

# Early Detection of Depression Indication from Social Media Analysis

Maharukh Syed<sup>1\*</sup>, Meera Narvekar<sup>2</sup>

<sup>1</sup>D.J. Sanghvi College of Engineering, mahrukhsyed11@gmail.com, Mumbai.

<sup>2</sup>Department of Computer Engineering, D.J. Sanghvi College of Engineering, meera.narvekar@djsce.ac.in, Mumbai.

**Abstract:** Depression that stems through social media has been steadily growing since the past few years but with the current inclination towards social media reliance it is highly imperative to detect the early signs. Continuous observation of a user's social media interests and activities may highlight suspicious and negative thoughts. This observation can help in understanding their future course of action and also indicate any suicidal thoughts and behaviors. By using the machine learning models, early indications of depression detection can be addressed. This work studies different word embedding techniques for early detection of depression from social media posts. Further, this work develops a model using various NLP processes in order to address the issue of early detection. The recommendations can be useful as a Decision Support System for counselors, psychologist and also can be of good use by the cyber-crime cell department for criminal investigations.

**Keywords:** — Social Media Analysis, Natural Language Processing, Embedding System, Tag Cloud and CNN.

## 1. INTRODUCTION

Social media is a very powerful tool for communicating ones thought process. The dependence on these platforms have grown to a point where an individual feels the inherent need to be sharing every thought and feeling they have with their social followers for a sense of validation. However, when an individual witnesses their peers doing better, it also causes a sense of envy leading to self-deprecating thoughts and low self-esteem. Studies have also shown that excess use of social media leads to stress and negative mental health. It also promotes negative experiences like depression & anxiety, inadequacy about your life or appearance, self-isolation, cyber bullying, fear of missing out, etc. While depression and other mental illnesses may lead to social withdrawal and isolation, it was found that social media platforms are indeed increasingly used by affected individuals to connect with others, share experiences, and support each other. Using the machine learning method, this report focuses on ways to classify early signs of depression in textual data in its premature stage. A complete analysis of the various algorithms is hence studied in detail and by describing the previous and current work with respect to text classification; this research can formulate the practical base of this work.

For text classification tasks in the analyzed papers, various Natural Language Processing's like GloVe, Word2vec, FastText [1] have been used to address the issue.

Elaborate results in word and content based tasks which are related to mental health behavior understanding, for example are achieved using the Convolution Neural Network [8] an efficient technique to model words and sentences.



Fig.1. A depressed user pretends to be happy.

The rest of the paper is organized as follows. Literature review is explained in section 2. Proposed system is explained in section 3. Experimental results are presented in section 4. Concluding remarks in section 5 and future scope are given in section 6.

## 2. LITERATURE REVIEW

Marcel Trotzek, Sven Koitka, and Christoph, M. Friedrich [1] predict early signs of depression using machine learning algorithms in particular the Convolution Neural Network based on various word embedding techniques. Further ERDE (Early Risk

Detection Error) as a metric is taken for detection of depression. Context free word embedding models such as Word2vec suggested by (Peters et al., 2018) and GloVe, Fast Text proposed by (Trotzek et al, 2019) are the popular Natural Language Processing techniques been used to address the issue.

Ranjana Jadhav, Shaurya Jadhav, Hitesh Sachdev, Vinay Chellwani [2] proposed a Decision Tree classifier for Mood Disorder Questionnaires. This research using Machine Learning Approach basically determines the most significant feature in the data set and makes that as a deciding factor for detection of Bipolar Mental Health Disorder.

Jane H. K. Seah, et al [3] have carried out the research based on the data collected from Reddit for detection of suicidal tendencies of users on social platforms. This is performed using LDA and data mining approach of topic modeling is used to discover different context in which suicidal thoughts occur.

Akkapon Wongkoblaph et al [4] have developed a prediction model to classify users with poor mental health using machine learning algorithm along with data from social media platforms.

Hao Wang et al [5] in their "Sentiment Expression via Emoticons on Social Media" have highlighted that emoticons play a vital role in understanding sentiments. They have analyzed the widely used emoticons along with performing feature extraction using word2Vec conversion and have classified sentiments with and without emoticons using Bayes Classification model along with Bag-of-Words.

Walter Gerych et al [6] explain that depression diagnosis is difficult as the percentage of afflicted users in most populations is small compared with those unaffected, leading to severe class imbalance. They have proposed a multi-stage machine learning pipeline. This is done using Auto-encoders and SVM Algorithm.

Sharath Chandra Guntuku, David B Yaden, Margaret L Kern, Lyle H Ungar, Joannes C Eichstaedt [7] in their findings have shown that social network data can be used to predict mental illness using various machine learning modules like Logistic Regression and Support Vector Machine (SVM).The prediction can be based on multiple scenarios where it can be self declared on twitter it can also be based on responses of online surveys, based on annotated posts, forum membership etc.

Y. Zhang et al [8] in their A Sensitivity Analysis of Convolution Neural Networks for Sentence Classification has helped understand CNN model in

detail for sentence classification. This also helps aim is to distinguish between important and comparatively inconsequential design decisions for sentence classification. Classifying negative and positive comments from the data set can be done using the CNN model.

Matthew Peters et al [9] in their research show how contextual word representation have provided significant improvements in the state of art for a wide range of NLP tasks. They have shown an empirical study of how Neural Architecture influences accuracy and qualitative properties of the representations that are learnt.

## 2.1 Gaps Identified

Having a left-to-right flow, the addressed models are unidirectional in their pre-training. This restricts their ability to understand the actual meaning of pre-trained representations. Models like OpenAI GPT (Trotzek et al, 2019.) [1], used left-to-right architecture and ELMo (bidirectional but shallow) concatenates forward and backward language models do not fulfill the precise needs as depression detection is an emotionally challenged condition which cannot be detected directly. Hence there is a need for Contextual language modeling to be taken into consideration. Contextual representations can further be unidirectional or bidirectional. A language model can be trained forward or backward, but it continues to be unidirectional as the prediction of future characters or words is supported by past seen data.

Analyzing the various papers it was seen that various word embedding techniques were used in order to classify text for further processing. A deep study of different models was carried out. A lot of research papers have been gone through for understanding depression detection but very few of them are applicable and appropriate to the particular topic.

## 3. PROPOSED SYSTEM

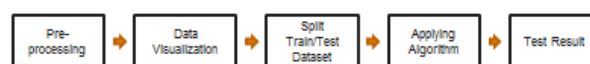


Fig.4.Steps involved in proposed system.

The proposed system aims at detecting depression indications using deep learning techniques applied on word2Vec embedding methods [1]. The steps include preprocessing of the data which means removing unwanted and utilizing only the information that is necessary. Data cleaning is performed in this step. Further the paper performs visualization of data where it creates word clouds and graphical representation of

data set is seen to understand data distribution. The next step includes dividing the training and testing dataset. A 60% - 40% ratio is taken for training and testing the data set. The last step includes applying the algorithm as shown in module II and lastly performing testing of the model.

### 3.1 Module I: Word2Vec

Pre-trained word descriptions, as seen in essay, can be context-free (i.e., word2vec, GloVe, fast Text) [1], which means that a single word description is formulated for each text in the vocabulary, or can also be contextual (i.e., ELMo and OpenAI GPT), [9] on which the word descriptions depend on the context where that word occurs, meaning that the same word in different formats can have different meanings. Let's understand the modules briefly. Trained to reconstruct linguistic descriptions of word, Word2Vec is a shallow, two-layer neural network. Typically taking in an oversized corpus of text, it creates a vector space of huge dimensions, and each unique word within the corpus is assigned a corresponding vector in the space. Words that share common contexts within the corpus are located in close presence of each other, within the space owing to the word vectors being positioned the certain way.

Being distinctly computationally-efficient predictive model for learning word embedding's from raw text; Word2Vec comes in two flavors viz. the Continuous Bag of Word (CBOW) & the Skip-Gram model. These models are similar in their algorithmic sense.

### 3.2 Module II: Convolution Neural Network

#### 3.2.1 System Design using CNN

This is a deep learning approach mostly used for analyzing data with grid-like structure other than image classification filtering [8]. It is a feed forward neural network which contains input data, multiple hidden layers and an output layer. The main aim is to learn a specific property from the given data and find out whether the user is depressed or not.

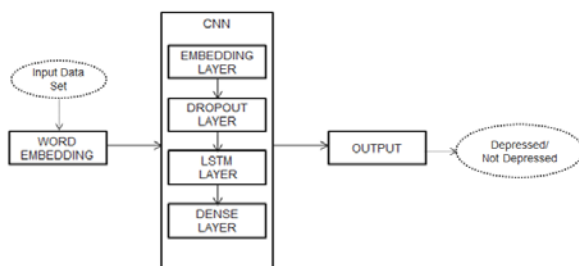


Fig.2. System Design using CNN.

Convolution neural networks have the following hidden layers:

#### I. Embedding Layer

The embedding layer will learn embedding for all the words from the dataset. It is basically a low dimensional space which in this case tries to take the large dataset and convert it into an output of 2D vector form, having 1 embedding for each word. Such kind of embedding makes it easier to perform machine learning on large social media data sets. It is also noted that such embeddings are flexible in nature and can be reused across the models. Let's consider a simple example to understand the use of Embedding Layer.

Suppose the input data contains various adjectives like joy,lonely,ugly,tensed,suicidal,horrible,gay,awful,beautiful,low,excited. From the above it can be categorized or clubbed the emotions into happy, sad or depressed. In this way a broader perspective for the emotion can be persuaded.

#### a. Dropout Layer

In deep learning neural network algorithms over fitting a training set is quiet common. For the same reason when you have a large network few nodes are dropped before every layer in order to avoid statistical noise. Prevention of over fitting is done here in order to remove unwanted data at a level of 50%. It is very important to remove the extra nodes from the neural network so as to get accurate predictions.

#### II. LSTM Layer

The Long Short Term Memory Layer has 2 layers known to detect complex layers. It is a type of recurrent network where it is capable of learning from its previous layer and using that information for sequence prediction. The LSTM layer supports active prediction. It also contains data from the previous layers hence drop out for recurrent data required here is with a level of 20%. When the CNN model is followed by the LSTM model it helps in defining two sub models namely the CNN model for feature extraction and LSTM model to predict the different features between the different layers of the model.

#### III. Dense Layer

This layer is always applied at the end in a CNN model which helps to connect all the layers to one another giving precise output and a fully connected layer. The neurons in each of the layers receive input from the previous layer in such format that all the layers are

tightly connected to each other. This helps in having a strong dense network.

#### 4. EXPERIMENT AND RESULT

The steps involved in implementation of CNN based approaches for detecting depression is shown below. Here the algorithm is applied on sentiment 140 1.5 Lac record of dataset. The characteristics include the context id, timestamp, username and comments. The implementation and test environment used is Microsoft Visual Studio IDLE 3.6(64-bit). System configuration includes Intel(R) core(TM) i7,8 GB RAM and 64 bit OS. Algorithm flow is shown as below:

1. Users tweet is taken; next pre-processing is performed to do data cleaning by replacing all NA values to zero.
2. A matrix with user, tweet and the timestamp is created.
3. Algorithms are applied to separate positive negative and neutral comments and to make prediction of whether the user is in depression or not.
4. Here cross validation splits the initial dataset into training and testing with 60-40 criteria. Next evaluation metrics are used for performance evaluation of algorithms.

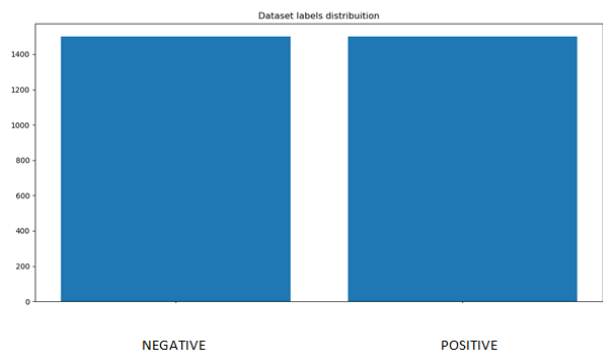


Fig.5. Figures shows execution of above mentioned algorithms to show positive and negative comments in the data set

##### I. Tag Cloud

Tags are single words which are of great importance and connection to each other generally used to detect metadata. It is a graphical representation of words which are of great prominence and appear more frequently in the dataset.



Fig .6.Word Cloud Generation for positive and negative words

Frequently occurring words appear in larger size compared to the words which occur rarely.

##### II. Accuracy and loss

It is the averaged squared difference between actual and estimated value. It is basically used to capture how accurate result the model has given. Lower value considered as best fit.

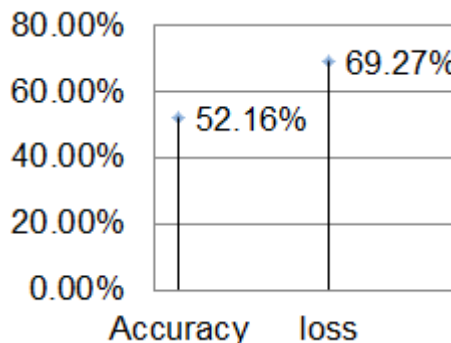


Fig.5.Accuracy and Loss using CNN.

## 5. CONCLUSION

In this research it is discussed that deep learning approach for early detection of depression from social media analysis can be done using deep learning algorithms. It has been observed that an accuracy of 52.1 % with loss of 69% is seen using Word2Vec along with CNN Model. Better results from future findings aims at using BERT as a word processing model along with Deep Learning models to find depression in a specific user over the period of time.

## 6. FUTURE WORK

This paper aims at doing a comparative study based on machine learning algorithms for prediction of depression using context-free model. In the future contextual bidirectional model can be considered for prediction in order to enhance the performance and which can provide better decision support systems to users.

## 7. REFERENCES

- [1] Marcel Trotzke, Sven Koitka, and Christoph, M. Friedrich, : "Utilizing Neural Networks and Linguistic Metadata for Early Detection of Depression Indications in Text Sequences", DOI 10.1109/TKDE.2018.2885515, IEEE Transactions on Knowledge and Data Engineering ,2019.
- [2] Ranjana Jadhav, Shaurya Jadhav, Hitesh Sachdev, Vinay Chellwani, : "Mental disorder detection. Bipolar disorder scrutinization using machine learning" IEEE, 978-1-5386-5933-5, 2019.
- [3] Jane H. K. Seah, et Al, : "Data Mining Approach to the Detection of Suicide in Social Media" IEEE-International Conference on Big Data, 2018.
- [4] Akkapon Wongkoblaph, Miguel A. Vadillo, Vasa Curcin, : "Detecting and Treating Mental Illness on Social Networks", IEEE-ICHI, DOI 10.1109/ICHI.2017.24, 2017.
- [5] Hao Wang, Jorge A. Castanon et Al, : "Sentiment Expression via Emoticons on Social Media" IEEE-International Conference on Big Data, 978-1-4799-9926-2/15/\$31.00, 2015.
- [6] Walter Gerych, Emmanuel Agu, Elke Rundensteiner, "Classifying Depression in Imbalanced Datasets using an Autoencoder-Based Anomaly Detection Approach" IEEE-ICSC DOI 10.1109/ICSC.2019.00028, 2019.
- [7] Sharath Chandra Guntuku, David B Yaden, Margaret L Kern, Lyle H Ungar, Joannes C Eichstaedt, : "Detecting depression and mental illness on social media : an intergrative review", IEEE-ICHI, DOI 10.1109/ICHI.2017.24, 2017.
- [8] Y. Zhang and B. Wallace, "A sensitivity analysis of (and practitioners' guide to) convolutional neural networks for sentence classification", IJCNLP, pages 253–263, 2017.
- [9] Matthew Peters, Mark Neumann, Luke Zettlemoyer, and Wen-tau Yih, : "Dissecting contextual word embeddings: Architecture and representation", pages 1499–1509, CEMNLP, 2018.