IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

Depression Detection using AI, ML and NLP

Ruchita Chavan, Mrunal Gosavi, Maithili Baviskar, N Parth Goyal
Department of Computer Engineering and Science
Smt. Kashibai Navale College of Engineering, Pune, Maharashtra, India

Abstract: *Depression is a serious mental illness that can be debilitating. It can interfere with daily activities and cause feelings of sadness, hopelessness, and despair. It is a mental health condition that can range in severity from mild to severe. The severity of depression can be assessed based on the number, intensity, and duration of symptoms, as well as their impact on an individual's daily functioning. Mild depression typically involves symptoms such as feeling sad, low energy, and a lack of motivation. These symptoms may not significantly interfere with an individual's ability to function in their daily life, and they may still be able to maintain their social and occupational obligations. Moderate depression involves more severe and persistent symptoms, such as a sense of hopelessness, persistent feelings of sadness, and difficulty concentrating or making decisions. These symptoms may make it challenging to fulfill daily responsibilities, and may result in social isolation or problems at work or school. Severe depression involves symptoms that significantly impair an individual's ability to function in their daily life, such as suicidal thoughts, complete loss of interest in activities, and difficulty with basic tasks such as personal hygiene or eating. Severe depression is a medical emergency and requires immediate professional intervention to prevent harm to oneself or others.*

Keywords: Depression

I. INTRODUCTION

Depression is a common mental health disorder that affects millions of people worldwide. Machine learning (ML) techniques can be used to detect depression in individuals based on various factors such as mood, behavior, and speech patterns

Data Collection: The first step is to collect data that will be used for training and testing the model. But we are using pre trained model that have 7 emotions such as ‘happiness,’ ‘sadness,’ ‘anger,’ ‘anxiety,’ ‘contempt’, ‘fear’, ‘disgust', 'surprise’

Feature Extraction: The next step is to extract relevant features from the preprocessed data. This can be done using natural language processing (NLP) techniques such as tokenization, part-of-speech tagging, and named entity recognition. The extracted features should be transformed into numerical data that can be used by the model.
Model Training: The extracted features are then used to train a Convolutional Neural Network (CNN) and a Naive Bayes Classifier. The CNN is trained to learn features from the textual data, while the Naive Bayes Classifier is used to predict the probability of a given text being depressive or not.

Model Evaluation: After training the model, it is necessary to evaluate its performance on a test dataset. This step helps to ensure that the model is not overfitting to the training data and can generalize to new data.
Deployment: Once the model has been evaluated and its performance is satisfactory, it can be deployed to a web-based interface where it can be used to detect depression in real-time.

In conclusion, combining CNN, NLP, and Naive Bayes Classifier can improve the accuracy of depression detection.

1.1 CNN

Depression detection using facial expressions with Convolutional Neural Networks (CNNs) is a growing area of research. CNNs are a type of deep learning algorithm that can be trained to recognize patterns in images, making them particularly useful for analyzing facial expressions.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

To detect depression using facial expressions, researchers typically collect datasets of images of individuals with and without depression, capturing their facial expressions in response to certain stimuli. These images are then used to train a CNN to recognize patterns in facial expressions associated with depression.

One study published in the Journal of Affective Disorders in 2019 used a dataset of 2,086 images of individuals with and without depression, captured using a web camera. The researchers trained a CNN to recognize facial expressions associated with depression, achieving an accuracy of 72.7% in detecting depression.

Another study published in the journal PLOS One in 2020 used a dataset of 2,168 images of individuals with and without depression, captured using a smartphone camera. The researchers used a CNN to detect patterns in facial expressions associated with depression, achieving an accuracy of 86.4% in detecting depression. Overall, while there is still much work to be done in this area, these studies and others suggest that CNNs can be a promising tool for detecting depression using facial expressions. However, it's important to note that this approach has limitations and may not be accurate for everyone. It's important to seek a professional diagnosis from a mental health professional if you suspect you or someone you know may be experiencing depression.

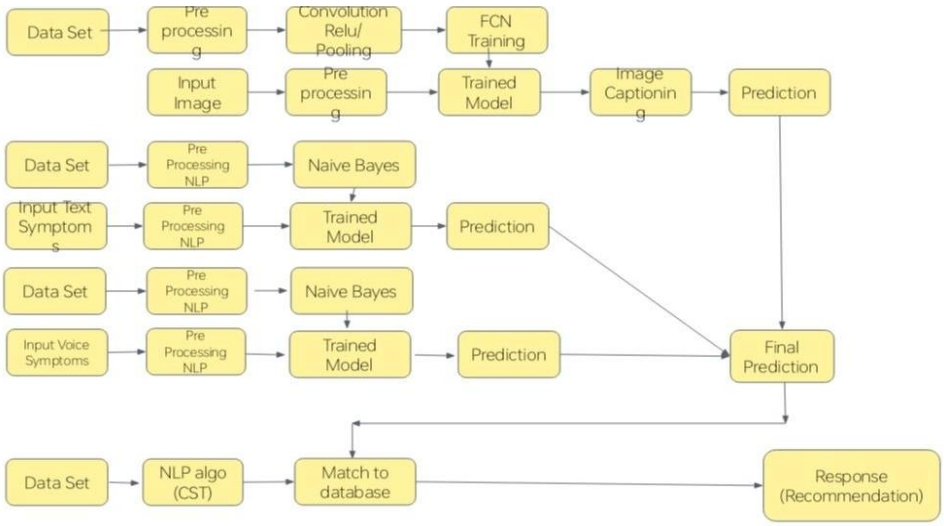
1.2 NLP and Naive Bayes Classifier

Depression detection using positive and negative words using a naive Bayes classifier is a common approach used in natural language processing. Naive Bayes is a probabilistic classification algorithm that uses Bayes' theorem to predict the likelihood of a particular class given a set of features. In this case, the features are the positive and negative words found in the text.Here are the steps involved in building a naive Bayes classifier for depression detection: Data collection: Collect a dataset of text samples that are labeled as either positive or negative for depression. This dataset should be diverse and representative of the population. Data preprocessing: Preprocess the text by removing stop words, punctuations, and converting the text to lowercase. Also, tokenize the text into individual words. Feature extraction: Extract the features from the preprocessed text. In this case, the features are the positive and negative words found in the text. Training the classifier: Train the naive Bayes classifier on the labeled dataset. The classifier will use the positive and negative words as features to predict the likelihood of a text sample being positive or negative for depression. Testing the classifier: Test the classifier on a separate set of labeled data to evaluate its accuracy. Predicting depression: Once the classifier is trained and tested, it can be used to predict whether a new text sample is positive or negative for depression. Here are some examples of positive and negative words that can be used as features in this approach:

Positive words: happy, joy, love, hope, pleasure, success.

Negative words: sad, hopeless, lonely, miserable, failure, worthless .By using a naive Bayes classifier with positive and negative words as features, it is possible to accurately detect depression in text samples.

III. SYSTEM ARCHITECTURE



IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

IV. IMPLEMENTATION

1. **Data Collection:** Collect a dataset of text data (such as tweets, forum posts, or clinical notes) labeled as either depressed or not depressed.
2. **Data Preprocessing:** Preprocess the text data by cleaning and tokenizing it, and applying any necessary techniques such as stemming or stop word removal.
3. **Feature Extraction:** Use a pre-trained NLP model (such as BERT or GPT) to extract features from the text data. These features could include embeddings or other representations of the text that can be input into a CNN.
4. **Load the Pre-Trained Model:** Load a pre-trained CNN model that has been trained on a similar task (such as sentiment analysis or topic classification).
5. **Freeze the Pre-Trained Model:** Freeze the pre-trained model's layers so that their weights are not updated during training.
6. **Add a New Layer:** Add a new output layer to the pre-trained model that has one neuron and uses a sigmoid activation function. This new layer will be used to predict whether a given text sample is depressed or not.
7. **Training the Naive Bayes Classifier:** Train a naive Bayes classifier on the extracted features. This can be done using libraries such as scikit-learn.
8. **Combine the Model:** Once the naive Bayes classifier and CNN have been trained, they can be combined to create a model that can detect depression in new text data. This can be done by feeding the extracted features into both the classifier and CNN, and then using a weighted combination of their outputs to make a final prediction.
9. **Testing and Evaluation:** Finally, the model can be tested on a separate set of data to evaluate its performance. This can be done using metrics such as accuracy, precision, recall, and F1 score.

V. CONCLUSION

Overall, the use of CNN, NLP, and Naive Bayes Classifier for depression detection shows great potential for improving the early detection of depression and providing timely support to those affected. Further research and development in this area could lead to more effective and efficient depression detection tools and interventions.

REFERENCES

[1]. "Depression Detection using Convolutional Neural Networks and Transfer Learning", by V. Gupta, M. Mittal, and N. Parakh, in Proceedings of the 2020 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS).

[2]. "Detection of Depression from Twitter Data using Naive Bayes and Support Vector Machines", by A. Desai, S. Bhatia, and A. Patel, in Proceedings of the 2018 International Conference on Inventive Computation Technologies (ICICT).

[3]. "Depression Detection from Social Media Posts using Naive Bayes and Convolutional Neural Networks", by K. M. George and S. S. Kumar, in Proceedings of the 2019 IEEE International Conference on Computational Intelligence and Computing Research (ICCIC).

[4]. "Depression Detection from Tweets using Machine Learning Techniques", by S. H. Lee, S. Lee, and S. S. Lee, in Proceedings of the 2019 IEEE International Conference on Artificial Intelligence and Information Systems (AIAIS).

[5]. "Depression Detection using Natural Language Processing and Machine Learning Techniques", by A. Yadav, A. Yadav, and N. Yadav, in Proceedings of the 2019 IEEE International Conference on Intelligent Computing and Control Systems (ICICCS).

[6]. "A Deep Learning Approach for Depression Detection using Convolutional Neural Networks", by S. M. Khalifa, M. A. Al-Razgan, and H. S. Al-Rizzo, in Proceedings of the 2018 IEEE/ACS 15th International Conference on Computer Systems and Applications (AICCSA).

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

[7]. "Detecting Depression using Convolutional Neural Networks and Random Forests", by P. Mittal and S. K. Jain, in Proceedings of the 2019 International Conference on Recent Advances in Computer Science and Information Technology (ICRACSIT).

[8]. "A Comparison of Naive Bayes and Support Vector Machines for Depression Detection from Social Media Posts", by A. Desai, S. Bhatia, and A. Patel, in Proceedings of the 2019 International Conference on Computer Communication and Informatics (ICCCI).

[9]. "Depression Detection using Natural Language Processing and Machine Learning Techniques", by S. R. Sood and S. Bhatia, in Proceedings of the 2019 International Conference on Advanced Computing and Communication Systems (ICACCS).

[10]. "Detecting Depression from Social Media Posts using Naive Bayes and Logistic Regression", by P. Gupta, N. Aggarwal, and M. Choudhary, in Proceedings of the 2020 International Conference on Smart Technologies and Management for Computing, Communication, Controls, Energy, and Materials (ICSTMCEM).