www.editada.org

International Journal of Combinatorial Optimization Problems and Informatics, 15(5), Dec 2024, 108-114. ISSN: 2007-1558.
https://doi.org/10.61467/2007.1558.2024.v15i5.564

Emotion detection using natural language processing

Antonio Álvarez Núñez, María del Carmen Santiago Díaz, Ana Claudia Zenteno Vázquez, Judith Pérez Marcial, Gustavo Trinidad Rubín Linares

Benemérita Universidad Autónoma de Puebla, Facultad de Ciencias de la Computación, 14 sur y Av. San Claudio, Col. Jardines de San Manuel. C.P. 72570. Puebla, Puebla, México.
antonio.alvarez@alumno.buap.mx, {marycarmen.santiago, ana.zenteno, judith.perez, gustavo.rubin}@correo.buap.mx.

Abstract. The analysis of human emotions is of great interest in data analysis, as it allows us to identify patterns and behaviors in people. Different techniques are used, such as linguistic rule-based approaches. Natural language processing (NLP) is a branch of AI that seeks to make machines understand language like humans do. It combines computational linguistics with machine learning and deep learning models. In emotional prosody, spoken words convey linguistic and paralinguistic information, where the emotional context influences the interpretation of the words. Alexithymia refers to difficulty identifying and expressing emotions. AI and NLP offer powerful tools for their study and application, which is why it was possible to develop an AI for the detection of emotions through natural language, resulting in a system that offers sentiment analysis for patients.

Keywords: Data Analysis, Emotion Recognition, Artificial Neural Networks, Natural Language Processing, Emotional Prosody, Alexithymia

Article Info
Received May 10, 2024.
Accepted Nov 20, 2024.

1 Introduction

Human emotions are an element of high interest in the world of data analysis. This is because from them we can identify behaviors and patterns in human beings and identify factors that can even affect their survival.

Emotion recognition has become a powerful tool to extract or identify valuable information from different human expressions [1], and it is also a field of study that focuses on identifying and understanding human emotions expressed in natural language, whether written or spoken. Its main objective is to analyze and classify the emotional content of a text or speech to infer the emotional state of the patient. To carry out emotion recognition, different techniques and approaches are used. First, rule-based approaches can be used, which involve the development of specific linguistic rules to associate words or phrases with particular emotions. However, this approach can be limited due to the complexity and subjectivity of language [2].

Both the recognition of emotions and many other applications have been enhanced with the use of Artificial Intelligence (AI), which is not something new that has just arrived in our lives, but its origins date back to the 1950s, when Alan Turing published an article called "Computing Machinery and Intelligence" [3], in which he questions whether a machine could think and learn to be able to interact with humans, for which he proposed the famous Turing test with which one can determine if a machine is intelligent or not. The test consists of 3 elements, two of which are people and the third is a computer designed to simulate responses as similar as possible to those of a human being, one of the people assumes the role of evaluator or interrogator and must maintain a conversation with the other two elements of the experiment trying to determine from the answers obtained, which of the two is the computer and which is the human being, in case the evaluator cannot distinguish them it is considered that the machine has passed the test. There are various AI techniques such as artificial neural networks (ANN), which over the years have received particular interest in technological applications, as they offer the means to effectively model complex problems [4]. These networks are capable of effectively finding patterns through a series of algorithms based on existing data. The main unit of the ANN is a processor called a neuron, which calculates the inputs and is activated to send a signal to the next neuron and so on until reaching an output, they are based on the behavior and functioning of the human brain, especially the nervous system.

Natural Language Processing (NLP) was born in the 1950s as an intersection between artificial intelligence and linguistics. It is a branch of AI that seeks to provide machines with the ability to understand written and spoken language, in the same way that humans do. NLP combines computational linguistics with machine learning and deep learning models [3].

In order to carry out NLP, it is necessary to consider several characteristics of language itself, which will be described below:

Emotional prosody: Spoken words contain both linguistic and paralinguistic elements. Linguistic information consists of the literal symbolic meaning of the word, while paralinguistic information consists of the context of the word. For example, the meaning of the word “crazy,” whether it is pronounced in the sense of “mental disorder,” “furious,” or “very excited,” can be disambiguated based on the evaluation of contextual paralinguistic information, such as the speaker’s current emotional state, as revealed by his or her tone of voice and facial expression [5].

Alexithymia: The concept of alexithymia was introduced by Sifneos to describe a group of symptoms observed in patients with psychosomatic illnesses. According to Sifneos, this term literally means absence of words to express emotions and denotes a difficulty in identifying and describing emotions, as well as an impoverished internal fantasy life [6].

Regarding speech emotion recognition, in recent years algorithms such as Support Vector Machines (SVM), Markov hidden layer models, mixed Gaussian models; and Deep learning [7] have been proposed and tested. In this last case, it has been shown that using Bidirectional Long Short-Term Memory (BiLTSM) is the best option to extract speech features seeking better results in emotion recognition [8], or it can be addressed on the Deep learning technologies used in Text Emotion Recognition (TER) where in general the “Word embedding” technology is introduced for the representation of language and then some neural network architectures are added [3].

The use of pre-trained Self Supervised Learning (SSL) architectures, similar to BERT, has also been proposed to represent both speech and text for emotion recognition tasks, where pre-trained models available in the Fair Seq tool [9] are used.

In the field of speech recognition, there are two options: developing the speech recognition model, or using software already specialized in the task through Application Programming Interfaces (API). Because speech recognition is only a means to convert speech to text, and is not the objective of this work, the existing API options to perform the task will be analyzed.

2 Methodology

The methodology used to perform sentiment analysis using artificial intelligence with an algorithm that can help in the detection and treatment of Alexithymia or with patients with psychosomatic illnesses, consists of a process in which the polarity and subjectivity of a given text is evaluated to determine whether its content is positive, negative or neutral.

Input text: Information provided to the program so that it can perform its task and process the data according to its logic and functionality [10].

Preprocessing: Before performing the analysis, the set of data provided or in any case the entire text, paragraph or document, is subjected to a cleaning preprocessing to make corrections such as: spelling mistakes, removing words that have no value to generate a more effective or precise result in the detection of feelings and emotions [11].

Tokenization: Tokenization is the process of dividing a text into smaller units, such as sentences or individual words, for which a tokenizer is used that is based on grammatical rules and regular expressions to perform this task [5], as can be seen in figure 1.

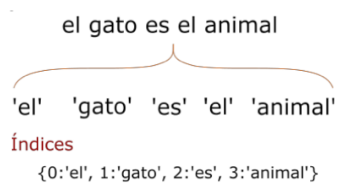

Figure 1. Tokenization of a sentence.

Grammatical analysis: The "POS TAGGING" algorithm learns grammatical patterns from large annotated data sets. It can identify features such as the use of suffixes, prefixes, endings, verb conjugations and inflections to determine the grammatical category of a word, all this to have an accurate and real result of the words provided [12], for example, in the sentence "My name is Fran and I live in Madrid", when performing the grammatical analysis, it understands a determiner, a noun, an auxiliary, and so on with each word that is assigned its value, as can be seen in figure 2.

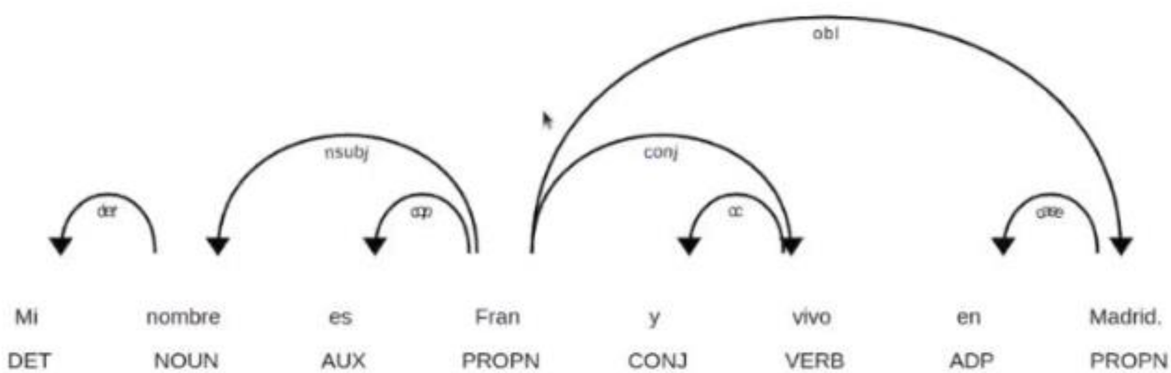

Figure 2. Grammatical analysis.

Feature extraction: Algorithms must be able to resolve this ambiguity and assign the most appropriate grammatical label to a word in a specific context. POS TAGGING algorithms may face challenges when dealing with words that are not present in their training dataset or misspelled words. They may employ strategies such as using statistical models or heuristic rules to infer the grammatical category of these words [12], in which the process is performed to determine unimportant words to remove them, since they are in a category that does not have the value to have a higher importance, as shown in Figure 3.

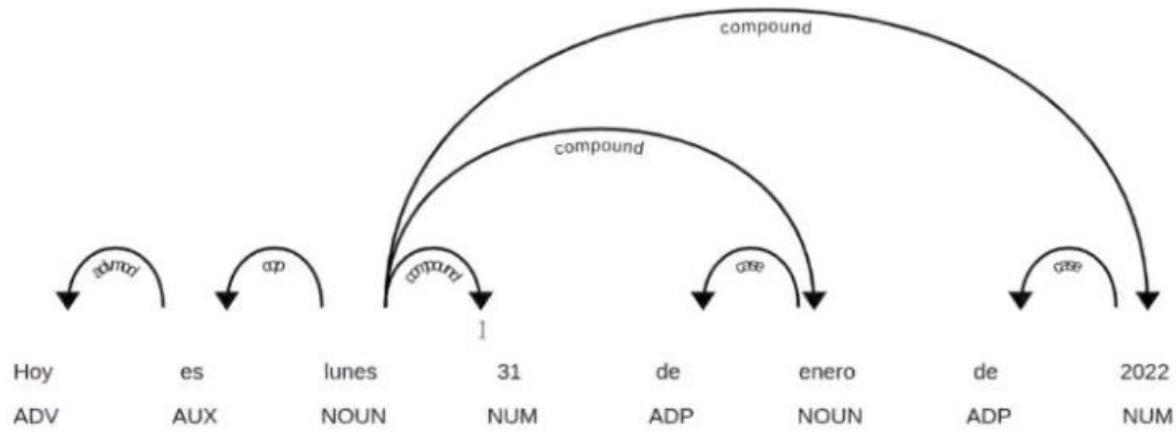

Figure 3. Extraction of features from a sentence.

Sentiment analysis: The sentiment of a sentence or a set of words can be analyzed. It uses a lexical polarity-based approach, where each word is assigned a polarity score (positive, negative, or neutral). These scores are combined to calculate the overall polarity of the text and the associated subjectivity [12].

Subjectivity analysis: Subjectivity is a measure that indicates the extent to which the text is objective or subjective. Subjectivity is expressed as a value between 0 and 1, where 0 represents high objectivity, meaning the text is more informative and factual, while 1 indicates high subjectivity, meaning the text is more opinionated and subjective. When the subjectivity value is close to 0, it implies that the text contains mostly objective and descriptive information, such as data, facts, or unbiased descriptions. In contrast, a value close to 1 suggests that the text is full of opinions, judgments, or subjective evaluations. We should consider that subjectivity is not necessarily related to the positive or negative sentiment of the text. A text may have high subjectivity without expressing a specific sentiment, while another text with low subjectivity may contain a strong positive or negative sentiment [13], as shown in Figure 4.

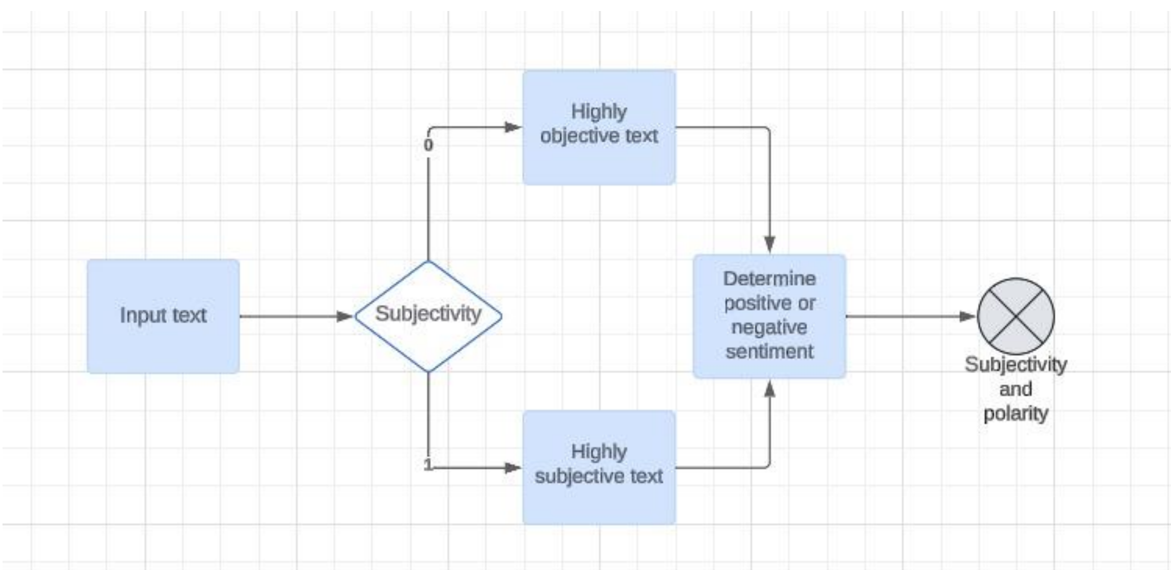

Figure 4. Flowchart for understanding subjectivity.

3 Results

The results obtained from the sentiment analysis using natural language are shown in the graphs generated for polarity, subjectivity and sentiment, and a summary is also shown on the screen, as well as the downloadable document with the information provided on the results screen, all of these elements concentrated on a web page where the system information and its start are presented, as shown in figure 5.

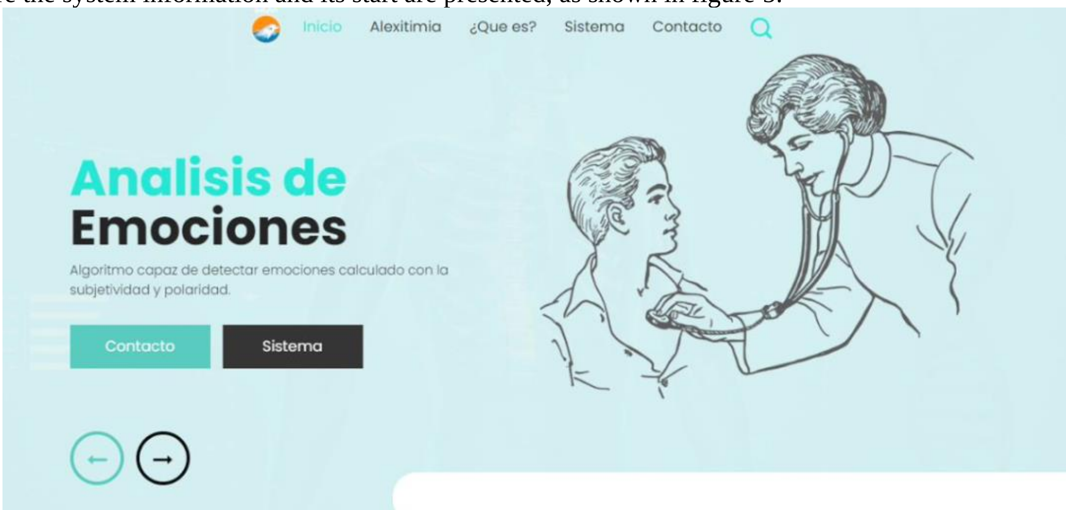

Figure 5. Main interface.

In the section called “System” the patient will be able to interact with the artificial intelligence through a user-friendly interface where the user can write the text he wants to analyze and, in this way, with the help of specialists, he can obtain a more precise and effective treatment, as shown in figure 6.

Figure 6. System with text.

In the first section of our results we have the sentiment analysis, where the sentiment that the provided text has is shown; the polarity is also shown along with the subjectivity, where negative numbers are observed, indicating that the system is doing its corresponding job correctly, as shown in figure 7.

Figure 7. Results obtained.

The system also provides an automatic summary of the text that the patient and user provided to the system for its respective analysis through artificial intelligence, where it additionally gives us a classification of the emotions to know what score we have in the given text, as shown below in figure 8.

Figure 8. Summary of emotion classification.

4 Conclusions

The implemented system offers a tool in the medical area for the treatment of emotion disorders. By analyzing sentiments and detecting emotions from text provided by patients, the system can help health professionals gain valuable insights into the emotional states and feelings expressed by individuals by using natural language processing and sentiment analysis techniques. The results obtained through the generated visualizations can provide a clear view of the polarity and subjectivity of the emotions experienced. Furthermore, while the system can be a valuable tool for the treatment of emotion disorders, it is important to note that it should not replace the clinical assessment and judgment of health professionals. Therefore, artificial intelligence can be a useful complement, but the main approach should be therapeutic, personalized, and holistic to address the emotional challenges of each patient. The degree of confidence with which the system predicts can be relatively high for general sentiment and subjectivity tasks, but its accuracy can vary depending on the model and the complexity of the analyzed text. Furthermore, the integration of high-resolution sensors and advanced cameras into mobile devices and surveillance systems has expanded the applications of emotion detection in mental health monitoring, human-computer interaction, and affective artificial intelligence research. Moreover, ethics and privacy in the collection and use of emotional data are also critical trends that need to be addressed in the development of emotion detection systems in the future.

References

1. Medina, P. C. F., & Navarrete, R. A. G. (2022). Análisis emocional para el reconocimiento de trastornos mentales mediante lenguaje natural. *GEEKS DECC- REPORTS*, *7*(2), 7–14. Available: https://doi.org/10.24133/gdr.v7i2.2851 [Accessed: 16/Junio/2023].
2. Gomez, M. A. C. (2021). Librería para Reconocimiento de Emociones: Procesamiento de Lenguaje Natural y Señales de Audio. Edu.co. Available: https://repositorio.unbosque.edu.co/bitstream/handle/20.500.12495/7058/Gomez_Camargo_Mario_Alejandro_2021.pdf?sequence=1&isAllowed=y [Accessed: 16/Junio/2023].
3. Santiago, A., Quevedo, M. Á. C., & Donoso, F. G. (2023). *Grado en Ingeniería Robótica Trabajo Fin de Grado*. Rua.ua.es. Available: https://rua.ua.es/dspace/bitstream/10045/135241/1/Aplicacion_de_reconocimiento_de_emociones_para_person_De_Anton_Santiago_Sara.pdf [Accessed: 17/Junio/2023].
4. Martinez, F., Portale, G., Klein, H., & Olmos, O. (2020). *Reconocimiento de voz, apuntes de cátedra para Introducción a la Inteligencia Artificial*. Edu.ar. Recuperado el 1 de agosto de 2023, de Available: https://www.frba.utn.edu.ar/wp-content/uploads/2021/02/IA1_IntroReconocimientoVoz.pdf [Accessed: 17/Junio/2023].
5. Martínez, H., & Rojas, D. (2018). *Prosodia y emociones: datos acústicos, velocidad de habla y percepción de un corpus actuado*. Redalyc.org. https://www.redalyc.org/pdf/5119/511951371017.pdf [Accessed: 17/Junio/2023].
6. Montalvo, J. F. (1994). *Vista de Alexitimia: concepto, evaluación y tratamiento*. Uniovi.es. Available: https://reunido.uniovi.es/index.php/PST/article/view/7222/7086 [Accessed: 19/Junio/2023].
7. Zawadzka, T., Wiercinski, T., Waloszek, W., & Wróbel, M. R. (2023). Ontological modeling for contextual data describing signals obtained from electrodermal activity for emotion recognition and analysis. *IEEE access: practical innovations, open solutions*, *11*, 32380–32398. Available: https://doi.org/10.1109/access.2023.3257573 [Accessed: 19/Junio/2023].
8. Zhang, H., Huang, H., & Han, H. (2021). A novel Heterogeneous Parallel Convolution Bi-LSTM for speech emotion recognition. *Applied Sciences (Basel, Switzerland)*, *11*(21), 9897. Available: https://doi.org/10.3390/app11219897 [Accessed: 20/Junio/2023].
9. Rodriguez, M., Uto, K., & Shinoda, K. (2021). *Multimodal Emotion Recognition with High-level Speech and Text Features*. Researchgate.net. Available: https://www.researchgate.net/publication/356427094_Multimodal_Emotion_Recog nition_with_High-level_Speech_and_Text_Features [Accessed: 20/Junio/2023].
10. Quintero, S. (2022). *Detección idioma texto de entrada con Python*. Stack Overflow en español. Available: https://es.stackoverflow.com/questions/515789/detecci%C3%B3n-idioma-texto-de- entrada-con-python [Accessed: 21/Junio/2023].
11. Loria, S. (2020). *API reference TextBlob 0.16.0 documentation*. Readthedocs.Io. Available: https://textblob.readthedocs.io/en/dev/api_reference.html [Accessed: 21/Junio/2023].
12. Loria, S. (2020). *TextBlob: Simplified Text Processing TextBlob 0.16.0 documentation*. Readthedocs.Io. Available: https://textblob.readthedocs.io/en/dev/index.html [Accessed: 21/Junio/2023].
13. Loria, S. (2020). *Tutorial: Quickstart — TextBlob 0.16.0 documentation*. Readthedocs.Io. Available: https://textblob.readthedocs.io/en/dev/quickstart.html [Accessed: 22/Junio/2023].
14. Mohammad, S. M., & Bravo-Marquez, F. (2017). Emotion intensities in tweets. In Proceedings of the 6th International Conference on Affective Computing and Intelligent Interaction (ACII) (pp. 27-36). IEEE.
15. Martínez, Fredy H., Hernández, César A., & Giral, Diego A.. (2023). Sistema integrado de reconocimiento de emociones para interacción hombre-máquina. Información tecnológica, 34(1), 117-128. Available: https://dx.doi.org/10.4067/S0718-07642023000100117 [Accessed: 22/Junio/2023].
16. Alenljung, B., Lindblom, J., Andreasson, R., & Ziemke, T. (2019). User Experience in Social Human-Robot Interaction. In I. Management Association (Ed.), Rapid Automation: Concepts, Methodologies, Tools, and Applications (pp. 1468-1490). IGI Global. Available: https://doi.org/10.4018/978-1-5225-8060-7.ch069 [Accessed: 22/Junio/2023].
17. Hung M, Lauren E, Hon E, Birmingham W, Xu J, Su S, Hon S, Park J, Dang P, Lipsky M, Social Network Analysis of COVID-19 Sentiments: Application of Artificial Intelligence, J Med Internet Res 2020;22(8):e22590 Available: https://www.jmir.org/2020/8/e22590 DOI: 10.2196/22590 [Accessed: 22/Junio/2023].