Depression Detection Through Hybrid Learning Algorithms Using Multi-Model Data

Mrs. M. Therasa M.E.,(PhD) Panimalar Enginering College
Merlin White.D, Nivetha.K, Varshini.B Department Of Computer Science And Engineering
Panimalar Institute Of Technology Chennai, India

***Abstract*—** Nowadays, most of the people are suffering from depression due to changes in lifestyle, increased academic expectations, genetic inheritance, improper diet and sleep, etc., Depression is a Bipolar disorder, where the feeling of sadness, emptiness, anger, helplessness and pessimism acts as an indicator. The seriousness of depression can be identified by the quantity of symptoms and their time period. Depression can also lead to serious medical illness like heart diseases, diabetes, cancer, etc., So, The earlier the diagnosis of depression, the better the chance of treating people who are affected. The detection is done using Data Science, which helps to encompass various methods and draw out insights from data. Data is collected in different forms like Text, Speech and Video. After getting the input, data will be processed using multiple Machine Learning Algorithms. Based on the emotions expressed by the people, the level of depression will be detected and displayed on the Browser.

Keywords-Depression, Insights, ML Algorithms

I. INTRODUCTION

Depression, commonly referred to as Clinical Depression, is a Bipolar condition that is common but significant. Type of depression like major depression, perinatal depression, etc., will vary among people due to different conditions such as mental strength, stability and physical conditions. Also, Mild, Moderate and Severe are the common levels of depression among people. So, there is no common treatment for depression and it differs according to their symptoms, severity, etc.,

The depression should be detected at the earlier stage for providing proper medication to the individuals. In this paper, the degree of depression will be detected using Data Science. The field of study known as “Data Science” works with enormous amount of data to gain insights using a variety of approaches and techniques. Data Science builds useful predictive and detective models using complex Machine Learning Algorithms.

The dataset utilised to identify depression is made up of text, audio signals and images that are gathered from social media platforms. The gathered Dataset will be used to train a model that helps us to detect the depression effectively. Different Machine Learning Algorithms are used for three different kinds of data such as text, audio and image. It helps us to detect the emotions from our stanza, facial expressions and voice modulation.

The HAAR Cascading Algorithm is used for recognizing the facial expressions from the image. The fuzzy color filter and histogram analysis approach are utilized in the face region extraction algorithm to recognize the skin tone. This technique will separate the entire image into three feature regions including the eye region, mouth region, and auxiliary region. Each zone yields a variety of information like geometric and shape information. By using this information, the degree of depression will be determined.

PyAudio is a Python library for working with audio streams that can be used in speech emotion recognition. After recording the audio signals, features are extracted from the audio signal that are indicative of emotions. Then by using machine learning algorithms such as Support Vector Machines (SVM) and Convolutional Neural Network – Long Short Term Memory Model (CNN-LSTM), we detect the depression.

Support Vector Machines (SVM) and NLP (Natural Language Processing) Algorithm can be used for text emotion recognition, which is the task of determining the emotional state of a writer from text. These algorithms can extract various features from text, such as word choice, sentiment, and syntax, that can be used to predict emotions, which in turn is used to detect the depression.

II. METHODOLOGY

In this paper, input is collected in three different forms such as text, audio and video. The system should be trained for those multi model data by using different datasets. For speech recognition, the data set used for training contains 24 experienced actors voicing two lexically-matched sentences in a neutral North American accent. For text recognition, a total of 2,468 daily writing contributions from 34 psychology students make up the text recognition dataset. For video recognition, the dataset consists of 48x48 pixel grayscale portraits of people. After collecting those dataset, the system is trained using various Machine Learning Algorithms.

A. *HAAR Cascading Algorithm*

The Haar algorithm is a feature-based object detection algorithm, which is used for video recognition. Object detection is identifying an object in an image. This algorithm can run in real-time and vast computation was not required. Regardless of the size and position of the image, it can detect objects. The algorithm deals with a lot of positive images and negative images to train the model. Positive image consists of faces and Negative image does not contain any face. These models are stored in XML files. Using this model, the gathered video input will be processed in the following way. In the HAAR feature, the sum of all the image pixels in the darker region and the sum of all the image pixels in the lighter region were discovered. Their difference was later found. Now, the HAAR value will be closer to 1 if the image has an edge separating dark pixels and light pixels, otherwise, the HAAR value will be far from 1.

B. *X-ception Model*

X-ception stands for Extreme Inception, is a Convolutional Neural Network. This model is used for video recognition which classify the images into object categories. This model is based on two major concepts namely, Depthwise Separable Convolution and Shortcuts between Convolutional block, similar to ResNet. A vast number of towers make up the inception module called Depthwise Separable Convolution. It performs convolutional computation one by one and preceded by pointwise convolution. It consist of a kernel that iterates through every single point. The pointwise convolution is used in integration with depthwise convolution to produce an effective Convolution called as depthwise separable convolution. ResNet stands for Residual Network, a kind of Convolutional Neural Network. It uses a concept called shortcut connection which skips some of the layers of the network by forming a residual network from a regular network.

C. *CNN-LSTM Algorithm*

Convolutional Neural Network is a deep learning algorithm, which is used for speech recognition. CNN uses a special approach known as Convolution, that involves the mathematical operation of two functions to create a third function. Convolutional Neural Network are fed with a group of signals to know some basics such as changes in frequency and amplitude. CNN has multiple layers, the first layer is provided with information. The second layer is provided with some recognizable features. By using these layer information, CNN decides either the signal has frequency change or amplitude change. Long Short Term Memory Model is designed to process or handle sequential data. The Recurrent Neural Network variant known as LSTM is effective in establishing long-term dependencies in sequential problem. LSTM creates a memory cell, which holds information for certain period of time. The three gates used to control the memory cell are Input gate which adds information to memory cell, Forget gate which removes information from memory cell and Output gate which provides output from the memory cell. The LSTM is used in combination with CNN for effective recognition of speech.

D. *Natural Language Processing*

Natural Language Processing is widely used to understand and elucidate the human language. The basic step performed in NLP is Data pre-processing which helps to get better results. This transforms raw data into a format that can be understandable by the NLP models. The Techniques such as tokenization, word stemming are used to resolve incomplete or inconsistent data. Removing inconsistent data involves, removing blank rows, stop words, non-alpha text, etc., The taken dataset will be split into two namely, train and test dataset. The test dataset is used to determine whether the model is performing as predicted after the model has been built using the training dataset. The model can understand only the numerical values so the categorical data should be converted. The result will be predicted using the Term Frequency and Inverse Document Frequency(TF-IDF) method, which turns text data into vectors. The two important models used in Natural Language Processing are Statistical model and Neural Language model. Natural Language Processing is popularly used for text recognition.

E. *SVM Algorithm*

A supervised learning technique called the Support Vector Machine may be utilised to resolve Classification and Regression issues. In SVM Algorithm, a decision boundary will be created that helps to separate n-dimensional space into classes. Due to the fact that this algorithm's extreme points are known as support vectors, it is known as the support vector machine and the decision boundary is known as hyperplane. The basic goal of SVM algorithm is to determine the optimum hyperplane with the largest margin distance between the closest points of two classes. This helps to minimize the error generalization in the classifier. An SVM algorithm produces a model that determines whether a new example falls into one of two classes given a collection of training examples, each labelled as belonging to one of two classes. Following that, new instances are mapped into the same area and assigned to the class according to which side of the gap they fall. In this paper, SVM Algorithm is used for both text and speech recognition.

F. *Naïve Bayes Algorithm*

Naïve Bayes Algorithm is a probabilistic classifier which is used for solving classification problems. It is called as Naïve because the occurrence of one feature is independent of other features and termed as Bayes as it follows the principles of Bayes Theorem. To determine the likelihood of the hypothesis, apply the Bayes Theorem. The Bayes Theorem has the following formula,

$$P(R|S) = \frac{P(S|R)\,P(R)}{P(S)}$$

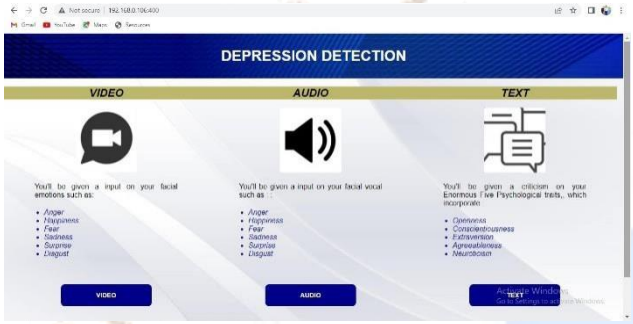
Where,
P(R|S) is the posterior probability

P(S|R) is the likelihood probability
P(R) is the prior probability
P(S) is the marginal probability

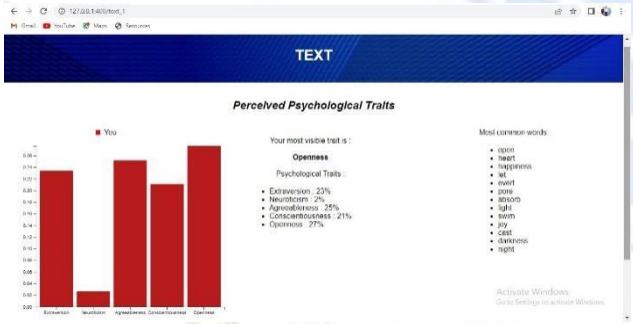
To solve a problem using Naïve Bayes Algorithm, first the dataset is converted into frequency tables. Then, by using the given features the likelihood table will be generated through the probabilities. Finally, the posterior probability was calculated using Bayes Theorem.

III. RESULTS AND DISCUSSION

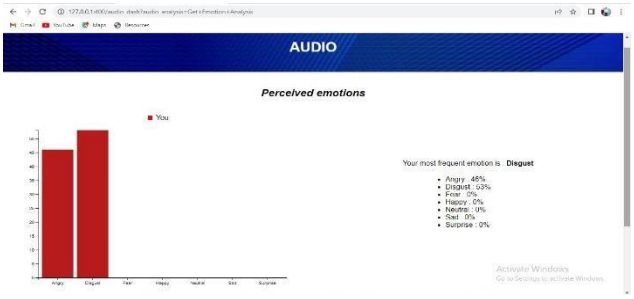
The goal of this paper is to detect the depression among the users using three distinct datasets. First, the input is collected in three different forms like text, audio and video. The outcomes and performance of the aforementioned Machine Learning Algorithms are discussed. The final result will be displayed in the web browser as follows,



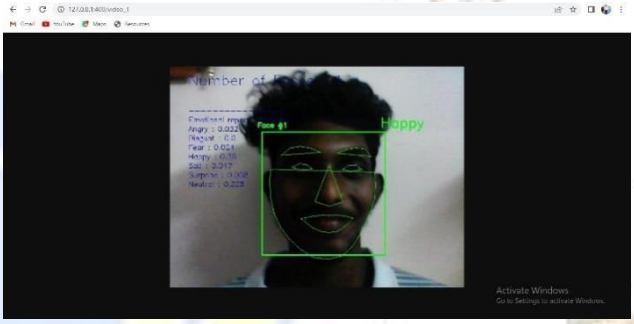
For text recognition, a total of 2,468 daily writing contributions from 34 psychology students make up the text recognition dataset and the system was trained using this dataset. With the help of Natural Language Processing, the text input will be converted into a vector format and it will be compared with the trained model to provide the depression level. The result will be displayed in percentage with the following psychological traits. They are openness, conscientiousness, neuroticism, agreeableness.



For speech recognition, the data set used for training the system contains 24 experienced actors vocalizing two lexically-matched sentences in a neutral North American accent. With the help CNN-LSTM and SVM Algorithms, essential features are extracted from the audio input that are indicative of emotions. These features are compared with the trained model and the detected depression will be displayed as graph with the following traits. They are surprise, sad, disgust, happy, fear, neutral and angry.



For video recognition the dataset contains 48x48 pixel grayscale portraits of people and this dataset is used to train the system. By using HAAR Cascading Algorithm, objects present in the image will be identified. During the image processing stage, the facial region and its facial components are extracted. This will partition the whole image into eye, mouth and auxiliary region. A variety of data, including geometric and shape information is taken from each area to determine the mood. The result will be displayed with some traits like angry, sad, happy, neutral, disgust, fear and surprise.



IV. CONCLUSION

Depression, a serious mental condition, is characterized by the feeling of sorrow, idleness, rage, helplessness and pessimism acts as an indicator. Through this project, the detection of depression will be performed by collecting input in various forms like Text, Speech and Video. The analytics will be performed on the selected dataset and it will provide some insight. Then by using these datasets, the system will be trained. After getting the input, we need to detect the emotions by identifying some essential features from the gathered input. Then, we apply some algorithms like HAAR Cascading Algorithm, NLP Algorithms, SVM Algorithm, CNN-LSTM Algorithm, X-ception Model and Naïve Bayes Algorithm to accurately detect the depression among people. Finally, the depression detected will be displayed on the web browser. This system helps the user to identify the depression at the earlier stage.

REFERENCES

[1] Georgescu, M.I., Ionescu, R.T. and Popescu, M., 2019. Local learning with deep and handcrafted features for facial expression recognition. IEEE Access, 7, pp.64827-64836.
[2] Huang, K.Y., Wu, C.H., Hong, Q.B., Su, M.H. and Chen, Y.H., 2019, May. Speech emotion recognition using deep neural network considering verbal and nonverbal speech sounds. In ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 5866-5870). IEEE.

[3] Burdisso, S.G., Errecalde, M. and Montes-y-Gómez, M., 2019. A text classification framework for simple and effective early depression detection over social media streams. Expert Systems with Applications, 133, pp.182-197.
[4] Katchapakirin, K., Wongpatikaseree, K., Yomaboot, P. and Kaewpitakkun, Y., 2018, July. Facebook social media for depression detection in the Thai community. In 2018 15th international joint conference on computer science and software engineering (jcsse) (pp. 1-6). IEEE.
[5] Stankevich, M., Latyshev, A., Kuminskaya, E., Smirnov, I. and Grigoriev, O., 2019, October. Depression detection from social media texts. In Elizarov, A., Novikov, B., Stupnikov., S (eds.) Data Analytics and Management in Data Intensive Domains: XXI International Conference DAMDID/RCDL (p. 352).
[6] Babu, N.V. and Kanaga, E.G.M., 2022. Sentiment analysis in social media data for depression detection using artificial intelligence: a review. SN Computer Science, 3, pp.1-20.
[7] Amanat, A., Rizwan, M., Javed, A.R., Abdelhaq, M., Alsaqour, R., Pandya, S. and Uddin, M., 2022. Deep learning for depression detection from textual data. Electronics, 11(5), p.676.
[8] Sardari, S., Nakisa, B., Rastgoo, M.N. and Eklund, P., 2022. Audio based depression detection using Convolutional Autoencoder. Expert Systems with Applications, 189, p.116076.
[9] Yang, K., Zhang, T. and Ananiadou, S., 2022. A mental state Knowledge–aware and Contrastive Network for early stress and depression detection on social media. Information Processing & Management, 59(4), p.102961.
[10] Shen, Y., Yang, H. and Lin, L., 2022, May. Automatic depression detection: An emotional audio-textual corpus and a gru/bilstm-based model. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 6247-6251). IEEE.
[11] Ansari, L., Ji, S., Chen, Q. and Cambria, E., 2022. Ensemble hybrid learning methods for automated depression detection. IEEE Transactions on Computational Social Systems.
[12] Liu, D., Feng, X.L., Ahmed, F., Shahid, M. and Guo, J., 2022. Detecting and measuring depression on social media using a machine learning approach: systematic review. JMIR Mental Health, 9(3), p.e27244.
[13] Loh, H.W., Ooi, C.P., Aydemir, E., Tuncer, T., Dogan, S. and Acharya, U.R., 2022. Decision support system for major depression detection using spectrogram and convolution neural network with EEG signals. Expert Systems, 39(3), p.e12773.
[14] Salas-Zárate, R., Alor-Hernández, G., Salas-Zárate, M.D.P., Paredes-Valverde, M.A., Bustos-López, M. and Sánchez-Cervantes, J.L., 2022, February. Detecting depression signs on social media: a systematic literature review. In Healthcare (Vol. 10, No. 2, p. 291). MDPI.
[15] Angskun, J., Tipprasert, S. and Angskun, T., 2022. Big data analytics on social networks for real-time depression detection. Journal of Big Data, 9(1), p.69.
[16] Nguyen, T., Yates, A., Zirikly, A., Desmet, B. and Cohan, A., 2022. Improving the generalizability of depression detection by leveraging clinical questionnaires. arXiv preprint arXiv:2204.10432.
[17] Joshi, M.L. and Kanoongo, N., 2022. Depression detection using emotional artificial intelligence and machine learning: A closer review. Materials Today: Proceedings, 58, pp.217-226.
[18] Ravi, V., Wang, J., Flint, J. and Alwan, A., 2022, May. Fraug: A frame rate based data augmentation method for depression detection from speech signals. In ICASSP 2022-2022 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 6267-6271). IEEE.
[19] Lia, R.J., Siddikk, A.B., Muntasir, F., Rahman, S.S.M.M. and Jahan, N., 2022. Depression detection from social media using Twitter's tweet. In Big Data Intelligence for Smart Applications (pp. 209-226). Cham: Springer International Publishing.
[20] Safa, R., Bayat, P. and Moghtader, L., 2022. Automatic detection of depression symptoms in twitter using multimodal analysis. The Journal of Supercomputing, 78(4), pp.4709-4744.
[21] Sardari, S., Nakisa, B., Rastgoo, M.N. and Eklund, P., 2022. Audio based depression detection using Convolutional Autoencoder. Expert Systems with Applications, 189, p.116076.
[22] Stankevich, M., Isakov, V., Devyatkin, D. and Smirnov, I.V., 2018, January. Feature Engineering for Depression Detection in Social Media. In ICPRAM (pp. 426-431).
[23] Hussain, J., Satti, F.A., Afzal, M., Khan, W.A., Bilal, H.S.M., Ansaar, M.Z., Ahmad, H.F., Hur, T., Bang, J., Kim, J.I. and Park, G.H., 2020. Exploring the dominant features of social media for depression detection. Journal of Information Science, 46(6), pp.739-759.
[24] Arora, P. and Arora, P., 2019, March. Mining twitter data for depression detection. In 2019 international conference on signal processing and communication (ICSC) (pp. 186-189). IEEE.
[25] De Melo, W.C., Granger, E. and Hadid, A., 2019, September. Depression detection based on deep distribution learning. In 2019 IEEE international conference on image processing (ICIP) (pp. 4544-4548). IEEE.