

# Depression Detection Using Sentimental Analysis

K.Chandan<sup>1</sup>, K. Kalyani<sup>2</sup>, A Swamy Goud<sup>3</sup>

<sup>1</sup> Student, Information Technology, B.V Raju Institute of Technology, Telangana, India

<sup>2</sup> Student, Information Technology, B.V Raju Institute of Technology, Telangana, India

<sup>3</sup> Assistant Professor, Information Technology, B.V Raju Institute of Technology, Telangana, India

## ABSTRACT

Datasets originating from social networks are valuable to many fields such as sociology and psychology. But the supports from technical perspective are far from enough, and specific approaches are urgently in need. This applies datamining to psychology area for detecting depressed users in social network services. Firstly, a sentiment analysis method is proposed utilizing vocabulary and man-made rules to calculate the depression inclination of each micro-blog. Secondly, a depression detection model is constructed based on the proposed method and different features of depressed users derived from psychological research. Then users and a classifier are used to verify the model, whose precisions are all around 80%. Also, the significance of each feature is analyzed.

Lastly, an application is developed within the proposed model for mental health monitoring. This study is supported by some psychologists, and facilitates them in data-centric aspect in turn.

**Keyword:** psychologists, Depression, Social networks networks, detection etc.....

## 1. Introduction

Depression has become a serious problem and affecting mental health. Depression is also considered an important illness that affected more than 264 million people worldwide [1]. There are several potential causes, including abrupt environmental changes, adjustments in neurotransmitter levels brought on by certain emotional assaults, and even hereditary characteristics .

Both counselling and medication are effective treatments for depression. Despite the fact that many individuals are receiving therapy, a sizable portion of the population still has depression that has not yet been identified. As a result, patients will experience negative ideas including self-isolation, erratic conduct, suicidal thoughts, and dependence on drugs like antidepressants. Due to a loss of enthusiasm and attention brought on by depression, daily activity performances might be slowed down. A person's health and brain may eventually suffer if depression is not identified [1].

The growth of online social networks and other communication technologies, in particular, has revitalised how individuals engage and communicate electronically. Applications like Facebook, Twitter, Instagram, and others allow users to express their feelings, emotions, and sentiments about a subject online in addition to hosting textual and video information.

This presents opportunities for those working in the health sector to gain insight into what might be going on at the mental state of someone who reacted to a topic in a particular way, which is great for users of social networking sites who can freely and openly contribute to and respond to any topic online.

One of the most prevalent mental diseases is depression, according to the global health organization's comprehensive mental health action plan 2013-2020 [2]. Chronic depression is the primary cause of acquired disability globally, affecting more than 300 million individuals. People use social media extensively in their daily lives.

Social media is constantly used to share a large amount of multimedia material, primarily brief words and photographs. Information posted on the Internet, as opposed to conventional human communication, may be swiftly shared by friends and read by total strangers. By avoiding face-to-face interaction, this technique increases the users' urge to communicate their emotions. As a result, researchers started to examine user social network features [2], and it has been demonstrated that Twitter is a useful platform for online emotion detection.

With the growth of online social networks, there are now more options than ever before to use information approaches to solve issues in a range of disciplines. For instance, traditional psychology study relies on surveys and formal interviews, but a lot of psychologists are increasingly focusing on digital media. They attempt to examine social network data from a psychological perspective. Unquestionably, this discipline integration gives psychology a boost, but the technical supports are insufficient.

## 2. Existing System

- Support Vector Machine (SVM) is a supervised machine learning technique that may be applied to both regression and classification applications.
- SVM is a sort of algorithm that solves binary class problems using a set of labelled training samples.
- SVM is a sort of algorithm that solves binary class problems using a set of labelled training samples.

### Disadvantages of Existing system

- It is computationally inefficient.
- Sensitive to the kernel hyper parameters.

## 3. Proposed System:

- Naive Bayes classifier to address a common NLP issue
- The classifiers for objectivity and polarity.
- To divide a tweet into objective and subjective categories, whereas the latter would do so into positive and negative categories.

### Advantages of Proposed system:

- Needs just a little bit of training data to figure out the parameters
- Can be learned more quickly than more complex models.



**True Positives (TP)** are the accurately predicted positive values, which indicate that both the actual and projected class values are true.

**True Negatives (TN)** are successfully predicted negative values, which indicates that both the actual and projected class values are negative.

**False Negatives (FN)** are when the expected class is no while the actual class is yes.

**Accuracy** - Accuracy is simply a ratio of correctly predicted observation to the total observations. For our model, we have got 0.87081 for Naïve Bayes algorithm which means our model are approximately 87% accurate.

The formula to find accuracy is given by

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+FN+TN}$$

**Precision** - Precision is the ratio of correctly predicted positive observations to the total predicted positive observations. Higher precision means lower false positives thus this parameter is importance we want our model to predict negative tweets correctly and has less false positives.

Precision is given by

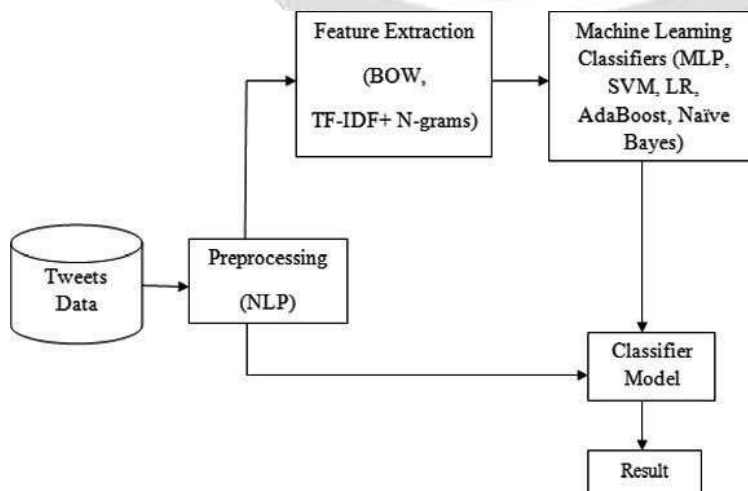
$$\text{Precision} = \frac{TP}{TP+FP}$$

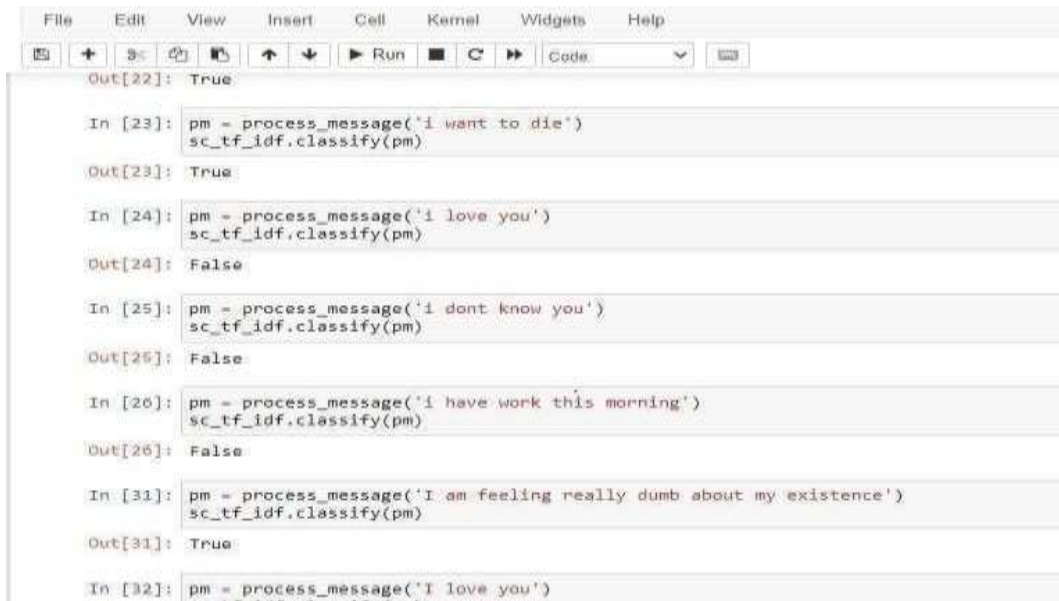
**Recall** – Recall is the ratio of correctly predicted positive observations to the all observations in actual class - yes. Recall is given by,

$$\text{Recall} = \frac{TP}{TP+FN}$$

### 5. Results

From above figure we can see the process involved and tweets data is collected in form of Dataset and its Preprocessed using NLP. Features from dataset are extracted, by using Machine Learning classifier. Efficient Model is used and result is been found.





```

File Edit View Insert Cell Kernel Widgets Help
+ 3c Run Code
Out[22]: True
In [23]: pm = process_message('i want to die')
sc_tf_idf.classify(pm)
Out[23]: True
In [24]: pm = process_message('i love you')
sc_tf_idf.classify(pm)
Out[24]: False
In [25]: pm = process_message('i dont know you')
sc_tf_idf.classify(pm)
Out[25]: False
In [26]: pm = process_message('i have work this morning')
sc_tf_idf.classify(pm)
Out[26]: False
In [31]: pm = process_message('I am feeling really dumb about my existence')
sc_tf_idf.classify(pm)
Out[31]: True
In [32]: pm = process_message('I love you')

```

In above figure we can see classifications based on the words, if its true it indicates depression, if its false it indicates not depressed. For example I love you is shown as false, which shows not depressed. I Want to die is shown as true, which means depressed.

## 6. Conclusion

To generate the greatest posterior probability value of the two review classes of sentiment, the Sentiment Analysis System for Depression Detection uses the Naive Bayes classifier approach. The first test was carried out by categorizing 209 tweets as random test data and training 10,105 tweets as the training data. The average accuracy generated by the Sentiment Analysis System for depression diagnosis is 87.74%. The accuracy value is directly related to the dataset's total number of reviews.

## 7. REFERENCES

- [1]. Govindasamy, K. A., & Palanichamy, N. (2021). *Depression Detection Using Machine Learning Techniques on Twitter Data. 2021 5th International Conference on Intelligent Computing and Control Systems (ICICCS)*. doi:10.1109/iciccs51141.2021.9432203
- [2]. Lin, C., Hu, P., Su, H., Li, S., Mei, J., Zhou, J., & Leung, H. (2020). *SenseMood: Depression Detection on Social Media. Proceedings of the 2020 International Conference on Multimedia Retrieval*. doi:10.1145/3372278.3391932