Anxiety detection using the AMAS-C test and feeling analysis on the Facebook social network

[1]Daniel Avila[0000-0001-7371-6104], [1]Ariel Altamirano[0000-0002-0677-2816], [2]Josué Avila[0000-0002-2086-053X]
and [1,3] Graciela Guerrero[0000-0002-0903-734X]
[1] Departamento de Ciencias de la Computación,
Universidad de las Fuerzas Armadas ESPE, Sangolquí, Ecuador.
[2] Facultad de Ciencias Psicológicas,
Universidad Central del Ecuador, Quito, Ecuador.
[3] Departamento de Sistemas Informáticos
Universidad de Castilla - La Mancha, Albacete, Spain.
[1]{rdavila, haaltamirano1, rgguerrero}@espe.edu.ec, [2]fjavila@uce.edu.ec, [3]rosagraciela.guerrero@alu.uclm.es

Abstract —**This research exposes the proposal of a web application based on Facebook, Meaning Cloud and AMAS-C test tools, with the purpose of predicting anxiety in university students. For this purpose, a prediction algorithm was developed that takes the processed information from the publications of the Facebook users and compares it with the information of the previous learning, giving as a result the prediction of total anxiety of the individual. Finally, this research exposes an average absolute error 6 of 84 possible points, corresponding to the AMAS-C test scale, where 5 correct predictions, 12 medium correct predictions and 1 erroneous predictions were obtained. The conclusion is that the tools selected together with the proposed algorithm could be used as an alternative to the AMAS-C test.**

Keywords - anxiety, feeling analysis, prediction, Facebook, AMAS-C.

I. INTRODUCTION

According to Wongkoblap, Vadillo & Curcina [1] mental illness is becoming a serious global health problem, with an increasing number of patients suffering from depression, anxiety and other disorders. This issue should be highlighted and given importance because of the fact that 350 million patients worldwide suffer from mental disorders, and it also impacts the world economy at a cost of US $2.5 trillion in 2010 and an estimated expenditure of US $6 trillion by 2030.

Anxiety is the most common mental health condition in children and adolescents, with prevalence rates ranging from 5% to 10% in children and up to 25% in adolescents [2]. Adolescents are more likely to suffer from anxiety disorders, which Gerez et al. [3] say are the most pervasive mental health problems, and despite the devastating effects of anxiety disorders (ADS) worldwide, debates about causes and remedies have not solved the psychologist's conundrum: should they be treated and how? Anxiety disorders have consequences for an individual's mental health, which leads to a certain insecurity and inferiority when it comes to offering professional advice. According to Deepali et al. [4] unlike physical illness, in mental illness the symptoms are not evident in the early stages, for this reason, psychologists look for different ways to understand the patient in treatment, so, with the popularity of social networks according to Jia [5] it is feasible to take advantage of these data for stress and the detection of depression and anxiety.

A university needs students to make it work, so institutions of higher learning should be aware of any factors that may affect the academic performance of their students, since low academic performance can lead to students missing classes or in the worst case, dropping out of school altogether. Anxiety is a determining factor in the academic performance of university students, especially when they are first-time students [6]. Although there is no direct external solution to this problem, teachers can warn their students about anxiety so that they can choose an assertive teaching methodology for this group of students [7].

The present research will seek to predict the anxiety of university students by combining an anxiety detection test for university students, an Application Programming Interface (API) provided by the Meaning Cloud tool to perform the sentiment analysis of the Facebook posts of the users to be evaluated. Finally, this information will be crossed to perform a learning and prediction of anxiety levels in university students.

This research is structured in 4 relevant sections, where: i) Section 1 indicates the related works in terms of sentimental analysis and sentimental analysis tools. ii) Section 2 presents the tools that have been used for the research proposal, the architecture and the applied development method. iii) Section 3 carries out the evaluation with users that are in the range recommended by the AMAS-C test, and also presents the results obtained. iv) Finally, the last section presents the conclusions and possible lines of future work.

II. RELATED WORKS

Three relevant databases were consulted for the selection of related works: Scopus, ACM and IEEE. A search string was formed: (Mental Health OR Anxiety OR Depression OR Mood) AND ((Social Network Analysis OR Feelings Analysis) AND (Application OR Tool)). By executing the search chain, different article numbers were obtained from each of them, making a total

2020 15th Iberian Conference on Information Systems and Technologies (CISTI)
24 – 27 June 2020, Seville, Spain
ISBN: 978-989-54659-0-3

of 4142. However, for the present research work only the 10 most relevant papers have been considered.

Over the past 10 years, social networks have become an important component of young people's lives, with the use of a peak in 16-24 years [2]. According to Miks & Aurel [8], social networks allow individuals to express their opinions, feelings and thoughts on a variety of topics in the form of short text messages or tweets. By posting the day-to-day feelings or thoughts of each user on social networks, this posted information becomes valuable information that according to Manikonda & De Choudhury [9] is self-disclosed by the user. Although there is a debate about the excessive use of social networks if they are for personal benefit or harm since it is not possible to control what type of information is true within social networks or if the personal information is completely secure; according to Park et al. [10] recent studies have used data from online social networks to predict the well-being of the individual.

For the present work of investigation, it has been considered to divide in three categories the search of related works because a work similar to the proposed one does not exist. In the first section, a description is made of what is a sentiment analysis versus social network analysis. The second section presents works related to our proposal on sentiment analysis, and finally the third section presents the tools that have been used in works that have performed sentiment analysis.

A. Sentiment Analysis versus Social Network Analysis.

The analysis of feelings is defined as the process of extracting the polarity of a subjective sentence [11]. Social network analysis is defined as the way to identify, analyze, visualize or simulate nodes and edges of various types of input [12]. In the selected related works, it was found that only sentiment analysis is used together with the detection of some phenomenon, for this reason it is recommended to use sentiment analysis tools.

B. Sentiment Analysis

Cao et al. [13] define sentiment analysis as the mining of affective information from data. Abhijit, Kshama & Kotrappa [14] recommend that to obtain a better result in the analysis of feelings, they should be captured in general feelings (positive, negative and neutral). Most of the tools and studies carried out in sentiment analysis are in English [13].

C. Sentiment Analysis Tools

Cerón & León [15] developed a sentiment analysis tool to analyze or detect a phenomenon in Colombia's 2014 presidential elections. Arriaga et al. [16] used Analyze Words as a sentiment analysis tool to analyze or detect a phenomenon of depression and anxiety. However, this tool was of Anglo-Saxon origin, so they attributed the erroneous results to the language support of this tool; since the phenomenon investigated was sought or detected in a Spanish-speaking context.

Finally, it was determined that the sentiment analysis tool to be selected for the present research project should respond to certain characteristics such as: Spanish language support, categorization of sentiments into three general sentiments (positive, negative and neutral) and free access.

III. Design and development

This section presents the tools used for the development of this research, the architecture used and the development method applied.

A. Tools

The following tools were used for the development of this research work:

1) Test AMAS-C

The test used is focused on the detection of anxiety in university students, and consists of 49 reagents concerning five subscales (restlessness/hypersensitivity, physiological anxiety, social concerns/concentration, test anxiety and lying) [17]. Interpretation of the results is made using tables provided by the test itself.

The AMAS-C test was used to identify the level of anxiety in university students, since they have an adequate level of reliability according to the alpha coefficient [18]. According to Reynolds et al. [19] the Total Anxiety Coefficient (TOT) is 0.94, while in the subscales the coefficients are close to 0.7 as in the case of hypersensitivity and physiological anxiety.

2) Meaning Cloud API

Meaning Cloud [20] is a software-as-a-service that offers text analytics. Among the various services it offers, the sentiment analysis API is the tool selected to analyze Facebook posts.

3) Firebase

Firebase [21] is a platform for the development of web applications that offers databases and hosting for each application. In the development of the present project, the database in real time and the hosting properly granted by this platform are used.

4) Learning model

Supervised machine learning techniques use the expected result of classification as an attribute involved in learning. This result should be provided by the expert in classifying problem patterns [22].

B. Architecture

For this research work, an architecture has been proposed (Fig. 1), where a web application developed by the authors connects to Facebook. If the user gives permissions to obtain his data (photo, name, email, publications), the API provided by Firebase will obtain them.

When the user enters the web application for the first time he must register with his active Facebook account, once registered he must log in to be able to take the AMAS-C anxiety test, when the survey is completed and sent by the user, the results obtained will be stored within the Realtime Database provided by Firebase.

Using the sentiment analysis tool Meaning Cloud, we will proceed to evaluate the publications made by the user (data obtained before registering in the application), where we will seek to compare and analyze four aspects: number of publications, publications grouped in neutral, publications grouped in negative, publications grouped in positive.

2020 15th Iberian Conference on Information Systems and Technologies (CISTI)
24 – 27 June 2020, Seville, Spain
ISBN: 978-989-54659-0-3

Contiguously, when there is a reasonable amount of records, the web application will compare the results obtained previously (saved in the Firebase database). In order that when a new user enters the application his publications are obtained and after having analyzed them it is possible to predict the results to be obtained in the AMAS-C anxiety test.

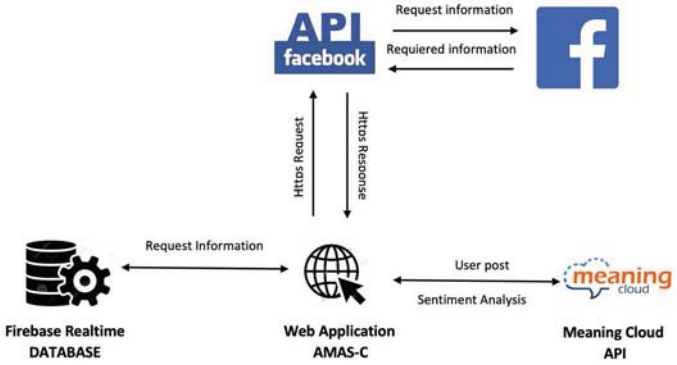

Figure 1. Application architecture.

C. Development

The development of the present work of investigation is conformed by 6 relevant aspects that are exposed next:

1) Creation of the project and obtaining public data from Facebook

In order to obtain public data permission from the Facebook user (Fig.2), an application was created and enabled within the Facebook for Developers platform in which there is an application identifier and a key, the same ones that were implemented in the project created by the Firebase web application platform.

Figure 2. Requesting permission for the web application on Facebook.

2) Implementation of the AMAS-C test and results

The implementation is done in a section called test within the web application which consists of all the questions and a send button which validates and enters the results previously analyzed to the database used, to achieve the identification of each user is obtained the public Facebook identifier that is unique to each user. Within the results section, a table is presented depending on the user's role. If the user's role is psychologist (Fig.3) at the institution, there is a table of each user's results (analysis of the test and the percentage of positive, negative and neutral publications); on the other hand, if the user is a student, there is a table with his or her results and a link to observe the results of the analysis of each of his or her publications.

Figure 3. AMAC-C test - role of the psychologist.

3) Request for permission from Facebook

Facebook publications not being of a public nature must make a request for permission the same that is requested from the application of Facebook for Developers previously created where you have to identify the identity of the owner of the application, its privacy policies, specify the case of use of the request and an attached video of the operation of the application, once sent the request response time of up to 5 days.

4) Posts Analysis

Previously accepted the permission request, through the MeaningCloud API we proceed to analyze each publication obtained at the moment the user enters the application. This analysis is saved in the database to perform the respective calculations and present them within the results table in the case of the psychologist; in the case of the student, the information of his publications is saved within the database in order to present the student with the analysis of each of his publications.

5) Learning Process

For the supervised learning, 34 students who are studying from the first to the third semester of their degree were considered. After entering the website (https://test-ansiedad.web.app/) with their Facebook credentials, the following algorithm was executed: i) Obtain a unique identifier (ID) that Facebook assigns to the developed application related to each user, accompanied by user data (including publications). ii) All the data previously obtained is stored in the database. iii) Once the user answers the AMAS-C test, the different scores will be obtained, respectively for each subscale of the test. iv)The scores obtained previously are stored in the database. v) All the publications obtained are filtered considering the exclusion criteria (table 1). vi)The sentimental analysis of the publications is carried out. vii) The total number of publications is categorized by user (positive, negative and neutral). viii) Finally, the scores obtained by answering the AMAS-C test, the categories of publications and the unique ID are linked.

6) Prediction Algorithm

As can be seen in Fig. 4, for the prediction of an individual, the amount in percentage of Facebook posts classified as positive, negative and neutral is needed as input data. Each of these data enter the processing for the prediction, where three

ISBN: 978-989-54659-0-3

results are obtained respectively to the amount of input data; these three predictions are averaged to obtain a final prediction. The prediction is related to predicting an individual's total anxiety.

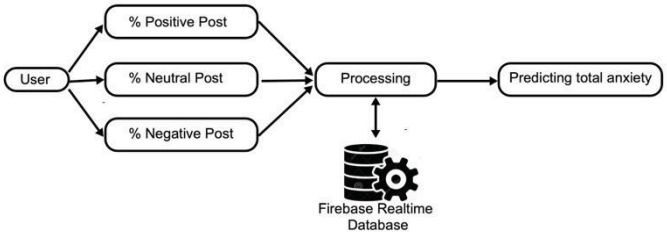

Figure 4. Flowchart Total Anxiety Prediction Algorithm

a) Data processing

The input data is compared with the previous learning data, where three different cases may occur:

- Case 1.- There is only one data completely equal to the input data. In this case, the total anxiety score corresponding to the detected data is obtained.
- Case 2.- There are several data completely equal to the input data. In this case, an average of the total anxiety score is obtained for all the data detected
- Case 3.- There is no data completely equal to the input data. For this particular case, a higher and a lower data is sought close to the input data. Where two notes of total anxiety will correspond (corresponding to the upper and lower data) which will finally be averaged to obtain a single note of total anxiety. It is highlighted that for these two data found, the procedure of case 1 and case 2 respectively will be applied as necessary.

This data procedure is repeated for each input data.

b) Prediction

At the end of the data processing, three total anxiety notes will be obtained, which will be averaged together to obtain a final prognosis of the individual's total anxiety.

IV. EVALUATION

For the evaluation, a population of approximately 200 university students was taken, which belong to the first to third level of the careers taught by the Computer Science Department of the University of the Armed Forces-ESPE. From this population a sample of 54 students was taken, of which 34 students were key to learning the application for future anxiety prediction. It is necessary to mention that this population of users are Spanish speaking.

A. Exclusion Criteria

In the sentiment analysis of the Facebook publications of the students belonging to the sample, exclusion criteria for these publications were determined. In Table 1, these criteria are shown with their respective reason.

TABLE I. FACEBOOK POSTING EXCLUSION CRITERIA AND REASONS

Exclusion Criteria for Facebook Posts		
Nº	***Criteria***	***Reason***
1	Publications about games, applications, or any service provided by Facebook.	The publications created by these services do not really reflect the feelings or thoughts of the user.
2	Publications containing only one link or several links.	A link or several links do not reflect any particular sentiment or the context of the sentiment may be lost when analyzing a URL.
3	Publications containing photos without any comments involved.	For sentiment analysis you need a specific text to analyze, so if it did not exist there would be no material for such a process.
4	Publications containing only labels.	A name or list of names does not reflect a particular feeling.
5	Publications containing only icons.	The API for sentiment analysis does not recognize icons so this does not reflect the user's feeling or thinking.
6	Publications containing only addresses.	These publications containing only addresses do not reflect an accurate feeling.

B. Proposed scenarios for predicting anxiety

To run the scenarios we worked with 54 students who took the AMAS-C test on the website of the proposed application, divided into two groups: i) with 34 students we performed the supervised learning, ii) while with 20 students we performed the prediction only with the Facebook comments, in order to make a comparison of the effectiveness of the proposed algorithm.

For this section, we worked with 20 university students who would have their total anxiety predicted. This group of students was randomly selected to create two scenarios for the prediction: Scenario 1: 18-25-year-old student from the population. Twelve students were selected for this scenario; Scenario 2: A student aged 18-25 years, who is studying from the fourth level onwards. Eight students were selected for this scenario.

C. Results

A total of seven parameters were defined to be considered to verify the performance of the proposed prediction algorithm, of which 2 parameters are related to error theory, of the others their importance and reason for choice is explained in Table 2.

TABLE II. PARAMETERS OF THE PROPOSED PREDICTION ALGORITHM

Prediction algorithm parameters		
Nº	***Parameter***	***Description***
1	Prediction Cases: Correct	Any case where the prediction error is 5% or less. Predictions very close to the real value.
2	Prediction Cases: Medium Correct	Any case where the prediction error is 15% or less. Predictions that are moderately far from the real value.
3	Prediction Cases: Incorrect	Any case where the prediction error is 16% or greater. Predictions far from their real value.
4	Probability of Getting a Correct Prediction	Know what the trend is to get a correct prediction based on the data obtained.
5	Probability of obtaining a Medium-Correct Prediction	Know what the trend is to obtain a moderately correct prediction based on the data obtained.

The results obtained from scenarios 1 and 2 are shown in table 3.

ISBN: 978-989-54659-0-3

TABLE III. PARAMETERS OF THE PROPOSED PREDICTION ALGORITHM

Exclusion Criteria for Facebook Posts		
Parameter	***Scenario 1***	***Scenario 2***
Prediction Cases: Correct	4	1
Prediction Cases: Medium Right	9	8
Prediction Cases: Incorrect	3 / *1	0
Probability of Getting a Correct Prediction	33.33 %	12.5 %
Probability of obtaining a Medium-Correct Prediction	75%	100%
Absolute Error Average	9 points / *6 points	6 points
Relative Average Error	17.81% / *8.63%	10.96%

* Table III, in scenario 1, presents the results (*) without the outliers, this happens in the prediction of incorrect cases, the average of the absolute error and in the average of the relative error.

The AMAS-C test is composed of 49 reagents, where the highest score is 84 points. As can be seen in Table 3, the correct prediction cases for Scenario 1 are 4, compared with Scenario 2. However, in Scenario 1 there is one case of incorrect predictions, while in Scenario 2 there are none, so the error increases in Scenario 1.

The error obtained is considered low (scenario 1: 8.63% and scenario 2: 10.96%) since in the categories of possible diagnoses for total anxiety of the AMAS-C test, an absolute error of 6 points means that the prediction of anxiety would obtain a diagnosis equal to the real one or with a maximum difference of one category (5 possible diagnostic categories).

V. CONCLUSIONS

From the analysis of the related works, information was extracted about development tools, where: API Meaning Cloud (sentiment analysis) supports the analysis in Spanish language, classifies the sentiments in three categories (positive, neutral, negative), free access; Fire Base provided a free hosting and database up to 10 GB of information (limiting the capacity of the registered data).

It is estimated that one of the main reasons why this research obtained positive results is because exclusion criteria were applied to the data obtained with respect to Facebook publications; these criteria allowed filtering and including publications that reflect the user's thinking, eliminating those publications related to games, music, etc.

With the results obtained, it was demonstrated that it is possible to predict the total anxiety of an individual with a low degree of error; however, it is proposed that in order to decrease the degree of error, it is recommended to increase the number of users for the learning stage, as well as for the prediction in order to evaluate if it is reliable.

Finally, one possibility for future work could be to include user comments on Facebook and other social networks, to obtain more data related to the feelings exposed by the user and to show if this action decreases the degree of error of the prediction.

REFERENCES

[1] A. Wongkoblap, M. Vadillo and V. Curcina, "Detecting and Treating Mental Illness on Social Networks," *IEEE Explore,* p. 330, 2017.

[2] J. Glover and S. Fritsch, "#KidsAnxiety and Social Media: A Review," SCOPUS, pp. 171-182, 2018.

[3] M. Gerez, E. Suárez, A. Fielding, A. Hendrick and R. Martin, "The crossroads of anxiety: Distinct neurophysiological maps for different symptomatic groups," SCOPUS, pp. 159-175, 2016.

[4] J. Deepali, J. Makhija, Y. Nabar and N. Nehet, "Mental Health Analysis Using Deep Learning for Feature Extraction," ACM, pp. 356-359, 2018.

[5] J. Jia, "Mental Health Computing via Harvesting Social Media Data," ACM, pp. 70-70, 2018.

[6] Tijerina, González, Gómez, Cisneros, Rodríguez and Ramos, "Depresión, ansiedad y estrés en estudiantes de nuevo ingreso a la educación superior," Revista Salud Pública y Nutrición, 2018.

[7] G. Zambrano, K. Rodriguez and L. Guevara, "ANÁLISIS DE LA DESERCIÓN ESTUDIANTIL EN LAS UNIVERSIDADES DEL ECUADOR Y AMÉRICA LATINA," Revista Pertinencia Academica, 2018.

[8] Q. Miks and J. Aurel, "Sentiment Analysis on Tweets with Punctuations, Emoticons, and Negations," ACM, pp. 266-270, 2019.

[9] L. Manikonda and M. De Choudhury, "Modeling and Understanding Visual Attributes of Mental Health Disclosures in Social Media," ACM, pp. 170-181, 2017.

[10] S. Park, I. Kim, S. Won and J. Yoo, "Manifestation of Depression and Loneliness on Social Networks: A Case Study of Young Adults on Facebook," ACM, pp. 557-570, 2015.

[11] M. Rathan, V. R. Hulipalled, P. Murugeshwari and H. M. Sushmitha, "Every Post Matters A Survey on Applications of Sentiment Analysis in Social Media," IEEE Xplore, pp. 709-714, 2017.

[12] N. Akhtar, "Social Network Analysis Tools," IEEE Explore, pp. 388-392, 2014.

[13] J. Cao, K. Zeng, W. Hui, J. Cheng, Q. Fengcai, W. Ding and Y. Gao, "Web-Based Traffic Sentiment Analysis Methods and Applications," IEEE Xplore, pp. 844-853, 2014.

[14] J. P. Abhijit, K. Kshama and S. Kotrappa, "Emotweet Sentiment Analysis tool for twitter," IEEE Xplore, pp. 157-159, 2016.

[15] J. A. Cerón Gúzman and E. León Gúzman, "A Sentiment Analysis System of Spanish Tweets and Its Application in Colombia 2014 Presidential Election," IEEE Xplore, pp. 250-257, 2016.

[16] T. C. Arriaga, J. L. Frías, S. H. Flores, O. I. Ayllon and A. S. Álvarez, "Escrutinio de depresión y ansiedad en twitter a través de un programa de análisis de palabras," Investigación en Educación Médica, vol. 4, no. 13, pp. 16-21, 2015.

[17] C. Reynolds, Richmond and Lowe, "AMAS. Escala de ansiedad manifiesta en adultos," Psicólogos Especialistas Asociados, 2006.

[18] L. J. Cronbach , Coefficient alpha and the internal structure of the test, Psychometrika, 1951.

[19] C. R. Reynolds, B. Richmond and P. Lowe, Escala de Ansiedad Manifiesta en Adultos, Mexico: El Manual Moderno, 2007.

[20] MeaningCloud LLC, "Meaning Cloud," 2020. [Online]. Available: https://www.meaningcloud.com/es. [Accessed 23 January 2020].

[21] Google Developers, "Firebase," 2020. [Online]. Available: https://firebase.google.com/?hl=es-419. [Accessed 12 November 2020].

[22] S. Marsland, Machine Learning An Algorithmic Perspective,, Boca raton: CRC, 2015.

ISBN: 978-989-54659-0-3