Models of automatic recognition of collaborative emotions in rural secondary education institutions in Puno, 2024.

Alejandro Apaza-Tarqui[1], Fidel Huanco-Ramos[2]

1 Professional School of Statistical and Informatics Engineering, National University of the Altiplano, Puno, Puno, Perú

2 Professional School of Systems Engineering, National University of the Altiplano, Puno, Puno, Perú.

Abstract

Nowadays, collaborative emotions expressed in social networks by teenagers are topics of educational and communication interest fundamental in human interactions, whose analysis has experienced an exponential growth in the digital era. The aim of the article is to determine a machine learning and/or deep learning model that performs better in the recognition of complex sequential text patterns of emotions expressed on social networking platforms by secondary school students in rural areas of Puno. The methodology comprises keyword sensing technique (KST) and natural language processing (NLP), and has been applied to a total of 2160 emotions expressed in comments on daily educational activities among friends and classmates on personal Facebook profile. Pre-processing, training, evaluation and implementation have been carried out in Google Colab with pandas, Scikit-learn, numpy, TensorFlow/Keras, Matplotlib and Seaborn libraries. As a result, the implementations of the ANN, SVM, Naive Bayes, XGBoost, Random Forest, KNN and LightGBM algorithms were compared and evaluated with the metrics of precision, recall, F1-score and AUC. ANN was found to be the best fitting model for the dataset, with an AUC of 97% correctly classifying emotions such as ‘Surprise’ and ‘Anger’, but has limitations with ‘Happiness’ and ‘Sadness’, emotions that are often confused with each other, especially with ‘Distress’. In conclusion, ‘Surprise’ highlights feelings expressed in the personal Facebook profile of students from rural Puno, where the Peruvian educational model does not meet the expectations of the highland youth who still preserve their own living culture of the Andean cosmovision, allowing them to survive in situations of poverty, a culture that instills in them the importance of solidarity and reciprocity, while ‘Anger’ reflects permanent frustrations due to marginalisation in basic sanitation services, motivating them to undertake migratory adventures to the cities of the coast. Once there, they suffer racial mistreatment, physical or psychological harassment and, emotionally, they lose self-esteem, unable to integrate into a full society with quality of life.

Keywords: emociones, aprendizaje profundo, cultura viva, redes sociales.

Author summary

The social network Facebook is massively used by adolescent secondary school students in rural communities in Puno. This practice has become very attractive for young people, especially for expressing their emotions on their personal profile walls, being an addiction that limits learning activities and is a reason for dropping out of school. For this reason, the use of artificial intelligence is essential to know in time the desires, wishes, expectations, inhibitions, fears, frustrations and inhibitions that are difficult to manifest in educational classrooms. In this society of the ayllu, an ancient term meaning a group

of families, with a living culture of the subsisting Andean cosmovision. There are studies on the analysis of feelings in other realities, but not in rural Andean societies in Puno. The result was the correct classification of emotions such as ‘Surprise’ and ‘Anger’, feelings expressed as rejection of the Peruvian educational model by the youth of the altiplano, who still preserve customs and traditions of the Andean cosmovision. The ayllu guarantees the survival of a very poor, culturally neglected reality, with permanent frustrations due to the lack of basic sanitation services. This population is motivated to undertake migratory adventures to the coastal cities, where they suffer racial abuse, physical or psychological harassment and, emotionally, they lose their self-esteem, without the possibility of being able to live in the cities. Therefore, the educational model in the Peruvian Andes is a complete failure.

Introduction

The study of emotions has become increasingly important in the field of academics and student well-being. The comments expressed in the personal social network profiles of students in educational institutions in rural Puno are diverse and have a direct impact on academic performance. Therefore, it is important to determine a machine learning and/or deep learning model that performs better in recognizing complex sequential and hidden textual patterns of emotions expressed on social networking platforms by secondary school students in rural Puno, in order to mitigate academic failure(Alcocer-Sánchez et al., 2023).

In rural secondary schools in Puno, students interact with technology and the strong roots of their living culture, to share experiences, traditions, customs, experiences of the Andean cosmovision, giving other ways to collaborate in their studies, homework and learning in general.

In secondary education centres in rural Puno, there are currently minimum conditions of access to the internet and social networks after the COVID-19 pandemic(Fidel & Alejandro, 2024), it is fundamental to understand how emotions affect personal, academic and social development. For which, it is important to classify, detect and predict such emotions in an adequate and timely manner, in order to vindicate the student's self-esteem, since, due to intercultural differences, such as the native language, they are subject to marginalization. All these are social taboos that must be discarded(Bermejo-Paredes et al., 2019).

In the Puno region, within the framework of the Andean cosmovision(Cruz, 2018), it is fundamental to understand how emotions affect personal, academic and social development. For which, it is important to classify, detect and predict such emotions in an adequate and timely manner, in order to vindicate the student's self-esteem, since, due to intercultural differences, such as the native language, they are subject to marginalization. All these are social taboos that must be discarded(Estrada et al., 2018).

The Andean cosmovision is based on the principle of good living, life in plenitude or knowing how-to live-in harmony and balance, in harmony with the cycles of Mother Earth, the cosmos, life and history, and in balance with all forms of existence. And this is precisely the path and the horizon of the community; it implies first knowing how to live and then to live together. It is not possible to Live Well if others live badly, or if Mother Nature is damaged (Gudynas, 2011). Living Well means understanding that the deterioration of one species is the deterioration of the whole. In other words, it is 'knowing how to eat, drink, dance, sleep, work, meditate, think, love, listen, speak, dream, walk, give and receive (ARCE-ROJAS, Rodrigo, 2024). Each of these knowledges is interconnected and reflects a profound relationship with nature and spirituality, promoting a balance between the individual and society', which constitutes another alternative form of social development that should be valued and taken into account in educational practices of teaching and learning in rural areas, reaching out to other traditional cultures.

Properly channelled, emotions collaborate in strengthening the Andean student's rebirth with his or her own cultural identity, sustainability and good living, and allow him or her to achieve an integral formation in the fusion of his or her living culture and technology (Gilar-Corbi et al., 2018).

In educational processes, the identification, classification and management of emotions improves group integration, motivation for social inclusion and active participation in life in solidarity, since emotional limitations due to low personal and cultural esteem restrict access to advanced technological and pedagogical resources(Nicastri et al., 2024). In this context, the use of classification algorithms for emotion recognition is presented as an innovative tool to improve educational interactions and school climate(CH & P, 2024). Artificial intelligence (AI) and machine learning (ML) technologies offer a wide set of techniques and applications to develop systems capable of performing tasks that traditionally require human intelligence(Jiang et al., 2020). Technologies such as natural language processing, computer vision and robotics have demonstrated their ability to learn and improve from large volumes of data, enabling optimal performance over time. In addition(Machová et al., 2023a), technologies have been developed that integrate emotions into the educational process, although there is little research that specifically addresses the rural context and cultural factors in education.

The aim of the article is to determine a machine learning and/or deep learning model to better recognise complex sequential text patterns of emotions expressed on social networking platforms by secondary school students in rural Puno(Das et al., 2023).

Emotions have their origin in neurochemical, physiological, biopsychological and cognitive sources, and fulfill an adaptive and associative function.(Belmer et al., 2016). Therefore, emotions and education are interconnected in the learning and teaching process (Imbir et al., 2015). Scientific research has shown that emotions affect fundamental aspects such as motivation, attention, memory and meaningful learning. According to John D. Mayer, emotions play a crucial role in the way students process information and engage in the classroom. According to John D. Mayer, emotions play a crucial role in the way students process information and engage in the classroom (Mayer et al., 2004). In addition, students' ability to regulate their own emotions has a significant influence on their academic success (Piedrahíta-Carvajal et al., 2021a).

The word 'emotion' derives from the Latin emotion, whose semantic nucleus motion is clearly related to motus (movement). It belongs to the same family as the word 'motivation' (Irrgang & Egermann, 2016), (Toivonen et al., 2012). Both terms have been developed in the field of psychology to explain the activity of the organism in relation to the ecosystem in which it operates.

According to Darwin (1873), emotions were proposed as universal adaptive mechanisms shared with other animals, not requiring cultural learning (Darwin & Darwin, 2009). To investigate this, he contacted missionaries in tribes with no exposure to Western culture, who showed photographs of English people expressing emotions, thus assessing the universality of emotional expression and recognition.

Emotion classification systems are organized sets of techniques that analyze human emotions from various perspectives (Ruiz & Delgado, 2023). They are generally divided into ‘Emotional Categories’ and ‘Emotional Dimensions’.

One of the best-known frameworks is the model of basic emotions proposed by Paul Ekman, which identifies six universal emotions: joy, sadness, anger, fear, surprise and disgust, which are recognizable through universal facial expressions (Ekman, 1992). On the other hand, Robert Plutchik's (1982) model of primary emotions suggests that, in order to properly label emotions, it is crucial to understand them within an evolutionary framework applicable to both humans and animals. Plutchik proposed eight primary emotions: fear, anger, joy, sadness, trust, disgust, surprise and anticipation (Plutchik, 1982), (Huang & Zaïane, 2019).

In the current field, emotion analysis relies on the use of advanced technologies, such as artificial intelligence (AI) (Azevedo et al., 2024), machine learning and deep learning (Janiesch et al., 2021), (Janiesch et al., 2021), to classify and analyze emotions in collaborative environments. This recognition applies not only to individual interactions, but also to group dynamics, exploring how the emotions of team members influence the overall outcome of their collaboration (Liu et al., 2020). In this context, the term ‘collaborative’ refers to an approach that considers emotions as a shared phenomenon among people working together on a project or task (Zhou et al., 2021).

In addition, social networking technologies express emotions and affect, as well as shaping the personal identity of young people, through various manifestations such as body language, facial expressions, tone of voice and textual patterns, making it possible to infer emotions in real time.
Emotion classification systems are organized sets of categories that analyze human emotions from different perspectives (Tanko et al., 2023). The models used are usually divided into ‘Emotional Categories’ and ‘Emotional Dimensions’"(Canales & Martínez-Barco, 2015).

- **Basic Emotions Model**: Paul Ekman proposed that there are six basic universal emotions: joy, sadness, anger, fear, surprise and disgust, recognizable through universal facial expressions (Ekman, 1992).
- **Primary emotions model**: Según Plutchik (1982), According to Plutchik, in order to identify and label primary emotions, it is necessary to understand them within an evolutionary framework applicable to both humans and animals. Plutchik suggested eight primary emotions: fear, anger, joy, sadness, trust, disgust, surprise and anticipation (Plutchik, 1982).
- n this context, the term ‘collaborative’ implies an approach that encompasses the emotions distributed among people working together on a project or task (Zhang et al., 2023), (Saisanthiya & Supraja, 2024). Social networking technology

expresses certain emotions and affect, as well as shaping the personal identity of young people, reflected in various manifestations such as body language, facial expressions, tone of voice and textual patterns, allowing emotions to be inferred in real time.

Machine learning algorithms are numerous and widely recognized for their efficiency and classification mechanisms (Płaza et al., 2022), (Siam et al., 2022). They include:

- **K-Nearest Neighbors (KNN):** a supervised non-parametric algorithm that uses proximity to classify data points (Xiong & Yao, 2021).
- **Naïve Bayes (NB):** a probabilistic classifier that assumes independence between features, being efficient for multi-class classification (Dart et al., 2002).
- **Random Forest (RF):** an ensemble algorithm based on multiple decision trees, known to be robust and efficient on large datasets (Talpur & O'Sullivan, 2020), (Awad & Khanna, 2015).
- **Support Vector Machine (SVM):** a supervised algorithm that searches for optimal separators (hyperplanes) to distinguish between classes (Awad & Khanna, 2015).
- **Extreme Gradient Boosting (XGBoost):** an optimized variant of the boosting algorithm, known for its high performance in classification tasks (Gono et al., 2023).
- **LightGBM:** es un método de ensamblaje de refuerzo de gradientes que se utiliza en la herramienta Entrenar con AutoML y se basa en árboles de decisión.

On the other hand, deep learning algorithms include:

- **Artificial Neural Networks (ANN):** computational systems inspired by the human brain, designed to model complex relationships between inputs and outputs (Al-Bakri & Sazid, 2021), (Haykin, 2008).

The evaluation of model performance is carried out using metrics such as accuracy, which measures the proportion of correctly classified cases out of the total number of instances; el AUC (area under the ROC curve), which evaluates the model's ability to differentiate between classes (Tanha et al., 2020), (Bradley, 1997); AUC (area under the ROC curve), which assesses the model's ability to differentiate between classes; and F1-Score, which is the harmonic average between Accuracy and recall. These metrics measure the proportion of correct classifications, the model's ability to distinguish between classes and the balance between accuracy and sensitivity, respectively.

Methodology

The procedure to be followed is illustrated in the flow chart describing the stages of the methodology used in this study. This methodology consists of six steps: Data Collection, Data Preprocessing, Data Set Partitioning, Training of Machine Learning and Deep Learning Algorithms, Model Evaluation, and Visualization and Analysis of Results. Each of these steps is detailed below in Fig 1.

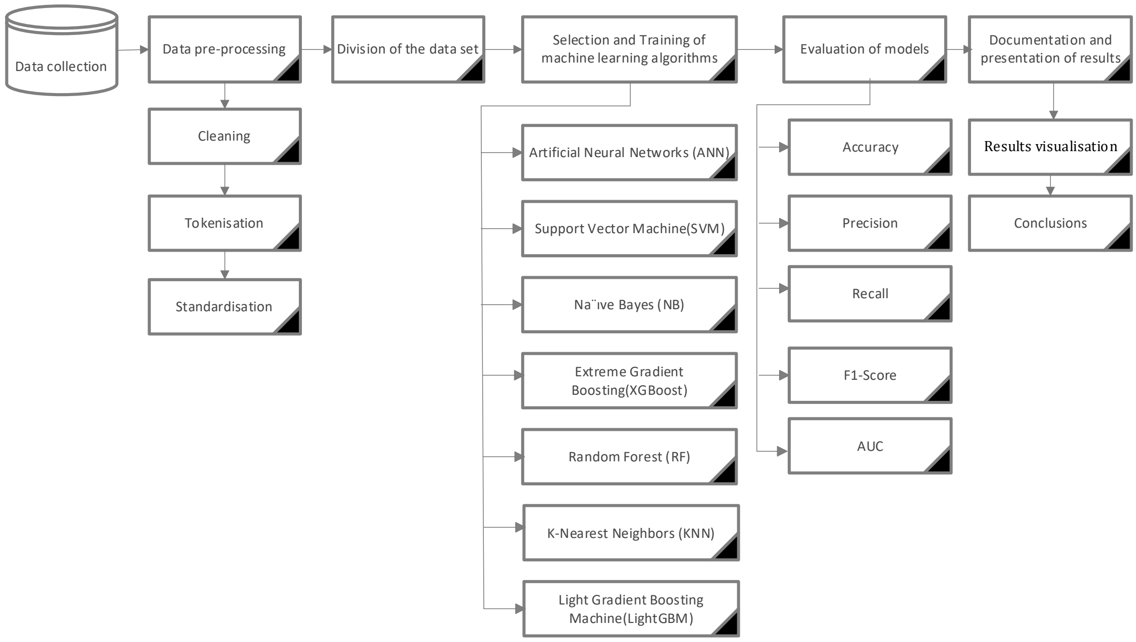

Fig 1. Methodology for Automatic Emotion Recognition

The data collection consisted of collecting the emotions expressed in the personal Facebook profile of each secondary school student from rural areas of Puno. A total of 2,160 emotions were obtained in Spanish.

Data pre-processing is a fundamental step in the development of machine learning models, especially when working with text or unstructured data. During this process, several data preparation tasks were performed in order to train and analyze machine learning models effectively. The procedure is detailed below: Carga de datos desde un archivo Excel.

- Loading data from an Excel file.
- Text cleaning, which consists of converting the text to lowercase, removing non-alphabetic characters and preparing the text for analysis.
- Emotion distribution analysis helps to examine the number of samples for each emotion category.
- TF-IDF is used for text vectorisation and to convert collaborative emotions into numerical representations.
- LabelEncoder is used to encode emotion labels and convert text classes into numerical values.

The splitting step is performed in two groups: one to train the model and one to test its performance. This process is necessary to ensure that the model fits the training data and can be generalised to new data. In this case, *train_test_split* from the *scikit-learn* library is used, whose function simplifies this process. For the study covering 2,160 emotions, *test_size=0.2* is used, which represents 20% of the test data, i.e. approximately 432 emotions used for evaluation. This amount is sufficient to obtain reliable metrics, while 80% (1,728 emotions) is intended for training, contributing to a good model fit. The corresponding Code is:

Python

```
from sklearn.model_selection import train_test_split

X_train, X_test, y_train, y_test = train_test_split

(X_tfidf, y_encoded, test_size=0.2, random_state=42, stratify=y_encoded)
```

The selection and training of machine learning and deep learning algorithms depends on many factors, such as the type of problem, the characteristics of the data and the objectives of the analysis. Different algorithms have been chosen to solve the problem of emotion classification in text, and each algorithm has particular characteristics suitable for different types of data. The selection of algorithms includes artificial neural networks (ANN), support vector machines (SVM), Naive Bayes, XGBoost, Random Forest, KNN (K nearest neighbours) and LightGBM. Each algorithm has its advantages depending on the characteristics of the dataset.

Training models involves adjusting their internal parameters to learn patterns in the data. During training, the model adapts to the input data (in this case, processed comments) and associated labels (emotions). To do this, each algorithm uses a different approach. For example, a neural network (ANN) adjusts its weights by backpropagation and gradient descent, while tree models, such as XGBoost and Random Forest, build multiple decision trees and combine them to improve accuracy and reduce overfitting.

Each algorithm requires hyperparameter settings to obtain the best performance. Hyperparameters are adjusted based on default settings or search parameters, such as the number of neighbours in KNN, the depth of the tree in Random Forest or the number of epochs in an artificial neural network (ANN). These parameters are necessary to control the trade-off between model over-fitting and generalisation.

Model evaluation is an important step in the process, as it allows determining how well a trained model generalises to unseen data. This evaluation is performed using several metrics, such as precision, recall, F1-score and AUC. Precision measures the accuracy of the model, while recall and F1-score provide a more complete picture in scenarios with unbalanced classes. The F1-score balances precision and sensitivity, which is useful when trying to reduce false positives and false negatives. In addition, the AUC (area under the curve) is used to assess the discriminative ability of the model through ROC curve analysis, especially in multi-class classification problems. An AUC close to 1 indicates that the model is efficient. These metrics help to compare the performance of different algorithms, allowing the most appropriate model to be selected and adjusted to improve its performance.

The documentation and presentation of the model results is done by means of graphs and tables, allowing visual comparison of the performance of different algorithms. Metrics such as accuracy, sensitivity, F1-score and AUC are presented in an ordered DataFrame, facilitating the interpretation of the results. In addition, Matplotlib is used to plot the ROC and AUC curves, providing a visual representation of the performance of the models. Based on the research results, the most effective model is determined based on the metric analysis. The documentation should highlight the most effective models and possible improvements that can be made to the process, such as optimising hyperparameters or exploring other algorithms. This ensures that the results are presented in a clear, understandable and useful way, allowing informed decisions to be made about which models to use in real-world applications.

Results

Analysis of collaborative emotions in adolescent secondary school students in rural Puno.

The Table 1 the different collaborative emotions expressed by the students, with a lower representation of some of them. Among these, emotions such as surprise, with 24\%, and anguish, with 23.7\%, followed by anger, which reached 22.8\%, stand out. These three emotions represent more than 70\% of the responses, which shows a tendency towards intense and reactive emotions. On the other hand, happiness, with 16.6\%, and sadness, with 13\%, reflect positive and melancholic emotions compared to the more explosive emotions. Also in Fig 2, the students' expressions reflect a variety of experiences and emotions.

Table 1. Distribution of Emotions in the Dataset

EMOTIONS	**QUANTITY**	**%**
Surprise	518	24
Anguish	511	23.7
Anger	492	22.8
Happiness	358	16.6
Sadness	281	13
Total	2160	100

```
Id                                         Comentarios emociones
 1  la identidad cultural est en relacin armnica c...   sorpresa
 2  siento emocin al ver nuestras fotos de viaje d...   Tristeza
 3  estudi con varios compaeros de clase estaba mu...      enojo
 4  a veces me pongo nervioso si nadie comenta mis...   Angustia
 5  la educacin es la socializacin e intercambio d...   sorpresa
```

Fig 2. Recording emotions in student responses

Model selection and training

The outcome of the research requires the selection of classical machine learning and deep learning algorithms, which, when applied to the dataset, allow the identification, classification and prediction of emotions. Common models include:

Classical models include SVM which allows probabilistic predictions; Random Forest based on decision trees; XGBoost optimises with the mlogloss metric for multiple categories; Naive Bayes Multinomial under efficient probabilistic approach; KNN for nearest neighbour-based classifications; and LightGBM, recognised for its speed and accuracy in gradient boosting. The deep learning model selected was an artificial neural network (ANN) with a dense layered architecture, ReLU and Softmax activations, and a 30% Dropout to prevent overfitting, optimised with the Adam algorithm. These models were chosen because of their ability to solve complex multi-class classification problems and their ability to identify emotional patterns in the analysed text.

The training of the models was performed using datasets preprocessed with TF-IDF, which converts texts into numeric vectors, and labels encoded with LabelEncoder. The data were divided into training dataset (80%) and test dataset (20%) and stratified to maintain the class distribution. Classical machine learning models, such as SVM, Random Forest, XGBoost, Naive Bayes, KNN and LightGBM, were trained using the standard fit method, while the artificial neural network (ANN) was trained for 10 epochs with an Adam optimiser and a categorical_crossentropy loss function, using a batch size of 32 and 20% cross-validation. Evaluation of all models was carried out using key metrics such as accuracy, precision, sensitivity, F1-Score and AUC shown in Table 1 to determine the best performing model for emotion classification.

Table 2 Model Evaluation Metrics

	Accuracy	**Precision**	**Recall**	**F1-Score**	**AUC**
ANN	0.828704	0.830108	0.828704	0.828864	0.971473
SVM	0.789352	0.799963	0.789352	0.78293	0.954887
Naive Bayes	0.773148	0.807327	0.773148	0.76866	0.93498
XGBoost	0.763889	0.77635	0.763889	0.763558	0.953807
Random Forest	0.761574	0.790986	0.761574	0.76012	0.958252
KNN	0.671296	0.719629	0.671296	0.67811	0.901213
LightGBM	0.631944	0.683252	0.631944	0.630449	0.891861

According to the results obtained in Table 2, the ANN showed the best performance with an accuracy of 0.8287, F1-Score of 0.8289 and an AUC of 0.9715, indicating a high class discrimination capability.Fig. 3 shows that artificial neural networks adequately handle complex data relationships. The SVM model also performed well with an AUC of 0.9549, although its F1-ScoreF1 of 0.7829 was lower than that of ANN, reflecting a poorer balance between accuracy and sensitivity. On the other hand, Naive Bayes and XGBoost show acceptable performance with an AUC above 0.93, but lower accuracy and F1-Score. Random Forest has an AUC of 0.9583, which is competitive, but the overall accuracy is lower. KNN and LightGBM had the lowest performance, 0.6713 and 0.6319 respectively, indicating that they are not suitable for this study. Overall, ANN was the most suitable emotion classification model in this study.

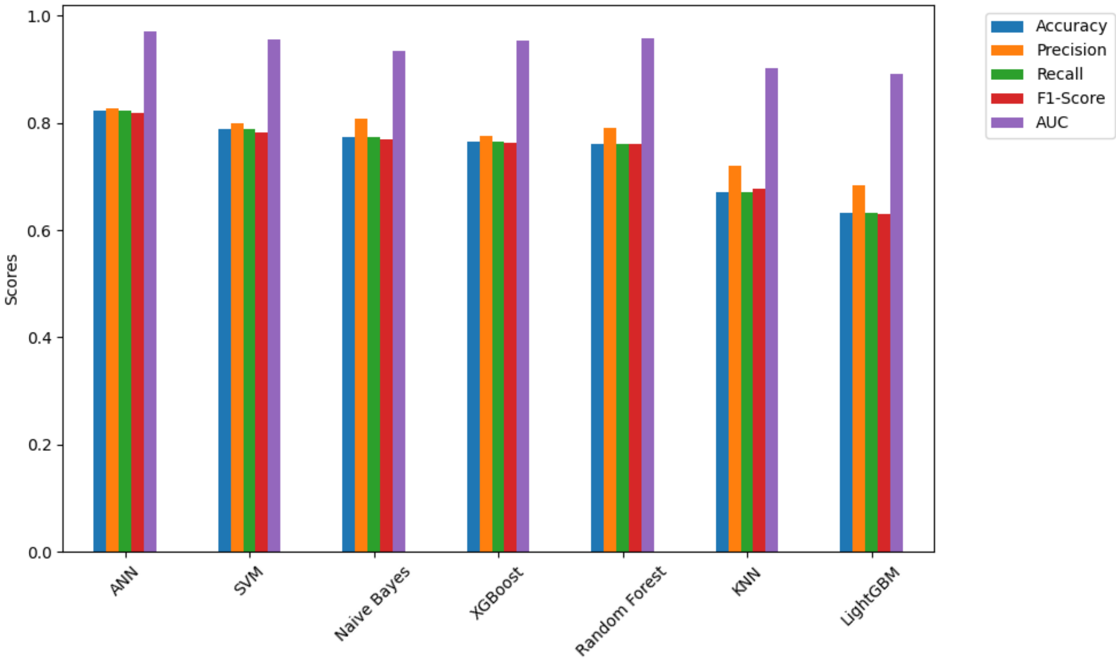

Fig 3 Evaluation of Machine Learning Models in Emotion Identification

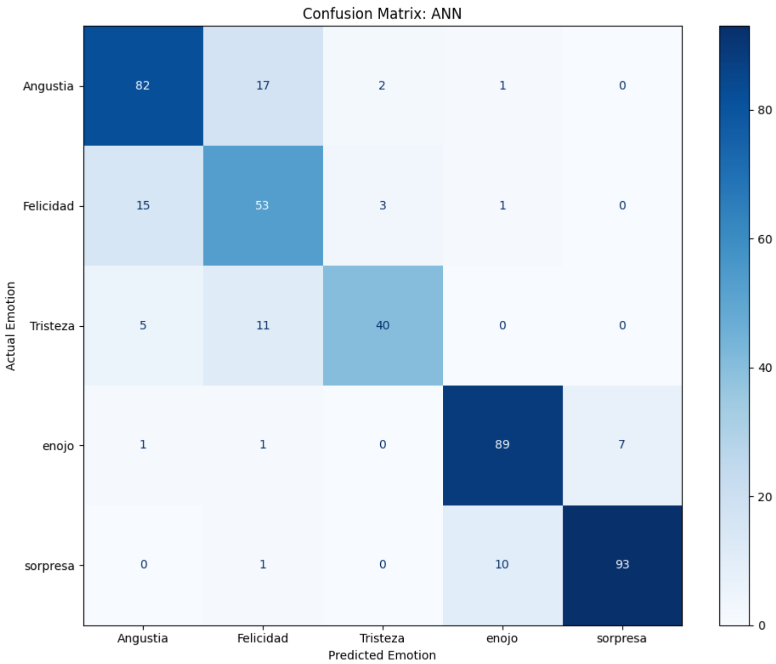

Fig 4. Multi-class confusion matrix

The Fig. 4 shows the multiclass confusion matrix, which presents the performance of the artificial neural network (ANN) model in classifying emotions such as anguish, happiness, sadness, anger and surprise in secondary school students from rural areas. It is observed that the model classifies anger emotions with high accuracy, achieving 89 correct predictions, and surprise, with 93 correct predictions. However, significant confusion is evident between anguish and happiness, where 17 cases were misclassified, and between sadness and happiness, with 11 cases of error.

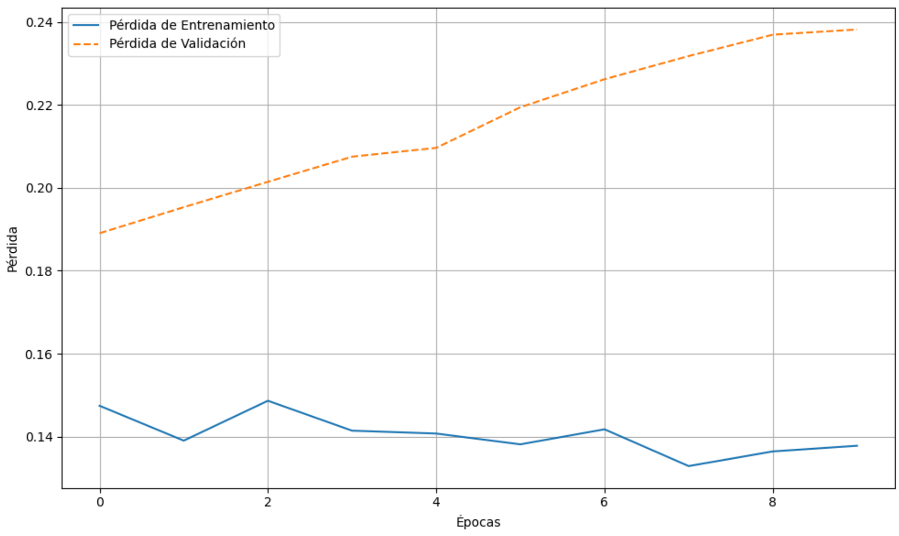

Fig 5. Loss during Training and Validation

Training and Validation Loss Fig. 5 shows the evolution of training and validation loss over 10 epochs. Initially, the training loss decreases steadily, indicating that the model is learning. However, the validation loss starts to increase after the first epoch, suggesting that the model might be overfitting the training data. Despite a good initial accuracy (93\% on the validation set in the first epoch), the validation accuracy gradually decreases as training progresses, reaching 90.46\% in the last epoch. These results indicate that the model would need to be fine-tuned.

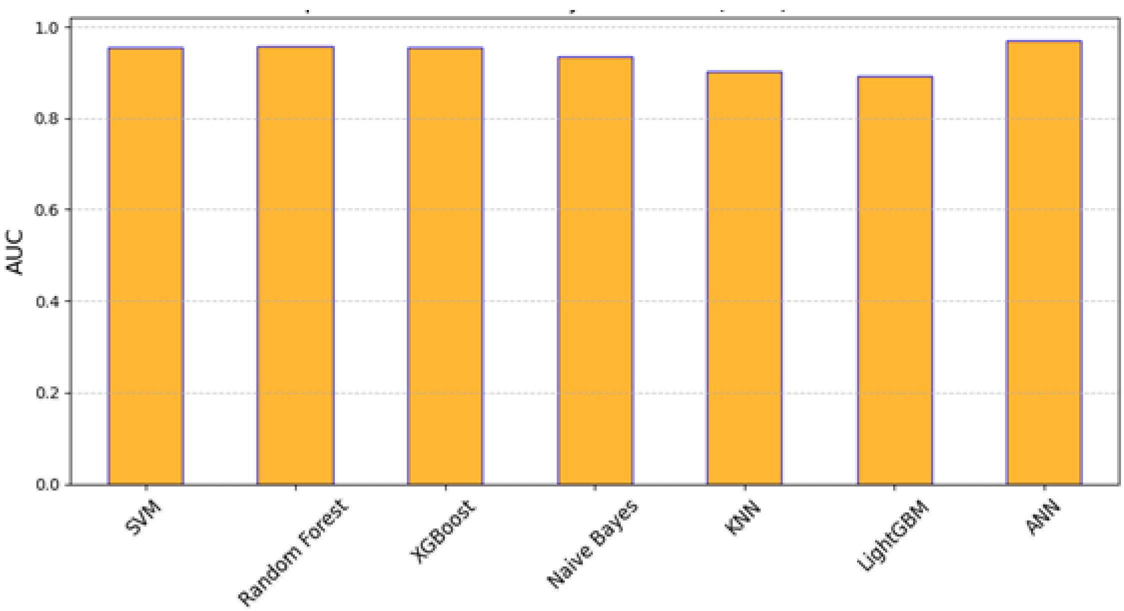

Fig 6. Comparison of models using the Area Under the Curve (AUC)

Fig. 6 compares (AUC) of different classification models. Of all the models, the ANN has a high performance, with values close to 1, being an excellent capacity to distinguish classes of emotions.

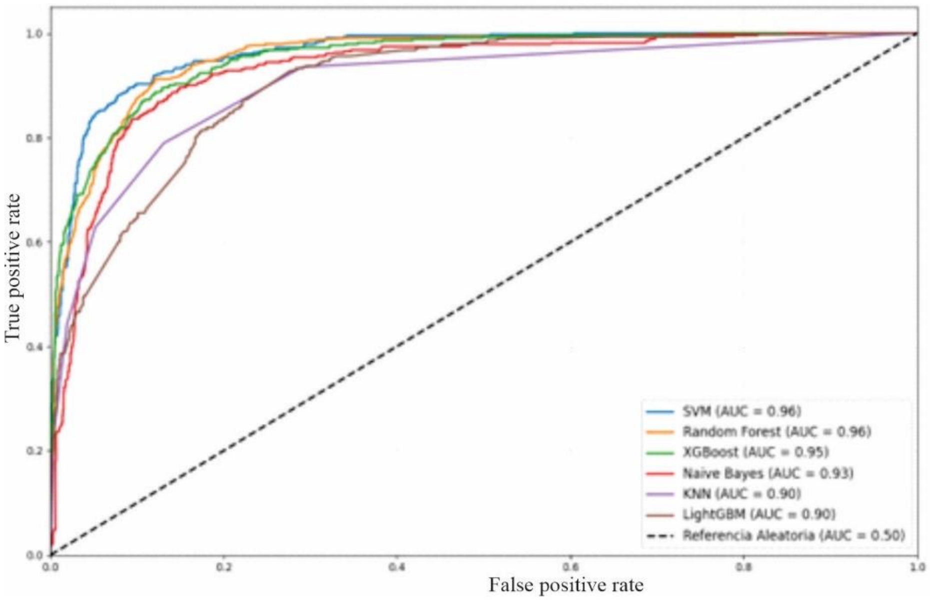

Fig 7. Comparison of Machine Learning Models Using Area Under the Curve (AUC)

La Fig. 7, shows the AUC curves for machine learning classification models such as SVM, Random Forest, XGBoost, Naive Bayes, KNN, and LightGBM, where SVM and Random Forest stand out with an AUC of 0.96, followed by XGBoost (0.95). In Fig. 8, the AUC of the artificial neural network (ANN) is estimated to be 0.97, which is higher than other models, indicating better performance in classifying positive and negative classes. This suggests that artificial neural networks may be the best option in this case..

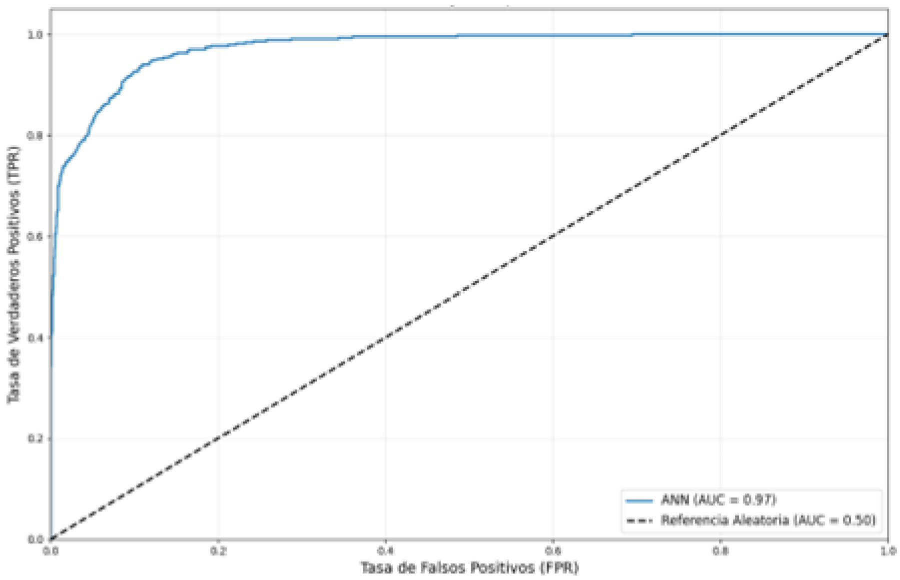

Fig 8. Area Under the Curve (AUC) using Deep Learning

Table 3. Comparison of collaborative emotions

Nro	REAL EMOTION	PREDICTED EMOTION
0	anguish	anguish
1	surprise	surprise
2	happiness	happiness
3	anger	anger
4	anguish	anguish
5	happiness	anguish
6	sadness	sadness
7	sadness	sadness
8	anguish	anguish
9	anger	anger
10	anger	anger
11	happiness	anguish
12	sadness	sadness
13	anguish	anguish
14	surprise	surprise
15	happiness	anguish
16	surprise	surprise
17	anger	anger
18	anguish	anguish
19	anger	anger

Table 3 shows the comparative analysis of the real emotions with the emotions predicted by the model, which identifies the following emotions well: "Anguish", "Sadness", "Anger" and "Surprise", because most of the predictions match the real emotion. However, this brings great difficulties to the classification of "Happiness", because in three out of four cases this emotion was confused with "Anguish". This indicates a bias towards wrongly classifying certain emotions as "Anguish", which warns of an imbalance in the data set or a limitation of the model in differentiating specific characteristics of some emotions.

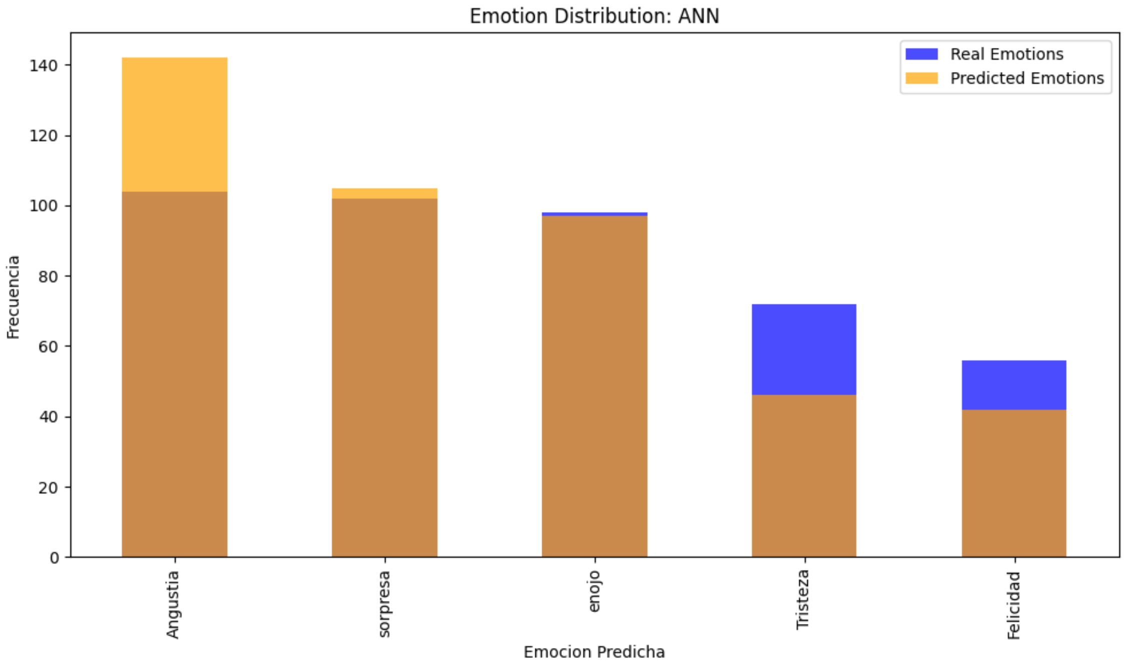

Fig 9. Actual and predicted emotions

The distribution of actual and ANN-predicted emotions in Figure 9 reflects that the model correctly classifies emotions such as “Surprise” and “Anger”, but has problems with “Happiness” and “Sadness”, which are often confused with other emotions, especially “Distress”, which is over-predicted. This indicates an imbalance in the model or the data that affects its performance.

Discussion

The findings on the predominant emotions among rural secondary school students in Puno, surprise (24\), anguish (23.7%) and anger (22.8%) are those that stand out as patterns of intense and reactive emotional responses observed in previous research. According to Ekman (1992) and his theories on basic emotions, it is natural to find that emotions such as anguish and anger predominate in adolescents, who are going through a critical period of emotional development. These classification models, based on ANN, have proven effective in identifying complex emotional patterns, as documented in previous studies (K & S, 2022).

The comparative Table 2 of model results positions the deep learning technique of Artificial Neural Networks ANN as the best performing model to classify emotions in texts of secondary school students from rural Puno, according to metrics such as Accuracy 82.87\% and AUC 97.15%, outperforming models such as SVM and Naive Bayes. The result is supported by, (Huang & Zaïane, 2019) to handle complex patterns in text data, even in culturally specific contexts.

.

The prevalence of negative emotions found such as anguish, surprise and anger is reinforced by Cáceres et al. (2020), who highlight the influence of socioeconomic and academic factors on the emotional well-being of students.

The results of negative emotions such as anger and surprise are natural manifestations of students from rural areas, coming from a multicultural society, where the educational model is not satisfactory to the students, while happiness, sadness and anguish are inconsistent.

The analysis of emotions and the monitoring of the level of attention of students in virtual environments allows teachers to take actions to improve the teaching-learning processes(Piedrahíta-Carvajal et al., 2021), so it is necessary to carry out more studies on the subject of emotion analysis in secondary education in rural areas.

Emotions are an integral part of human life, which directly influences the success of the student's teaching and learning processes. Being a complex pattern of moods, a conscious mental reaction expressed in different ways.(Machová et al., 2023b), the classification of which is essential for a variety of applications, including sentiment analysis, student feedback analysis, and mental health monitoring(Dayananda et al., 2023).

By analyzing the performance of the ANN model, which showed an accuracy of 0.8287 and an AUC of 0.9715, the model's ability to handle complex relationships in large volumes of data is confirmed (Lecun et al., 2015). The consistency between these performance metrics and contemporary machine learning theories suggests that the selection of appropriate algorithms is crucial for emotion analysis. a study (He et al., 2016), the difficulty in differentiating between similar emotions, such as confusion between distress and happiness, highlights the need to improve data features and explore more complex network architectures, which can capture emotional subtleties.

Además, los resultados del estudio complementan la literatura más reciente sobre el análisis emocional en entornos educativos. Por ejemplo, un estudio realizado por (Chowanda et al., 2022) utilizó algoritmos de aprendizaje profundo para identificar emociones en jóvenes, encontrando resultados similares en cuanto a la confusión entre emociones como felicidad y angustia. Esto coincide con los hallazgos presentados aquí, donde se reportó una sobrepredicción de la angustia en comparación con la felicidad, lo que sugiere un fenómeno recurrente en la clasificación emocional.

Furthermore, the results of the study complement the most recent literature on emotional analysis in educational settings. For example, a study (Chowanda et al., 2022) used deep learning algorithms to identify emotions in young people, finding similar results regarding confusion between emotions such as happiness and distress. This is in line with the findings presented here, where an overprediction of distress compared to happiness was reported, suggesting a recurring phenomenon in emotional classification.

Another study (Chutia & Baruah, 2024) highlights the importance of data quality in emotion prediction, arguing that an imbalanced dataset can significantly decrease the accuracy of models. This finding is pertinent, given the imbalance observed in the emotions classified in this study, particularly in the struggle to differentiate happiness from distress. This emphasizes that, although the ANN proved to be the most suitable model for this analysis, greater attention is required in dataset preparation to avoid bias in emotion classification, an aspect also mentioned in recent studies on machine learning and its applicability in educational contexts.

Conclusions

The findings of this study reinforce the relevance of using deep learning models in understanding emotions in students, while underlining the need to improve datasets and preprocessing techniques to achieve a more accurate and effective analysis of emotions in educational contexts.

Finding ANN as the best performing model with AUC of 97.1% to correctly classify emotions such as "Surprise" and "Anger", but with difficulties regarding the emotions "Happiness" and "Sadness", which are often confused with "Anguish".

The emotion surprise indicates mixed feelings of an educational model imposed by the Peruvian Ministry of Education, which is not up to the expectations of a student population, which still lives with its living culture, in a scenario of life under principles of the Andean worldview, which still guarantees survival in situations of poverty, culturally postponed in its development.

The emotion anger means that the rural adolescent student population has permanent frustrations due to marginalization in basic sanitation services, it motivates them to migrate to coastal cities, being in these cities, they encounter racial abuse, physical or psychological harassment, emotionally they lose personal self-esteem without conditions to integrate into a full society with quality of life.

Acknowledgments

I thank God for granting us the strength to overcome adversity and for guiding our path at every step. I also express my sincere gratitude to the National University of the Altiplano for its contribution to the development of this article. Furthermore, I extend my deep gratitude to all the individuals who participated in this research, whose contributions were essential for its development.

References

Al-Bakri, A. Y., & Sazid, M. (2021). Application of Artificial Neural Network (ANN) for Prediction and Optimization of Blast-Induced Impacts. *Mining*, *1*(3), 315–334. https://doi.org/10.3390/MINING1030020

Alcocer-Sánchez, D. J., Palmero Castillo, A., Muñoz, D., & Canto Herrera, P. J. (2023). Digital Competencies and Emotions in University Students in the Dominican Republic Dominicana. *Publicaciones*, *53*(1). https://doi.org/10.30827/publicaciones.v53i1.27986

ARCE-ROJAS Rodrigo. (2024). El buen vivir como alternativa al desarrollo y como respuesta a la megacrisis. *Scientia*, *26*(26). https://doi.org/10.31381/SCIENTIA.V26I26.6943

Awad, M., & Khanna, R. (2015). Efficient learning machines: Theories, concepts, and applications for engineers and system designers. *Efficient Learning Machines: Theories, Concepts, and Applications for Engineers and System Designers*, 1–248. https://doi.org/10.1007/978-1-4302-5990-9/COVER

Azevedo, N., Kehayia, E., Jarema, G., Le Dorze, G., Beaujard, C., & Yvon, M. (2024). How artificial intelligence (AI) is used in aphasia rehabilitation: A scoping review. *Aphasiology*, *38*(2), 305–336. https://doi.org/10.1080/02687038.2023.2189513

Belmer, A., Lanoue, V., Patkar, O. L., & Bartlett, S. E. (2016). Excitatory/inhibitory balance of serotonergic axon connectivity in the brain. *J Neurol Neuromedicine*, *1*(9), 18–22. www.jneurology.com

Bermejo-Paredes, S., Maquera-Maquera, Y. A., Bermejo-Paredes, S., & Maquera-Maquera, Y. A. (2019). Interpretación de la escuela rural andina encomunidades aimaras de Puno-Perú. *Revista Electrónica Educare*, *23*(2), 66–80. https://doi.org/10.15359/REE.23-2.4

Bradley, A. P. (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. *Pattern Recognition*, *30*(7), 1145–1159. https://doi.org/10.1016/S0031-3203(96)00142-2

Cáceres, M. L. M., García, R. C., & García, O. R. (2020). El manejo de la inteligencia emocional en los estudiantes de secundaria. Un estudio exploratorio en una telesecundaria en México. *Conrado*, *16*(74).

Canales, L., & Martínez-Barco, P. (2015). *Emotion Detection from text: A Survey*. 37–43. https://doi.org/10.3115/V1/W14-6905

CH, S., & P, V. (2024). Ameliorate grasshopper optimization algorithm based long short term memory classification for face emotion recognition system. *Multimedia Tools and Applications*, *83*(13), 37961–37978. https://doi.org/10.1007/S11042-023-16837-1

Chowanda, A., Iswanto, I. A., & Andangsari, E. W. (2022). Exploring deep learning algorithm to model emotions recognition from speech. *Procedia Computer Science*, *216*, 706–713. https://doi.org/10.1016/J.PROCS.2022.12.187/EXPLORING_DEEP_LEARNING_ALGORITHM_TO_MODEL_EMOTIONS_RECOGNITION_FROM_SPEECH.PDF

Chutia, T., & Baruah, N. (2024). A review on emotion detection by using deep learning techniques. *Artificial Intelligence Review 2024 57:8*, *57*(8), 1–80. https://doi.org/10.1007/S10462-024-10831-1

Cruz, M. A. (2018). *Andean Cosmovision and Interculturality: A look at sustainable development from the sumak kawsay*. *5*(2550–6722), 119–132. https://www.mendeley.com/catalogue/cfb584b3-1800-3dc4-848e-eff637e8cecb/

Dart, A. M., Du, X. J., & Kingwell, B. A. (2002). Gender, sex hormones and autonomic nervous control of the cardiovascular system. *Cardiovascular Research*, *53*(3), 678–687. https://doi.org/10.1016/S0008-6363(01)00508-9

Darwin, C., & Darwin, F. (2009). The expression of the emotions in man and animals. *The Expression of the Emotions in Man and Animals*, 1–401. https://doi.org/10.1017/CBO9780511694110

Das, R. K., Islam, M., Hasan, M. M., Razia, S., Hassan, M., & Khushbu, S. A. (2023). Sentiment analysis in multilingual context: Comparative analysis of machine learning and hybrid deep learning models. *Heliyon*, *9*(9). https://doi.org/10.1016/J.HELIYON.2023.E20281/SENTIMENT_ANALYSIS_IN_MULTILINGUAL_CONTEXT_COMPARATIVE_ANALYSIS_OF_MACHINE_LEARNING_AND_HYBRID_DEEP_LEARNING_MODELS.PDF

Dayananda, C., Hemashree, P., Devamane, S. B., Aravind, E., Impu, D., & Janavi, S. (2023). Text Emotion Detection Using Machine Learning Algorithms. *Proceedings - 2023 International Conference on Computational Intelligence for Information, Security and Communication Applications, CIISCA 2023*, 304–307. https://doi.org/10.1109/CIISCA59740.2023.00065

Ekman, P. (1992). Facial expressions of emotion: an old controversy and new findings. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *335*(1273), 63–69. https://doi.org/10.1098/RSTB.1992.0008

Estrada, M. L. B., Cabada, R. Z., Ávila, S. L. R., Bustillos, R. O., & Guerrero, M. G. (2018). *Use of Emotion Analyzer in Intelligent Educational Systems*. *147*(6), 179–188. https://www.mendeley.com/catalogue/37a0726f-b602-31b2-8b38-4f7f8b9dfda4/

Fidel, H.-R., & Alejandro, A.-T. (2024). Covid 19 identification model using Deep Learning techniques from thorax of lung X-ray images | Modelo de identificación de Covid 19 usando técnicas de Deep Learning a partir de imágenes de Rayos X de torax de pulmones. *Proceedings of the LACCEI International Multi-Conference for Engineering, Education and Technology*. https://doi.org/10.18687/LACCEI2024.1.1.1491

Gilar-Corbi, R., Pozo-Rico, T., Pertegal-Felices, M. L., & Sanchez, B. (2018). Emotional intelligence training intervention among trainee teachers: a quasi-experimental study. *Psicologia: Reflexao e Critica*, *31*(1). https://doi.org/10.1186/S41155-018-0112-1

Gono, D. N., Napitupulu, H., & Firdaniza. (2023). Silver Price Forecasting Using Extreme Gradient Boosting (XGBoost) Method. *Mathematics*, *11*(18). https://doi.org/10.3390/MATH11183813

Gudynas, E. (2011). *Buen vivir: Germinando alternativas al desarrollo*. *462*(Febrero), 1–20. https://www.mendeley.com/catalogue/f981ff05-44a8-3fd3-b591-6d59f754f5e1/

Haykin, S. (2008). Neural Networks and Learning Machines. *Pearson Prentice Hall*, *936*, undefined-undefined. https://doi.org/978-0131471399

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. *Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, *2016-December*, 770–778. https://doi.org/10.1109/CVPR.2016.90

Huang, C., & Zaïane, O. R. (2019). Generating responses expressing emotion in an open-domain dialogue system. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *11551 LNCS*. https://doi.org/10.1007/978-3-030-17705-8_9

Imbir, K. K., Jarymowicz, M. T., Spustek, T., Kuš, R., & Zygierewicz, J. (2015). Origin of Emotion Effects on ERP Correlates of Emotional Word Processing: The Emotion Duality Approach. *PLOS ONE*, *10*(5), e0126129. https://doi.org/10.1371/JOURNAL.PONE.0126129

Irrgang, M., & Egermann, H. (2016). From Motion to Emotion: Accelerometer Data Predict Subjective Experience of Music. *PLOS ONE*, *11*(7), e0154360. https://doi.org/10.1371/JOURNAL.PONE.0154360

Janiesch, C., Zschech, P., & Heinrich, K. (2021). Machine learning and deep learning. *Electronic Markets*, *31*(3), 685–695. https://doi.org/10.1007/S12525-021-00475-2

Jiang, D., Wu, K., Chen, D., Tu, G., Zhou, T., Garg, A., & Gao, L. (2020). A probability and integrated learning based classification algorithm for high-level human emotion recognition problems. *Measurement: Journal of the International Measurement Confederation*, *150*. https://doi.org/10.1016/J.MEASUREMENT.2019.107049

K, S., & S, J. (2022). An Efficient AP-ANN-Based Multimethod Fusion Model to Detect Stress through EEG Signal Analysis. *Computational Intelligence and Neuroscience*, *2022*, 1–18. https://doi.org/10.1155/2022/7672297

Lecun, Y., Bengio, Y., & Hinton, G. (2015). Deep learning. *Nature*, *521*(7553), 436–444. https://doi.org/10.1038/NATURE14539

Liu, J. M., Su, Y. Q., Wei, P., & Liu, Y. H. (2020). Video-EEG Based Collaborative Emotion Recognition Using LSTM and Information-Attention. *Zidonghua Xuebao/Acta Automatica Sinica*, *46*(10), 2137–2147. https://doi.org/10.16383/J.AAS.C180107

Machová, K., Szabóova, M., Paralič, J., & Mičko, J. (2023a). Detection of emotion by text analysis using machine learning. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/FPSYG.2023.1190326/PDF

Machová, K., Szabóova, M., Paralič, J., & Mičko, J. (2023b). Detection of emotion by text analysis using machine learning. *Frontiers in Psychology*, *14*. https://doi.org/10.3389/fpsyg.2023.1190326

Mayer, J. D., Salovey, P., & Caruso, D. R. (2004). Emotional intelligence: Theory, findings, and implications. *Psychological Inquiry*, *15*(3), 197–215. https://doi.org/10.1207/S15327965PLI1503_02

Nicastri, M., Dinçer D'Alessandro, H., Giallini, I., D'Amico, A., Geraci, A., Inguscio, B. M. S., Guerzoni, L., Cuda, D., Vestri, A., Fegatelli, D. A., & Mancini, P. (2024). Emotional abilities in preadolescents and adolescents with long-term cochlear implant use. *International Journal of Pediatric Otorhinolaryngology*, *177*. https://doi.org/10.1016/J.IJPORL.2024.111866

Piedrahíta-Carvajal, A., Rodríguez-Marín, P. A., Terraza-Arciniegas, D. F., Amaya-Gómez, M., Duque-Muñoz, L., & Martínez-Vargas, J. D. (2021a). Aplicación web para el análisis de emociones y atención de estudiantes. *TecnoLógicas*, *24*(51), undefined-undefined. https://doi.org/10.22430/22565337.1821

Piedrahíta-Carvajal, A., Rodríguez-Marín, P. A., Terraza-Arciniegas, D. F., Amaya-Gómez, M., Duque-Muñoz, L., & Martínez-Vargas, J. D. (2021b). Aplicación web para el análisis de emociones y atención de estudiantes. *TecnoLógicas*, *24*(51). https://doi.org/10.22430/22565337.1821

Płaza, M., Trusz, S., Kęczkowska, J., Boksa, E., Sadowski, S., & Koruba, Z. (2022). Machine Learning Algorithms for Detection and Classifications of Emotions in Contact Center Applications. *Sensors*, *22*(14). https://doi.org/10.3390/S22145311

Plutchik, R. (1982). A psychoevolutionary theory of emotions. *Social Science Information*, *21*(4–5). https://doi.org/10.1177/053901882021004003

Ruiz, M. C., & Delgado, R. P. (2023). Dependencia emocional y su relación con el riesgo suicida en adultos jóvenes. *LATAM Revista Latinoamericana de Ciencias Sociales y Humanidades*, *4*(1), 329–339. https://doi.org/10.56712/LATAM.V4I1.247

Saisanthiya, D., & Supraja, P. (2024). Neuro-facial Fusion for Emotion AI: Improved Federated Learning GAN for Collaborative Multimodal Emotion Recognition. *IEIE Transactions on Smart Processing and Computing*, *13*(1), 61–68. https://doi.org/10.5573/IEIESPC.2024.13.1.61

Siam, A. I., Soliman, N. F., Algarni, A. D., Abd El-Samie, F. E., & Sedik, A. (2022). Deploying Machine Learning Techniques for Human Emotion Detection. *Computational Intelligence and Neuroscience*, *2022*. https://doi.org/10.1155/2022/8032673

Talpur, B. A., & O'Sullivan, D. (2020). Cyberbullying severity detection: A machine learning approach. *PLoS ONE*, *15*(10 October). https://doi.org/10.1371/JOURNAL.PONE.0240924

Tanha, J., Abdi, Y., Samadi, N., Razzaghi, N., & Asadpour, M. (2020). Boosting methods for multi-class imbalanced data classification: an experimental review. *Journal of Big Data*, *7*(1). https://doi.org/10.1186/S40537-020-00349-Y

Tanko, D., Demir, F. B., Dogan, S., Sahin, S. E., & Tuncer, T. (2023). Automated speech emotion polarization for a distance education system based on orbital local binary pattern and an appropriate sub-band selection technique. *Multimedia Tools and Applications*, *82*(26), 40839–40856. https://doi.org/10.1007/S11042-023-14648-Y

Toivonen, R., Kivelä, M., Saramäki, J., Viinikainen, M., Vanhatalo, M., & Sams, M. (2012). Networks of Emotion Concepts. *PLOS ONE*, *7*(1), e28883. https://doi.org/10.1371/JOURNAL.PONE.0028883

Xiong, L., & Yao, Y. (2021). Study on an adaptive thermal comfort model with K-nearest-neighbors (KNN) algorithm. *Building and Environment*, *202*. https://doi.org/10.1016/J.BUILDENV.2021.108026

Zhang, P., Fu, M., Zhao, R., Wu, D., Zhang, H., Yang, Z., & Wang, R. (2023). ECMER: Edge-Cloud Collaborative Personalized Multimodal Emotion Recognition Framework in the Internet of Vehicles. *IEEE Network*, *37*(4), 192–199. https://doi.org/10.1109/MNET.003.2300012

Zhou, J. J., Phadnis, V., & Olechowski, A. (2021). Analysis of Designer Emotions in Collaborative and Traditional Computer-Aided Design. *Journal of Mechanical Design*, *143*(2). https://doi.org/10.1115/1.4047685