Lecture Notes in Networks and Systems 576

Jezreel Mejia
Mirna Muñoz
Álvaro Rocha
Víctor Hernández-Nava *Editors*

New Perspectives in Software Engineering

Proceedings of the 11th International Conference on Software Process Improvement (CIMPS 2022)

Springer

Lecture Notes in Networks and Systems

Volume 576

The series “Lecture Notes in Networks and Systems” publishes the latest developments in Networks and Systems—quickly, informally and with high quality. Original research reported in proceedings and post-proceedings represents the core of LNNS.

Volumes published in LNNS embrace all aspects and subfields of, as well as new challenges in, Networks and Systems.

The series contains proceedings and edited volumes in systems and networks, spanning the areas of Cyber-Physical Systems, Autonomous Systems, Sensor Networks, Control Systems, Energy Systems, Automotive Systems, Biological Systems, Vehicular Networking and Connected Vehicles, Aerospace Systems, Automation, Manufacturing, Smart Grids, Nonlinear Systems, Power Systems, Robotics, Social Systems, Economic Systems and other. Of particular value to both the contributors and the readership are the short publication timeframe and the world-wide distribution and exposure which enable both a wide and rapid dissemination of research output.

The series covers the theory, applications, and perspectives on the state of the art and future developments relevant to systems and networks, decision making, control, complex processes and related areas, as embedded in the fields of interdisciplinary and applied sciences, engineering, computer science, physics, economics, social, and life sciences, as well as the paradigms and methodologies behind them.

Indexed by SCOPUS, INSPEC, WTI Frankfurt eG, zbMATH, SCImago.

All books published in the series are submitted for consideration in Web of Science.

For proposals from Asia please contact Aninda Bose (aninda.bose@springer.com).

More information about this series at https://link.springer.com/bookseries/15179

Jezreel Mejia · Mirna Muñoz ·
Álvaro Rocha · Víctor Hernández-Nava
Editors

New Perspectives in Software Engineering

Proceedings of the 11th International Conference on Software Process Improvement (CIMPS 2022)

Springer

Editors
Jezreel Mejia
Centro de Investigación en Matemáticas, A.C.
Unidad Zacatecas
Zacatecas, Mexico

Mirna Muñoz
Centro de Investigación en Matemáticas, A.C.
Unidad Zacatecas, Zacatecas
Zacatecas, Zacatecas, Mexico

Álvaro Rocha
ISEG
Universidade de Lisboa
Lisbon, Portugal

Víctor Hernández-Nava
Universidad Hipócrates
Acapulco de Juárez, Guerrero, Mexico

ISSN 2367-3370 ISSN 2367-3389 (electronic)
Lecture Notes in Networks and Systems
ISBN 978-3-031-20321-3 ISBN 978-3-031-20322-0 (eBook)
https://doi.org/10.1007/978-3-031-20322-0

This Springer imprint is published by the registered company Springer Nature Switzerland AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Organization

Conference

General Chairs

Jezreel Mejía	Mathematics Research Center, Research Unit Zacatecas, Mexico
Mirna Muñoz	Mathematics Research Center, Research Unit Zacatecas, Mexico

The general chairs and co-chair are researchers in computer science at the Research Center in Mathematics, Zacatecas, México. Their research field is software engineering, which focuses on process improvement, multi-model environment, project management, acquisition and outsourcing process, solicitation and supplier agreement development, agile methodologies, metrics, validation and verification and information technology security. They have published several technical papers on acquisition process improvement, project management, TSPi, CMMI and multi-model environment. They have been members of the team that has translated CMMI-DEV v1.2 and v1.3 to Spanish.

General Support

CIMPS general support represents centers, organizations or networks. These members collaborate with different European, Latin America and North America Organizations. The following people have been members of the CIMPS conference since its foundation for the last 10 years.

Gonzalo Cuevas Agustín	Politechnical University of Madrid, Spain
Jose A. Calvo-Manzano Villalón	Politechnical University of Madrid, Spain
Tomas San Feliu Gilabert	Politechnical University of Madrid, Spain
Álvaro Rocha	Universidade de Lisboa, Portugal
Yadira Quiñonez	Autonomous University of Sinaloa, Mexico
Gabriel A. García Mireles	University of Sonora, Mexico

Local Committee

CIMPS established a Local Committee from the Mathematics Research Center, Research Unit Zacatecas, MX, the University Hippocrates, of Acapulco, Guerrero, México, CAMSA Corporation, Cybersecurity Systems and Intelligence Enterprise and CANIETI Guerrero, México. The list below comprises the Local Committee members.

CIMAT UNIT ZACATECAS:

Hipócrates University (HU):

CAMSA Corporation:

Cybersecurity Systems and Intelligence Enterprise:

Hugo Montoya Diaz (CEO and CTO), Mexico

CANIETI Guerrero, México:

Luis Felipe Monroy Álvarez (President of the Office), Mexico

Scientific Program Committee

CIMPS established an international committee of selected well-known experts in software engineering who are willing to be mentioned in the program and to review a set of papers each year. The list below comprises the Scientific Program Committee members.

Adriana Peña Pérez-Negrón	University of Guadalajara, Mexico
Alejandro Rodríguez González	Politechnical University of Madrid, Spain
Alejandra García Hernández	Autonomous University of Zacatecas, Mexico
Álvaro Rocha	Universidade de Lisboa, Portugal
Ángel M. García Pedrero	Politechnical University of Madrid, Spain
Antoni Lluis Mesquida Calafat	University of Islas Baleares, Spain
Antonio de Amescua Seco	University Carlos III of Madrid, Spain
Baltasar García Pérez	Schofield, University of Vigo, Spain
Benjamín Ojeda Magaña	University of Guadalajara, Mexico
Carlos Abraham Carballo Monsivais	CIMAT Unit Zacatecas, Mexico
Carla Pacheco	Technological University of Mixteca, Oaxaca, Mexico
Carlos Alberto Fernández y Fernández	Technological University of Mixteca, Oaxaca, Mexico
Cesar Guerra-García	Autonomous University of San Luis, Mexico
Claudio Meneses Villegas	Catholic University of North, Chile
Edgar Alan Calvillo Moreno	Technological University of Aguascalientes, Mexico
Edgar Oswaldo Díaz	INEGI, Mexico
Eduardo de la Cruz Gamez	Technological Institute of Acapulco, Mexico
Eleazar Aguirre Anaya	National Politechnical Institute, Mexico
Fernando Moreira	University of Portucalense, Portugal
Francisco Jesús Rey Losada	University of Vigo, Spain
Gabriel A. García Mireles	University of Sonora, Mexico
Giner Alor Hernández	Technological Institute of Orizaba, Mexico
Gloria Monica Martinez	Technological University of Torreón, Mexico
Gloria P. Gasca Hurtado	University of Medellin, Colombia
Gonzalo Cuevas Agustín	Politechnical University of Madrid, Spain
Gonzalo Luzardo	Higher Polytechnic School of Litoral, Ecuador

Gustavo Illescas	National University of Central Buenos Aires Province, Argentina
Héctor Cardona Reyes	CIMAT Unit Zacatecas, Mexico
Himer Ávila George	University of Guadalajara, Mexico
Hugo Arnoldo Mitre Hernández	CIMAT Unit Zacatecas, Mexico
Hugo O. Alejandrez-Sánchez	National Center for Research and Technological Development, CENIDET, Mexico
Iván García Pacheco	Technological University of Mixteca, Oaxaca, Mexico
Jezreel Mejía Miranda	CIMAT Unit Zacatecas, Mexico
Jorge Luis García Alcaraz	Autonomous University of Juárez City, Mexico
José Alberto Benítez Andrades	University of Lion, Spain
Jose A. Calvo-Manzano Villalón	Politechnical University of Madrid, Spain
José Antonio Cervantes Álvarez	University of Guadalajara, Mexico
José Antonio Orizaga Trejo	University of Guadalajara CUCEA, Mexico
José-Eder Guzmán-Mendoza	Polytechnic University of Aguascalientes, Mexico
Jose-Francisco Gazga-Portillo	Technological Institute of Acapulco, Mexico
José Guadalupe Arceo Olague	Autonomous University of Zacatecas, Mexico
José Luis David Bonilla	University of Guadalajara, Mexico
José Luis Hernández Hernández	Institute of Chilpancigo, Mexico
José Luis Sánchez Cervantes	Technological Institute of Orizaba, Mexico
Juan Francisco Peraza Garzón	Autonomous University of Sinaloa, Mexico
Juan Manuel Toloza	National University of Central Buenos Aires Province, Argentina
Lisbeth Rodriguez Mazahua	Technological University of Orizaba, Mexico
Lizbeth A. Hernández González	University of Veracruz, Mexico
Lohana Lema Moreta	University of the Holy Spirit, Ecuador
Luis Omar Colombo Mendoza	Technological Institute of Teziutlán, Mexico
Luz Sussy Bayona Ore	Autonomous University of Peru
Magdalena Arcilla Cobián	National Distance Education University, Spain
Manuel Pérez Cota	University of Vigo, Spain
María de León Sigg	Autonomous University of Zacatecas, Mexico
María del Pilar Salas Zárate	Technological Institute of Teziutlan, Mexico
Mario Andrés Paredes Valverde	University of Murcia, Spain
Mario Hernández Hernández	Technological Institute of Chilpancingo, Mexico

Mary Luz Sánchez-Gordón	Østfold University College, Norway
Miguel Ángel De la Torre Gómora	University of Guadalajara CUCEI, Mexico
Miguel Hidalgo-Reyes	Superior Technological Institute of Xalapa, Mexico
Mirna Muñoz Mata	CIMAT Unit Zacatecas, Mexico
Miriam Martínez Arroyo	Technological Institute of Acapulco, Mexico
Omar S. Gómez	Higher Polytechnic School of Chimborazo, Ecuador
Patricia María Henríquez Coronel	University Eloy, Alfaro de Manabi, Ecuador
Perla Velasco-Elizondo	Autonomous University of Zacatecas, Mexico
Ramiro Goncalves	University Tras-os Montes, Portugal
Raúl Aguilar Vera	Autonomous University of Yucatán, Mexico
Ricardo Colomo Palacios	Østfold University College, Norway
Roberto Solis Robles	Autonomous University of Zacatecas, Mexico
Santiago Matalonga	University of the West, Scotland
Sergio Galván Cruz	Autonomous University of Aguascalientes, Mexico
Sodel Vázquez Reyes	Autonomous University of Zacatecas, Mexico
Sonia López Ruiz	University of Guadalajara, Mexico
Stewart Santos Arce	University of Guadalajara, Mexico
Sulema Torres Ramos	University of Guadalajara, Mexico
Tomas San Feliu Gilabert	Politechnical University of Madrid, Spain
Ulises Juárez Martínez	Technological Institute of Orizaba, Mexico
Uziel Trujillo Colón	Technological Institute of Acapulco, Mexico
Uziel Jaramillo Ávila	CIMAT Unit Zacatecas, Mexico
Vianca Vega	Catholic University of North Chile, Chile
Víctor Saquicela	University of Cuenca, Ecuador
Viviana Y. Rosales Morales	University of Veracruz, Mexico
Yadira Quiñonez	Autonomous University of Sinaloa, Mexico
Yasmin Hernández	CENIDET, Mexico
Yilmaz Murat	Çankaya University, Turkey

Contents

Knowledge Management

Software Systems, Applications and Tools

Implementation of Sentiment Analysis in Chatbots in Spanish to Detect Signs of Mental Health Problems

Eduardo Aguilar Yáñez, Sodel Vazquez Reyes(✉), Juan F. Rivera Gómez, Perla Velasco Elizondo, Alejandro Mauricio Gonzalez, and Alejandra García Hernández

Autonomous University of Zacatecas, 98000 Zacatecas, Zac, Mexico
{37180547,vazquezs,jf_riverag,pvelasco,amgdark, alegarcia}@uaz.edu.mx

Abstract. The detection of mental health problems such as depression has become increasingly important, nowadays its increase due to various factors such as the appearance of social networks, a global health crisis that has forced people to isolate themselves in their homes, among others. Unfortunately, these problems are often not detected in time because many people do not have the confidence to express their problems to doctors, psychologists or relevant authorities, causing them to advance to a more critical stage of their condition. However, many people feel more confident sharing their thoughts through social media platforms or chatbots these days, which are a form of liberation where you can lessen the social pressure to express your thoughts through anonymity. That is why this presented study proposes a system with which possible signs of mental health can be detected. This is in a scenario of high school students (young people from 14 to 18 years old on average) through interaction with the chatbot of the high school of the Autonomous University of Zacatecas, and using conversation flows, the chatbot was deployed during a period where the implementation of sentiment analysis algorithm presented an accuracy of 0.86 showing optimistic data for the detection of signs of problems of mental health.

Keywords: Chatbot · Conversational bot · Sentiment analysis · Mental health problems

1 Introduction

Nowadays, due to the social isolation that is carried out preventively by the SARS-Cov2 pandemic virus, the number of cases of mental health problems has been increasing. A study of 1,210 inhabitants of 194 cities in China, post COVID, showed that 53.8% of the participants had moderate to strong psychological impact; 16.5% had moderate to strong depressive symptoms; 28.8% had moderate to strong anxiety symptoms; and 8.1% had moderate to strong stress levels [1].

J. Mejia et al. (Eds.): CIMPS 2022, LNNS 576, pp. 196–214, 2023.
https://doi.org/10.1007/978-3-031-20322-0_14

China, being the country in which the SARS-Cov2 virus pandemic originated, had an increase in mental health illnesses, which is also reflected worldwide, since according to the WHO (World Health Organization) it was calculated that the COVID-19 pandemic had caused a 27.6% increase (95% uncertainty interval (IU): 25.1–30.3) in cases of major depressive disorder (MDD) and an increase of 25, 6% (95% II: 23.2–28.0) of anxiety disorders (AD) cases worldwide in 2020 [2].

Taking into account the above statistics, we can deduce that there is clearly an increase in cases of anxiety, stress and depression or symptoms that are linked to these diseases. This is a consequence of various factors, such as social isolation, pressure and danger to medical personnel or civilians, changes in people's routine, among others.

Technology in the field of mental health has advanced significantly in recent years, for example, speech recognition, Natural Language Processing (NLP) and Artificial Intelligence (AI) have been used to support the detection of symptoms of psychiatric disorders [3]. Research has shown that artificial intelligence~~s~~ such as Alexa, Siri or the Google Assistant is often considered by people as a friend or family member [4].

Currently, the use of chatbots is becoming more and more normalized, because they allow them to automate time-consuming tasks such as the attention provided by trained personnel in different areas and companies. In addition, chatbots are used to maintain conversations with other people, this for research purposes or even as an alternative to combat loneliness and depression.

The problem with most chatbots, in the case of customer service or for the detection of certain patterns, is that very few of them offer a personalized experience for the user, offer a poor understanding of the context or a solution to problems with little complexity. They are simply programmed to understand certain sentences and respond with limited options, ignoring the context or simplifying the interaction to a limited set of questions.

On the other hand, as for conversation with chatbots, they present the problem that most are programmed to respond only with sentences that the algorithm has determined correspond to the message that the user has inserted, making the experience with the chatbot short, besides that many times the coherence of the conversation is lost or simply the response is not very adequate to what the user has expressed.

Tuva Lunde Smestad, in his master's thesis regarding improving the user experience of chatbot interfaces, mentions that most chatbots are not doing their task properly and result in faulty interfaces that fail to predict the simplest of questions. Furthermore, he mentions that chatbot interactions are unintelligent, unhelpful and ineffective [4]. This tells us that chatbots only satisfy certain basic needs of a conversation, but in a limited and not personalized way, creating the user a feeling of lack of understanding or "tact" on the part of the bot.

A negative scenario is that 57% of companies have implemented or are planning to implement a chatbot as part of the services they provide soon [5]. And if this problem persists, it could cause a loss of reliability and even revenue for companies that implement chatbots as customer service. Not to mention that it may also cause a loss of interest in this type of application, in addition to the fact that this completely takes it away from the goal of simulating a conversation with a real person.

In the area of health and psychology, this problem means that the objective of helping people to detect any disease or disorder is affected due to the simplicity of the chatbot conversation flow and a lack of processing of user responses. For this reason, this work aims to use sentiment analysis in a conversational chatbot, in the area of psychology, to detect signs of mental health problems.

Next in this article you will find a State of the art section (Sect. 2) where different papers are compared about where sentiment analysis is used in chatbots and its use in the area of mental health. Then there will be a Sect. 3 where the entire development process of the prototype and its architecture will be described, a Sect. 4 where the results obtained will be shown and a Sect. 5 where the conclusions and future work of the investigation will be found.

2 State of Art

Next, we describe what sentiment analysis is, its use in chatbots, improvements that can be made to chatbots, and how it has been used in the mental health context.

2.1 Chatbots

Chatbots or conversational bots are applications that use natural language processing and/or rules defined in a question-and-answer system [6]. These have the objective of simulating a conversation with a person through text. This type of conversation can have the objective of just chatting with someone or it can also be to provide support and help to customers in a company that practices e-commerce, among others.

Chatbots use natural language processing (NLP), which is the way a computer can interpret human language based on reasoning, learning, and understanding [7]. On the other hand, NLP is not only implementing an algorithm to respond to a text input or process and classify the received inputs, but it is also combining different algorithms and techniques to achieve a better understanding of the received inputs and the context they refer to by a computer.

2.2 Sentiment Analysis in the Detection of Depression

Sentiment analysis or emotion recognition is the study of extracting specific information such as opinions or sentiments from texts by using PLN [8].

But… What are the sentiments? Feelings are a state of mind which is given in relation to external inputs and these are produced when the emotion is processed in the brain and the person is aware of that emotion and the mood it produces [9].

In other words, emotions are reactions of our body based on the context that the person is living in, while feelings are a state of mind that is produced when a person presents an emotion. Although nowadays many uses them as synonyms.

Now, feelings have been classified in several different ways over time, but the most common is that they are divided into positive, negative and neutral; in addition, these three classifications are divided into more. For example, in the positive ones we find happiness, euphoria and love, while in the negative ones we find sadness, anger or fear and the neutral ones are the only way in which it refers to expressions without any emotion or feeling involved [9].

It is because of the above that, by means of sentiment analysis and the use of PLN, it is possible to extract from texts or audio the sentiments involved in them and also the type of opinion that a certain person has, all this by classifying the entries into positive or negative sentiment, even with a more specific classification.

A paper in which a way to analyze sentiments in a specific way to detect depression is proposed is the article "EmoCure" by Basantani et al. [10], which uses formulas available on github.com for the classification of depressive texts. These formulas are the following, in (1) it uses a for loop to go through each word in the text to be analyzed and will validate if a word is within the list of positive words by adding one unit to the variable sum or within the list of negative words by subtracting one unit to the variable sum In (2) it calculates the sentiment score by dividing the variable by the number of words in the analyzed text and finally in (3) based on the score obtained the text will receive a label by validations where 1 is a positive text, 0 is a neutral message and −1 is a negative message:

```
For each word in the text do
    if the word is in the positive list then
        sum = sum +1
    else
        if the word is in the negative list then
            sum = sum -1
        End if
    End if
End for
```
(1)

$$score = (\Sigma\ sum) / n \tag{2}$$

```
If score >= 0.2 then
    label = 1
else
    if -0.5 < score < 0.2 then
        label = 0
    else
        if score <= -0.5 then
            label = -1
        End if
    End if
End if
```
(3)

The formulas shown above show a way of classifying possible depressive texts and assigning a numerical value based on a list of words related to depression. These same formulas were used by Yeow B, et al. to make diagnoses of people with possible depression, comparing them with other implementations where he concludes that this is the one with the highest classification accuracy, with a value of 78.95% [11].

Another study that contrasts with the work carried out by Basantani et al. is the one carried out by Wang et al., where, using sentiment analysis, a method is proposed that uses vocabulary and rules created to calculate the inclination to depression of microblogs. In addition, a depression detection model is built based on the proposed method and 10 characteristics of depressed users derived from psychological research. Then 180 users and 3 types of classifiers were used to verify the model, whose accuracy is around 80%. For later in this work, this model was used to implement it in a web application [12].

The papers [10] and [12] show two different ways of analysis where, in [10] the use of a list of words used in the proposed function is described, while in [12] a deeper linguistic analysis is shown where not only a vocabulary is defined but also rules which should be followed to classify texts as depressive or not. And on the other hand, in the study carried out by Hassan et al., an analysis of how to find the pressure level is presented based on the implementation of the algorithms of Support vector Machine (SVM), Naïve Bayes (NB) and Maximum Entropy (ME), which all presented a classification accuracy greater than 80% [13] as well as the function proposed in the article by Basantani et al. [10], only that in [13] was implemented in different social networks and with a preprocessing based on tokenization, where the SVM presents the best results with an accuracy of 84% but with an accuracy of 91% with the test data.

2.3 Implementations of Sentiment Analysis in Chatbots

Currently, chatbots and sentiment analysis are implemented in different ways. As mentioned above, these can go hand in hand to provide a more pleasant and empathetic conversation for users.

These are achieved by analyzing user responses and, based on them, classifying them and responding according to their feelings. One of the first where a sentiment analysis application was clearly observed was with Parry, a chatbot developed by Keneth Mark Colby. This chatbot was designed in 1975 to behave like a paranoid person based on the rules and structure of Eliza, but with more language understanding abilities and emotion detection, where if the anger level of a response was high, the bot responded with hostility [14].

Currently, sentiment analysis has been applied to different chatbots or devices, such as Alexa and Google Assistant, both of which have an SDK open to anyone who wants to use it for personal projects [15].

There has also been a great interest in developing chatbots and implementing sentiment analysis in the government sector. A clear example of this has been the chatbot (CADELA) developed by the Spanish National Statistics Institute (SNSI) in 2019, which is a bot in charge of conversing with elderly people to prevent depression by classifying the answers as good, bad, or neutral sentiment [16].

All the above shows that sentiment analysis has been applied in several chatbots and for different objectives, such as providing a better interface, analyzing messages, or giving answers more in line with the inputs received by the chatbot.

2.4 Improving Chatbots Through Sentiment Analysis

As described above, the joint use of chatbots and sentiment analysis brings several improvements in the interaction with them, such as: improving the accuracy of chatbots' responses, managing the emotions of a group of people, redirecting the conversation based on feelings, etc.

A clear example of this is presented by C. Antony et al., which shows how to integrate emotion recognition in chatbots to make them more empathetic by responding based on the feelings detected in the user's messages. They mention three important stages to respond appropriately to the detected sentiment: preprocessing, sentiment recognition and response generation. In the preprocessing stage, normalization, cleaning and tokenization of the sentences sent by the user is performed, where tokenization is the key part of this stage as it improves the machine learning rate and inferences to a great extent [17].

Then follows the sentiment recognition stage where the identified tokens are embedded into binary vectors containing the meaning corresponding to each word. Then, these word vectors are combined and form a matrix representation of a sentence. Attention mechanisms are applied to filter out crucial information, thus compacting the matrix and preserving contextual data [17]. This process is similar to the one mentioned in the article "Combating depression in students using an intelligent chatbot" where tokenization is also used to parse sentences [18].

The article explains that the classification of the sentences will depend on the model of the chatbot, whether it is rule-based, PLN (automated) or a hybrid. This is because if it is rule-based, it will submit the classification to a decision tree, if it is automated, it will be based on a machine learning algorithm and if it is hybrid, it will combine rules and machine learning.

In the response generation phase there are two paths: the most common is the retrieval model, where there is a predefined set of queries and responses and the response that best fits the context and takes into account the emotion that is chosen. While the other model, called generation, creates responses from scratch, it is difficult to develop bots using this model, as it requires a massive human-human interaction dataset with emotional variants and a large amount of training time to achieve contextually accurate responses [17].

Now, the responses sent by the chatbot can not only be text, but can go beyond that. For a bot's response to correspond with people's feelings, the response should also be reflected in the bot's image by humanizing it, as a "human appearance can foster natural human interaction with chatbots" [5]. In Tuva Lunde Smestad's thesis, she created a chatbot named Bella, to which she gave personality, causing users to have a better experience when interacting with it, since a high level of humanity in bots provides a more natural interaction with humans.

This is why sentiment-based responses are not only text but, for example, GIFs, emojis and images. In the article "Chatbot-based Emotion Management for Distributed Teams: A Participatory Design Study", the chatbot uses these elements to send messages about the situation of a working team and also try to reverse this emotional situation in case there is some anger or displeasure in the team.

Another answer could be audio, Sofia Pizarro mentions that the chatbot she developed recommends music through the Spotify API so that the user has a more pleasant and relaxing experience when interacting with it [15].

The use of sentiment analysis brings an improvement in the accuracy of chatbot responses as long as the right algorithms, models and architecture are chosen, and also improves not only the accuracy but also the way users perceive the bots, because with the right responses, the bots appear more empathetic and with a more human attitude.

2.5 Sentiment Analysis in Chatbots to Detect Mental Health Problems

The detection of mental disorders such as anxiety or depression has been sought to be detected through chatbots. An example is that proposed by Ana Chieng Cueva where attention is provided via the Internet in a timely and automated manner to a user concerned about their mental health who wants to mitigate, control or mitigate any level of anxiety or depression disorder that may affect their well-being and/or current or future behavior in society [19].

This chatbot performs an assessment on the user through a series of questions, with limited options, which help define whether the user has any symptoms of depression, anxiety, or other mental disorder. This tells that, through the closed flow of a chatbot, it can evaluate symptoms of disorders and suggest actions such as advice or schedule appointments with psychologists to prevent the disorder from progressing.

Another proposal to use chatbot for mental health, is using deep learning and PLN not only to provide conversational assistance in the form of a friendly chatbot, but also to provide a toolbox of useful features such as (a) SRQ (Self Reporting Questionnaire) assessment scale, (b) K-10 KESSLER assessment scale and (c)PHQ9 (Patient Health Questionnaire) assessment scale to maintain mental health [20]. This chatbot aims to make the user not feel uncomfortable, with comforting conversations, friendly responses and making the chatbot non-judgmental to the user, thus allowing a trusting environment.

The proposal by [19] presents an analysis through a questionnaire that prompts users to answer specific questions and thus facilitates the application of sentiment analysis for the detection of any mental health problem, but the proposal by [20] uses NLP and deep learning so that through the messages received by the chatbot, not only analyzing the message, but also give depth to the analysis of feelings carried out.

Another similar paper is that of Ouerhani et al. where the construction of a chatbot called COVID-chatbot is mentioned, which through the use of an initial questionnaire and the use of a Depression Detector model (DDM) that detects anxiety in text input through a sentiment analysis model of deep inclination to help make the decision to send a reassuring message if a bad behavior is detected [21], this makes the chatbot send messages according to the situation but everything originates from the initial questionnaire.

The interesting thing about the papers [19, 20], and [21], is that the conversation flow is initiated by a questionnaire that delimits the user's responses to increase the probability that the user sends messages, which can be analyzed and applied by classification algorithms, NLP and sentiment analysis to detect problems or a level of mental health or illness related to this field. That is why this will be taken as a basis to implement it in the prototype of this investigation.

3 Chatbot Prepas UAZ

It is for this reason that this work implements sentiment analysis to detect possible signs of mental health in high school students of the Autonomous University of Zacatecas by processing the messages generated through the interaction of the chatbot.

Based on the implementation described by Ana Chieng Cueva [19] of a chatbot to detect symptoms of mental health diseases, it has been decided to develop a prototype whose purpose is to detect signs of mental health problems through sentiment analysis and with emphasis on depression using a classification algorithm based on word lists as proposed by Basantani et al. [10].

3.1 Current Status of the Chatbot

The Autonomous University of Zacatecas has 13 high schools throughout the state and, due to the need of those interested in obtaining information, at the beginning of 2021, Chatbot Prepas UAZ has been implemented in the WhatsApp application to answer the most frequently asked questions regarding students, teachers and procedures to be carried out.

The Chatbot Prepas UAZ, is a bot made through the WhatsAuto application. The flow of the chatbot is simple, it starts the conversation with a greeting, the chatbot responds by showing the main menu with five options, which are: Login, Student, Applicant, Teacher and Contact.

These options show a default message, with the suggestion to get more information by requesting another specific entry, but also the bot lets you know that you can return to the main menu with the word “Home”. The above described, shows the need to implement improvements.

3.2 Chatbot Architecture

To implement the new functions concerning sentiment analysis, the following two modules were identified:

Chatbot Module
This module is composed of the graphical interface of the chatbot and the answers assigned to the user’s responses. This module is in charge of allowing the user to interact with the system.

Sentiment Analysis Module
This module is in charge of processing user messages, implementing sentiment analysis for their classification and, based on this, to be able to detect any indication of a problem in the mental health of users.

In Fig. 1 you can see a diagram of components which shows all that makes up the chatbot, where you can also see that through WhatsApp messages are received from users and these pass through the WhatsAuto application. Then a request is sent to Dialogflow and these requests are sent to the server and the server sends the personal data of the users and both the sentiment score determined by Dialoglow and the depression index determined by the classification algorithm based on a list of words in Spanish obtained in the requests received to an email handler to subsequently send emails with sentiment analysis reports to the administrators of the psychology area of the university.

These sentiment analysis reports will contain the user’s information provided in the interactions, the sentiment score determined by Dialogflow accompanied by a traffic light based on the score obtained and the calculated depression index.

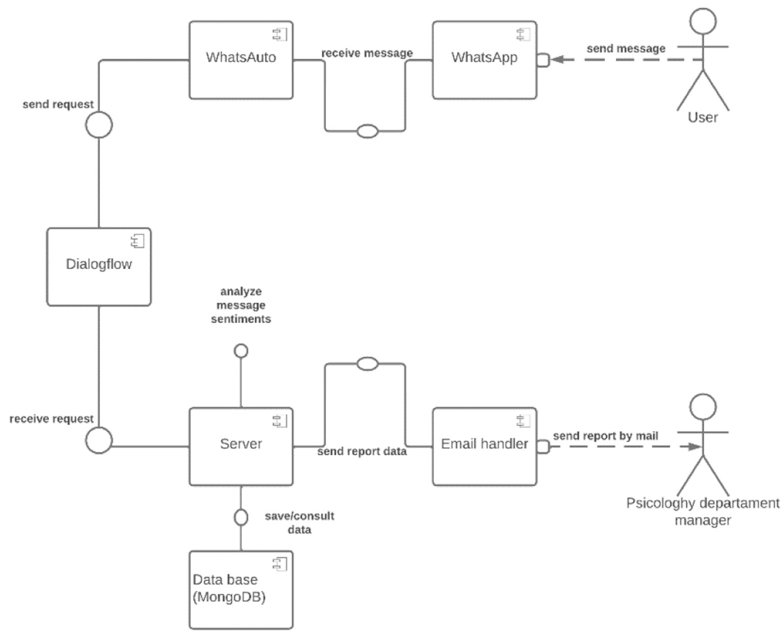


Fig. 1. Prepas UAZ Chatbot component diagram

The calculation of the depression index will be calculated using the functions mentioned by Basantani et al. [10], where the message will be received by the function and this will take care of dividing it into words. Each word will be compared with the list of depressive words and if a word is found within this list, one will be added to the counter and at the end the score will be obtained by dividing the number of words in the analyzed text.

The list of words was built by collecting words from different websites that talk specifically about words related to depression, these sites are [22, 23] and [24].

In the sentiment analysis, a corpus in Spanish was not used, since now such a deep linguistic analysis is not needed and the Dialogflow classification algorithm is being used, but it is planned to include this in the future. But as for the depression classification algorithm, it will use a list of words in Spanish.

3.3 Functions Implemented to the Chatbot

To promote the interaction of students with the chatbot, some functions were proposed so that the users can interact with the chatbot in a continuous way and can express how they feel in a textual and numerical way. The four functions implemented to achieve this objective are as follows:

A. *Initial Question*

The function of the initial question of the chatbot, is that the chatbot when receiving a greeting from the user, will answer asking how he/she feels at that moment and also asks how he/she feels on a scale of 1 to 5 (where 1 is very bad and 5 is very good).

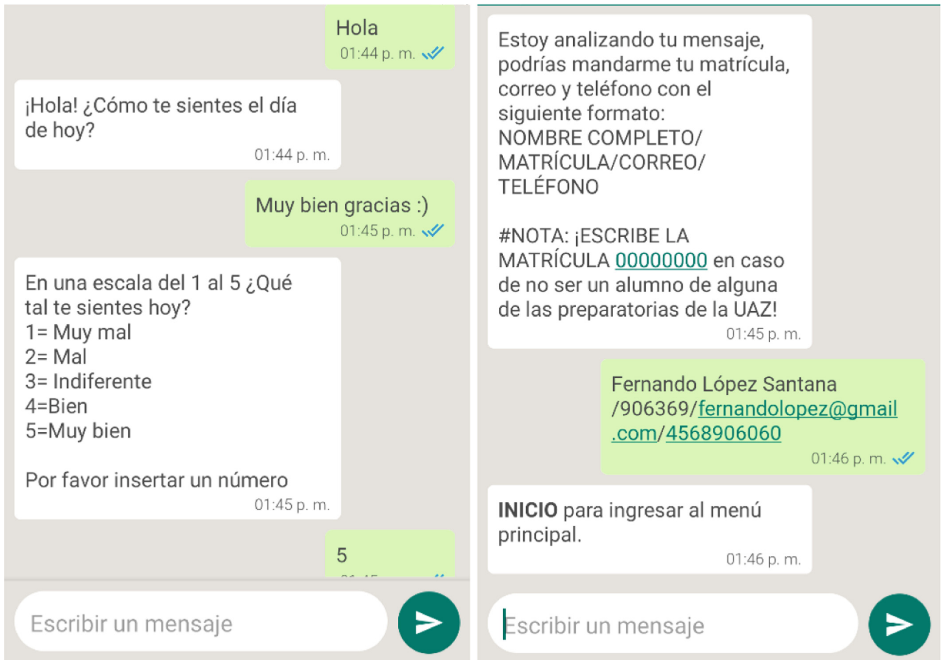


Fig. 2. Conversation flow of the initial question of the Prepas UAZ chatbot.

Figure 2 shows how interactions are carried out on WhatsApp, interactions which will serve to analyze the texts sent by users and by allowing students to express themselves to know if they have some kind of negative feeling or mental health problem.

B. *Complaints mailbox*

The complaint mailbox, asks what your complaint is and also asks if you also feel on the same scale above from 1 to 5. This is in order to see if the complaint also involved some mental health problem or just some negative comment.

The flow of this functionality can be seen in Fig. 3 and Fig. 4, where you can see how the flow develops in the WhatsApp interface.

C. *Mentor assignment*

This functionality asks the user how he/she feels on the mentioned scale from 1 to 5, and then, through a series of questions, will obtain the campus, semester to which he/she belongs, the subject in which he/she needs mentoring, the mentor of his/her choice and at the end he/she will be asked what problem he/she has with the subject, as well as other contact information.

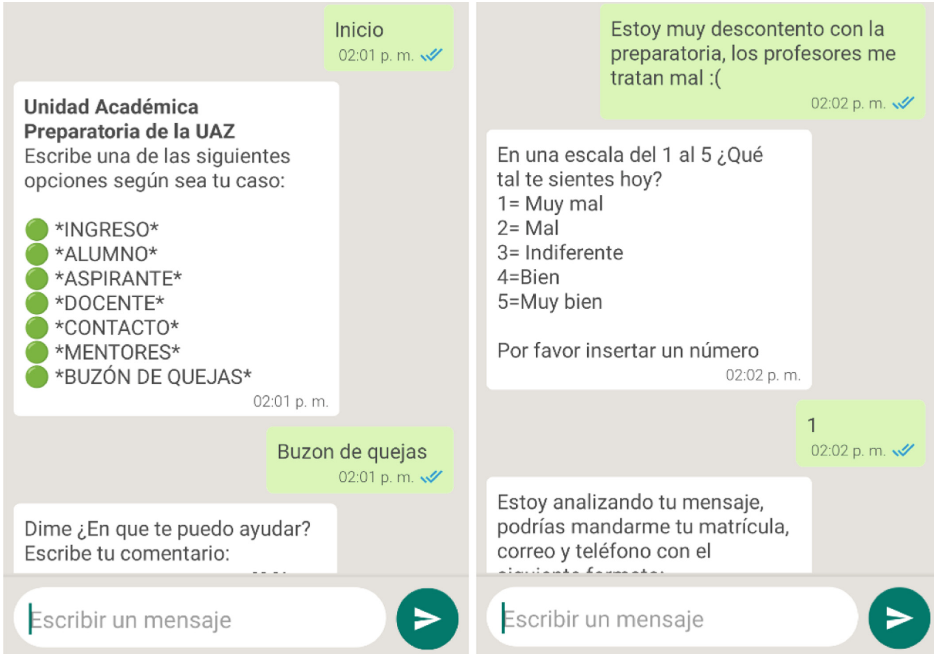


Fig. 3. Conversation flow of the Prepas UAZ chatbot complaint mailbox.

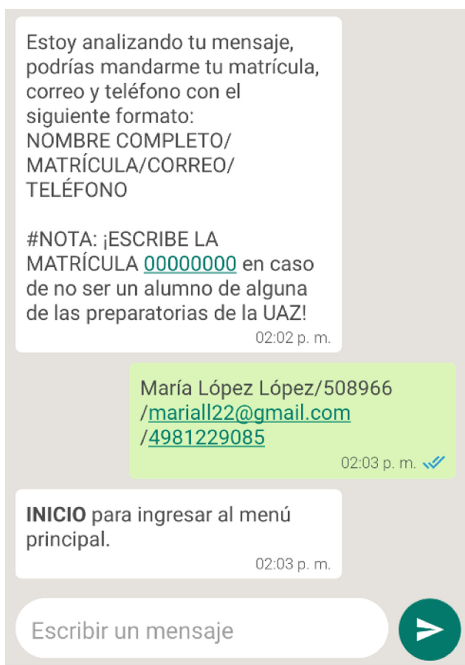


Fig. 4. Conversation flow of the Prepas UAZ chatbot complaint mailbox 2.

This flow, which can be seen in Fig. 5, Fig. 6, Fig. 7, and Fig. 8, will allow us to analyze the text where the student's problem with the subject is expressed and thus verify if the student has a possible mental health problem due to his comments.

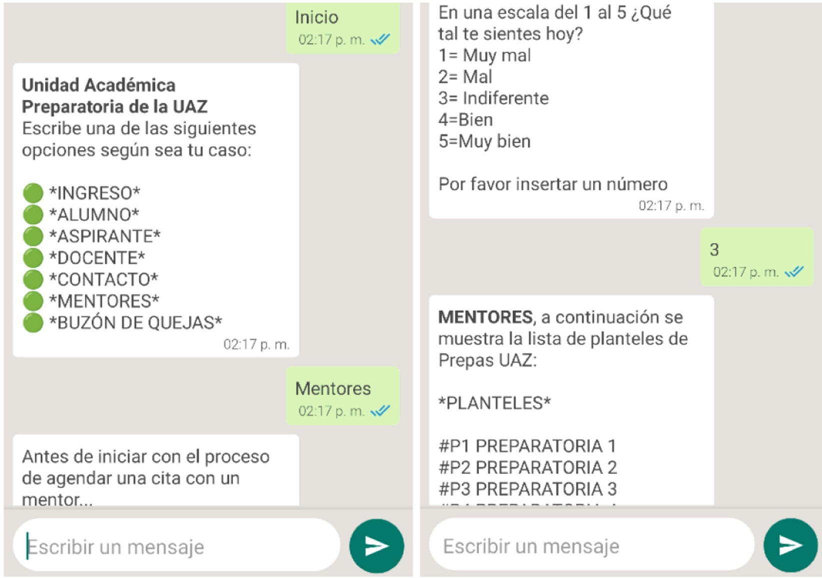


Fig. 5. Conversation flow of the Prepas UAZ chatbot mentoring assignment.

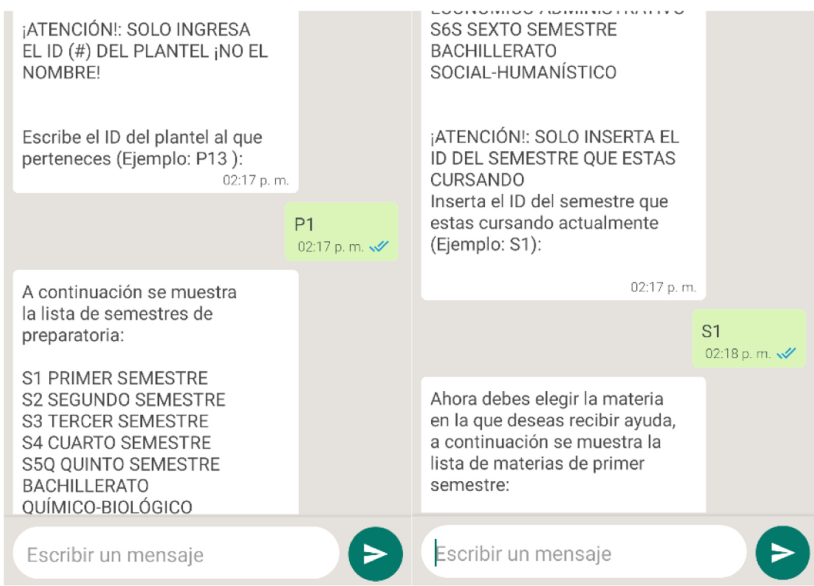


Fig. 6. Conversation flow of the Prepas UAZ chatbot mentoring assignment 2.

D. *End-of-semester control question*

The end-of-semester control question function asks how you felt at the end of the semester and also asks the value of your feelings on the scale mentioned before.

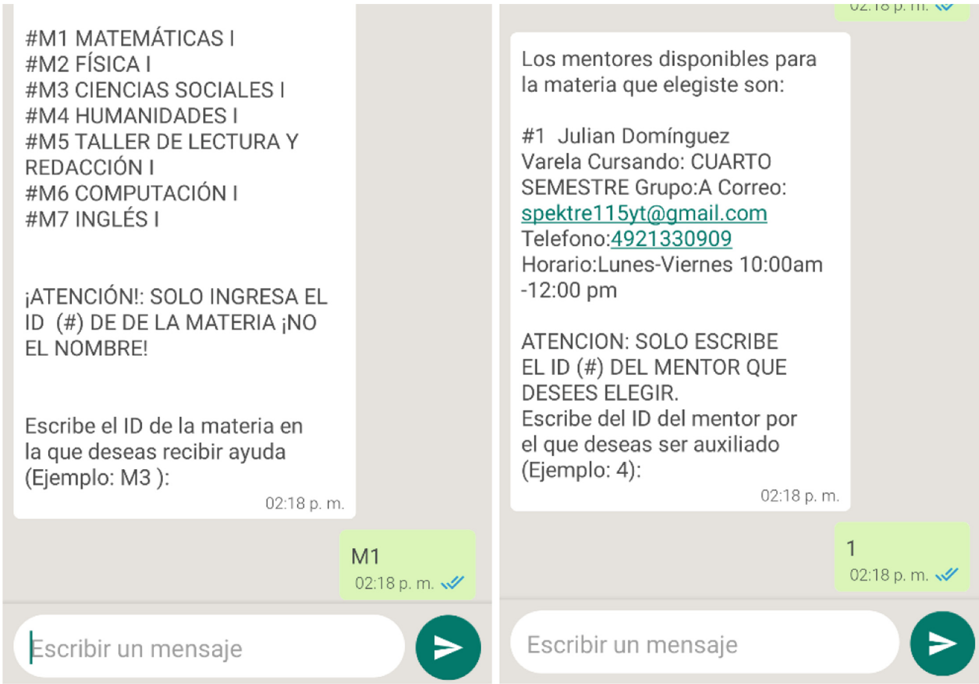


Fig. 7. Conversation flow of the Prepas UAZ chatbot mentoring assignment 3.

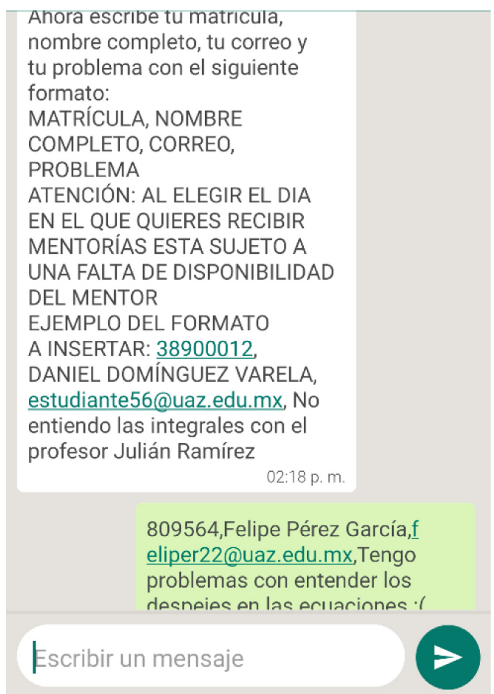


Fig. 8. Conversation flow of the Prepas UAZ chatbot mentoring assignment 4.

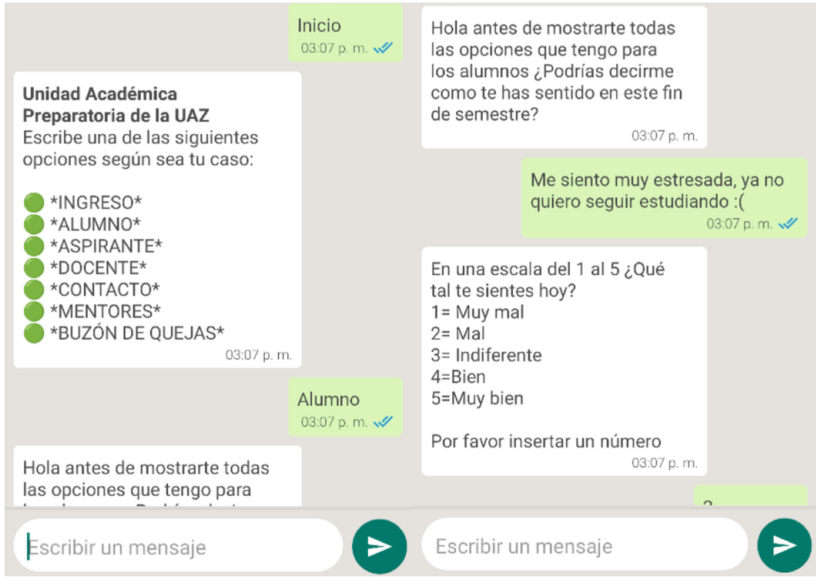


Fig. 9. Conversation flow of the end-of-semester control question of the Prepas UAZ chatbot

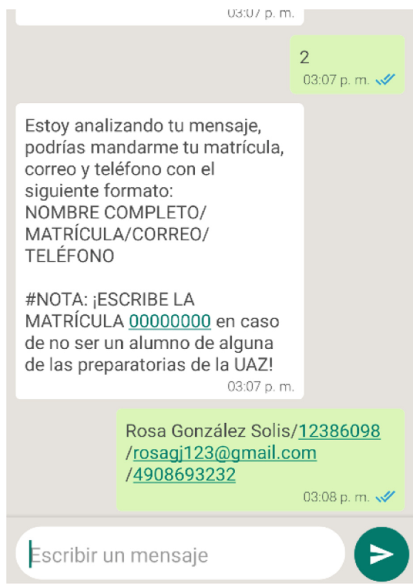


Fig. 10. Conversation flow of the end-of-semester control question of the Prepas UAZ chatbot 2

The flow shown in this functionality, which can be seen in Fig. 9 and Fig. 10, will allow collecting information about the students' feelings at the end of the semester, encouraging interaction to express their feelings and thus detect mental health problems.

3.4 Sentiment Analysis, Depression Index and Message Classification

All the functions mentioned in Sect. 3.3, its intention is that the chatbot receives a message in order to classify it through an urgency of attention traffic light based on the extraction of the sentiment score and based on this send a mental health report to the high school psychologists in order to provide timely attention to students in need.

Dialog flow is in charge of evaluating each message received by the conversational agent and gives a score through its classification algorithm on a scale of −1 to 1 for the labeling of these messages as positive (corresponds to a positive score) or negative (corresponds to a negative score).

Table 1. Sentiment scoring system.

Sentiment score range	Signal (color)	Attention urgency
−0.7 to −1	**Red**	High
−0.1 to −0.6	**Yellow**	Medium
0 to 1	**Green**	Low

Table 1 shows the signals used with respect to the sentiment score determined by Dialogflow. If a message has a score between −0.7 and −1, it will be marked with red, indicating a high urgency of attention to the student, from −0.6 to −0.1 it will be marked with yellow, indicating a medium urgency of attention to the student, and green is indicated when the score is between 0 and 1, representing a low urgency of attention to the student.

With this signal system, it is expected that when used in health reports, high school psychologists will focus on the most urgent cases to prevent or detect mental health problems in students.

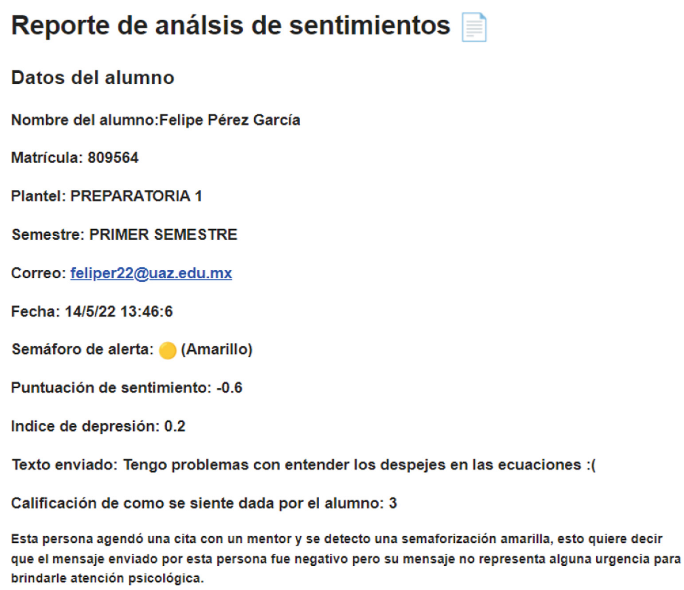
In addition to the signal and the sentiment score, another data to be provided will be the depression index based on formulas (1) and (2) of Sect. 2.2, to compare this index with the sentiment score, so that psychologists can have an additional reference to rule out or detect depression in students who present a negative sentiment score. This index is calculated based on the average number of words matching a list of depression-related words collected from different web pages [22–24].

3.5 Mental Health Reports

The mental health reports that will be sent to the psychology area of the UAZ high schools contain the contact information provided by the student, such as name, enrollment, phone number and email. In addition, in the case of the mentor assignment function, data such as semester and campus will also be sent.

Additionally, the report includes the texts sent by the student where he/she expresses his/her problem with the subject, his/her complaint or how he/she feels at that moment, depending on the function with which the user has interacted.

As can be seen in Fig. 11, the reports will show the sentiment score, the related traffic lights, a description of the results in terms of their traffic lights and the depression index, data necessary for the psychologists who will be in charge of providing care to the students who need it based on the data contained in these reports.



Reporte de análsis de sentimientos

Datos del alumno

Nombre del alumno:Felipe Pérez García

Matrícula: 809564

Plantel: PREPARATORIA 1

Semestre: PRIMER SEMESTRE

Correo: feliper22@uaz.edu.mx

Fecha: 14/5/22 13:46:6

Semáforo de alerta: (Amarillo)

Puntuación de sentimiento: -0.6

Indice de depresión: 0.2

Texto enviado: Tengo problemas con entender los despejes en las ecuaciones :(

Calificación de como se siente dada por el alumno: 3

Esta persona agendó una cita con un mentor y se detecto una semaforización amarilla, esto quiere decir que el mensaje enviado por esta persona fue negativo pero su mensaje no representa alguna urgencia para brindarle atención psicológica.

Fig. 11. Sentiment analysis report sent by mail by the Prepas UAZ chatbot.

4 Results

The chatbot was deployed for a week so that high school students of the Autonomous University of Zacatecas could interact with it and, through the functionality of the end of semester control question, 73 interactions were obtained, with which by comparing the sentiment score of the Dialogflow algorithm and the users' response based on a numerical scale from 1 to 5 to describe how they feel (where 1, 2 and 3 is negative and 4 and 5 are positive) the following confusion matrix was obtained, shown in Table 2.

Table 2. Confusion matrix of chatbot results.

	True	False
Positive	57	9
Negative	4	3

As can be seen, a total of 57 true positive data were obtained, which indicates that 57 positive messages were correctly classified and that these represent 78.08% of the interactions received by the chatbot. So, it says that most of the messages sent by the students were positive, since in total, there are a total of 66 messages (90.41% of the messages) that should have been classified as positive by the algorithm and only 7 (9.59% of the messages) should have been classified as negative.

Table 3 shows the classification metrics, an accuracy of 0.86, which indicates that most of the positive messages were classified correctly, while the accuracy was 0.84, which means that 84% of the messages sent were classified.

Table 3. Metrics of the Dialogflow classification algorithm.

Metric	Formula	Results
Precision	TP/(TP + FP)	0.86
Accuracy	(TP + TN)/(TP + TN + FP + FN)	0.84
Recall	TP/(TP + FN)	0.95
F-Score	TP/(TP + ((FP + FN)/2))	0.90

As for the completeness, a value of 0.95 was obtained, which represents that 95% of the relevant data was successfully classified, and on the other hand, an F-score of 0.90 was obtained, which confirms that there is precision and completeness between 90%.

As for the depression index calculated on all interactions with the chatbot of all messages sent, they received a negative depression index which indicates that, based on the word list created for this prototype, it was determined that no message contained enough words related to depression to think that a student had any indication of depression. This indicates that the word list should be enriched to cover a larger number of phrases and its evaluation should be broader and this explains why the implementation of the classification algorithm increased from 0.78 to 0.86 in precision with respect to the implementation of Basantani et al. [10].

5 Conclusion and Future Work

In conclusion, we can say that the classification algorithm implemented in the chatbot achieved high accuracy in classifying the messages received, although it would be necessary to analyze a greater number of negative messages to see the behavior of this algorithm. And in the case of the depression index, it can be mentioned that the list of words with which the classification algorithm works to obtain this index should be enriched with more words, since no index obtained could be greater than 0.

These aspects mentioned above will be improved to improve the sentiment analysis reports that are defined by the chatbot.

Greater accuracy in the calculation of the depression index could help to have greater certainty in the information sent in the sentiment analysis reports and this could be used in educational institutions or in the workplace to detect signs of depression or any other mental health problem, and provide timely psychological help to those people who need it.

Finally, when verifying that the prototype has obtained promising results, we will be working with mental health specialists to evaluate the results manually to give greater depth to the chatbot classification.

Acknowledgements. We are grateful for the support provided to High School No. 1 of the Autonomous University of Zacatecas for the evaluation of the chatbot.

References

1. Ramírez-Ortiz, J., Castro-Quintero, D., Lerma-Córdoba, C., Yela-Ceballos, F., Escobar-Córdoba, F.: Mental health consequences of the COVID-19 pandemic associated with social isolation. Colomb. J. Anesthesiol. **48**(4) (2020). https://doi.org/10.5554/22562087.e930
2. Organización Mundial de la Salud (OMS): Salud mental y la pandemia de Covid-19. N. Engl. J. Med. **383**(6), 510–512 (2020). https://doi.org/10.1056/NEJMp2008017
3. Edward, J., Loboa, C., de Cali, U.S., de Ingeniería, F., de Maestría, P.: Chatbots para apoyar la detección de síntomas de trastornos psiquiátricos, pp. 1–11 (2019)
4. Skjuve, M., Følstad, A., Fostervold, K.I., Brandtzaeg, P.B.: My chatbot companion - a study of human-chatbot relationships. Int. J. Hum. Comput. Stud. **149**(December), 2021 (2020). https://doi.org/10.1016/j.ijhcs.2021.102601
5. Smestad, T.L.: Personality matters! Improving the user experience of chatbot interfaces, Mater Thesis. Norwegian University of Science and Technology (NTNU) (2018)
6. Rodríguez, J.M., Merlino, H., Fernández, E.: Comportamiento adaptable de chatbots dependiente del contexto. Revista Latinoamericana de Ingenieria de Software **2**(2), 115 (2014). https://doi.org/10.18294/relais.2014.115-136
7. Mendez, D.C.: Historia de la Inteligencia artificial relacionada con los Chatbots (2017). https://planetachatbot.com, https://planetachatbot.com/historia-de-la-inteligencia-artificial-relacionada-con-los-chatbots-41a6cda22906. Accessed 06 Mar. 2020
8. Alissa, M., et al.: Sentiment Analysis for Open Domain Conversational Agent, no. 2016 (2021). http://arxiv.org/abs/2101.00675
9. Thomen Bastdardas, M.: Qué son los sentimientos: lista, tipos y ejemplos (2019). https://www.psicologia-online.com/
10. Basantani, A., Kesarwani, Y., Bhatia, S., Jain, S.: EmoCure: utilising social media data and smartphones to predict and cure depression. IOP Conference Series: Materials Science and Engineering **1110**(1), 012010 (2021). https://doi.org/10.1088/1757-899x/1110/1/012010
11. Yeow, B.Z., Chua, H.N.: A depression diagnostic system using lexicon-based text sentiment analysis. Int. J. Percept. Cogn. Comput. (IJPCC) **8**(1), 29–39 (2022). https://journals.iium.edu.my/kict/index.php/IJPCC/article/view/250%0A, https://journals.iium.edu.my/kict/index.php/IJPCC/article/download/250/174
12. Wang, X., Zhang, C., Ji, Y., Sun, L., Wu, L., Bao, Z.: A depression detection model based on sentiment analysis in micro-blog social network. In: Li, J., Cao, L., Wang, C., Tan, K.C., Liu, B., Pei, J., Tseng, V.S. (eds.) PAKDD 2013. LNCS (LNAI), vol. 7867, pp. 201–213. Springer, Heidelberg (2013). https://doi.org/10.1007/978-3-642-40319-4_18
13. Hassan, A.U., Hussain, J., Hussain, M., Sadiq, M., Lee, S.: Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression. In: International Conference on Information and Communication Technology Convergence: ICT Convergence Technologies Leading the Fourth Industrial Revolution, ICTC 2017, vol. 2017-Decem, pp. 138–140 (2017). https://doi.org/10.1109/ICTC.2017.8190959
14. Yeung Shum, H., Dong He, X., Li, D.: From Eliza to XiaoIce: challenges and opportunities with social chatbots. Front. Inform. Technol. Electron. Eng. **19**(1), 10–26 (2018). https://doi.org/10.1631/FITEE.1700826
15. Pizarro, S.: Bot conversacional detector de sentimientos (2019). http://riull.ull.es/xmlui/handle/915/13468. Accessed: 05 March 2021
16. Regueiro Janeiro, J.Á., de Arriba Pérez, F., García Méndez, S., González Castaño, F.J.: Análisis de sentimiento aplicado a un chatbot conversacional. In: IV Machine Learning Workshop Galicia 2019, p. 77. Vigo, España (2019 Oct.). https://wgml.es/files/actas_wgml2019.pdf#page=46

17. Antony, C., Pariyath, B., Safar, S., Sahil, A., Nair, A.R.: Emotion recognition-based mental healthcare chat-bots: a survey. SSRN Electronic Journal no. Icicnis, 69–80 (2021). https://doi.org/10.2139/ssrn.3768304
18. Patel, F., Thakore, R., Nandwani, I., Bharti, S.K.: Combating depression in students using an intelligent ChatBot: A cognitive behavioral therapy. In: 2019 IEEE 16th India Council International Conference, INDICON 2019 - Symposium Proceedings, pp. 1–4 (2019). https://doi.org/10.1109/INDICON47234.2019.9030346
19. Chieng-Cueva, A.-I.: Evaluación De Trastornos Mentales De Ansiedad Y Depresión Vía Chatbot. Repositorio Institucional - Ulima (2020). http://repositorio.ulima.edu.pe/handle/20.500.12724/13253
20. Tewari, A., Chhabria, A., Khalsa, A.S., Chaudhary, S., Kanal, H.: A Survey of Mental Health Chatbots using NLP. SSRN Electron. J., 1–6 (2021). https://doi.org/10.2139/ssrn.3833914
21. Ouerhani, N., Maalel, A., Ben Ghézala, H., Chouri, S.: Smart Ubiquitous Chatbot for COVID-19 Assistance with Deep learning Sentiment Analysis Model during and after quarantine, pp. 1–9 (2020)
22. Bhandari, S.: Glossary of depression terms (2020). https://www.webmd.com, https://www.webmd.com/depression/depression-glossary-terms. Accessed 14 May 2022
23. Marks, J.L.: Speaking Depression: A Glossary of Terms Used to Describe the Disorder (2020). https://www.everydayhealth.com, https://www.everydayhealth.com/depression/speaking-depression-a-glossary-of-terms-used-to-describe-the-disorder/. Accessed 14 May 2022
24. Eichstaedt, J.C., et al.: Facebook language predicts depression in medical records. Proc. Natl. Acad. Sci. USA **115**(44), 11203–11208 (2018). https://doi.org/10.1073/pnas.1802331115