JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1751 - 1756
https://publishoa.com
ISSN: 1309-3452

Mental Health Diagnostic System using Machine Learning Model

Shubhangi Vaikole[1], Smaran Ummadishetty[2], Anushka Vaidya[3], Chaitanya Keerthi[4]

[1]Associate Professor, Department of Computer Engineering, DattaMeghe College of Engineering, Airoli, India,
shubhangi.vaikole@dmce.ac.in

[2]UG student, Department of Computer Engineering, DattaMeghe College of Engineering, Airoli,India,
smaran.umadishetty@dmce.ac.in

[3]UG student, Department of Computer Engineering, DattaMeghe College of Engineering, Airoli, India,
alvaidya00@gmail.com
[4]UG student, Department of Computer Engineering, DattaMeghe College of Engineering, Airoli, India,
keerthichaitanyaa@gmail.com

Abstract

Mental health detection website that can artificially discover and detect mental disorders with the help of ML algorithms. A set of questionnaires and face with voice recognition will be used for the same. It will also contain a counseling section where manual counseling can be done which includes registration, payments, discounts and offers, chatbots, etc.

Keywords— Mental Health, Machine Learning, Speech Identification, Diagnosis, Analysis and Prediction.

I. INTRODUCTION

A. *Motivation*

The rate of increase in mental health illnesses has reached an all-time high. Early warning indicators, if addressed in a timely manner, can help you avoid severe losses in the future. These issues require 'From Home' which can be used to detect one's mental condition from home at their fingertips.

B. *Need of a Mental Health Analysis System*

Mental health is a condition of well-being in which a person is aware of his or her own potential, is able to cope with everyday challenges, is able to work successfully and fruitfully, and is a valuable asset for society. Mental health problems are on the rise all across the world in today's world of intense competition and few well-wishers. Mental health problems have increased by 13% in the recent decade.

A 13% increase is a big increase that cannot be overlooked; one in every five years is now spent disabled. It has severe repercussions, including anxiety, personality etc. A mental well being detection website that uses ML algorithms to automatically uncover and detect mental problems. For this, a collection of questionnaires as well as face and voice recognition will be employed.

II. DATA

Two major dataset used in this research are REVDESS Dataset and a local dataset by a counseling agency named Guiding Star. Guiding Star is a counseling firm situated in Mumbai, India, that specializes counseling in mental health difficulties and challenges. The dataset contains demographic, socioeconomic, nutritional, and health-related questions, as well as responses from a variety of people.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1751 - 1756
https://publishoa.com
ISSN: 1309-3452

The data and demographics from the questionnaires were used to learn more about the mental health, alcoholism, and drug use of the participants. Using correlation scores, features were chosen from many perspectives, including laboratory and inspection. There are 7356 recordings containing acted-emotional content. There are 3 types, one is complete audio visual. Second is only video and third is only audio. There are 2 audio files for song and speech. Every data has a one actor who represents 1 out of 8 emotions listed below: furious, cool, afraid, neutral, astonished, happy, distgused and sad excluding the emotion that shows neutrality, which comprises of regular intensity, these emotions are created at two degrees of emotional intensity, one is strong and other is regular.

We employed the whole Audio video data and the speech data in our studies as audio visual emotion are considered for identification in speech not music. The number of files was limited to 1440, with a largest and smallest period of 5.31 and 2.99 seconds, respectively. The recordings had 24 actors based on gender who said just 2 matching phrases based on lex in a neutral accent of North America, because of this it is ideal for studying the language of emotions, excluding the lexical, decreasing the cultural biaeds in emotional responses. It had a balanced amount of files present for each emotion, which eliminated issues that might arise when training algorithms with unbalanced data.

Apart from this we also used datasets from NIMH (National Institute of Mental Health) and Hopkins Medicine for training our model and to generate statistical analysis for our research. The data in the questionnaire dataset was primarily categorical, but the data in the other datasets was continuous. Each dataset included an explanation for each of the questions answered, making it easier for us to interpret the results.

A. Extraction

We had 150 datasets with between 10 and 48 questions in each. We combined these datasets into a single column, which is the Sequence number, which is unique to each person questioned. The dataset covers 2 years and contains over 100 characteristics for each row, with 2000 unique observation points for each year. Using the interviewee's unique identification, the datasets were combined as needed. The attributes that enable to merge information from multiple perspectives into a single representation are improved by using Multiview techniques on the data.

To better comprehend the data, categorical values were converted to integers and scaled.

Missing values were dealt with by deleting or imputing them based on why they were missing.

III. LITERATURE SURVEY

Chlastaa,. presented a method for detecting depression in voice using deep convolutional neural networks. We investigated 5 network topologies and found ResNet 50 and ResNe 34 to be the most effective at classifying data. The findings point to a possible new avenue in the use of audio spectrograms for early screening of depressed people utilizing brief speech samples. The spectrograms demonstrated the ability to generate CNN learnable features. Using the TTA approach, the algorithm achieved accuracies of 70% and 77 percent. Despite the difficulty of using voice for predicting depression, this was the case. Less size for sample (fifteen seconds) were employed in the solution, which, in our opinion, reduced the influence of noise.

There have been a few studies that have used Reddit to look at linguistic elements of mental health. Geographies, communities, gender, and age groups all have different cultural contexts in which depression is expressed. The linguistic aspects of the output language must be considered for text measures for depression that may be aggregated over population for inter lingual, as demonstrated by their examination of personal pronoun incidence in two distinct languages. The incapacity of the features to take word order into account is one of the work's limitations. It would be interesting to run a similar experiment with attention-based processes and see what happens. Their sadness and non-depression ML and interpretable Artificial Intelligence text data present in English as well as translated English using Google and its Urdu form shows: (1) When person is depressed then increased use of pronouns used, (2) Difference in Natural language processing outcome and text used in social media text in lingual text internally and in position (3) Large ratio of posts of depression to non depression, and (4) Sensitivity with respect to data processing techniques.

IV. Proposed Methodology

We use survey-based ground truth metrics from the Guiding star dataset and RAVDESS dataset to construct techniques for measuring subjective well-being, which is quantified in the form of graphical demographic which indicates the possibility of the certain mental condition.
We ran some Empirical Data Analysis with the preprocessed data to better understand the data and determine whether itactually helps us reach our aim. By adjusting for age, gender, and sexual orientation, we can see substantial connections.Principal Component Analysis (PCA) was implemented for exploratory data analysis,. Females appear to be more sad than guys, according to research. While the 18-25 age group has higher rates of mental illness than other age groups. When age categories are controlled for, mid-aged Non-Hispanic Black persons and elderly Non-Hispanic White people exhibit greater indicators of mental health conditions. We developed a hypothesis and evaluated it with a T test, compensating for multiple comparisons by adding a Benjamini—Hochberg false discovery rate adjustment to the significance level (p.05).

A. Questionnaire

This part comprises multiple-choice questions (MCQs) with 4-5 potential responses. These professionally developed 24 MCQs can help you understand a person's mental health. The responses will then be examined and evaluated to find a pattern that corresponds to the behavioral style or habits of a person suffering from mental illness. The findings, as well as counseling ideas and tips to help them improve their mental health, will be shown on the interface.

On our data, we employed a feature selection process. We use PySpark and HDFS to run Linear regression on 160 datasets (with the number of features ranging from 10 to 48) on a 2-year pooled dataset after preprocessing the data. Using the Spark framework to load and analyze the multistage dataset allowed for the generation of test statistics and beta values through linear regression. We may choose the most linked characteristics and relevant datasets for the prediction job by looking at the top 50 beta values. Tensorflow is used to execute multivariate LR on the Questionnaire dataset with the following hyperparameters (10 iterations, 0.001 rate of learning,0.9 for momentum).

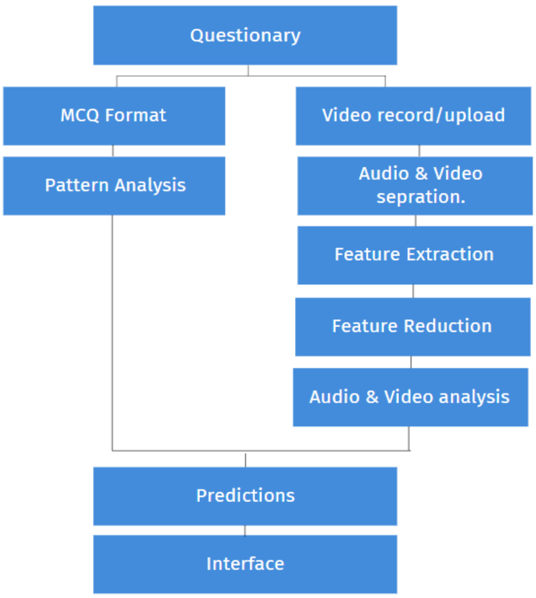

Fig 1: Flowchart for Prediction of Mental Health Conditions.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1751 - 1756
https://publishoa.com
ISSN: 1309-3452

We were then able to get the data using node js frameworks and display a graphical demographic on the UI displaying the likelihood of a certain mental disease depending on the responses to the questionnaire.

B. Analysis using text, audio and video

The spoken emotion recognizer and the face emotion recognizer were two systems in our proposed framework for mental health detection and analysis. We used a late fusion method to aggregate the outcomes of these subsystems.

Figure 1 depicts a block schematic of our systems' primary components as well as a synopsis of the comparison of the 2 techniques: atype of models, in the face emotion recognizer, the use of maximum voting or a bi-LSTM

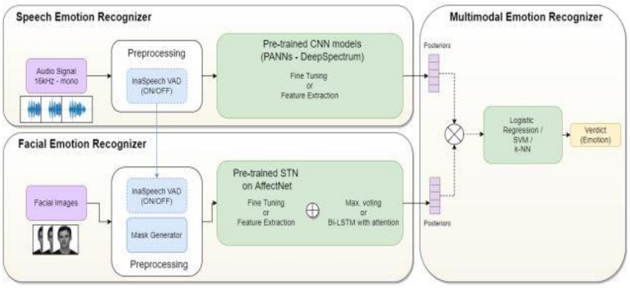

Fig 2: Block diagram of the implemented systems.

We employed the whole Audio video data and the speech data in our studies as audio visual emotion are considered for identification in speech not music. The number of files was limited to 1440, with a largest and smallest period of 5.31 and 2.99 seconds, respectively. The recordings had 24 actors based on gender who said just 2 matching phrases based on lex in a neutral accent of North America, because of this it is ideal for studying the language of emotions, excluding the lexical, decreasing the cultural biaeds in emotional responses. It had a balanced amount of files present for each emotion, which eliminated issues that might arise when training algorithms with unbalanced data.

This dataset poses significant hurdles in the area of emotion identification. The human accuracy rate achieved is confirmation of this: 67 percent utilizing only voice inputs and 75 percent using visual information.

A Long Short-Term Memory network is used in this Recurrent Neural Network. These methods are used to process inputs in order, being performed on each

Given the many components that made up our input sequence, where ht is the concealed state, t denotes the time step, while W denotes the network's weights.We fed the embeddings or posteriors generated by the model as an input. STNs for each video frame (a1, a2,..., aN) to discover the temporal relationships from the video.

The STN finds spatial similarities and makes a final prediction at the video level.

Each bi-LSTM layer functions in a bidirectional manner, allowing us to acquire sequence information in both directions from the LSTMs' hidden states h1, h2,..., hN. A Bi-LSTM, in instance, is made up of two LSTMs: a forward LSTM for analysing frames from x1 to xN, and an inverse or reverse LSTM for analysing frames in the opposite manner, from aN toax1.

We combined the outputs from the analysis done in each specific direction for each embedding of a frame to produce the emotional tag from our bi-LSTM layer relates to the concatenation operator and L to the size of each LSTM).

hi = −→hi ||←−hi, where hi ∈ R 2L ……. Equation (1)

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1751 - 1756
https://publishoa.com
ISSN: 1309-3452

V. EXPERIMENTAL RESULTS

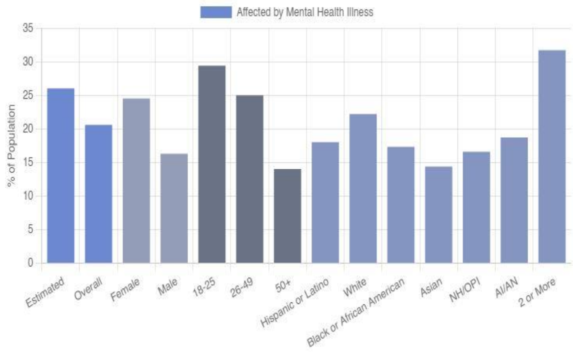

Fig 3 : Statistics of Mental Health Conditions among different categories.

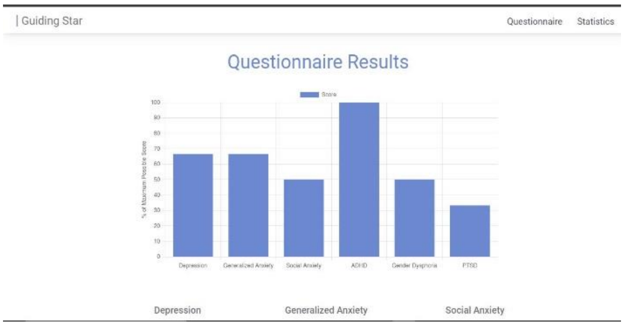

Fig 4 : Maximum marks possible for each mental health issue
Questionnaire results

VI. CONCLUSION

The system strives to make society into a more positive place with fewer mental health issues, more encouragement, and support, all of which will enhance people's quality of life. Better learning and working performance will come from improved attention, more creativity, reduced absenteeism, and higher productivity. It also reduces our demand for certain medical treatments in the future.

There are several problems, such as resources, scope, and so on. Multilanguage collaboration and a high-quality dataset are two of the most significant hurdles.

As human emotions can be acted or say forged, for later enhancement of the system we can eliminate the acted or forged samples from the real/authentic ones. We can create a chatbot that converts between human and computer and answers the user's inquiry. Any user may utilize the chatbot and can ask their questions at any time. The same topic was posed in several ways to test the chatbot's accuracy, including modifying the phrasing of phrases and attaching various special characters and extraneous terms. Artificial intelligence and machine learning are fields that have yet to be fully explored. Developers will have an easier time adding additional complications to the analytic system and producing more precise and well-defined findings.

REFERENCES

[1] J. F. Curtis, (Ed.), Deep Neural Networks for Depression Recognition Based on 2D and 3D Facial Expressions Under Emotional Stimulus Tasks. New York: Harper and Row, 1978. II.

JOURNAL OF ALGEBRAIC STATISTICS
Volume 13, No. 2, 2022, p. 1751 - 1756
https://publishoa.com
ISSN: 1309-3452

[2] [n.d.].https://www.who.int/mental_health/management/depression/wfmh_
paper_depression_wmhd_2012.pdf.

[3] Farah Adeeba and Sarmad Hussain. 2011. Experiences in Building Urdu WordNet. In Proceedings of the 9th Workshop on Asian Language Resources. Asian Federation of Natural Language Processing, Chiang Mai, Thailand, 31– 35. IV.

[4] Sonnenschein AR, Hofmann SG, Ziegelmayer T, and Lutz W. 2018. Linguistic analysis of patients with mood and anxiety disorders during cognitive behavioral therapy. CognBehavTher. (2018) G.

[5] 6 Chlasta, Karol &Wołk, Krzysztof &Krejtz, Izabela. (2019). Automated speech-based screening of depression using deep convolutional neural networks. Procedia Computer Science. 164. 618-628. 10.1016/j.procs.2019.12.228.

[6] Srimadhur, N.S. &Sreeram, Lalitha. (2020). An End-to-End Model for Detection and Assessment of Depression Levels using Speech. Procedia Computer Science. 171. 12-21. 10.1016/j.procs.2020.04.003.

[7] Graham, Sarah & Depp, Colin & Lee, Ellen &Nebeker, Camille & Tu, Xin & Kim, Ho-Cheol &Jeste, Dilip. (2019). Artificial Intelligence for Mental Health and Mental Illnesses: an Overview. Current Psychiatry Reports. 21. 10.1007/s11920-019-1094-0.

[8] Karol Chlastaa, Krzysztof Wołka, Izabela Krejtz. (2019). Automated speech-based screening of depression using deepconvolutional neural networks. Karol Chlasta et al. / Procedia Computer Science 00 (2019) 000–000.

[9] Sarah Graham, Colin Depp, Ellen E. Lee, Camille Nebeker, Xin Tu, Ho-Cheol Kim, Dilip V. Jeste. Artificial Intelligence for Mental Health and MentalIllnesses: an Overview. Current Psychiatry Reports (2019) 21:116 https://doi.org/10.1007/s11920-019-1094-0.

[10] Steven R. Livingstone, Frank A. Russo. (2018). The Ryerson Audio-Visual Database of Emotional Speech and Song (RAVDESS): A dynamic, multimodal set of facial and vocal expressions in North American English.