SENTIMENT ANALYSIS FOR DEPRESSION DETECTION

Article · January 2022

DOI: 10.54473/IJTRET.2022.6305

5 authors, including:

Swarada Jalukar

1 PUBLICATION **0** CITATIONS

SEE PROFILE

Arati Ratnaparkhi

Bharati Vidyapeeth College Of Engineering Women

1 PUBLICATION **0** CITATIONS

SEE PROFILE

Vinaya Kulkarni

Bharati vidyapeeth college of engineering for women

19 PUBLICATIONS **3** CITATIONS

SEE PROFILE

doi.org/10.54473/IJTRET.2022.6305
International Journal of Trendy Research in Engineering and Technology
Volume 6 Issue 3 June 2022
ISSN NO 2582-0958

SENTIMENT ANALYSIS FOR DEPRESSION DETECTION

Swarada Jalukar , Arati Ratnaparkhi , Priyanka Shinde , Simran Kunkulol , Vinaya Kulkarni
Department of Computer Engineering, Bharti Vidyapeeth's College of engineering for Women, Pune, India
Corresponding author jalukarswarada@gmail.com*)*

ABSTRACT

The Covid-19 pandemic has dramatically changed the way we have used to live. The pandemic has been causing significant devastations in economy, and health, inter alia. Mental health, especially, has become a growing concern due to employment terminations, income loss, family stress and other uncertainties. The pandemic disproportionally affected mental health of younger population. Nowadays risk of early death is increasing due to mental illness which is mostly caused due to depression. Depression creates suicidal thoughts causing serious impairments in daily life. Sentiment analysis is a hot topic that's been on research for decades, which intends to find the nature of text and classifies into positive, negative and neutral. In today's digital world lot of data can be made available for sentiment analysis.

Hence, our aim is to focus on creating a depression detection system from text, video & audio analysis. Sentiment Analysis and Natural Language Processing methods will be used to develop this system. The system will classify text, audio and video cues as positive or negative depending on the emotions inferred from user's input.

I. INTRODUCTION

Depression has been recognized as a significant health concern worldwide. Depression is one of the most common and disabling mental disorders, and has a relevant impact on society [1]. According to the World Health Organization (WHO), more than 300 million people suffer from depression in their daily lives. A complex mental disorder that could not be solely captured from one single modality is called as 'Depression' [2]. Various researchers have shown that features integrating acoustic, textual and visual biomarkers to analyze psychological distress have shown great performances for depression detection.

Modern society has made human life so busy, making it vulnerable to mental disorders like depression, anxiety etc[2]. Psychological health proves a vital role on their overall personal and social life [3]. Neglecting psychological problems results in rise of issues such as stress, anxiety, depression etc. [4]. Detection and controlling of these problems at the initial stages itself is necessary to achieve better mental health. So, an automated system is required that will pick out the people who are dealing with depression. A system proposed, captures frontal face videos, extracts the facial features from each frame and analyses these facial features to detect signs of depression in them [5].

Sentiment analysis is a hot topic that's been on research for decades. Sentiment analysis (SA) represents a computational study of opinions, sentiments, emotions, and attitudes expressed in texts or other media about a specific topic [6]. An innovative solution to monitor and to detect potential users with emotional disturbances, based on the

classification of sentences with depressive or stressed content can be done [7]. Further ideas can be investigated to improve the system performance. The performance may improve if additional facial expression images are added into the training process. Therefore, in the future work, more videos of the same person, taken at different time duration can be considered [8].

Artificial intelligence is making computers smarter and take decision based on programming done. It is systematizing automation in a very broad sense and contributing in almost every field of research. Numerous machine learning algorithms are available for the classification of the texts, some of them are Naïve Bayes, Support vector machine algorithms, Random forest algorithms etc. But the documented experiments have showned better result for Naïve Bayes and support vector machine algorithms, the study is limited to these algorithms [3].

Our system will be using real time video capture using image processing techniques. These will surely prove beneficial for scanning image features such as extracting sentiments from face; mainly happy, sad, neutral, etc. So, after detecting depression, such a person can be given good counselling of how to deal with mental stress and can be guided to follow the right path to success. In sentiment analysis, each acquisition framework is denoted as a modality (modal quality) and is associated with our dataset. As audio, video, and text are very useful for characterizing human interaction and communications, they have been used in many applications such as speech processing techniques, safety and security applications and human-machine interaction for robots too. Also, many research efforts have been directed towards using these datasets for depression prediction and evaluation. There is tremendous interest in automatically determining valence in sentences and text through supervised AI systems. Over the last decade this has become evident from the large number of research papers, textual datasets, shared task competitions, and machine learning systems developed for valence prediction.

II. OBJECTIVE

Increase in more and more pressure, which people have to face with the increase of the pace of work and life, increases the possibility of person suffering from depression. Depression, now is becoming very alarming situation. It is one of the major mental disorders that can affect a person's daily life, his habits, and he may lose interest in everything around him and feel isolated from the world. He may lose over his sentiments and hence can affect his life. Hence, in order to detect and take necessary precautions at an early stage, with the help of 'Artificial Intelligence', we can determine Depression using Sentiment Analysis.

However, due to the serious imbalance in the doctor-patient ratio in the world many patients may fail to get a timely diagnosis. Consequently, to improve current medical care, we use sentiment analysis to extract a representation of depression cues in text, audio and video for automatic depression detection.

III. METHODOLOGY

Sentiment analysis: The sentences are filtered and scored by the sentiment metric by giving one threshold value. This range was tested and validated. The sentiment intensity of the sentence will determine the level of depression then. There are three levels of text sentences: extreme, intermediate and lower. The text intensity levels were determined according to the users opinions. Examples of very positive texts include intensity adverbs, such as

much, very, strongly, among others. The proposed method can be divided into five segments. These segments are, negative keywords, search and extract texts from user entered sentences, and feature extraction, introducing Naive Bayes, model

efficiency evaluation can be done further. Emphasis is given to the approaches utilizing various artificial intelligent methods for detection of depression. The system will read the sentimental signs from the user using NLTK and process the result as positive, negative or neural.

Currently, video-based systems for depression assessment have not been applied in the general population to evaluate their feasibility and only been found in research-related projects. Although currently limited to research applications, the field has been very popular, although with some challenges. More work is going on in improving the existing systems or replace it with the new efficient. The facial expressions are captured by the real-time camera and then the prediction will be done for video-based system. The system will read the sentimental signs from the user expression using OpenCV and process the result as positive, negative or neural.

Now, through user audio the input is converted into text using google speech API and using NLTK the depression is detected. Proper data processing and analysis will be done beforehand. After it, artificial intelligent techniques and other APIs will be used to display the final output. All the applied algorithms on python and libraries are NumPy, Pandas, Scikit- Learn, and Matplotlib. Additionally, this will cover strategies for research, data collecting, research topics, pre-processing, processing, statistical analysis, and implementation. Design and develop a web-based application for detecting depression level using textual, audio and visual cues with the help of sentiment analysis and artificial intelligent techniques. The system will read the sentimental signs from the user and process the result as positive, negative or mild. Hence, our project: 'Sentiment analysis for depression detection' will be a web based application with the feasibility to almost everyone. Application, with the help of Naïve bayes classifier for textual detection, google speech to text API and OpenCV library will detect the depression level followed by suggestions of instant actions to be taken by patient.

IV. BACKGROUND AND ANALYSIS

DATASET

TEXT data:

Here, we use our own dataset mainly, 'POSITIVE.TXT' and 'NEGATIVE.TXT' which contains list of words that may be classified as positive and negative in our natural language. Dataset 'POSITIVE.TXT' contains words such as, Happy, Joyful, Wonderful, etc. The dataset 'NEGATIVE.TXT' contains words such as sad, unhappy, dull, lonely, etc. There words will prove a helping hand in identifying a person's symptoms and classifying it as 'Depressed' or 'Not depressed'.

USER data:

To store user information, we use SQLite database at the backend. One of SQLite's greatest advantages is that it can run nearly anywhere. SQLite lets you store data in structured manner. SQLite has higher performance. SQLite databases can also be queried and the data retrieval is much more robust. SQLite has been ported to a wide variety of platforms: Windows, MacOS, Linux, iOS, Android, and more.

MATHEMATICAL MODEL

NAIVE BAYES CLASSIFIER ALGORITHM

Naïve Bayes algorithm is a supervised machine learning algorithm, based on Bayes theorem and used for solving classification problems. It is mainly used in text classification that includes a high-dimensional training dataset. Naïve Bayes Classifier is one of the simple and most effective Classification algorithms which helps in building the fast machine learning models that can make quick predictions. It predicts on the basis of

the probability of an object as it is probabilistic classifier.

In our project, Naive Bayes algorithm is used to predict whether a text entered by user or the audio which is converted into text using google speech API is negative or positive based on text alone. The objective is to train a classifier that given a text entered by user determines if it's positive or negative. In this test data is user response. Bayesian classification is used to predict the occurrence of any event. Statistical classifiers with the Bayesian probability understandings are bayesian classifiers. The level of belief, expressed as a probability is expressed by Bayesian classifiers. Bayes theorem came into existence after Thomas Bayes, who first utilized conditional probability to provide an algorithm that uses evidence to calculate limits on an unknown parameter.

The formula for Bayes' theorem is given as:

$$P(A|B) = \frac{P(B|A).\ P(A)}{P(B)}$$

Where,

P (A|B) is Posterior probability: Probability of hypothesis A on the observed event B.

P (B|A) is Likelihood probability: Probability of the evidence given that the probability of a hypothesis is true.

P (A) is Prior Probability: Probability of hypothesis before observing the evidence.

P (B) is Marginal Probability: Probability of Evidence.

OPENCV

OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products. The library has more optimized algorithms, which includes a comprehensive set of both classic and state-of-the-art computer vision and machine learning algorithms. These algorithms can be used to detect and recognize faces, identify objects, classify human actions in videos, track camera movements, find similar images from an image database, follow eye movements, recognize scenery and establish markers to overlay it with augmented reality, etc.

So, here we are using OpenCV for Visual detection of depression. We are also using the Haar - Cascade Algorithm for face detection. Haar-cascade has a benefit that they're very fast at computing 'Haar-like' features due to the use of integral images (also called summed area tables).

CONVOLUTIONAL NEURAL NETWORK

A Convolutional Neural Network (ConvNet/CNN) is a Deep Learning algorithm which can take in an input image, assign importance (learnable weights and biases) to various aspects/objects in the image and be able to differentiate one from the other. The pre-processing required in a ConvNet is much lower as compared to other classification algorithms. While in primitive methods filters are hand-engineered, with enough training, ConvNets have the ability to learn these filters/characteristics.

We, in our system use CNN for extracting facial features and recognize user's sentiment with the help of Webcam and hence predict the depression level.

SYSTEM ARCHITECTURE

A system architecture is the systematic model that defines the structure, behavior, and more views of a system. A system architecture consist of system components and the sub-systems, that will work together to model the overall system. A system architecture diagram is **the distribution of the functional correspondences**. These are formal elements, the embodiment of concepts and

information. Architecture defines the relations between elements, amongst features, and the surrounding elements.
Below is our Proposed System Architecture:

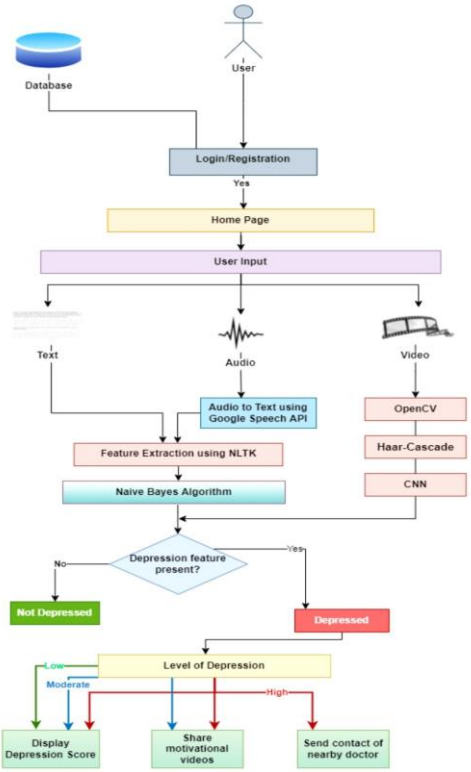

Fig 1: Architectural diagram

V. FUNCTIONAL REQUIREMENTS:

All of the functional and quality requirements of the system is contained under functional requirements. A detailed description of the system and all its features is covered.

Patient data through sign up is gathered (Login Requirement)

- Purpose: Provides user authentication
- Inputs: Inputs are through the keyboard and mouse clicks
- Processing: The input is verified by checking if the user already exists in the database.
- Outputs: The correct input will result in the next page i.e., the depression analysis page being loaded. If the input is incorrect then an error message will be displayed.

Registration Form Requirement (Sign up)

- Purpose: Registration of a new user, who don't have account.
- Inputs: Inputs are through the keyboard and mouse clicks.
- Processing: The input is validated using client side as well as server side validation. The client side validation will include checks for missing information in the required fields and other text fields like email and phone numbers will be checked for validity. The server side validation will involve checking if the username entered is already used by a member in the database. The appropriate error messages are displayed if the input is not acceptable
- Outputs: The member is directed to the main page on successful registration.

Analysis Requirement

- Purpose: To perform depression analysis through visual, textual or audio methods.
- Inputs: Input will be the answer entered by the user and consequently the data that the user wants to provide for the analysis.
- Processing: Depending on analysis, the appropriate statistical algorithms are used to calculate depression level on backend. The analysis being conducted is correct and that all invalid inputs are not accepted.

- Outputs: The output will be the analysis results (depression level) displayed on the web browser page.

Account Information Requirement

- Purpose: Display Profile page
- Inputs: Input will be viewing history and account data by user. Page should load on mouse click.
- Processing: At the back-end user should be registered and history for detection is present.
- Outputs: On the profile page a user can view his/her information.

VI. CONCLUSION

We have introduced an artificially intelligent system that was proposed for depression detection using Audio, Video and Text features. The system will read the sentimental signs from the user and process the result as positive, negative or mild. Application, with the help of Naïve bayes classifier for textual detection, google speech to text API and OpenCV library along with CNN will detect the depression level followed by suggestions of instant actions to be taken by patient.

VII. FUTURE SCOPE

- Virtual psychology
- Surveys & Research purposes
- Personalized mental-health assistance
- Create awareness and reduce stigma around depression
- Can be improved for complex mental health issues detection system.

REFERENCES

[1] J. Joshi, R. Goecke, S. Alghowinem, A. Dhall, M. Wagner, J. Epps, G. Parker, and M. Breakspear, "Multimodal assistive technologies for depression diagnosis and monitoring," Journal on Multimodal User Interfaces, vol. 7(3) (2013) 217–228

[2] A practical guide to sentiment analysis by Erik Cambria, Dipankar Das, Sivaji Bandyopadhyay, Antinio Feraco (eds.) Springer (2017)

[3] Hemanthkumar M, Latha A, Depression Detection With Sentiment Analysis Of Tweets International Research Journal of Engineering and Technology (IRJET) 6(2019) 1197-1201

[4] Johann Hari, Lost connections uncovering the real causes of depression-and the unexpected solutions (2018) Published by Bloomsbury USA

[5] Namboodiri Sandhya Parameswaran, D. Venkataraman, 'A computer vision based image processing system for depression detection among students for counseling,' Indonesian Journal of Electrical Engineering and Computer Science 14(2019)503-512

[6] B. Liu, Sentiment analysis: Mining opinions, sentiments, and emotions,. Cambridge University Press, 2015.

[7] Renata L. Rosa, Gisele M. Schwartz, Wilson V. Ruggiero, and Demostenes Z. Rodríguez',A Knowledge-Based Recommendation System that includes Sentiment Analysis and Deep Learning', IEEE Transactions on Industrial Informatics 15(2019)2124-2135

[8] Mihai Gavrilescu and Nicolae Vizireanu Predicting Depression, Anxiety, and Stress Levels from Videos Using the Facial Action Coding System, Sensors(Basel) 19(2019)3693