

# Depression Prediction using BERT and SVM

Ritu Menon<sup>1</sup>, Sanchit Gide<sup>2</sup>, Swajit Ghatte<sup>3</sup>, Sahil Mendon<sup>4</sup>, Dr. Madhu Nashipudimath<sup>5</sup>

<sup>1,2,3,4</sup>UG Student, Dept. of Computer Engineering, Pillai College of Engineering, New Panvel, India

<sup>5</sup>Professor, Dept. of Information Technology, Pillai College of Engineering, New Panvel, India

\*\*\*

**Abstract** - As there has been an increase in the number of mental illness cases, there is a need to curb this problem. Due to the complexity of traditional techniques based on clinical diagnosis, there should be an automated system for the detection and prevention of illness and hence there comes the need to develop a depression prediction system. The data is collected from Reddit, followed by preprocessing and cleaning which includes removal of stop words, URL and HTML tags, expanding abbreviations etc. Following this, the process of feature extraction is carried out. N-grams and TF-IDF (Term Frequency - Inverse Document Frequency) will be used to extract word count, pronouns, negations and other features from the comments made by the users following selection. Bidirectional Encoder Representations from Transformers (BERT) and Support Vector Machine (SVM) classifiers are applied to obtain the results. Further, the tokenized words are embedded into a vector and passed to the encoder in BERT and then to perform depression prediction, a classification layer is added on the top of the transformer output. Similarly the tokenized input is embedded into a vector of SVM, which is mapped to a class label, followed by classification into depressed and non-depressed classes. Both methods are compared for depression analysis. This helps in providing an early detection of depression in people.

**Key Words:** Depression, BERT, SVM, TF-IDF, N-grams

## 1. INTRODUCTION

Mental illness is a very serious issue which has its impact all over the world. About thirteen percent of the world's total population are grieving from depression and anxiety. Nowadays, due to the stigma of mental illness, it becomes difficult for the affected ones to confide in their peers and hence such people tend to use social media as a way to express their feelings through posts, comments, etc and therefore, social media has become a way to analyze a mentally ill person. Depression prediction systems when used by users can help them in early detection of this illness which may help them in timely diagnosis and treatment which is very essential. The primary aim of this system is to predict if a user is depressed or not.

## 2. LITERATURE SURVEY

### A. BERT and SVM:

In this paper, from vectors trained on Twitter data, several word embeddings were generated for each tweet. The vectors and the averaged embedding generated have a dimensionality of 200. Prior to giving these embeddings as an input to the SVM model, text preprocessing was carried out. In a particular series of the SVM experiment performed, only tokenization and lowercasing was carried out for all tweets, but in the subsequent one, specific preprocessing of tweets was performed. SVM model used a linear kernel.[1]

### B. CNN,RNN and BiLSTM:

The data collected is divided into three main broad categories, namely image, text and behavioral. The image is sent to CNN for classification. For text, word embeddings were produced using a technique called Word2vec, in which the embeddings had a dimension factor of 300, which acts as a feature to be further sent to BiLSTM. Deep learning methods (e.g. CNN, RNN) and Scikit-learn (aid to apply traditional machine learning) are used. [2]

### C. KNN and LIWC:

Here, KNN is applied to the Facebook dataset. NCapture is used to collect data. KNN extracts paraphrases to detect emotions. LIWC (Linguistic Inquiry and Word Count) is used for analyzing user comments. [3]

### D. Multinomial Naive Bayes, Support Vector Regression:

Tweets are taken up from Twitter and they are classified on the basis of keywords like Depression, Anxiety, Mental Illness etc. Various processes namely Sentiment Extraction, Tokenization, Stemming, and Part Of Speech tagging is done for feature extraction, followed by which both these techniques are applied to the data for classification.[4]

**E. Word2Vec, Rapid Automatic Keyword Extraction Algorithm:**

Depression symptoms were extracted from Twitter by carrying out text mining followed by which a semantic graph was generated that represented a correlation between depression and the symptoms leading to depression. A hybrid method was proposed that integrates the statistical analysis and various techniques of natural language processing which differed from the frequently used statistical methods.[5]

**Table -1:** Summary of literature survey

| Sr No | Paper  | Advantages and Disadvantages  |
|-------|--|---|
| 1     | DavidOwen Jose, Camacho Collados, Luis Espinosa Anke [1]   | Advantages:<br>1. The accuracy of SVMs in TFIDS+Word Embs is 74%.<br>2. The accuracy of BERT is between 70-88%<br>Disadvantages:<br>1. LMs worked fairly well on the concerned task with a balanced set, yet they do not master the lighter-weight methods in case of unbalanced training data.   |
| 2     | Chun Yueh Chiu, Hsien Yuan Lane, Jia Ling Koh, Arbee L. P. Chen [2]                                    | Advantages:<br>1. An accuracy of 84.2% was achieved by AlexNet, which outperformed the other models.<br>2. The accuracy of BiLSTM are slightly better than the others with a different number of layers<br>Disadvantages:<br>1. Experimental results show that detection performance can be largely improved by aggregation method.       |
| 3.    | Md Rafiqul Islam, AbuRaihan M.Kamal, Naznin Sultana, Robiul Islam, Mohammad Ali Moni, Anwaar ulhaq [3] | Advantages:<br>1. Out of The KNN classifiers used,Coarse KNN proved to be the best model.<br>2. Performances are measured using precision recall and F-Measure<br>3. Accuracy of around 60-70% was obtained.<br>Disadvantages:<br>1. Other available techniques to extract paraphrases from more types of emotional features can be used. |

|    |                                 |  |
|----|---------------------------------|--|
| 4. | Priyanka Arora, Parul Arora [4] | Advantages:<br>1. Support Vector Regression and Multinomial Naive Bayes gave better results than NaiveBayes, K-means and SVM.<br>2. Accuracy of Multinomial Naive Bayes was 78% and Support Vector Regression was 79.7%<br>Disadvantages:<br>1. Accuracy of the SVM classifier on textual reviews available on the internet is 78.8%, which is certainly better than the multinomial naïve bayes, but is found to be lower than support vector regression. |
| 5. | Long Ma, Yan Wang [5]           | Advantages:<br>1. The semantic graph gives the most pertinent entities, specifically for depression.<br>2. Recurrent symptoms of depression have very well scored to the “depression” class.<br>Disadvantages:<br>1. More accuracy can be obtained on building a classifier by combining various symptoms of depression and the tweets generated by depressed people.  |

**3. PROPOSED WORK**

Support Vector Machine (SVM) and Bidirectional Encoder Representations Transformers (BERT) can be used to predict depression which is done broadly on the basis of text. The results obtained from the individual SVM and BERT model suggest that SVM performs better classification and BERT model produces comparatively better word vector embeddings. An ensemble model would be developed in order to provide a fair result for depression detection. In the Ensemble architecture BERT is used for producing a vector embedding for all the sentences in the data set and SVM is used to classify sentences into respective classes on the basis of the embedding vectors produced by BERT.[1]

**3.1 SYSTEM ARCHITECTURE**

The system architecture is given in Figure 1. Each block is described in this Section.

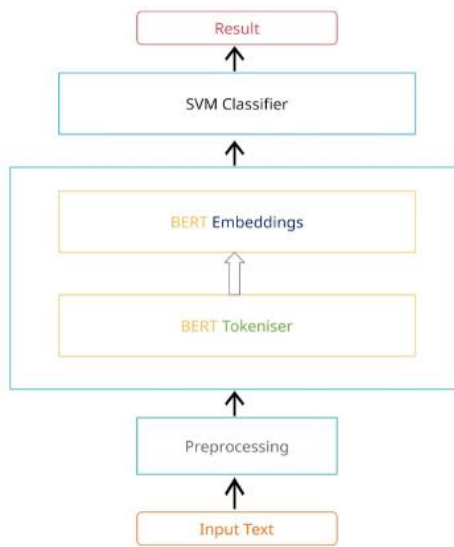


Fig. 1 Proposed system architecture

**A. Input Processing :** The model receives pairs of sentences as input and learns to predict if the second sentence is the subsequent sentence, out of which 50% was trained and from the remaining, a random sentence from the corpus will be chosen as the second sentence. A [CLS] token and a [SEP] token are inserted at the beginning and the end of each sentence respectively.

**B. BERT :** BERT stands for Bidirectional Encoder Representations from Transformers. It is designed to read a text in such a way to obtain both left and right context. The BERT model has its own feature extraction method called BERT tokenizer which uses BERT embeddings to prepare the input. Input data is to be converted into an appropriate format so that each sentence can be sent to the pre-trained model to obtain the corresponding embedding.

**C. SVM:** Support vector machine algorithm generates a hyperplane in an N-dimensional space where N is the number of features that distinctly classifies the data points available. The goal is to find a plane that has the maximum distance between data points belonging to both the classes. Support vectors are data points that influence the position and orientation of the hyperplane. These support vectors, thus help in maximizing the margin of the classifier.

**D. N Grams :** N-gram is a sequence of n items like phonemes, syllables, letters, words or base pairs according to the application, from a given sample of text or speech.

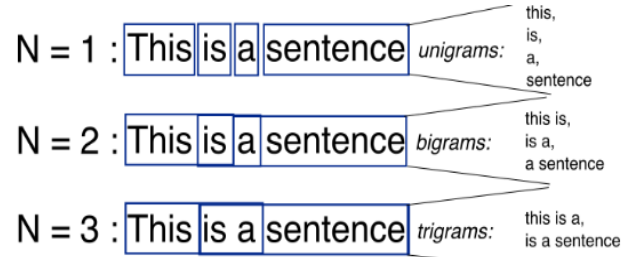


Fig. 2 N-grams

**E. TF-IDF :** It is a technique that is used to quantify words in a set of documents. It signifies the importance of a word by computing a score for each word in the document. The tf-idf value rises as the number of occurrences of a word in the document increases. It is computed as :

$$TF-IDF = (TF) * (IDF)$$

- **Term Frequency (TF) :** It gives the measure of the frequency of a word in a document.
- **Document Frequency :** It measures the importance of documents in the entire set of the corpus.
- **Inverse Document Frequency (IDF) :** It is basically the inverse of the document frequency which measures the informativeness of term t.

### 3.2 REQUIREMENT ANALYSIS

#### A. Software

|                      |            |
|----------------------|------------|
| Operating System     | Windows 10 |
| Programming Language | Python     |

#### B. Hardware

|           |             |
|-----------|-------------|
| Processor | 2 GHz Intel |
| HDD       | 180 GB      |
| RAM       | 2 GB        |

#### C. Dataset

The dataset contains posts and text features for time frames from 28 mental health and non-mental health subreddits (15 mental health support groups) from 2018-2020.

Fig 3 : Reddit Mental Health Dataset

#### 4. CONCLUSION

This depression detection model can help detect people suffering from depression. According to the literature survey, there are separate BERT and SVM models for depression prediction, which achieved a fair accuracy. Hence to enhance the accuracy more, in the proposed work, an ensemble model (BERT and SVM) was developed. With the user giving the inputs, a decision is made whether the user is under depression or not. Thus, the model for depression prediction successfully categorizes depressed and non-depressed people.

#### ACKNOWLEDGEMENT

We would like to put forth our sincere gratitude towards our project guide, Dr Madhu Nashipudimath, who guided us and provided us with her valuable knowledge and suggestions on this project and helped us improve our project beyond our limits. Secondly, we would like to thank our Project Coordinator, Prof Payel G. Thakur, who helped us finalize this project within the stipulated time frame and by constantly supporting us. We would also like to propound our genuine thanks to Head of the Department (Computer Department) Dr. Sharvari S. Govilkar, for being a pillar of support and providing us with a platform where we can try to work on developing projects and demonstrate the practical applications of our academic curriculum. We would like to express our gratitude to our Principal, Dr. Sandeep Joshi, who provided us with a golden opportunity to do this wonderful project on the topic of 'Depression Prediction and Analysis using BERT and SVM', which has also helped us in undergoing a lot of research and extracting deep insights towards the problem statement.

#### REFERENCES

- [1] Owen, David, Jose Camacho Collados, and Luis Espinosa-Anke. "Towards Preemptive Detection of Depression and Anxiety in Twitter." arXiv preprint arXiv:2011.05249 (2020),
- [2] Chiu, Chun Yueh, et al. "Multimodal depression detection on instagram considering the time interval of posts." *Journal of Intelligent Information Systems* 56.1 (2021): 25-47.
- [3] Islam, Md Rafiqul, et al. "Detecting depression using k-nearest neighbors (knn) classification technique." *2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2)*. IEEE, 2018.
- [4] Arora, Priyanka, and Parul Arora. "Mining twitter data for depression detection." *2019 International Conference on Signal Processing and Communication (ICSC)*. IEEE, 2019.
- [5] Ma, Long, and Yan Wang. "Constructing a semantic graph with depression symptoms extraction

from twitter." *2019 IEEE Conference on Computational Intelligence in Bioinformatics and Computational Biology (CIBCB)*. IEEE, 2019.

- [6] Low, D. M., Rumker, L., Torous, J., Cecchi, G., Ghosh, S. S., & Talkar, T. (2020). Natural Language Processing Reveals Vulnerable Mental Health Support Groups and Heightened Health Anxiety on Reddit During COVID-19: Observational Study. *Journal of medical Internet research*, 22(10), e22635 "https://zenodo.org/record/394138"

#### BIOGRAPHIES



Ritu Menon



Sanchit Gide



Swajit Ghatte



Sahil Mendon