Model to Assess the Level of Depression by Analyzing Facial Images and Voice of Patients

Alexander Ramos-Cuadros, Luis Palomino-Santillan and Willy Ugarte[a]

Universidad Peruana de Ciencias Aplicadas (UPC), Lima, Peru

Keywords: Depression, Facial Detection, Audio Analysis.

Abstract: Depression is considered as a common mental disorder, which is present in people of all ages causing a negative impact on different aspects of life such as mood, vitality, and interests in the enjoyment of activities, making them impossible in the long term, and in the most chronic cases can lead to suicide. Giving rise to the opportunity for collaboration between mental health specialists and the use of technological tools to support the evaluation of the level of depression to provide an optimal clinical diagnosis of the patient and an adequate referral to start treatment. In Peru, the COVID-19 epidemic has reduced physical contact and accessibility to health professionals in a timely manner, causing the patient's mental health to not be recognized or treated properly, which leads to the chronicity of the disease, to the psychological suffering, and the high costs that are required for special care. Thus, one of the challenges of this research is to implement a technological model that evaluates levels of recurrent depression by analyzing facial images and voice to detect the chronicity of depressive symptoms in young Peruvians. Our results show that in a simulated scenario, young patients were disposed to execute a self-administered questionnaire for depression having an optimal perception of satisfaction and usability on the mobile application based on the functionalities of the model.

1 INTRODUCTION

According to the World Health Organization (WHO)[1], it was estimated that depression was one of the most common mental disorders that affected around 264 million people of all ages, being one of the main causes of disability in the world. Depression is characterized by affecting the mood of the sufferer, which is why it is also known as a mood disorder or affective disorder which causes suffering and disability in family, work and social environments, and this can be classified between the levels: mild, moderate and severe, depending on the amount and severity of symptoms presented[23]. For the most serious cases of depression, this disease can lead to suicide, and it is estimated that about 800 thousand people conclude this act, considering it the fourth leading cause of death in people aged 15 to 29 years[3].

Currently, it is estimated that in Peru, 80% of suicides are related to severe depression, and the Peruvian entities in charge of treating mental health are not sufficiently of people who suffer from the disorder[4]. Either, due to the limited number of mental health specialists available to attention to these cases or technological deficiencies to provide efficient health services, which have been important barriers to accessing mental health services during the COVID-19 pandemic, which has led to a deterioration in mental health and the vital functions of people who already suffered from a mental disorder previously deteriorate due to the multiple factors experienced during the state of health emergency[5]. The percentage of cases that received some treatment per year from 2014 to 2018 is approximately 14%, the remaining 86% do not usually receive some type of treatment for depression symptoms (Villarreal-Zegarra et al., 2020). According to the National Institute of Mental Health (INSM)[6], it is reported that among the population with this disability, young people between the ages of 17 and 25 in Peru suffer the most from this dis-

[a] https://orcid.org/0000-0002-7510-618X

[1] "Adolescent mental health" - WHO

[2] Depression - National Ministry of Health (in Spanish)

[3] "Depression" - WHO

[4] "Severe depression is the principal cause of death by suicide" - MINSA (2019)

[5] "COVID-19 and the need of act in relation with mental health" - UN (2020)

[6] Ending stigma towards people with mental health problems, the challenge of psychiatry

Ramos-Cuadros, A., Palomino-Santillan, L. and Ugarte, W.
Model to Assess the Level of Depression by Analyzing Facial Images and Voice of Patients.
DOI: 10.5220/0011034500003188
In *Proceedings of the 8th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE 2022)*, pages 26-36
ISBN: 978-989-758-566-1; ISSN: 2184-4984

order, seeing a considerable increase in mental health problems in Peru's children and adolescents, so it became a priority to address these cases in specialized centers in order to provide the corresponding mental health services. However, the care gaps continued to be high, considering that not only the prescription for a medical drug was enough but also the support of mental health professionals for comprehensive recovery and reintegration into society.

Artificial intelligence (AI) and cloud services are progressively focused on supporting the health sector, which are used to make better decisions based on the large amount of data that they analyze. For example, in (Graham et al., 2019) point out that AI generates benefits in terms of the detection and diagnosis of mental disorders due to its algorithms and the ability to extract information from a data source, which provide a better understanding of the prevalence of these disorders in the population allowing health professionals to focus mainly on the human aspects of medicine and doctor-patient treatment while AI would focus on cases of self-administered smart health treatments that improve the limited time of patient care. The mental health sector can obtain various benefits from this technology, as well as how it benefited in the digitization process to improve patient care, now AI must be used to make a more efficient, accurate and personalized diagnosis or treatment selection in less time.

The analysis of emotions is very important in these cases, since many people do not receive adequate attention because they believe that the symptoms, they present do not need medical attention, when in fact these first symptoms are essential to identify depression in time. On the one hand, facial perception is one of the key indicators of social interactions allowing to determine clues about thoughts and emotions through the facial expressions of an individual, likewise, mental disorders can be determined through negative facial expressions, without However, it is a challenge to be able to differentiate between the facial recognition of a person suffering from a mental disorder and someone with optimal health controls (Simcock et al., 2020). On the other hand, speech has the potential to provide characteristics that help detect a mental disorder, since the vocal anatomy is a unique structure that provides the ability to vocalize various acoustic signals in a coordinated and meaningful way, making it a marker suitable for detecting health conditions (Cummins et al., 2018). This research will focus on the evaluation of the level of depression considering the implementation of a technological model whose components can detect the chronicity of depressive symptoms and address them with the help of technological tools and experts in mental health. The main contributions of the proposed model are the following:

- We propose a technological model to support the treating mental health professional to optimize the diagnosis and level of depression by analyzing facial images and voice, signals that will help detect the chronicity of depressive symptoms.
- Our technological solution benefits from the facial recognition characteristics that will be obtained through the camera of a mobile device, and these will be evaluated with algorithms in the Azure Cognitive Services cloud.
- Our technological solution benefits from the voice recognition features that will be obtained through the microphone of a mobile device and these will be transcribed from audio to text with the IBM Watson Speech to Text service and then analyzed with IBM Watson Tone Analyzer.
- We propose to obtain a better diagnosis through the support of mental health specialists so that the patient begins with the corresponding treatment, which will be presented in our experiments.

This paper is organized as follows. In section 2, we will describe the differences and comparisons with other works about the evaluation of the level of depression; in section 3 we will address the key concept for the core of our approach in the evaluation of depression level with facial and voice analysis and the aggregated value of the our work according to the evaluation of the level of depression. Subsequently, in section 4 we will present the validation of the technological model functionalities in a simulated scenario.

Finally, in section 5 we will specify our main conclusions and results of the finished application.

2 RELATED WORKS/DISCUSSION

In (Li et al., 2021), the authors describe that the problem related to the persistence of anxiety and depression in the population due to the COVID-19 pandemic was addressed. They present a technique related to the analysis of potential risk factors in different types of population associated with the symptoms of the mental disorders mentioned above, where self-administered medical instruments were used to measure levels of depression (Zung SDS) and anxiety (Zung SAS) to measure the severity of symptoms. In contrast to this research, we contemplate the use of the Zung SDS test, also used in public health entities in Peru, to identify the level of depression based on the symptoms evidenced in the patient considering the quality that allows it to perform the test without the mandatory need for the accompaniment of a mental health professional.

In (Khanal et al., 2018), addressed the risk of suffering from health problems in older adults, and the limitations to implement technological solutions for routine surveillance. Therefore, they developed an intelligent model that detects emotions in real time using facial images using the Microsoft Azure Face API cognitive service. We extend this work with the idea of analyzing facial emotions with the help of the Face API cognitive service to obtain the main emotions that are related to depressive disorder and provide a greater number of characteristics to specialists.

In (Ralston et al., 2019), the authors addressed the need to provide a better user experience and a better understanding of complex behaviors in different conditions through the integration of emotional capabilities in Chatbots. They provide a comparison of different Chatbot APIs where interactivity with these could support different languages and analyze the tone of voice and mood of the user with services from IBM, Amazon, and Google. Unlike his work, we adapted the IBM Watson Tone Analyzer technology tool to detect emotions in texts, but we complemented it with the IBM Watson Speech-To-Text speech-to-text transcription tool to be able to perform the analysis of the patient's voice.

In (Williamson et al., 2019), the authors addressed the limitations of capacity for clinical office visits by patients by automatically estimating depression from facial and voice analysis. The technique used by the authors is to develop an algorithm that estimates articulatory coordination of speech from audio and video signals and uses these coordination characteristics to allow you to learn the prediction model and track the severity of depression using the scale of Hamilton (HAM-D). Unlike the authors, our approach is aimed at obtaining a presumptive result of the level of depression in the first instance where facial and voice analysis are used during the execution of a medical questionnaire (Zung SDS), in this way we couple the facial and speech analysis tools from cognitive services to reinforce Zung SDS test results.

3 DEPRESSION DETECTION WITH IMAGES AND VOICE

3.1 Preliminary Concepts

In this section, the main concepts involved in our research will be developed. We propose that, for each concept, there is a definition and a respective example based on a review of the literature on depressive disorder and facial and voice recognition.

3.1.1 Facial Recognition

Definition 1 (Facial Recognition (Li et al., 2017))**.** *When compared to other bio-metric characteristics such as fingerprints, palms, etc., it has several advantages in obtaining these characteristics, since they can be extracted through the images of the cameras in a non-intrusive way.*

Example 1 (Face Detection)**.** *Given the Fig. 1, the procedure for obtaining multiple face detection is displayed. Fig. 1a shows the detected face-like regions, Fig. 1b shows the rough face detection result, Fig. 1c shows the spatial distribution of facial features, and Fig. 1d shows the refined result.*

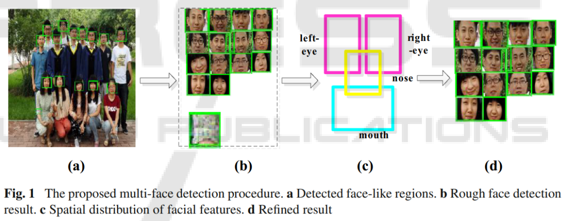

Figure 1: Face recognition (Li et al., 2017).

Definition 2 (Facial Expression (Zhang et al., 2016))**.** *This is one of the key social indicators which allows us to determine clues about thoughts and emotions from the movements and positions of the facial muscles under the skin of the face. These movements are a form of non-verbal communication and transmit the emotional state to an observer.*

Example 2 (Detection of Facial Expressions)**.** *Given the Fig. 2, the features to obtain the emotional expressions of are: i)The distance between the two eyes is identified, ii) The width of the nose is estimated, iii) The vertical distance between the eyes and the center of the mouth is calculated and iv) The distance between the eyes and the eyebrows is measured.*

3.1.2 Voice Recognition

Definition 3 (Voice Recognition (Cummins et al., 2018))**.** *Speech has the potential to provide charac-*

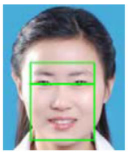

Figure 2: Facial expressions detection (Zhang et al., 2016).

teristics that help detect a mental disorder, since the vocal anatomy is a unique structure that provides the ability to vocalize various acoustic signals in a coordinated and meaningful way, making it a suitable marker to detect health conditions.

Example 3 (Detection of Facial Expressions)**.** *Given the Fig. 3, the muscles and structures that produce the voice signal are shown, which supports being an identifier of different health conditions.*

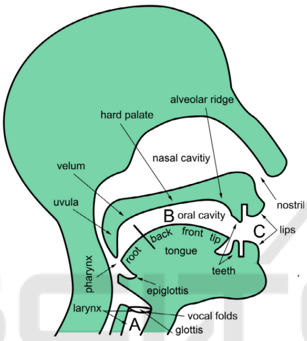

Figure 3: Muscular key groups to produce an speech.

3.1.3 Depression

Definition 4 (Depression[2])**.** *Depression is a common mental disorder considered one of the main causes of disability worldwide that can lead to suicide, where the individual who suffers from it experiences a depressed mood, losing enjoyment and interest in developing activities. Likewise, depressive episodes can be categorized by levels: mild, moderate, or severe, depending on the severity of the symptoms and the impact on the person's functionality (WHO, 2021).*

Example 4. *Some typologies of mood disorder indicated by the WHO are: i) Single episode depressive disorder, ii) Recurrent depressive disorder and iii) Bipolar disorder.*

According to the National Ministry of Health (MINSA)[3], depressive disorder is considered a disease that mainly affects the mood of an individual, which is why it is known as a mood disorder or affective disorder. In addition, individuals who experience the disorder often experience deep feelings of sadness, which can hinder their family relationships and work responsibilities, due to the loss of desire to perform activities.

In Fig. 4, the depressive episode screening flow is shown, using the PHQ-9 instrument. Where it depends on the score obtained and the medical criteria evaluated, the patient can start a treatment of the depressive episode level or in cases of moderate or severe episodes be referred to a psychiatrist.

3.2 Method

In this section, we will present the design of a model to assess the level of depression by analyzing facial images and a patient's voice. To explain the design process, it will be divided into three sections: component analysis and benchmarking, technology model design, and solution architecture.

3.2.1 Benchmarking and Analysis of Components

First, looking for components in the technological model to assess the level of depression using facial and voice analysis with three layers:

- Front Office Layer: It will make it possible to identify the way in which customers will go to acquire the product and service[7].
- Middle Office Layer: It will be related to the intermediate section of the business architecture focused on the execution of external and / or internal rules that satisfy the needs of the business logic[8].
- Back Office layer: It will be related to the software in charge of managing the core functions of the solution[9].

According to the National Institute of Statistics (INEI)[10], in the first quarter of 2020, 93.3% of households have at least one member who has a mobile phone, and of the total number of people who have internet access, 87.9% do so through this device.

Due to these facilities provided by the mobile device, in addition to its components such as the integrated camera and microphone, it is for this reason that various scientific investigations on depression rely on to carry out the evaluation of depressive symptoms or the level of depression that occur in people. Likewise, the essential tools used to carry out the depressive disorder detection procedure in patients were a minimum recording camera of 30 FPS and a microphone with a frequency of 16 kHZ (Zeghari et al.,

[7] "Front Office scanning: from the back room to the counter" - IBM

[8] "Data office as part of enterprise architecture"

[9] "Back office software" - FinancialForce

[10] Statistics of information and communication technologies in households (in Spanish)

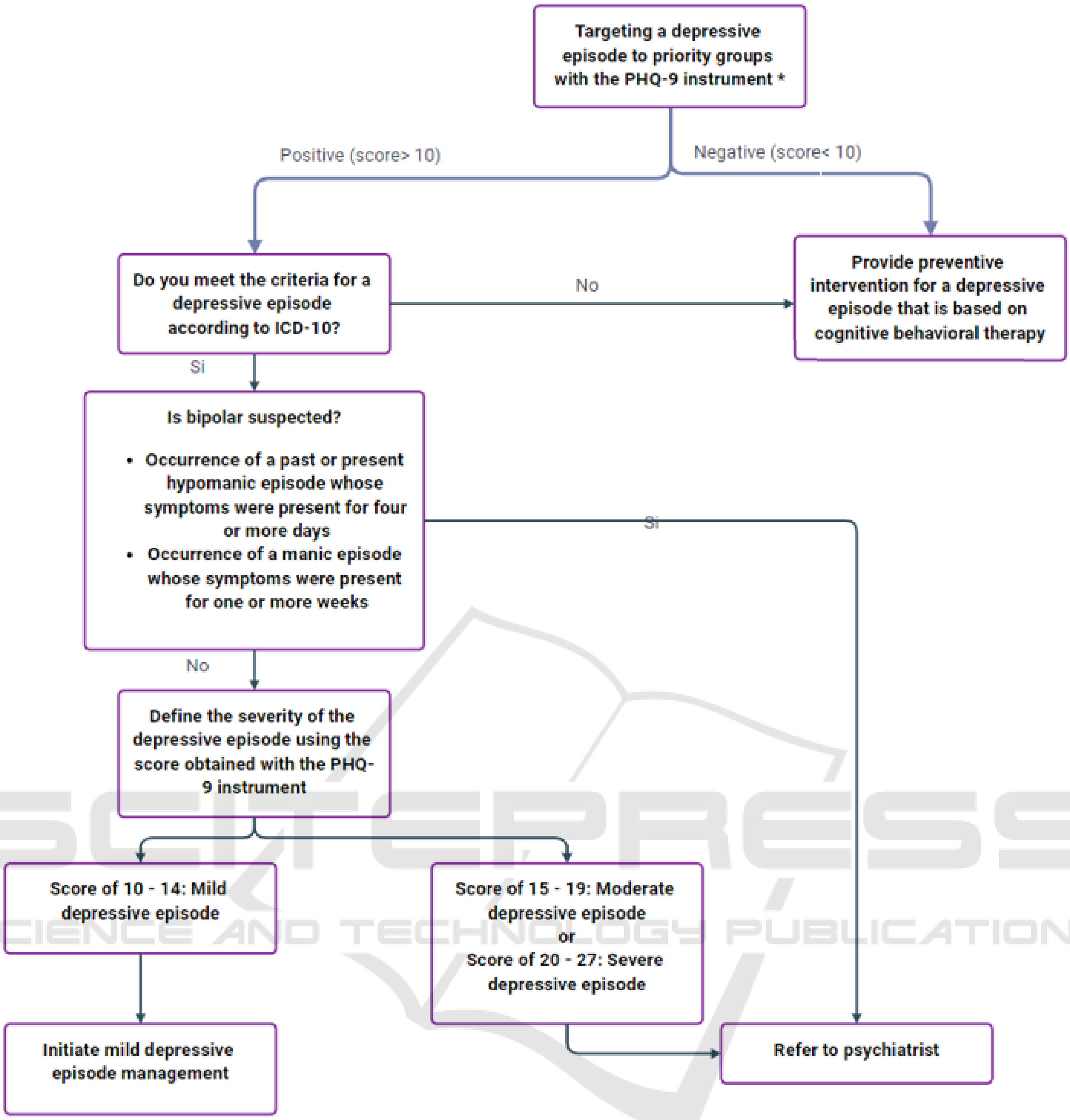

Figure 4: Screening and diagnosis flowchart for mild depressive episode (Macciotta-Felices et al., 2020).

Table 1: Analysis of key characteristics in self-administered medical questionnaires.

	PHQ-9	CES-D	BDI-II	Zung SDS	MDI
Estimated time (minutes)	5 to 10	5	5 to 10	10	5 to 10
Target Audience	General Public	Older than 12 years	Older than 13 years	General Public	General Public
Number of items (questions)	10	20	21	20	12
Sensitivity	.88	.98	.88	.79	.86
Specificity	.88	.57	.98	.72	.82

2021). In the Middle Office layer, the analysis and benchmarking of self-administered medical questionnaires (Screening Test) to evaluate the level of depression, facial image analysis cloud services and voice analysis cloud services were considered. Regarding the Screening Tests to assess the level of depression, five medical tools were considered:

- Patient Health Questionnarie (PHQ-9)

Table 2: Analysis of key characteristics in cloud services for facial images analysis.

	Face API (Microsoft Azure)	Amazon Rekognition (AWS)	Vision AI (Google Cloud)
Purpose	Artificial intelligence service that analyzes faces in images and videos	Automate image and video analysis with machine learning	Detect emotions, interpret text and images
Supported image format	JPEG, PNG, GIF and BMP	JPEG and PNG	JPEG, PNG8 and PNG24
Real-time analysis	Yes	Yes	Yes
Service availability	.99	.99	.99
Face attributes that it detects	Emotions, age, blur, exposure, gender, head position,hair, noise, occlusion and smile	Gender, smile, emotions, face landmarks, posture, beard, glasses	Emotions, age, nose, ears, mouth, face position and blur

Table 3: Analysis of key characteristics in cloud services for voice recognition.

	Watson Speech to Text (IBM Cloud)	Amazon Transcribe Medical (AWS)	Speech to Text (Google Cloud)
Purpose	Use speech recognition to convert a language to text	Add speech-to-text functionality for the medical field	Convert speech to text with precision using AI technology
Supported audio size	<=100MB	<=2GB	<=10MB
Detect voice in real time	Yes	Yes	Yes
Complementary services to analyze the text	Tone Analyzer (IBM Cloud)	Twinword API	Natural Language API
Service availability	.99	.99	.99
Classification of feelings	By type of feeling	By type of feeling	Positive, negative or neutral

Table 4: Analysis of key characteristics in NoSQL databases.

	Mongo DB	Cassandra	Redis	Couchbase
Data storage	Documents (BSON, XML, etc)	Oriented to flexible columns	Data structures such as lists, ordered sets, strings, bitmaps	Documents (JSON, XML, etc)
Use cases	Real-time analysis, mobile applications	E-commerce, fraud detection, IOT	Chat or messaging, real-time analysis, cache storage	Mobile apps Open Source Yes Yes Yes Yes
DBaaS	ScaleGrid, MongoDB	-	Redis Enterprise e Cloud	-
Provider	AWS, Google cloud Platform o Microsoft Azure	AWS, Google cloud Platform o Microsoft Azure	AWS, Google cloud Platform o Microsoft Azure	AWS, Google cloud Platform o Microsoft Azure

- Center for Epidemiologic Studies-Depression scale (CES-D)
- Beck Depression Inventory (BDI-II)
- Zung Self-Rating Depression Scale (Zung SDS)
- Major Depression Inventory (MDI)

Table 1 shows the analysis of key characteristics about the self-administered medical questionnaires mentioned to measure the level of depression. As a result of the benchmarking estimation, the Screening Test with the best fit for our proposal was the Zung Self-Rating Depression Scale (Zung SDS), since it covers a General Public, and it has a good trade-off between sensitivity and specificity. This was considered as the component that will support the evaluation of the level of depression.

Table 2 shows the evaluated criteria according to cloud services of analysis of facial images. As a result of the benchmarking estimate, the facial image recognition and analysis cloud service with the highest score was Face from Microsoft Azure. This service will be considered as part of the solution due to its functionalities.

Table 3 shows the evaluated criteria of the voice recognition cloud services. As a result of the benchmarking estimate, the speech recognition and text analysis cloud services with the highest scores were the Speech to Text and Tone Analyzer services. NoSQL databases store data in documents rather than relational tables. NoSQL database technology stores information in JSON documents instead of columns and rows used by relational databases. In the Back Office layer, the analysis and benchmarking of NoSQL databases was considered, since these are designed specifically for specific data models and you have flexible schema to create applications, compared to a relational SQL database that is a collection of Predefined data elements among them, being organized as a set of tables with columns and rows (Khasawneh et al., 2020). Regarding NoSQL databases, non-relational databases were considered table: i) Mongo DB, ii) Cassandra, iii) Redis and iv) Couchbase.

Table 4 shows the evaluated criteria of the NoSQL databases. As a result, we have that the databases that satisfy the most these criteria were MongoDB and Redis, this is because both present facilities for the migration of the database to a cloud environment. However, a higher score was obtained by MongoDB since the data storage is document oriented and this will be essential for better data collection.

3.2.2 Design of the Technological Model

Second, we carry out the design of the technological model, incorporating the components of the first section of component analysis and benchmarking. In Fig. 5. the proposed model is shown.

To explain the attributes of the technological model in greater detail, we will indicate the model input, phases of the model and the model output.

Model Input: It begins when a young patient seeks institutions or mental health professionals due to the recurrent presence of depressive symptoms that hinder their daily activities. We consider the age between 18 and 29 years, considered as the stage of youth[11].The initial inputs would be, on the one hand, the demographic data of the young patient, and on the other hand, the self-administered Screening Test that will allow knowing presumptively the level of depression suffered by the patient. Both the demographic data and those of the Screening Test are planned to be hosted on a Backend of a mobile application deployed in the Microsoft Azure cloud environment with its Service App service.

[11]"Health Situation of Adolescents and Young People in Peru" (in Spanish) - MINSA

Phases of the Model

Devices. The model contemplates the use of mobile devices, both smartphones and tablets, which will allow the young patient to access a mental health counseling application. Likewise, the patient will be able to perform the self-administered Screening Test to assess the level of depression, and the device's camera and integrated microphones will be used to capture emotions in the face and voice to reinforce the results obtained during filling.

Internet Connection. An important aspect for accessing the mental health counseling application is having an internet connection, for this reason it is estimated that patients can contact it, through their WI-FI network or through their mobile data from net. This to be able to access the services within the application, such as facial and voice analysis, data storage or information listing in real time, which will be deployed in the Microsoft Azure Service App cloud environment.

Depression Level Assessment Process. To begin with the process of evaluating the level of depression, it is necessary to indicate that the levels of depression that will be considered in this project are those that are granted by the Screening Test of the Zung Self Depression Scale (SDS), which based on the score obtained by the patient depression is classified into levels: normal, mild, moderate, and severe. Regarding data hosting, this will be done in the Backend deployed in the Microsoft Azure cloud environment with its Service App service, which is planned to be designed so that the mental health advisory application can be run on cross-platform mobile devices (Android and iOS). In addition, regarding the software, the tools that will allow the construction of the Backend would be C# and the .NET CORE framework for the construction of code and business logic, integrating through libraries the cognitive cloud services of facial analysis (Face-Microsoft Azure) and speech analysis (Speech to Text and Tone Analyzer-IBM Watson), and the Mongo DB Atlas which will allow the storage of non-relational data; Regarding the

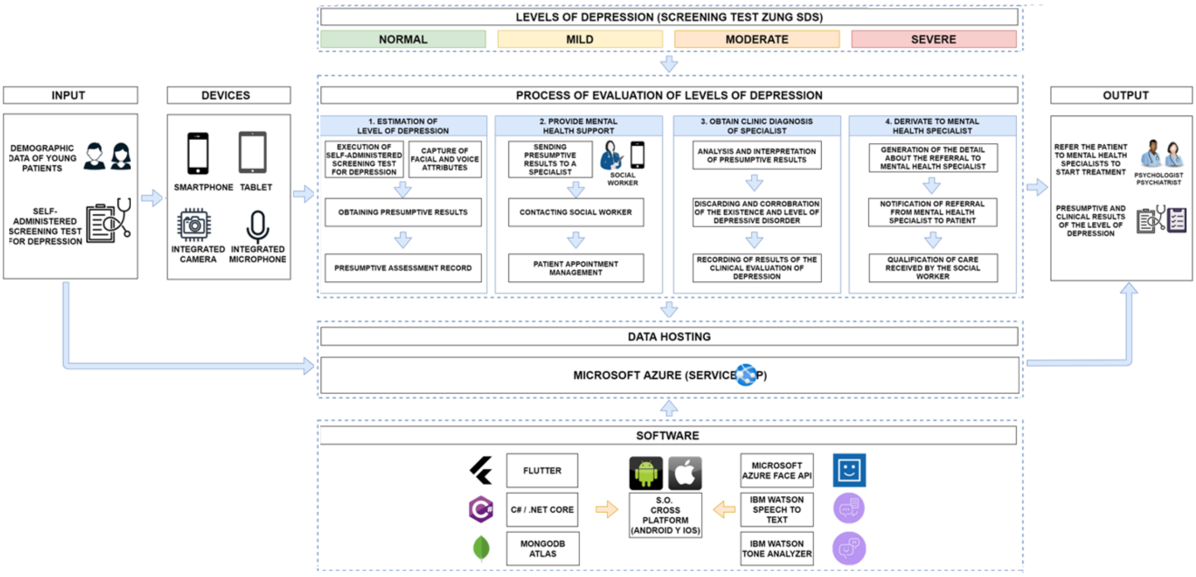

Figure 5: Our proposal.

Frontend construction tools, they would correspond to the Flutter framework for the user interface.

Estimation of the Level of Depression: The first step to assess the level of depression has been the estimation of the level of depression, where only the young patient intervenes, who will perform the Zung SDS self-administered Screening Test, which consists of 20 questions and 4 alternatives, to obtain a result presumptive level of depression. It is planned that the young patient can carry out this activity within the proposed mobile application, where the capture of facial and voice attributes will also be executed at the same time, to reinforce the results obtained from the Screening Test. After that, it is planned that the data obtained from the presumptive evaluation can be registered within the platform, so that a history of evaluations carried out and the option of being able to contact See a mental health specialist of your choice to receive mental health support.

Provide Mental Health Support: The second step to assess the level of depression takes place when the young patient sends the presumptive results to a mental health specialist, who will provide the necessary support based on their knowledge and the battery of psychological or medical evaluations. that you consider necessary to carry out. For this scenario, it is planned that the young patient will have the possibility of contacting a social worker who will help as a first filter in the evaluation of the level of depression. It is proposed that the management of the appointment can be carried out within the mental health counseling application where the details of the meeting are indicated so that both the patient and the specialist can follow up.

Obtain a Clinical Diagnosis from a Specialist: The third step to assess the level of depression relies mainly on the analysis and interpretation of presumptive results by the social worker in the meetings that have been held together with the patient. In this way, the mental health specialist will have the possibility to rule out symptoms and corroborate the existence of the depressive disorder. Since the mental health of the patient is a very delicate subject, the best option he chose for the research is that the social worker as the first mental health filter can determine if the result of the level of depression obtained from the Screening Test is related with the true level of depression experienced by the patient.

Refer to a Mental Health Specialist: The fourth step to assess the level of depression has been the referral to mental health specialists, where the social worker will generate a detail that involves the presumptive results and clinical diagnosis of the patient. For these cases, in (Macciotta-Felices et al., 2020) where the authors point out that patients suffering from mild depression should be referred to a psychologist, while cases of moderate or severe depression are linked to referral to a psychiatrist. In addition, it is planned that once the patient receives the details of the referral notification, they will have the option of being able to rate the care received by the specialist from the first mental health filter.

Model Output. To finalize the flow of the proposed model, the outputs generated would be to refer the patient to a mental health specialist to initiate optimal treatment, thus depending on the level of depression identified by the social worker as the first mental health filter, the referral will allow timely attention from psychologists or psychiatrists according to the level of depression experienced.

3.2.3 Solution Architecture

Once the technological model was mentioned, the proposed integrated architecture sketch of the mobile solution that would be built was carried out to validate the functionalities of the model.

In the Frontend, the tools to be used are mainly based on Flutter which is a tool provided by Google to design the user interface of mobile, web or desktop applications. The reason for the use of this tool can be seen reflected in (Faisol et al., 2021), where they used the Flutter cross-platform framework to create a voice recognition mobile application.

In the Backend, Microsoft's open-source framework, .NET Core, was incorporated, which will allow the creation of cross-platform applications in Android and iOS environments, using the C# programming language to perform the business logic of a mobile application[12]. Also, to achieve the development of the validation solution, the Integrated Development Environment (IDE) with which it would work would be Visual Studio or Visual Studio Code.

4 EXPERIMENTS

In this section, the procedure and the necessary tools will be shown to carry out the deployment of the prototype of the technological model of the research that will support the evaluation of the level of depression of a patient and the corresponding referral.

4.1 Experimental Protocol

For this study, a prototype was developed that served as support to validate the technological model, which is a mobile application that has the following services.

Table 5: Services used in the development.

Service	Provider
App Service	Microsoft Azure
Service Cognitive Face API	Microsoft Azure
Non-Relational Database	Mongo DB
Service Cognitive Speech-To-Text	IBM Watson
Service Cognitive Tone Analyzer	IBM Watson

[12]"Introduction to .NET" - Microsoft

First, an account was created in the MongoDB Atlas database and the "Free & Hobby" cloud database implementation plan was chosen. After that, the connection to the cluster was made where the connection to the IP address and the creation of the database user were selected, also the selected driver is C# .Net.

Second, the Backend was deployed, that is, the cognitive services of Microsoft Azure and IBM Watson. To do this, a Microsoft Azure App Services is created in the .NET Core 3.1 LTS environment and Windows Operating System where they gave us a Basic B1 plan with a total size of 1.75 GB of memory.

Finally, to generate the APK of the mobile application called "Help +" a key.properties file is created, the signature is configured in gradle, the application is changed to relread mode and finally the command to finish flutter build apk is executed to generate the file APK to be installed on a mobile device. The APK to install is in the following path: https://github.com/LuisPA-ui/AYUDA-.

4.2 Results

4.2.1 Participants

The research focused on supporting young Peruvians between 17 and 25 years old, for that reason we were able to contact 60 young people between the age ranges to participate and interact with the prototype made. These were selected randomly by sending invitations by mail to faculty students. Besides, mental health specialists were contacted, a psychiatrist and a social worker that were contacted by mail, they gave us their support to validate the technological model.

4.2.2 Validation

For validating of the prototype, we aim to prove these strategies: desirability, usability, and satisfaction.

Desirability: Identifying if a problem worth solving is being solved, surveys were conducted for young people and specialists, and a Show & Tell was held with both stakeholders. Regarding the survey designed in Google Forms to validate the functionalities of the Technological Model aimed at the role of:

- Patient: which was carried out by 57 young Peruvians:
  - 75.4% would be willing to look for a mobile App to be able to address depressive disorder.
  - 63.2% intend to contact a specialist to deal with depression.

 - 80.7% would be willing to run a medical questionnaire (Zung SDS Test) with the functionalities of facial and voice analysis.
 - 89.5% would like to receive the clinical diagnosis from the specialist in the same App
- Specialist: which was carried out by 18 mental health professionals:
 - 44.4% indicate that the largest number of cases are related to adolescents and young people.
 - 50% indicate that the moderate level of depression is mostly evident in the Peruvian population.
 - 75% agree that a worker or social worker is a first filter to diagnose symptoms and refer the patient to the specialist.
 - Zung SDS, BDI-II and PHQ-9 are some depression diagnostic questionnaires that specialists consider to be most effective.

Regarding the show & tell for:

- Participants: the meeting was held with 20 users, between 18 to 29 years old, through videoconference rooms, where:
 - The research project was explained
 - The depressive disorder context was explained.
 - A demo of a mobile app based on the technological model function was presented from the patient's perspective (Video format).
- Specialists: the meeting was held with two mental health professionals, through videoconference rooms, where:
 - The research project objective was explained.
 - The prototype of a mobile app based on technological model functions was presented.
 - The main benefits of the technological model presented were identified.

Usability: Unit Tests were performed on the young patients who used the prototype during the previously mentioned interviews, where they were given the APK of the mobile application so that they could perform the respective tests. After, they gave their opinions regarding its functionalities and usability, where it was obtained as a result that 70% of users identified that it is easy to use, as shown in Fig. 6.

Satisfaction: The level of satisfaction was also carried out during the interviews with the young patients and at the end of the tests 85% of the users indicated that they were satisfied with the results and the ease of contacting a mental health specialist to support the depression, as shown in Fig. 7.

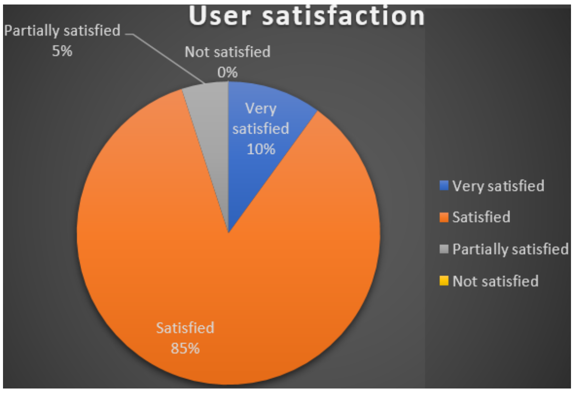

Figure 6: Usability of the prototype.

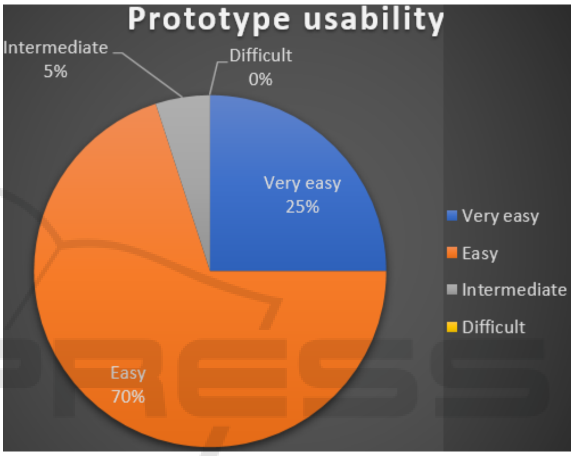

Figure 7: User satisfaction.

Regarding the precision of effectiveness of the cognitive services of Microsoft Azure and IBM Watson of facial and voice recognition used respectively in the research project, the following is presented:

Table 6: Precision of the cognitive services.

Service	Accuracy
Microsoft Azure Face API	90% – 95%
Watson Tone Analyzer	41% - 68%

5 CONCLUSIONS AND PERSPECTIVES

The research carried out on the methodologies that are currently used to evaluate the level of depression of a patient allowed us to know the current situation of this process and the deficiencies that it presents. The improvement opportunities focus on not only using technological tools to assess the level of depression, but also including faster and more effective contact with a specialist. Thanks to the validation strategy that was based on desirability, feasibility, usability and satisfaction, it allowed us to identify that young Peruvians are willing to use technological tools in order to be supported to improve their mental health, as well as mental health specialists identify a great opportunity to improve in this sector.

A technological solution to monitor the depressive state of a patient by analyzing social media posts in order to monitor the signs of depressive symptoms that a patient is going through by analyzing their daily posts on their social media to obtain the evolution of the chronicity of symptoms in each time range. Evenmore, using Genetic information to seek for historical data about a patient depression (Arroyo-Mariños et al., 2021) or monitoring symptoms with a technological solution similar to other disease (Jorge-Lévano et al., 2021).

Preventive model to address suicidal depressive episodes with the help of a virtual assistant that seeks to prevent suicidal ideas caused by severe episodes of depression using strategies that promote positive coping in people with the help of virtual assistants. Since it is considered that a depressive episode can occur at any moment in an individual's life, it is planned to develop a virtual assistant that can accompany and provide mental health support in severe episodes of depression, and that this allows to recommend or contact directly to a mental health professional after the level of depression subsides.

REFERENCES

Arroyo-Mariños, J. C., Mejia-Valle, K. M., and Ugarte, W. (2021). Technological model for the protection of genetic information using blockchain technology in the private health sector. In *ICT4AWE*.

Cummins, N., Baird, A., and Schuller, B. W. (2018). Speech analysis for health: Current state-of-the-art and the increasing impact of deep learning. *Methods*, 151.

Faisol, M., Ramlan, S. A., Hafizah, A., Mozi, A., and Zakaria, F. F. (2021). Mobile-based speech recognition for early reading assistant. *Journal of Physics: Conference Series*, 1962.

Graham, S., Depp, C., Lee, E., Nebeker, C., Tu, X., Kim, H.-C., and Jeste, D. (2019). Artificial intelligence for mental health and mental illnesses: an overview. *Current Psychiatry Reports*, 21.

Jorge-Lévano, K., Cuya-Chumbile, V., and Ugarte, W. (2021). Technological solution to optimize the alzheimer's disease monitoring process, in metropolitan lima, using the internet of things. In *ICT4AWE*.

Khanal, S. R., Reis, A., Barroso, J., and Filipe, V. (2018). Using emotion recognition in intelligent interface design for elderly care. In *WorldCIST*, volume 746 of *Advances in Intelligent Systems and Computing*.

Khasawneh, T. N., AL-Sahlee, M. H., and Safia, A. A. (2020). Sql, newsql, and nosql databases: A comparative survey. In *ICICS*.

Li, C., Wei, W., Li, J., and Song, W. (2017). A cloud-based monitoring system via face recognition using gabor and CS-LBP features. *J. Supercomput.*, 73(4).

Li, X., Yu, H., Yang, W., Mo, Q., Yang, Z., Wen, S., Zhao, F., Zhao, W., Tang, Y., Ma, L., Zeng, R., Zou, X., and Lin, H. (2021). Depression and anxiety among quarantined people, community workers, medical staff, and general population in the early stage of covid-19 epidemic. *Frontiers in Psychology*, 12.

Macciotta-Felices, B., Moron-Corales, C., Luna-Matos, M., Gonzales-Madrid, V., Melgarejo-Moreno, A., Zafra-Tanaka, J. H., Goicochea-Lugo, S., Martinez-Rivera, R. N., Nieto-Gutierrez, W., Fiestas-Saldarriaga, F., Taype-Rondan, A., Timana-Ruiz, R., and Garavito-Farro, H. (2020). Clinical practice guideline for the screening and management of the mild depressive episode at the first level of care for the peruvian social security (essalud). *ACTA MEDICA PERUANA*, 37(4).

Ralston, K., Chen, Y., Isah, H., and Zulkernine, F. H. (2019). A voice interactive multilingual student support system using IBM watson. In *IEEE ICMLA*.

Simcock, G., McLoughlin, L. T., Regt, T. D., Broadhouse, K. M., Beaudequin, D. A., Lagopoulos, J., and Hermens, D. F. (2020). Associations between facial emotion recognition and mental health in early adolescence. *International Journal of Environmental Research and Public Health*, 17.

Villarreal-Zegarra, D., Cabrera-Alva, M., Carrillo-Larco, R. M., and Bernabe-Ortiz, A. (2020). Trends in the prevalence and treatment of depressive symptoms in peru: a population-based study. *BMJ Open*, 10(7).

Williamson, J. R., Young, D., Nierenberg, A. A., Niemi, J., Helfer, B. S., and Quatieri, T. F. (2019). Tracking depression severity from audio and video based on speech articulatory coordination. *Comput. Speech Lang.*, 55.

Zeghari, R., König, A., Guerchouche, R., Sharma, G., Joshi, J., Fabre, R., Robert, P., and Manera, V. (2021). Correlations between facial expressivity and apathy in elderly people with neurocognitive disorders: Exploratory study. *JMIR Form Res*, 5(3).

Zhang, Y., Yang, Z., Lu, H., Zhou, X., Phillips, P., Liu, Q., and Wang, S. (2016). Facial emotion recognition based on biorthogonal wavelet entropy, fuzzy support vector machine, and stratified cross validation. *IEEE Access*, 4.