Identifying depression and anxiety on Twitter through NLP

Cristian Alzate Urrea[1]

Advisor(s):
Omar Alexis Becerra[2]

Research practice 1
Final Report
Mathematical Engineering
Department of Mathematical Sciences
School of Sciences
Universidad EAFIT

November 20, 2022

[1]Mathematical Engineering Student at Universidad EAFIT
[2]Mathematical Engineer oabecerras@eafit.edu.co

Abstract

In this paper we explore the identification of metal diseases in Spanish Tweets by topic modeling. To do so, we use the BERTopic tool and analyze the concentration of depressive comments in each topic, which allows us to identify when a comment has a high probability of being depressive. Finally, we tested an enrichment technique on the vectors representing the comments, to see that it improves classification performance.
Keywords: BERTopic, tweets, Spanish, depression.

1 Introduction

Mental disorders such as depression are difficult to detect, and in some cases, it is too late. Usually, people do not go for help, and identifying a person with a disorder is complicated in everyday interactions, as people often hide their true feelings. Depression affects different aspects of people, such as face-to-face interaction, contributing to being lonely and affecting people's well-being. According to Penninx *et al.* (2013), depression is the most common psychiatric disorder in the world, and around 6% of society has depression at some moments.

New generations use social networks more and more, and it has become part of life for them. It is common for many people to use social networks to share different personal information and even their moods. For example, many people use Twitter to express their position on politics and defend against other people. The growing use of social networks in recent years supports this fact. For example, Facebook had 1.5% of the world's population in 2008 as users and went to about 30% in 2018 (Our World in Data, 2019). Thus, social networks can be a powerful tool for identifying people with mental disorders.

Now, we use the techniques of NLP to process large amounts of written language. However, a limitation to analyzing the information generated on the Web is the time since a person would not be able to process it Chowdhary (2020), making NLP the best option for this case. Here, NLP helps us to obtain the meaning of different texts; but in some cases, achieving this is complex since there are several ways of saying something, and the expressions change depending on the geographical location. Nevertheless, topic modeling is a great tool to segment what people write on topics, and we will analyze it later to determine their characteristics.

In this work, we will use public information extracted from Twitter users to process with NLP and perform an analysis with Topic modelling. To complete the data extraction will be using the Twitter API with special Academic Research access. In this order of ideas, one of the contributions of this work will be the construction of the dataset, which will be available to the general public or for those researchers in the region who can work with it. Then, we will preprocess the information and build some Topics relative to the research using Topic modelling. Thus, we will try to identify people with mental disorders through the topics. In addition, another contribution is to apply Topic modeling for the analysis of Tweets in Spanish, which, according to the literature revision has not been done.

Prompt treatment in cases of mental disorders is critical to people's well-being. As seen in Zalsman *et al.* (2016), anti-suicide campaigns in schools reduce suicide attempts. Ionescu *et al.* (2016) found that rapid doses of Ketamine in patients with a low likelihood of suicide significantly reduce suicidal thoughts. Thus, we see the impact of early treatment in helping to reduce suicidal tendencies in individuals.

In this sense, the benefits of early identification of mental illness will be evident in individuals and healthcare systems. The above is because prompt treatment in cases of mental disorders helps to reduce the likelihood of people committing suicidal acts. Also, a mechanism that contributes to the early identification of cases of mental illness can help reduce overcrowding in health systems, as well as provide time to plan the relevant procedures and resources used for certain patients.

2 State of the art

Mental illnesses, mental disorders also called psychiatric disorders, are common among people. This illness affects people physically and mentally, who can vary their moods between being sad, irritable, or even feeling empty, losing interest in the activities they perform in their daily lives for about two weeks (World Health Organization, 2021a). In addition, 1 out of every 100 deaths are due to this disorder (World Health Organization, 2021b), seeing the seriousness of the matter.

Actually, mental illness has a complex treatment this type of problem is made difficult by the existing stigmas surrounding these diseases (Ran *et al.*, 2021). According to Mascayano *et al.* (2020), the different types of stigmas, whether public stigma, consumer stigma, or family stigma, considerably affect people with mental disorders since because of cultures of "machismo" or "honour culture" they tend to repress their feelings and their state of mind. All these attitudes towards mental illness in the region only aggravate the phenomenon, as they force people to hide their sentiments, identifying too late in some cases.

Recently, is noting the importance that social media campaigns have begun to acquire nowadays. The use of these media for health and information campaigns intensified during the last few years, with the crisis generated by Covid 19 pandemic. In this sense, Abbas *et al.* (2021) analyzed that social networks played an important role in providing health information during the crisis. Likewise, this study also highlights the importance of mutual support in social networks since, in many cases of depression, anxiety, etc. induced by preventive isolation, social interactions of this type alleviated people's health disorders. Likewise, Alonzo & Popescu (2021) describes the so-called 5×5 campaign, where we see the feasibility and benefits of carrying out these efforts in social networks. It shows the potential and impact that a social media campaign can have in encouraging people to seek mental health help in underserved populations.

The problem of mental illness can be tackled in different manners. A manner that shows better results has been the development of a support system through social networks for people with mental disorders, mainly in the prevention of fatalities and people's well-being. Naslund *et al.* (2020) mentioned that sometimes people who have said they have depression seek help on social networks. Thus, by identifying people with some disorder, we can suggest options to help and, in extreme cases, avoid suicide. In this way, strategies like this can be implemented in the future by governments as new forms of anti-suicide campaigns.

3 Methodology

The main stages of this work is summarized on figure 1.

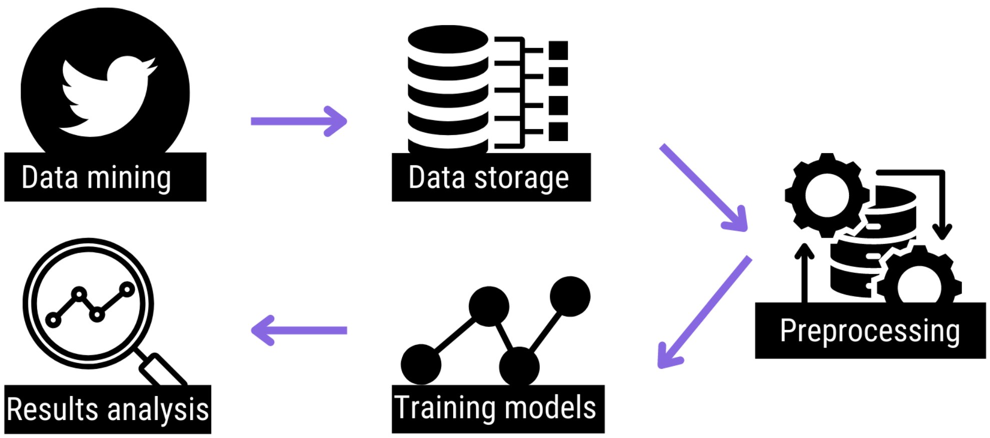

Figure 1: Project methodology.

3.0.1 Data collection

With Twitter API v2 we extracted Tweets using Tweepy. It was the best option for its features, in addition to being the recommended one to work. The idea is to search for Tweets related to mental disorders and collect the data. Then, we manually classify potential users who suffer a mental illness by explicitly manifesting themselves. Finally, we extract tweets from random users to get balanced categories. In this order of ideas, we created two datasets, one with balanced classes and the other with unbalanced classes, but with more comments to see diversity in topics.

3.0.2 Preprocessing

In this part, we start preparing the texts to pass them to our models. To do so, we perform the following steps.

- **Noise elimination** We removed mentions, links, and misspelled words.
- **Tokenization** We divide the documents into the words that make up the document.
- **Normalization** We convert all words to lowercase letters.

After this, we have to pass numerical data to our Machine Learning models, so we vectorize. For vectorization, we use Term Frequency–Inverse Document Frequency (TF-IDF), which scikit-learn library has already implemented.

3.0.3 Models

BERTopic is a technique that uses Transformers and TF-IDF to extract relevant topics in a document. It takes the comments, creates embeddings, reduces dimensionality, creates clusters, and extracts topics with TF-IDF, as seen in image 2. This technique is ideal for an unsupervised approach.

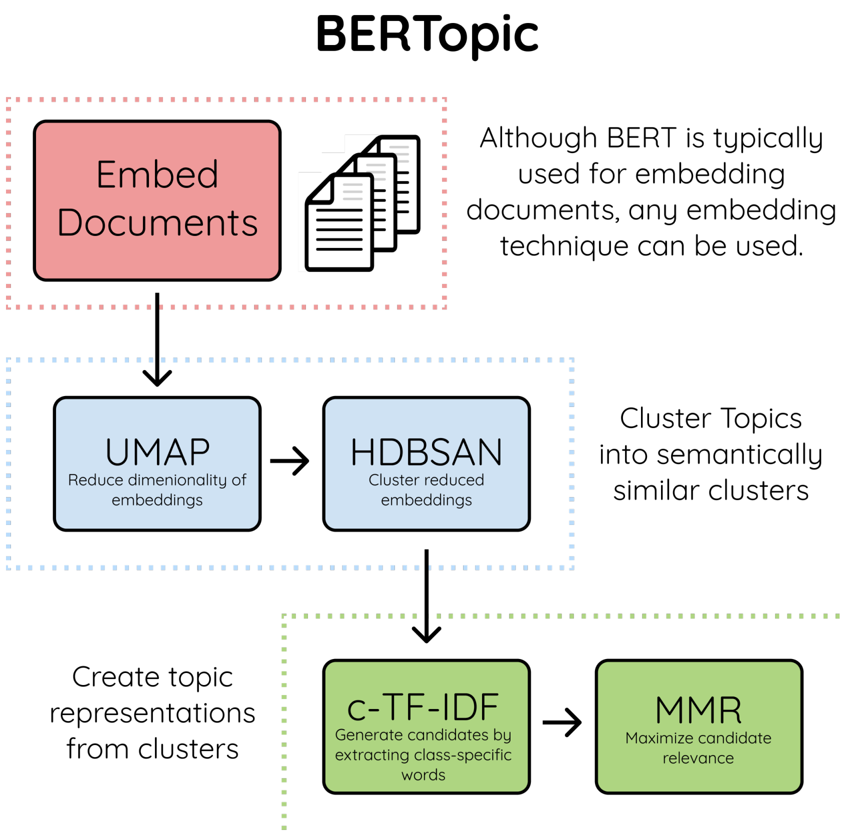

Figure 2: Components of BERTopic David Davis (2021).

First, we used BERTopic to analyze the comments with an unsupervised approach. Then, it gives us a vector of probabilities distribution that we used to enrich the embeddings of the comments. We did it because those vectors are scattered data, so enriching them improves the performance of models (Alhaj *et al.*, 2022). Our procedure is the same as shown in that article.

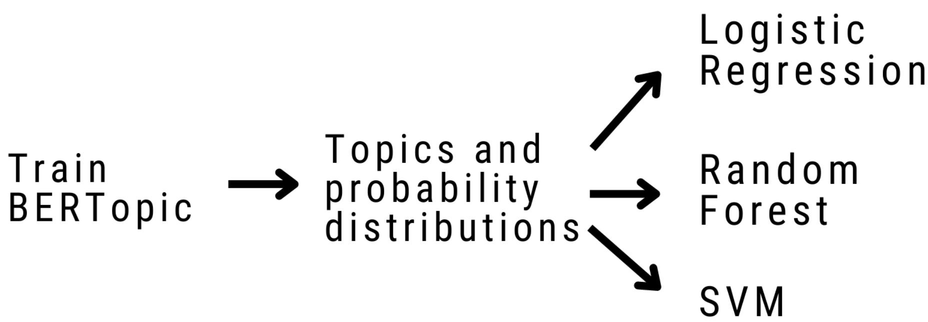

Figure 3: Implemented models.

4 Results

Here we show the exploratory analysis and the model results. We made it for both datasets. First, we have a word distribution image of the dataset, then a co-occurrence graph, and a word cloud. Then, we trained machine learning models to identify depressive people and analyzed their results.

4.1 Exploratory analysis

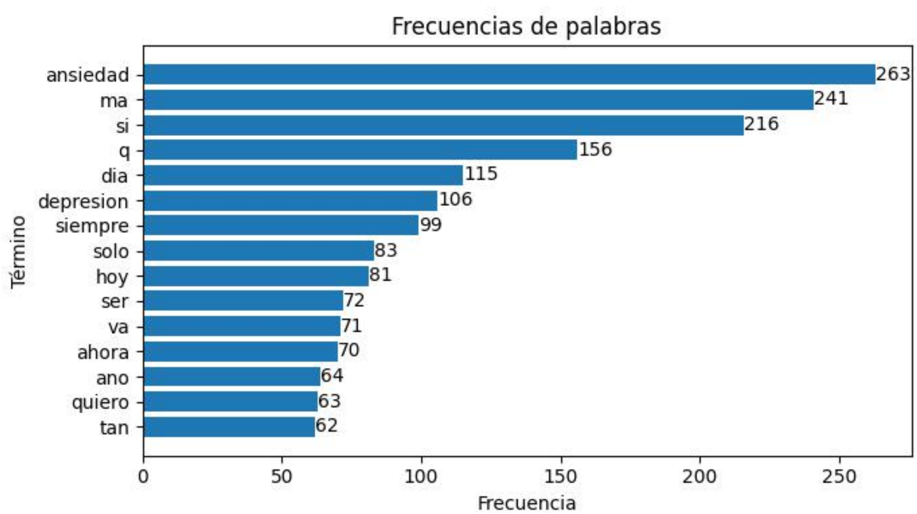

Figure 4: Word distribution of balanced dataset.

The most frequent words are *ansiedad*, *ma*, *si*, *depresion*, etc. and it makes sense since we chose a query to extract depressive comments, while with the second query we extracted random tweets.

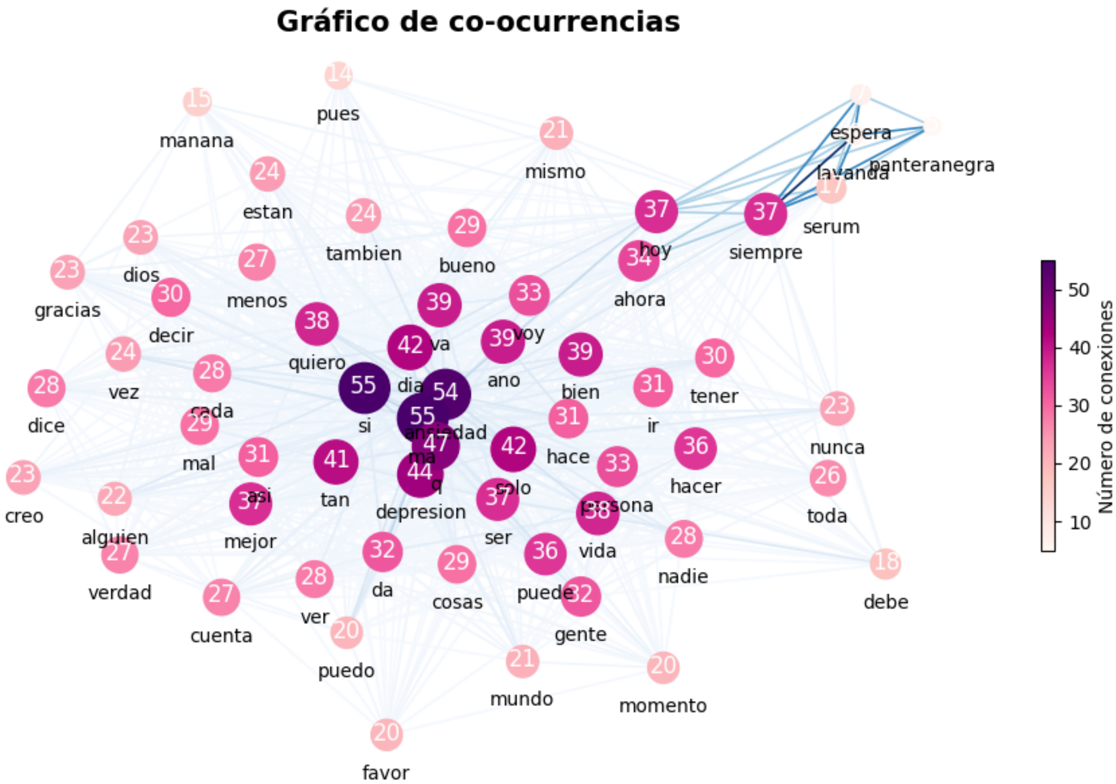

Figure 5: Co-occurrence graph of balanced dataset.

With this graph we can know how many times two words appear together, the thicker the line that joins them the more times they appear together. Also, it is useful to see how many connections a word has, which is the number inside the circle. So, it is interesting to see that *depression*, *anxiety*, *day* appear a lot, and *waiting*, *black panther*, *always* are strongly connected, which tells us that people are referring to the movie *Black Panther: Wakanda Forever*.

4.1.1 BERTopic

Then we use BERTopic to see the most relevant subjects in the comments. First, we train it with the balanced dataset, and with the unbalanced dataset. After this, we can see some topics bellow. However, there are few topics because we don't have many comments.

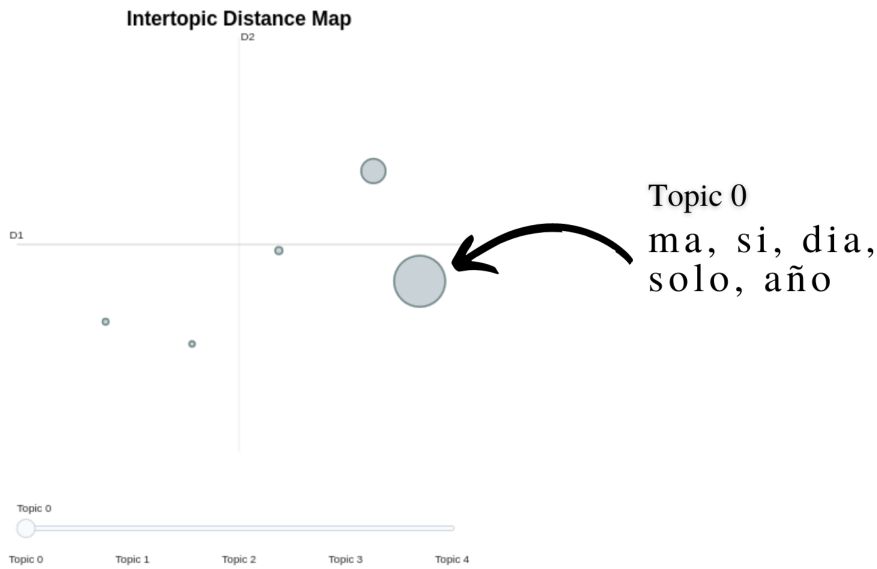

Figure 6: Words in topic 0 balanced dataset.

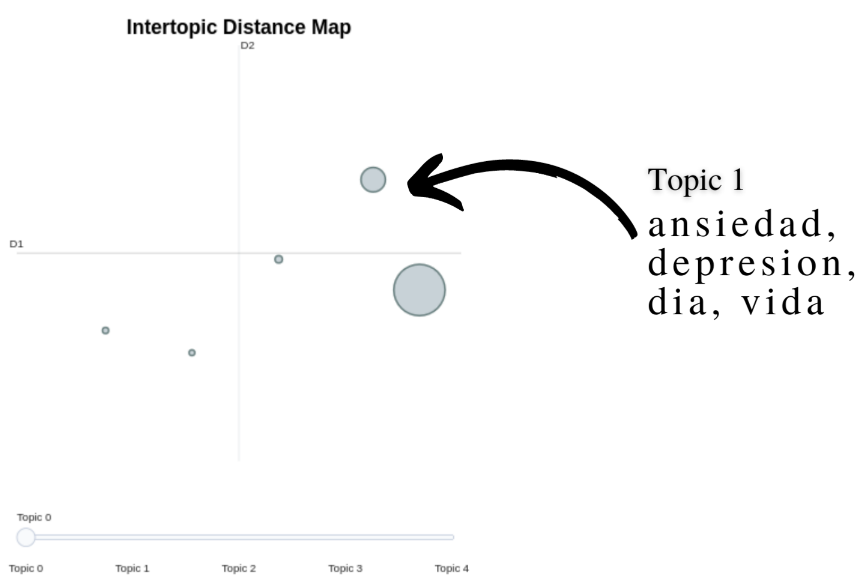

Figure 7: Words in topic 1 balanced dataset.

After having the topics, we want to know what is the percentage of depressive topics in each topic, so we count the number of depressive comments and divide by the total number of comments in that topic. Only topics with one or more depressive comments were taken into account. We summarized the results in the following table.

Topic	**Percentage**
0	0.74%
1	78.89%
4	9.52%
2	5.71%

Table 1: Percentage of depressive topics in balanced dataset.

Here we can say that if any new comment is placed in topic 1 it has a high probability of being depressive. In addition, we did the same procedure above with an unbalanced dataset, and we see more topics because we have many comments.

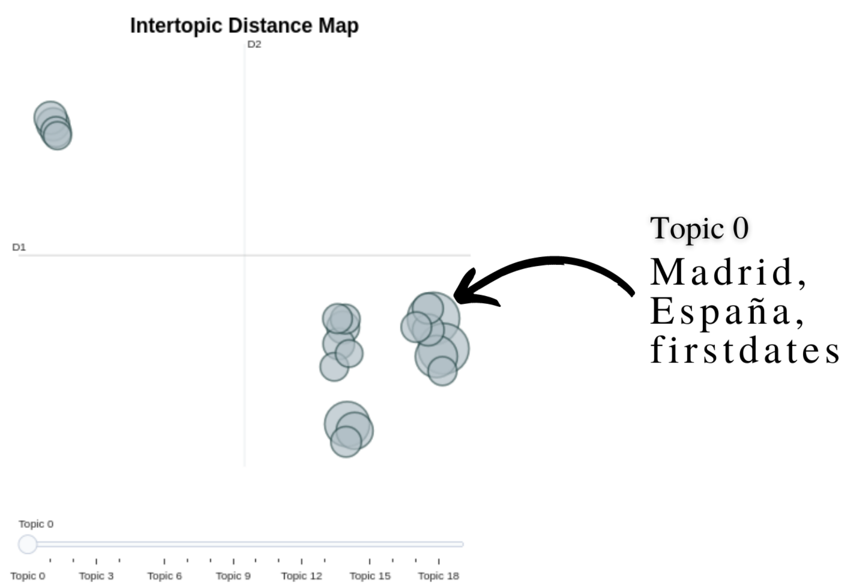

Figure 8: Words in topic 0 unbalanced dataset.

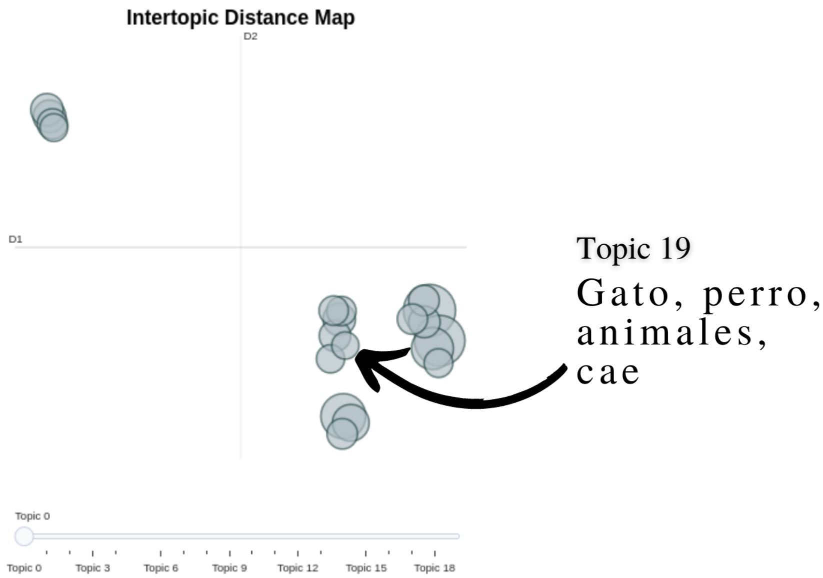

Figure 9: Words in topic 1 unbalanced dataset.

The table with percentage of comments in each topic is bellow

Topic	Percentage
17	1.69%
0	95.54%
7	91.03%
9	2.74%
15	1.56%
1	0.53%
14	3.12%

Table 2: Percentage of depressive topics in unbalanced dataset.

With this, we can say that if any new comment is placed in topic 0 or 7 it has a high probability of being depressive.

4.1.2 Machine learning models

Finally, we tested different machine learning models to classify depressive comments. This compares the performance between the original vectors and the enriched vectors. Below two graphs are comparing the f1-score in the classification of each class.

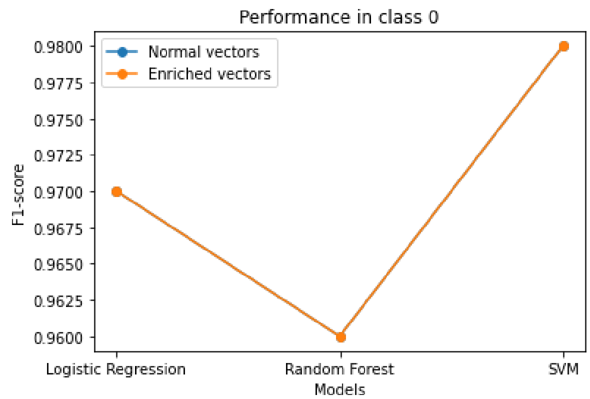

Figure 10: Comparison of classification performance of the models for class 0.

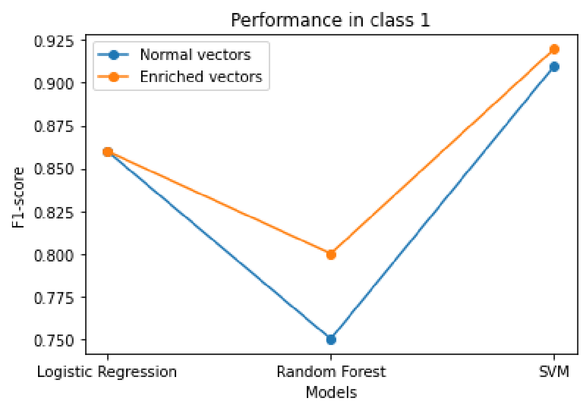

Figure 11: Comparison of classification performance of the models for class 1.

In general, we obtained better results in models classifying depressive comments when we enrich the vectors.

5 Conclusions and future research

Labeling data is hard and extensive work. This is because we have to read each comment and the majority extracted by the Twitter API are not depressive comments. For this reason, is not a good way to develop a project of classification. Instead, an unsupervised approach can be considered, like search depressive people through topics. We saw that some topics have the majority of depressive comments, and we can think that if a new comment is inside that topic then it is depressive.

Finally, using enriched vectors improve the performance of ML models. Since vectors of comments are sparse data, enriching them with the information generated by BERTopic helps us to improve the classification. We can see that improves the performance on depressive data, and it has sense because if a comment is not depressive so it is inside of any other topic, but if it is depressive, it is inside a depressive topic, which is a low number.

Acknowledgements

Special thanks to Omar Alexis Becerra who was present at all times during the project and contributed to its success.

References

Abbas, Jaffar, Wang, Dake, Su, Zhaohui, & Ziapour, Arash. 2021. The role of social media in the advent of COVID-19 pandemic: crisis management, mental health challenges and implications. *Risk management and healthcare policy*, **14**, 1917.

Alhaj, Fatima, Al-Haj, Ali, Sharieh, Ahmad, & Jabri, Riad. 2022. Improving Arabic cognitive distortion classification in Twitter using BERTopic. *International Journal of Advanced Computer Science and Applications*, **13**(1), 854–860.

Alonzo, Dana, & Popescu, Marciana. 2021. Utilizing social media platforms to promote mental health awareness and help seeking in underserved communities during the COVID-19 pandemic. *Journal of Education and Health Promotion*, **10**.

Chowdhary, KR1442. 2020. Natural language processing. *Fundamentals of artificial intelligence*, 603–649.

David Davis. 2021. *Tutorial de PNL: Modelado de temas en Python con BerTopic*. https://hackernoon.com/es/nlp-tutorial-tema-modelado-en-python-con-bertopic-372w35l9. Accessed: 2022-11-01.

Ionescu, Dawn F, Swee, Michaela B, Pavone, Kara J, Taylor, Norman, Akeju, Oluwaseun, Baer, Lee, Nyer, Maren, Cassano, Paolo, Mischoulon, David, Alpert, Jonathan E, *et al.* 2016. Rapid and sustained reductions in current suicidal ideation following repeated doses of intravenous ketamine: secondary analysis of an open-label study. *The Journal of clinical psychiatry*, **77**(6), 10114.

Mascayano, Franco, Toso-Salman, Josefina, Ho, Yu Chak Sunny, Dev, Saloni, Tapia, Thamara, Thornicroft, Graham, Cabassa, Leopoldo J, Khenti, Akwatu, Sapag, Jaime, Bobbili, Sireesha J, *et al.* 2020. Including culture in programs to reduce stigma toward people with mental disorders in low-and middle-income countries. *Transcultural psychiatry*, **57**(1), 140–160.

Naslund, John A, Bondre, Ameya, Torous, John, & Aschbrenner, Kelly A. 2020. Social media and mental health: benefits, risks, and opportunities for research and practice. *Journal of technology in behavioral science*, **5**(3), 245–257.

Our World in Data. 2019. *The rise of social media*. https://ourworldindata.org/rise-of-social-medianote-5. Accessed: 2022-07-29.

Penninx, Brenda WJH, Milaneschi, Yuri, Lamers, Femke, & Vogelzangs, Nicole. 2013. Understanding the somatic consequences of depression: biological mechanisms and the role of depression symptom profile. *BMC medicine*, **11**(1), 1–14.

Ran, Mao-Sheng, Hall, Brian J, Su, Tin Tin, Prawira, Benny, Breth-Petersen, Matilde, Li, Xu-Hong, & Zhang, Tian-Ming. 2021. Stigma of mental illness and cultural factors in Pacific Rim region: a systematic review. *BMC psychiatry*, **21**(1), 1–16.

World Health Organization. 2021a. *Depression*. https://www.who.int/news-room/fact-sheets/detail/depression: :text=Depression%20is%20a%20common%20mental,affected%20by%20depression%20than%20men. Accessed: 2022-07-21.

World Health Organization. 2021b. *One in 100 deaths is by suicide*. https://www.who.int/news/item/17-06-2021-one-in-100-deaths-is-by-suicide. Accessed: 2022-08-08.

Zalsman, Gil, Hawton, Keith, Wasserman, Danuta, van Heeringen, Kees, Arensman, Ella, Sarchiapone, Marco, Carli, Vladimir, Höschl, Cyril, Barzilay, Ran, Balazs, Judit, *et al.* 2016. Suicide prevention strategies revisited: 10-year systematic review. *The Lancet Psychiatry*, **3**(7), 646–659.