. Following this, the model proceeds through feature extraction, regularization, classification, and evaluation stages, each playing a crucial role in enhancing the model's accuracy and reliability.

In the subsequent subsections, we will delve into each stage, detailing the methodologies and techniques employed to construct our predictive model.

4.1. Pre-Processing

We implemented a thorough cleaning process to ensure the quality and suitability of the text data from the IberLEF 2024 competition. One key step in this process was the removal of stop words.

Stop words, such as articles, prepositions, and conjunctions, are common in any language but generally do not add significant meaning. Removing these words reduces noise in the data and improves the accuracy of natural language processing tasks.

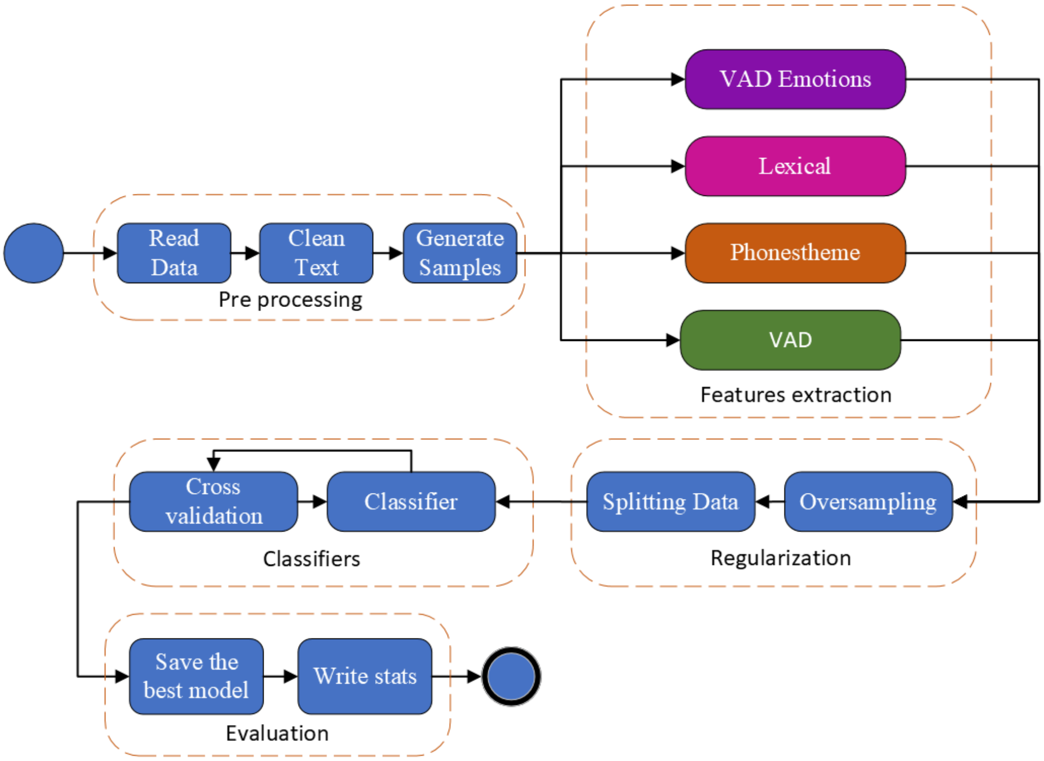

Figure 1: System Pipeline.

By eliminating these non-essential words, we aimed to enhance the dataset's focus on relevant linguistic features related to mental health. This pre-processing step was crucial for improving the model's ability to identify patterns and extract meaningful information pertinent to detecting depression and anxiety.

4.2. Feature Extraction

In sentiment analysis, extracting key linguistic features from the text is crucial for accurate analysis. This section describes the different feature extraction methods implemented in our model.

4.2.1. Lexical Features

In sentiment analysis, capturing key linguistic elements is essential, and lexical features play a crucial role. Our model leverages several important lexical features to analyze text data effectively. By including first, second, and third-person pronouns (both singular and plural), we gain insights into self-reference and interpersonal dynamics, which may signal depressive tendencies. For example, a frequent use of first-person pronouns might suggest an increased self-focus, often linked with depression. Additionally, adverbs denoting time, negation, place, manner, and quantity provide vital contextual information, aiding in the accurate interpretation of the text's temporal, spatial, and emotional nuances. Words like "never" or "always" can indicate extreme viewpoints commonly associated with depressive or anxious expressions.

Furthermore, our model incorporates negative and positive adjectives to capture the emotional tone of the text, reflecting either pessimistic or optimistic sentiments. Adjectives such as "sad" or "happy" directly convey the user's emotional state. We also consider specific lexical patterns, including mentions, URLs, hashtags, emojis, and retweets, which reveal communication styles and trends linked to depressive symptoms. For instance, frequent use of sad emojis or negative hashtags can highlight a negative emotional state. By integrating these diverse lexical features, our model provides a comprehensive analysis of the sentiment conveyed in the text, offering valuable insights into potential mental health conditions.

4.2.2. Phonestheme Embedding

Phonestheme embedding served as a feature extraction technique to enhance the sentiment analysis of text data. This method captures the phonetic characteristics of words, enabling the model to grasp additional linguistic subtleties. We employed pre-trained Word2Vec models that were specifically trained on phonesthemes, which are the fundamental sound units in a language. These models produce vector representations that accurately reflect the phonetic structure of words, adding a layer of linguistic information that enriches the overall analysis.

4.3. VAD

We utilized the NRC-Lexicon to extract Valence, Arousal, and Dominance (VAD) for each sentence in the user's messages. These features quantify the emotional content across the user's text.

- Valance: Indicate how positive or negative is the emotional content of the text.
- Arousal: Refers to the level of excitement or activation associated with the text
- Dominance: Captures the degree of control or power expressed in the text.
  For each sentence, we calculated the VAD scores for individual words and then averaged these scores to obtain a mean VAD value for the sentence. This averaging process involved summing the valence, arousal, and dominance scores for all words in a sentence and dividing by the total number of words.

4.4. VAD Emotion

The NRC-Lexicon also includes an emotion lexicon that associates specific words with particular emotions. We used the Spanish version of this lexicon to identify emotions in the text, capturing the following emotions: anger, anticipation, disgust, fear, joy, sadness, surprise, and trust. For each word in the text, we recorded the associated emotion scores and then averaged these scores to determine the overall emotional content of the user's messages. This averaging process involved summing the emotion scores for all words and dividing by the total number of words in the sentence.

4.4.1. Unified Feature Vector Construction

We combined the feature vectors generated from lexical features, Phonestheme embeddings, VAD, and VAD emotions to consolidate the diverse information extracted. This process results in a single, unified feature vector that enhances the robustness of our analysis and prediction capabilities. By merging these various features, our model captures a broad spectrum of linguistic signals and patterns, allowing for a comprehensive understanding of the sentiment expressed in the text and providing valuable insights into mental health conditions.

This integrated feature vector is then used as the input for the subsequent stages of our predictive model, which involve classification algorithms or regression models. These models leverage the consolidated features to predict whether an individual is experiencing depression, anxiety, or neither, and to estimate the probability of these mental health conditions.

4.5. Regularization

To increase the diversity and balance of our dataset, we applied data augmentation using the Synthetic Minority Over-sampling Technique (SMOTE). SMOTE creates synthetic examples by interpolating between existing samples from minority classes, which helps to mitigate class imbalances and improve the representation of these groups.

In addition, we used class-balancing methods to achieve a more equitable distribution of instances across the different categories. This step was crucial for boosting the model's performance and reducing biases due to an uneven dataset.

By combining data augmentation with regularization techniques, we effectively addressed class imbalance issues. This integrated approach significantly enhanced the model's ability to predict mental health conditions accurately.

4.6. Classifiers

This subsection offers an overview of the classifiers employed for the multiclass classification in Task 1. We selected various Scikit-learn [25] classifiers with default settings, each with different characteristics, to ensure both accuracy and reliability in our results.

4.6.1. Multiclass Classification Classifiers

To evaluate their performance and predictive capability, we implemented a variety of classifiers for the multiclass classification in Task 1. The classifiers used include Random Forest (RF), Decision Tree (DT), Naive Bayes (NB), Logistic Regression (LR), Support Vector Machine (SVM), and Multi-layer Perceptron (MLP).

4.6.2. Cross-Validation

Cross-validation is a widely used technique for evaluating the performance of machine learning models, with the primary aim of estimating how well a model will generalize to new, unseen data. In this study, we employed cross-validation to reliably assess the classifiers and regressors used for detecting depression in text data.

We specifically utilized k-fold cross-validation, which involves dividing the dataset into k equally sized subsets (folds). The model is then trained and evaluated k times, with each subset serving as the test set once, while the remaining k-1 subsets are used for training. This methodology ensures that all instances in the dataset are utilized for both training and testing, providing a thorough evaluation of the model's performance.

4.7. Evaluation

In this phase, we compared the results of various models to select the one with the best performance across multiple metrics.

We used Accuracy, Precision, Recall, and F1 scores as evaluation metrics for determining the classifiers' effectiveness in classifying instances and balancing true positives, false positives, and false negatives.

5. Experiments Conducted and Training

In our experiments, we followed a systematic approach to classify text data. First, we evaluated classifiers using individual feature extraction methods. Then, we assessed the performance of concatenated features, combining phonestheme, VAD, VAD Emotions, and lexical features. These steps were repeated for multiclass classification in Task 1.

Pre-processing and feature extraction were conducted sequentially. At the classification stage, all classifiers were run in parallel for a comprehensive comparison. We used various metrics to measure their performance, providing insights into the effectiveness of different feature combinations and classifiers for the disorder detection task.

5.1. Results of Training Process for Multiclass Classification Task 1

For the multiclass classification Task 1, we evaluated several models using different feature combinations. The models tested included Random Forest (RF) with 200 estimators, Decision Tree (DT) with a maximum depth of 4, Gaussian Naive Bayes (NB), Logistic Regression (LR), and Multi-Layer Perceptron (MLP). Each model was assessed based on its performance with various feature sets.

The performance metrics for these models are summarized in Table 1. This table presents the accuracy, precision, recall, and F1 score for each feature combination and model.

Table 1
Performance of Different Feature Combinations

Approach	Model	Accuracy	Precision	Recall	F1
VAD, Phonesthemes, Lexical	RF	0.78	0.78	0.78	0.78
VAD, VAD Emotions, Phonesthemes, Lexical	RF	0.81	0.81	0.81	0.81

The results indicate that the Random Forest model consistently outperformed other models across multiple feature combinations. Specifically, the RF model achieved the highest accuracy and precision using the combined VAD, VAD Emotions, phonesthemes, and lexical features. This highlights the effectiveness of the RF model for the task, making it the most suitable choice for detecting depression and anxiety based on text data.

The superior performance of the RF model with combined features demonstrates its capability to leverage diverse linguistic cues and patterns, providing a comprehensive analysis for multiclass classification in mental health detection.

6. Results of VerbaNex AI in MentalRiskEs Task Evaluation

In this section, we present the results achieved by the VerbaNex AI team in evaluating the MentalRiskES task for the current year. This task focused on detecting mental disorders, specifically depression and anxiety, in Spanish comments from Telegram users. For each subtask, the same model was submitted multiple times for evaluation by the VerbaNex AI team.

The VerbaNex AI team approached Task 1 using various classification models to determine if users exhibit signs of depression or anxiety. Table 2 displays the classification-based evaluation results for Task 1, including the team's rank, run number, accuracy, macro-precision (Macro-P), macro-recall (Macro-R), and macro-F1 scores. These metrics provide a comprehensive assessment of the accuracy and overall performance of the team's models in detecting mental disorders.

Table 2
Classification-based evaluation in Task 1 - VerbaNex AI

Rank	Team	Run	Accuracy	Macro-P	Macro-R	Macro-F1
25, 26	VerbaNex AI	1, 2	0.527	0.598	0.372	0.303
27	VerbaNex AI	0	0.512	0.551	0.353	0.271

Additionally, Table 3 presents the team's performance in terms of latency and speed of detection for Task 1, including metrics such as ERDE5, ERDE30, latencyTP, speed, and latency-weighted F1 scores. These metrics evaluate the team's efficiency in promptly identifying mental disorder risks.

Table 3
Latency-based evaluation in Task 1 - VerbaNex AI

Rank	Team	Run	ERDE5	ERDE30	latencyTP	speed	latency-weightedF1
27, 28	VerbaNex AI	1, 2	0.440	0.439	1	1	0.221
29	VerbaNex AI	0	0.458	0.458	1	1	0.164

Our analysis of failed predictions revealed that messages with ambiguous or less explicit emotional content were frequently misclassified. This highlights the need for more robust contextual understanding in future models. Additionally, runs 1 and 2 used the same system configuration, resulting in identical outcomes, as reflected in Table 2.

Furthermore, the carbon emissions associated with our experiments were measured. For the VAD-based approach, the total emissions were approximately 0.0037 kg CO_2, with an energy consumption of 0.0082 kWh. For the VAD and emotion-based approach, the total emissions were approximately 0.0015 kg CO2, with an energy consumption of 0.0111 kWh. These metrics are crucial for understanding the environmental impact of our computational processes.

7. Conclusion

In this study, we developed and evaluated a model for detecting depression and anxiety from text data, leveraging a combination of lexical features, phonesthemes, VAD, and VAD Emotions. Our approach demonstrated promising results, with the Random Forest classifier consistently outperforming other models across multiple feature combinations. The integration of diverse linguistic features enabled our model to capture a broad spectrum of cues and patterns, providing a comprehensive analysis for multiclass classification in mental health detection.

Despite these encouraging results, our evaluation in the MentalRiskES task identified areas for improvement, particularly concerning latency and real-time detection speed. Addressing these challenges is crucial for enhancing the practical application of our model in real-time monitoring and intervention scenarios.

Additionally, we assessed the carbon emissions associated with our experiments, highlighting the environmental impact of our computational processes. For the VAD-based approach, the total emissions were approximately 0.0037 kg CO2, with an energy consumption of 0.0082 kWh. For the VAD and emotion-based approach, the total emissions were approximately 0.0015 kg CO2, with an energy consumption of 0.0111 kWh.

Overall, our research contributes to the field of automated mental health assessment by demonstrating the potential of advanced text analysis techniques for detecting mental disorders. We remain committed to refining our methodology and improving our model's performance to better support individuals affected by depression and anxiety.

8. Future Work

Future work will focus on addressing the limitations identified in this study. A key area of improvement is reducing latency and enhancing the real-time detection capabilities of our model, which involves optimizing the processing pipeline and exploring more efficient algorithms.

Another significant direction for future research is the refinement of our feature extraction methods. Incorporating more sophisticated techniques, such as advanced embeddings and deeper neural networks, may further enhance the model's performance and its ability to detect subtle linguistic cues.

We also plan to integrate additional linguistic features and improve regularization techniques to better handle class imbalances. This will help to enhance the robustness and accuracy of our model in detecting depression and anxiety from text data.

Furthermore, we aim to reduce the environmental impact of our computational processes by optimizing the energy efficiency of our models. The assessment of carbon emissions has provided us with insights into the environmental cost of our experiments, guiding us towards more sustainable practices in future work.

Finally, collaborating with mental health professionals will be crucial to validate our model's predictions in real-world settings. This collaboration will ensure that our approach aligns with clinical needs and ethical considerations, ultimately improving the support provided to individuals affected by these mental health conditions.

Acknowledgments

The authors would like to acknowledge the support provided by the master's degree scholarship program in engineering at the Universidad Tecnologica de Bolivar (UTB) in Cartagena, Colombia.

References

[1] D. Wasserman, M. Iosue, A. Wuestefeld, V. Carli, Adaptation of evidence-based suicide prevention strategies during and after the covid-19 pandemic, World psychiatry 19 (2020) 294–306.
[2] H. Kim, S. Lee, S. Lee, S. Hong, H. Kang, N. Kim, et al., Depression prediction by using ecological momentary assessment, actiwatch data, and machine learning: observational study on older adults living alone, JMIR mHealth and uHealth 7 (2019) e14149.
[3] J. M. Bolton, D. Gunnell, G. Turecki, Suicide risk assessment and intervention in people with mental illness, Bmj 351 (2015).
[4] J. Hirschberg, C. D. Manning, Advances in natural language processing, Science 349 (2015) 261–266.
[5] E. Puertas, L. G. Moreno-Sandoval, F. M. Plaza-del Arco, J. A. Alvarado-Valencia, A. Pomares-Quimbaya, L. Alfonso, Bots and gender profiling on twitter using sociolinguistic features, CLEF (Working Notes) (2019) 1–8.
[6] L. G. Moreno-Sandoval, E. Puertas, F. Plaza-Del-Arco, A. Pomares-Quimbaya, J. Alvarado-Valencia, A. Ureña-López, Celebrity profiling on twitter using sociolinguistic features notebook for pan at clef 2019, CLEF (Working Notes) (2019).
[7] E. Puertas, L. G. Moreno-Sandoval, J. Redondo, J. A. Alvarado-Valencia, A. Pomares-Quimbaya, Detection of sociolinguistic features in digital social networks for the detection of communities, Cognitive Computation 13 (2021) 518–537.
[8] M. Kabir, T. Ahmed, M. B. Hasan, M. T. R. Laskar, T. K. Joarder, H. Mahmud, K. Hasan, Deptweet: A typology for social media texts to detect depression severities, Computers in Human Behavior 139 (2023) 107503.
[9] M. Shoaib, B. Shah, S. Ei-Sappagh, A. Ali, A. Ullah, F. Alenezi, T. Gechev, T. Hussain, F. Ali, An advanced deep learning models-based plant disease detection: A review of recent research, Frontiers in Plant Science 14 (2023) 875.
[10] L. Chiruzzo, S. M. Jiménez-Zafra, F. Rangel, Overview of IberLEF 2024: Natural Language Processing Challenges for Spanish and other Iberian Languages, in: Proceedings of the Iberian Languages Evaluation Forum (IberLEF 2024), co-located with the 40th Conference of the Spanish Society for Natural Language Processing (SEPLN 2024), CEUR-WS.org, 2024.
[11] Alba María Mármol-Romero, Adrián Moreno-Muñoz, Flor Miriam Plaza-del-Arco, M.

Dolores Molina-González, María-Teresa Martín-Valdivia, L. Alfonso Ureña-López, Arturo Montejo-Ráez, Overview of mentalriskes at iberlef 2024: Early detection of mental disorders risk in spanish, Procesamiento del Lenguaje Natural 73 (2024).

[12] A. M. Mármol Romero, A. Moreno Muñoz, F. M. Plaza-del Arco, M. D. Molina González, M. T. Martín Valdivia, L. A. Ureña-López, A. Montejo Ráez, MentalRiskES: A new corpus for early detection of mental disorders in Spanish, in: N. Calzolari, M.-Y. Kan, V. Hoste, A. Lenci, S. Sakti, N. Xue (Eds.), Proceedings of the 2024 Joint International Conference on Computational Linguistics, Language Resources and Evaluation (LREC-COLING 2024), ELRA and ICCL, Torino, Italia, 2024, pp. 11204–11214. URL: https://aclanthology.org/2024.lrec-main.978.

[13] E. A. Puertas Del Castillo, et al., Análisis de elementos fonéticos y elementos emocionales para predecir la polaridad en fuentes de microblogging, 2023-12-01. URL: http://hdl.handle.net/10554/63548.

[14] J. Cuadrado, E. Martinez, J. C. Martinez-Santos, E. Puertas, team utb-nlp at finances 2023: financial targeted sentiment analysis using a phonestheme semantic approach, Procesamiento del Lenguaje Natural (2023).

[15] L. G. Moreno-Sandoval, E. A. P. Del Castillo, A. P. Quimbaya, J. A. Alvarado-Valencia, Assembly of polarity, emotion and user statistics for detection of fake profiles., CLEF (2020).

[16] J. M. Pérez, D. A. Furman, L. A. Alemany, F. Luque, Robertuito: a pre-trained language model for social media text in spanish, arXiv preprint arXiv:2111.09453 (2021).

[17] A. Paul, D. P. Mukherjee, P. Das, A. Gangopadhyay, A. R. Chintha, S. Kundu, Improved random forest for classification, IEEE Transactions on Image Processing 27 (2018) 4012–4024.

[18] E. Puertas, J. C. Martinez-Santos, Phonetic detection for hate speech spreaders on twitter, CLEF (2021).

[19] M. De Choudhury, M. Gamon, S. Counts, E. Horvitz, Predicting depression via social media, in: Proceedings of the international AAAI conference on web and social media, volume 7, 2013, pp. 128–137.

[20] S. G. Burdisso, M. Errecalde, M. Montes-y Gómez, A text classification framework for simple and effective early depression detection over social media streams, Expert Systems with Applications 133 (2019) 182–197.

[21] R. Chiong, G. S. Budhi, S. Dhakal, F. Chiong, A textual-based featuring approach for depression detection using machine learning classifiers and social media texts, Computers in Biology and Medicine 135 (2021) 104499. URL: https://www.sciencedirect.com/science/article/pii/S0010482521002936. doi:https://doi.org/10.1016/j.compbiomed.2021.104499.

[22] A. Amanat, M. Rizwan, A. R. Javed, M. Abdelhaq, R. Alsaqour, S. Pandya, M. Uddin, Deep learning for depression detection from textual data, Electronics 11 (2022). URL: https://www.mdpi.com/2079-9292/11/5/676. doi:10.3390/electronics11050676.

[23] N. V. Babu, E. G. M. Kanaga, Sentiment analysis in social media data for depression detection using artificial intelligence: a review, SN Computer Science 3 (2022) 1–20.

[24] R. U. Mustafa, N. Ashraf, F. S. Ahmed, J. Ferzund, B. Shahzad, A. Gelbukh, A multiclass depression detection in social media based on sentiment analysis, in: S. Latifi (Ed.), 17th

International Conference on Information Technology–New Generations (ITNG 2020), Springer International Publishing, Cham, 2020, pp. 659–662.

[25] F. Pedregosa, G. Varoquaux, A. Gramfort, V. Michel, B. Thirion, O. Grisel, M. Blondel, P. Prettenhofer, R. Weiss, V. Dubourg, J. Vanderplas, A. Passos, D. Cournapeau, M. Brucher, M. Perrot, E. Duchesnay, Scikit-learn: Machine learning in Python, Journal of Machine Learning Research 12 (2011) 2825–2830.