

Detecting Depression in Tweets Using DistilBERT

U. Yaraswini¹, Y. Sasidhar², P. Siva Sai³, P. Eswar⁴, and V. Swathi⁵

^{1,2,3,4}Students, Department of Computer Science & Engineering, Dhanekula Institute of Engineering and Technology, Vijayawada, Andhra Pradesh, India

⁵Assistant Professor, Department of Computer Science & Engineering, Dhanekula Institute of Engineering & Technology/JNTUK/, Vijayawada, Andhra Pradesh, India

Correspondence should be addressed to U. Yaraswini; yaraswiniuppaluri000@gmail.com

Copyright © 2021 Made U. Yaraswini et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

ABSTRACT- Depression is a mood disorder that will affect a person's daily life. Depression can change life with all the suicidal thoughts, and the youth is straining a lot. Social media is growing tremendously day by day, and the applications like Twitter, Facebook are being used by youngsters. They share their opinions about their mood, and we can analyze a person's state of mind with those tweets written on Twitter. Our paper aims to detect that the person is depressed or not by using the Tweets using distilBERT, the distilled version of BERT. This distilBERT model is used to train the data and helps to achieve higher accuracy.

KEYWORDS- Depression, Social Media, BERT, distilBERT, Tweets.

I. INTRODUCTION

Depression leads to thoughts where one cannot share themselves with another, becomes introverted day by day, and finally get to have suicidal thoughts, which are more dangerous. But most people tend to share their feelings more actively on social media platforms like Twitter as tweets. It is a magic tool for scraping and fetching the data from Twitter based on the desired keywords in a few lines of code to execute in any terminal [5].

So, we can examine them using the tweets for the detection of depression. We use the latest distilBERT model, a distilled version of BERT, which helps us train the dataset. BERT's key technical innovation is applying the bidirectional training of Transformer, a popular attention model, to language modeling [4]. This model makes the dataset train and helps in the prediction of the tweets which are not in the dataset. Knowledge distillation (sometimes also referred to as teacher-student learning) is a compression technique in which a small model is trained to reproduce the behavior of a larger model [3].

II. LITERATURE SURVEY

To detect depression in tweets, researchers have presented many machine learning algorithms and their variants. Many of the proposed projects make use of text features extraction

techniques like TF-IDF vectorizer and word embedding. And the models like SVM[1] are used for the vectors tokenization, preprocessing. The simple transformers library is used for the Simple Transformers library, which provides the Application Programming Interface, which in turn provides access to the BERT, ALBERT. In some cases, the BERT, ALBERT classifiers, perform their tokenization. The other project of detecting depression is with the usage of Logistic Regression and SVM, Word-based RNN Model, Word-based GRU Model, Character-based GRU Model. Logistic Regression and SVM are used to perform Binary Classification on the Tweets. RNN's are popular for tackling NLP tasks because they sequentially process text [2] and obtain (tweet, label) pairs.

III. PROPOSED WORK

Our system is developed using transformer in machine learning which contains a variety of techniques like Bert. DistilBERT uses Natural language processing. DistilBERT is a technique developed based on Bert but enhanced. DistilBERT takes a less loss in the data which leads to the increment of the efficiency in accuracy.

Advantages:

- DistilBERT contains less loss to the data when compared to the BERT and it is also much faster than Bert.
- Language Modelling in DistilBERT is surpassing the Bert.

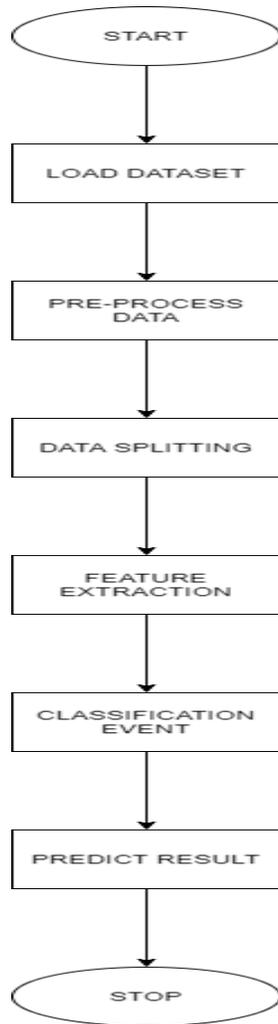


Fig. 1: Flow Chart of proposed work

A. Working

The goal of this project is to predict that the given tweet is depressed or not. This project is built on the DistilBERT model using Python programming language. As the BERT models are one of the transformers which are in turn under the Deep Learning models. We require the ktrain which is a library used to help build, train, debug, and deploy neural networks in the deep learning software framework. So it is used in our program for training the dataset. The preprocessing also works after assigning the DistilBERT model to the function involved in the preprocessing model. It is declared and initialized in the project for the preprocessing of the dataset. The next step of the preprocessing is the training. Training of the dataset helps the model to understand the detection of the depressed tweets which can identify the new tweets given by the user to detect them Figure 1 Flow Chart determines the procedure for detecting depression in tweets using distilBERT, which contains the flow of tasks. The path is followed for the prediction of tweets as depressed or not.

IV. RESULT

The data required for detecting the depression, such as the tweets, are given to the method which predicts the tweet as it is depressed or not. The method given for the prediction of the data for depression predicts the depressed tweet and gives the output as it is depressed or not. After the training process to our dataset, we are achieving the val_accuracy of 0.9949. It will be displayed as a label and a not_label, which means that which means that the label indicates that the given tweet is 'depressed' and the not_label defines that the tweet is 'not depressed'.

```
data=['Sleeping can reduce depression', 'Today is a big day', 'In pain, since depressed']
```

```
predictor.predict(data)
```

```
['label', 'not_label', 'label']
```

Fig. 2: Result

V. CONCLUSION

Here, we presented a novel approach to optimize word-embedding for classification tasks. We performed a comparative evaluation on some of the widely used deep learning models for depression detection in tweets. We built a predictive model to predict whether a user's tweet is depressed or not. We have observed that people suffering from depression are more socially isolated by determining how they interacted with trending hashtags, popular emojis used in their tweets.

CONFLICTS OF INTEREST

The authors declare that they have no conflicts of interest.

ACKNOWLEDGEMENT

We would like to take this opportunity to express our deep gratitude to the members who assisted us directly and indirectly for the completion of this project work. We would like to thank Mrs. V. SWATHI the project guide for her esteemed guidance and support, especially the valuable ideas and thoughts provided during this project work. She/he is expertise in the area of Python Programming, this helps us in solving the problems encountered during the project work. We would like to express our immense pleasure in expressing immeasurable sense of gratitude to Mr. K. SANDEEP, Project Coordinator for giving opportunity to make this project a successful one. We wholeheartedly acknowledge SURESH SUNDARADASU, Professor & Head of the Department in Computer Science & Engineering and Dr. RAVI KADIYALA, Principal for the valuable suggestions in the completion of the project. We also extend our thanks to all the faculty members of Computer Science & Engineering department for their valuable contributions in this project.

REFERENCES

- [1] David Owen, Jose Camacho-Collados, Luis Espinosa-Anke, "Towards Preemptive Detection of Depression and Anxiety in Twitter", December 12, 2020.
- [2] Diveesh Singh, Aileen Wang," <https://web.stanford.edu/class/archive/cs/cs224n/cs224n.1184/reports/6879557.pdf>".
- [3] Victor Sanh," Smaller, faster, cheaper, lighter: Introducing DistilBERT, a distilled version of BERT", August 28, 2019.
- [4] Rani Horev, "BERT Explained: State of the art language model for NLP", Nov 10, 2018
- [5] Tulasiram, "Detecting Depression in Social Media via Twitter usage", Oct 30, 2020
- [6] Cristian Bucila, Rich Caruana, Alexandru Niculescu-Mizil, "Model Compression".
- [7] Geoffrey Hinton, Oriol Vinyals, Jeff Dean, "Distilling the Knowledge in a Neural Network".

ABOUT THE AUTHORS



Uppaluri Yasaswini, Under Graduate Final year Student seeking her degree in Computer Science Engineering at Dhanekula Institute of Engineering & Technology, Vijayawada, Andhra Pradesh, India



Yalamanchili Sasidhar, Under Graduate Final year Student seeking his degree in Computer Science Engineering at Dhanekula Institute of Engineering & Technology, Vijayawada, Andhra Pradesh, India



P. Siva Sai, Under Graduate Final year Student seeking his degree in Computer Science Engineering at Dhanekula Institute of Engineering & Technology, Vijayawada, Andhra Pradesh, India



P. Eswar, Under Graduate Final year Student seeking his degree in Computer Science Engineering at Dhanekula Institute of Engineering & Technology, Vijayawada, Andhra Pradesh, India