ICEST-2K24 [International Conference on Engineering, Science and Technology]
In Association with
International Journal of Scientific Research in Science, Engineering and Technology
Print ISSN - 2395-1990 | Online ISSN : 2394-4099 (www.ijsrset.com)

Sentiment Analysis of Dysthymia Form of Depression Using Multimodal Approach

Prof. D.R. Kamble[1], Mr. Vishal K. Madane[2], Mr. Sandip M. Jagtap[2], Mr. Kshitij S, Kale[2], Mr. Yogesh S. Thorat[2]
*[1]Assistant Professor, Department of Computer, SBPCOE Indapur, Maharashtra, India
[2]B E Computer Student, Department of Computer, SBPCOE Indapur, Maharashtra, India

ABSTRACT
To investigate the correlation between linguistic traits and dysthymia symptoms, researchers have compiled a comprehensive set of speech features, drawing on previous research findings. These features encompass various aspects of speech, including prosody, pitch, intonation, speech rate, and pauses, among others. Analysing these linguistic cues can provide valuable insights into an individual's emotional state and mental well-being. One commonly used technique for assessing the severity of depression is the Beck Depression Inventory (BDI). This approach involves individuals answering specific questions that help gauge their level of depression. By incorporating such self-report measures into the analysis of linguistic traits, a more holistic assessment of dysthymia can be obtained. The combination of objective linguistic analysis and subjective self-assessment measures can enhance the accuracy and reliability of dysthymia detection. We have one solution to solve this issue in terms of machine learning.

Keywords: Machine Learning, Video Processing, Facial Expressions, Sentiment Analysis, fusion Algorithm, KNN, Depression Detection, Bag-of-Words.

I. INTRODUCTION

Every Human being in day to day life is being diagnosed with depression due to affection of different parameters. It disturbed mental state of the human being, So as consider to technology we have one solution to solve this issue in terms of machine learning. Machine learning is a process which leams from past experience and provide the best result when the same issue or event occurs in the future.[2] It considers different parameters like user emotions. Depression is a leading cause of mental ill health. It is a major cause of suicidal ideation and leads to significant impairment in daily life. Machine Learning can help detection and can generate possible solutions to tackle depression. Suicide is one of the most serious social health issues that exists in today's culture. Suicidal ideation, also known as suicidal thoughts, refers to people's plans to commit suicide. Our motivation is to find out a speech feature set to detect, evaluate and even predict Dysthymia. For examining the correlation between Dysthymia and speech, we extract features as many as possible according to previous research to create a large voice feature set. It can be used as a suicide risk measure. India Depression is a mental illness that is not taken seriously in some countries that can cause us depression.[1] among the top countries among in the world to have annual suicide rate. Depression is a psychiatric disorder that needs to be addressed with medication. According to Our World in Data Website, Depressive disorders occur with varying severity. The WHO'S International Classification of Diseases defines this set of disorders ranging from mild to moderate to severe. The Institute of Health Metrics and Evaluation adopt such definitions by disaggregating to mild, persistent depression (dysthymia) and major depressive order (severe). Though this project can use varieties of techniques such as facial expression detection, social media feeds. Question naire, etc. to target and identify users depression levels. We limit the scope of this project by using only facial expression detection[10] and questionnaire based solution to tackle depression.

II. LITERATURE SURVEY

"A Survey of Multimodal Sentiment Analysis"- Mohammad Soleymani, David Garcia, Brendan Jou, Bjorn Schuller, Shih-Fu Chang, Maja Pantic* In this paper, it represent an overview of concept and the goal of multimodal sentiment analysis and discussed about the challenges and perspectives related to above field. Sentiment analysis is a promising approach to complementary channels of information for sentiment analysis such as recognition and subjective analysis.

"Hierarchical Attention Network for Document Classification"- Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E* In this paper, it represents Hierarchical Attention Network (HAN) for classify documents. For better visualization use highly informative components of a documents. kind of pagination anywhere in the paper. Donot number text heads-the template will do that for you. Picking out important sentences and words from the documents.

"Multimodal Sentiment Analysis Based on Deep Learning"-J. Bian, L. Rajamanickam, Z. Nopiah*. In this paper, it shows on the basis of characteristics of deep learning algorithm. There are various fusion methods that implements sentiment analysis. The CNN model is used which is pre-trained on large scale image data set and then send to train model to train the text emotions classification model. Result is obtained by decision fusion.

"Multimodal Sentiment System and Method Based on CRNN-SVM"- Y. Zhao, M. Mamat, A. Aysa, K. Ubul*. This paper proposed as an AI deep learning method that is used in sentiment analysis for multimodal approach. It is improving recognition rate and analysis accuracy of sentiments. In this paper system performance optimization method is tested, and remarkable result achieved.

"Multimodal Emotion Recognition Model Based on a Deep Neural Network with Multi-objective Optimization"- M. Li, X. Qiu, S. Peng, L. Tang, W. Yang, y. Ma* This paper represents a multimodal emotion recognition model based on multimodal objective algorithm. It will gives accuracy and result at a same time. It is best improvement for emotion recognition model.

"Understanding and measuring psychological stress using social media"- Guntuku, S.C., Buffone, A., Jaidka, K., & Eichstaedt, J.C* In this paper, as a sample take an Social media account like twitter or instagram, check the stress individually. The result show in LIWC. The result also shows that psychological survey data by deep understanding the environment.

"A Survey on Multimodal Sentiment Analysis"- S.J. Fulse* In this paper it is show that the sentiment analysis multi-modal problem is a problem occurs in machine learning. There are many difficulties to sentiment analysis as Cultural influence, linguistic variation and it is difficult to derive sentiment.

"Multimodal sentiment analysis: Addressing key issues and setting up the baselines."- Poria, S., Cambria, E., & Bajpai, R.* This paper shows useful baseline for the multimodal sentiment analysis for emotion detection. It having different aspects as multimodal sentiment analysis problem like cross-dataset, unknown speaker etc.

"An MLP-based Model for Multimodal Sentiment Analysis and Depression Estimation"- Hao Sun, Hongyi Wang, Jiaqing Liu, Yen-Wei Chen, Lanfen Lin *. In this paper, treat multimodal fusion as feature mixing and propose the MLP-based Cube MLP for unified multimodal feature processing. In Cube MLP, we perform the mix-up at all axis of multimodal features. Cube MLP can reach the state-of-the-art performance for sentiment analysis and depression detection while keeping the computational burden low. We analysed Cube MLP's components and compared it to other techniques.

"Detecting depression of microblog users via text analysis"- Sihua Lyu, Xiaopeng Ren, Yihua Du and Nan Zhao*. This study found that depression could be detected solely through word frequency features by machine learning methods. This model could have potential value in the screening for depression and be able to generalized across platforms. Furthermore, our study demonstrated that in addition to LIWC, which was commonly used in previous studies, lexicons related to cultural psychology and suicide risk.

III.PROPOSED SYSTEM

A. Problem Statement

To develop a user centric application program which addresses growing problem of depression in teenagers. Basically, to design and develop an application which can be helpful to the normal user where machine learning is playing a big role to calculate the depression level of the user according to the user input or face expression detection (parameters like face edges and audio).

B. Block Diagram

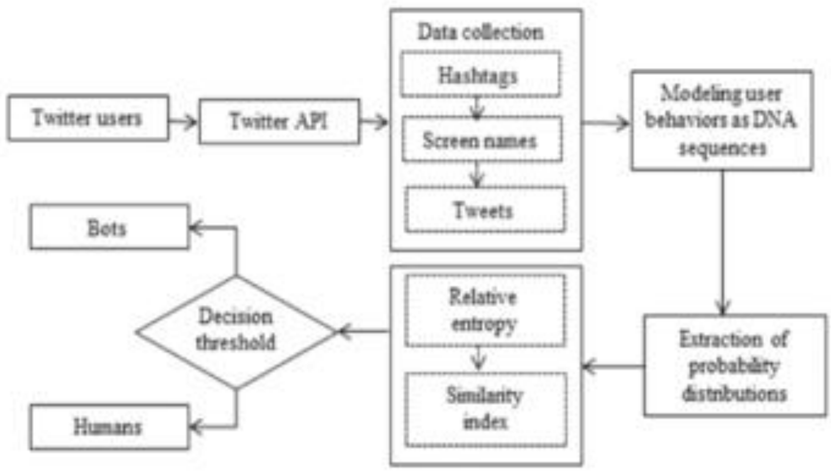

Figure 1: Block Diagram

C. Software Requirement

- Operating System – Windows
- Application Server - Apache Tomcat
- Back End – Python
- Database - My SQL
- IDE – VS Code

D. Hardware Requirement

- Processor - Intel i3/i5/i7
- Speed - 3.1 GHz
- RAM - 4 GB(min)
- Hard Disk - 30 GB

E. Algorithm for System

Sentiment analysis for the detection of dysthymia, a form of depression, using a multimodal approach involves analysing various types of data, such as text, audio, and possibly visual information. Here are the algorithmic steps along with descriptions for a multimodal approach to dysthymia detection:

1) Input Data Collection: Gather multimodal data from individuals, including textual data (e.g., social media posts, messages), audio recordings (speech patterns, intonations), and potentially visual data (facial expressions, body language).
2) Preprocessing: Clean and preprocess each modality of data. This includes text cleaning (removing stop words, stemming), audio preprocessing (feature extraction, normalization), and visual data preprocessing (if applicable).
3) Textual Analysis: Utilize natural language processing (NLP) techniques to extract features from textual data. This may involve sentiment analysis, emotion detection, and other relevant NLP tasks.
4) Audio Analysis: Extract features from the audio data using signal processing techniques. Focus on aspects such as pitch, tone, speech rate, and other acoustic features indicative of emotional states.
5) Multimodal Integration: Combine the features extracted from different modalities into a unified representation. Techniques such as feature concatenation, fusion, or attention mechanisms can be employed for effective integration.
6) Model Training: Train a machine learning or deep learning model using the preprocessed and integrated multimodal data. Consider using algorithms capable of handling multimodal inputs, such as multimodal neural networks or ensemble models.
7) Cross-Validation: Validate the model's performance using cross-validation techniques to ensure robustness and prevent overfitting.
8) Output: Deploy the trained model for real-world applications, ensuring that it integrates seamlessly with the target environment, whether it be a mental health app, social media monitoring system, or other platforms.

IV. RESULT

A. Input Location

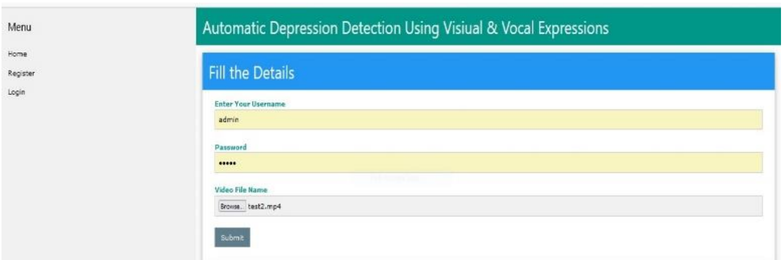

Figure 2: Input Page

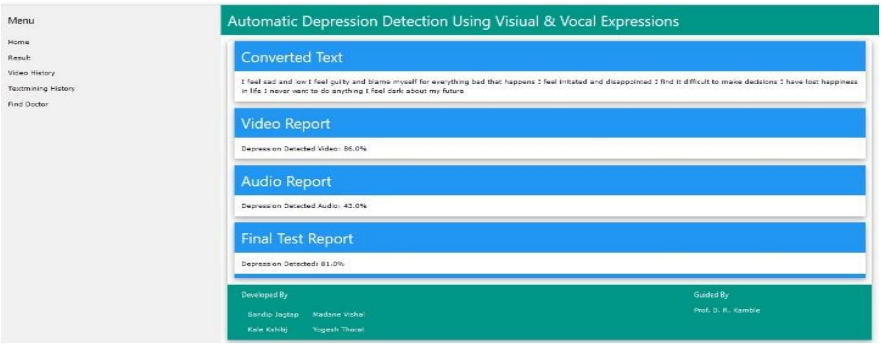

Figure 2: Output Page

V. RESULT DISCUSSION

Pre-Processing the Emotion Recognition Challenge dataset is impacted by anomalous circumstances including light fluctuation, facial occlusion, and more. Pre-processing is in our architecture. Our first step is to align, normalise, and resize the to 224x224 px, the face. Next, we select frames from video every 2sec interval to use as inputs for the facial image model. Similar to this, after employing 98 facial landmarks initially, we entered 3 frames into the facial landmarks model. For training, we utilize Cloud Speech-to-Text API to convert the audio into text format to extract features from the audio model.

B. Conclusion and Discussion We use several neural networks to extract complementing features for the Multimodal based emotion identification challenge in order to get better performance. We have primarily used two types of fusion approaches for our research: fusion by segments and fusion by video. Additionally, and include a variety of fusion techniques, such as voting on the highest value and the score produced by each model. Experiments reveal that the latter has superior accuracy. The outcomes of the various networks that we presented are shown the Facial Landmarks Model achieved an accuracy of 70.00%. while the Facial Expression Model was 70.71 percent accurate, that of the Audio Model is 49.29%, It's vital to note that although while the audio fusion model performs less accurately than other networks, it still has a significant impact because of the complementing data it provides. Finally, the top fusion framework scored 76.43% on the database of validation. In the past, accuracy was determined using segment-based data. As a result, we suggest a novel approach to determining accuracy rate: computation by video unit. Each video's five clips are counted, and the forecast with the greatest number of occurrences is chosen as the final outcome.

VI. CONCLUSION

Finally, we conclude that Sentiment Analysis Of Dysthymia Form Of Depression Using Multimodal Approach in the field of The questionnaire covers various aspects of mental health, including symptoms, duration, and impact on daily functioning. By systematically assessing these factors, we can gauge the severity of depression and tailor interventions accordingly. Additionally, our system incorporates advanced facial expression recognition technology. By utilizing artificial intelligence algorithms, we can accurately detect and analyse facial expressions associated with different emotional states.

VII. REFERENCES

[1]. Mohammad Soleymani, David Garcia, Brendan Jou, Bjorn Schuller, Shih-Fu Chang, Maja Pantic (2017). A Survey of Multimodal Sentiment Analysis.

[2]. Yang, Z., Yang, D., Dyer, C., He, X., Smola, A., & Hovy, E. (2016). Hierarchical attention networks for document classification. Proceedings of the 2016 Conference of the North American Chapter of the Association for Computational Linguistics: Human Language Technologies, 1480-1489.

[3]. J. Bian, L. Rajamanickam, Z. Nopiah (2023), Multimodal Sentiment Analysis Based on Deep Learning.

[4]. Y. Zhao, M. Mamat, A. Aysa, K. Ubul (2023), Multimodal Sentiment System and Method Based on CRNN-SVM.

[5]. M. Li, X. Qiu, S. Peng, L. Tang, W. Yang, y. Ma (2021), Multimodal Emotion Recognition Model Based on a Deep Neural Network with Multiobjective Optimization.

[6]. Poria, S., Cambria, E., & Bajpai, R. (2017). Multimodal sentiment analysis: Addressing key issues and setting up the baselines. IEEE Intelligent Systems, 32(3), 74-80.

[7]. Hao Sun, Hongyi Wang, Jiaqing Liu, Yen-Wei Chen, Lanfen Lin (2022), An MLP-based Model for Multimodal Sentiment Analysis and Depression Estimation.

[8]. Sihua Lyu, Xiaopeng Ren, Yihua Du and Nan Zhao (2023), Detecting depression of microblog users via text analysis.

[9]. Upadhye, P. A., Ghule, G., Tatiya, M., Shirkande, S. T., Kashid, S. M., & Bhamare, D. Optimizing Communication Systems with Applied Nonlinear Analysis Techniques, Communications on Applied Nonlinear Analysis, https://doi.org/10.52783/cana.v30.277.

[10]. Parlewar, P, Jagtap, V., Pujeri, U, Kulkarni, M. M. S, Shirkande, S. T., & Tripathi, A. (2023). An Efficient Low-Loss Data Transmission Model for Noisy Networks. International Journal of Intelligent Systems and Applications in Engineering, 11(9s), 267–276.

[11]. Nalawade, V. S., Ashok, G. K., Hanumant, B. A., & Reshma, G. (2021). ENCRYPTION THEN COMPRESSION BASED SYSTEM USING GRAYSCALE BASED IMAGE ENCRYPTION FOR JPEG IMAGES.