ANALYSIS OF TEXT FOR DEPRESSION DETECTION USING DEEP LEARNING

Akshay Doke*1, Shantanu Pawar*2, Shridhar Kengar*3, Sonali Pakhmode*4

*1,2,3Student, Department Of Information Technology, Vasant Dada Patil Pratishthan's College Of Engineering And Visual Arts, Mumbai, Maharashtra, India.
*4Professor, Department Of Information Technology, Vasant Dada Patil Pratishthan's College Of Engineering And Visual Arts, Mumbai, Maharashtra, India.

ABSTRACT
Depression is the leading cause of mental illness, which has been found to increase the risk of death. During the covid-19 epidemic depression has risen worldwide. In this paper we have suggested analyzing the Depressive sentiment in tweets of the twitter user to predict user stress level. The current study aims to use twitter feeds to perform stress analysis using Deep Learning (DL). Individual tweets are classified as neutral or negative based on a list of terms to determine the tendency of stress. The main contribution of this study is to help twitter users to know about their depression in order to improve their quality of life.

Keywords: Twitter, Depression Analysis, Natural Language Processing, Deep Learning (DL), Depression Detection.

I. INTRODUCTION
Depression is a mental disorder caused by persistent emotions or a loss of interest in daily activities, leading to serious problems in daily life. During the covid-19 epidemic, depression increased worldwide. According to WHO worldwide, about 5% of adults suffer from this disorder. It is characterized by persistent sadness and a lack of interest or enjoyment of rewarding or exciting things in the past. It can also interfere with sleep, loss of appetite, fatigue and mental instability often increasing this number in the 15–25-year group. Today most adults use social media platforms such as twitter, Facebook and Instagram to talk about their daily lives. Twitter is a social media platform that allows users to express their emotions to the world through tweets. In this paper we will use Deep Learning and the LSTM algorithm to detects stress from social media platforms such as twitter.

II. LITERATURE SURVEY
[1] Early depression detection from social network using Deep learning technique, 2020
In this paper they have identify the most effective deep neural network architecture among a few of selected architectures that were successfully used in natural language processing tasks to detect depression in given limited unstructured text data extracted from the Twitter social media platform.

In the research paper they have used question based analysis and then with help machine learning algorithms like regression, random forest and XG Boost classifiers with accuracy 83% is most accurate.

[3] Depression detection by analysing social media post of user, 2019
This paper proposed a model that takes a username and analyzes the social media posts of the user to determine the levels of vulnerability to depression using naïve bayes and SVM algorithm.

[4] Depression detection on social media network using Depression analysis, 2021
In this paper a predictive model to predict whether a user's tweet is depressed or not based on detecting depressed users using a supervised learning approach to Depression analysis using SVM and naïve bayes algorithm

In this data is collected in the form of text from reddit social media. Machine learning algorithms were used. To identify a lexicon of terms those are more common among depressed accounts.

In this paper they have used deep learning technique Bi-LSTM for detecting the depression. In this research work, a hybrid model has been proposed that can detect depression by analyzing user's textual posts.
learning algorithms were trained using the training data and then performance has been evaluated on the test data of the dataset of reddit which was published for the pilot study.


This paper aims to apply natural Language processing on Twitter feeds for conducting emotion analysis focusing on depression. Individual tweets are classified as neutral or negative, based on a curated word-list to detect depression tendencies. Multinomial Naïve Bayes has performed the best with the F1 score of 83.29 whereas SVM has achieved a lower F1 score of 79.73.

III. SYSTEM REQUIREMENTS

HARDWARE REQUIREMENTS
System: Pentium Dual Core.
HDD: 120 GB.
Screen: 15” LED
Input Devices: Keyboard, Mouse
Ram: 4 GB.

SOFTWARE REQUIREMENTS:
Coding Language: Python
Tool: Python 3.6.1

IV. METHODOLOGY

EVALUATION DATASET
To identify depression, we train our models on the dataset of ‘Twitter’ social media tweets data. The dataset used in our model are:
- ‘Positive tweets’ from Kaggle.com Website
- ‘Depressive’ from Kaggle.com Website

![Figure 1: Depressive words](image1)

![Figure 2: Positive words](image2)
DATA PRE-PROCESSING

We use the NLP tools to pre-process the dataset before it is proceeded to the tokenization and training stage.

1. Stemming: Helps to group similar words together

2. Stop Words Removal: Used to remove stop words.
   Example: the, is, and etc.

Part of speech (POS) tagger: it is used to extract only adjectives, noun and adverbs since other parts of speech are not of much significance.

Example: "I am happy" – Happy noun is extracted rest words are removed.

3. Tokenization: To divide the posts into individual tokens
4. **Word2vec**: Word2Vec is a word embedding technique which uses neural networks. Synonyms and words with similar meanings are grouped together.

V. **ALGORITHM**

**Long Short Term Memory [LSTM]**

Long Short Term Memory is a kind of recurrent neural network. In RNN output from the last step is fed as input in the current step. It tackled the problem of long-term dependencies of RNN in which the RNN cannot predict the word stored in the long-term memory but can give more accurate predictions from the recent information. As the gap length increases RNN does not give an efficient performance. LSTM can by default retain the information for a long period of time. It is used for processing, predicting, and classifying on the basis of time-series data.

---

**Figure 4: LSTM Architecture**
Comparision of Previous System and Proposed System

Table 1: Comparision of previous system and proposed system

<table>
<thead>
<tr>
<th>Algorithm</th>
<th>Precision</th>
<th>Recall</th>
<th>F1</th>
<th>Accuracy</th>
<th>Paper</th>
</tr>
</thead>
<tbody>
<tr>
<td>SVM</td>
<td>0.804</td>
<td>0.793</td>
<td>0.79</td>
<td>0.79</td>
<td>Depression Detection using Emotion Artificial Intelligence CFP17M19-ART, ISBN:978-1-5386-1959</td>
</tr>
<tr>
<td>CNN</td>
<td>0.81</td>
<td>0.77</td>
<td>0.79</td>
<td>0.87</td>
<td>Applying Deep Learning Technique for Depression Classification in Social Media Text</td>
</tr>
<tr>
<td>LSTM (Proposed system)</td>
<td>0.83</td>
<td>0.84</td>
<td>0.87</td>
<td>0.89</td>
<td>Proposed Model</td>
</tr>
</tbody>
</table>

VI. RESULTS

Figure 5: Accuracy of Model

In this work different approach has been taken to perform detection on depression using twitter datasets and while working on the dataset with help of LSTM (long short term memory) model we got accuracy of 89%.

VII. CONCLUSION

For this project, we have tried to identify the stress of the Twitter user using Deep Learning method and tried to provide a solution for the user to get out of his stressful situation. We noted the close relationship between depression and language used NLP (Natural Language Processing) and LSTM. According to our findings, language predictions for depression contained words related to anxiety, feelings of sadness, anxiety, anger, and suicidal thoughts.

ACKNOWLEDGEMENTS

We would like to thank our project guide, Prof. Sonali Pakhmode and our Project coordinator (Professor at Faculty of Information Technology, Vasantdada Patil Pratishthan's College of Engineering and Visual Arts) for guiding us in every stage of the project development starting from Requirement Analysis to the actual Implementation of the system. We would also like to thank (Head of Information Technology Department, Vasantdada Patil Pratishthan's College of Engineering and Visual Arts) for sharing their perspectives and domain experience which helped in many aspects of this study. Moreover, we would like to express our gratitude and appreciation to our friends who have helped and supported us in this study.
VIII. REFERENCES


