IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⪤ Impact Factor 8.066 ⪤ Peer-reviewed / Refereed journal ⪤ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10594

DEPRESSION DETECTION IN SOCIAL MEDIA BY MACHINE LEARNING ALGORITHMS

Dr Sharath Kumar Y H[1], Monica V A[2], Rajath K[3], Ruchitha H N[4], Rakshith B[5]

Professor & Head, Department of Information Science, Maharaja Institute of Technology Mysore, India[1]

Student, Department of Information Science, Maharaja Institute of Technology Mysore, India[2-5]

Abstract: Depression is a major contributor to the overall global burden of disease. Traditionally, doctors have diagnosed depression in people face-to-face using criteria for clinical depression. However, more than 70% of patients would not consult a doctor in the early stages of depression, which leads to further worsening of their condition. Meanwhile, people increasingly rely on social media to reveal emotions and share their daily lives, so social media has been successfully used to help detect physical and mental illnesses. Based on these inspirations, our work focuses on the early detection of depression through social media data collection. We create a well-labelled depression and non-depression dataset on Twitter and extract six clusters of depression-related features that cover not only criteria for clinical depression but also online social media behavior. With these feature groups, we propose a multimodal depression dictionary learning model to detect depressed users on Twitter. Finally, we analyse a large dataset on Twitter to uncover underlying online behaviors between depressed and non-depressed users.

Keywords: Depression Detection , Machine learning, Twitter dataset.

I. INTRODUCTION

In our current society, depression is a common disease that affects many people. Depression is a medical condition that can negatively affect the way a person thinks, feels, and behaves. According to the data provided by the World Health Organization, about 322 million cases were registered in 2015, of which about 788,000 cases ended in suicide. However, as serious as it sounds, there is still a stigma in society that having a mental disorder is a sign of weakness and can often lead to social exclusion. A study found that people tend to view depression as a poor treatment option, even though it is actually a serious problem. This can make people with mental disorders reluctant to seek professional help and reduce access to appropriate treatment. About 75-85% of people with depression do not have enough support to fight depression. In the current situation where people tend to turn to social media to solve their problems, this will not only provide psychiatrists and psychologists with more information before using the collected data, but also the possibility of early diagnosis. Topic social media platforms. This idea is supported by research showing that students with depressive symptoms use the Internet significantly more than those without. So researchers are under pressure to find the best way to suppress early detection.

Recently, there has been a movement to use social media data to identify, estimate and track changes in disease outbreaks. The proliferation of social media provides many opportunities to increase the data available to mental health practitioners and researchers, enabling a more informed and well-equipped mental health field. In addition, negative emotions that are contagious in social networks have a negative effect on people and lead to depression and other mental illnesses. Mental illness is recognized as a major risk factor for suicide. Approximately 80% of those who commit or attempt suicide suffer from some form of mental illness. Depression is considered the most common mental illness, but it is underdiagnosed and undertreated because it is under-recognized or dismissed. Major depression can be prevented by early recognition of its symptoms and timely intervention and treatment. Many studies have identified physical and mental illnesses caused by the large amount of information on social media, and there have even been studies on depression. D. Chowdhury and colleagues used tweets posted by people with major depressive disorder and their social media activity to categorize and categorize whether they suffer from depression or are likely to develop depression in the future. I found it predictable. Nadeem et al., Dredze et al., Benton et al. Tsugawa et al. focused on whether user tweets are of a depressing nature. and Coopersmith et al. Focus on user activity on Twitter. The study also aims to identify whether users are depressed by the nature of their tweets and activities on the network. In addition, it can be used to identify other psychiatric disorders and may form the underlying infrastructure for new mechanisms to help identify and limit the spread of depression in social networks.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588
International Advanced Research Journal in Science, Engineering and Technology
ISO 3297:2007 Certified ⁙ Impact Factor 8.066 ⁙ Peer-reviewed / Refereed journal ⁙ Vol. 10, Issue 5, May 2023
DOI: 10.17148/IARJSET.2023.10594

II. LITERATURE SURVEY

A. Park et al.shows that depressed Twitter users tend to post negative emotional tweets More emotions than healthy users.

B. D Chowdhury etc. It seemed depressing Signals are noticeable in tweets posted by users with major depressive disorder. So far, various features have been used to detect depression from Twitter data

C. Sugawa et al. shows that word usage frequencies, along with subject modeling, are useful features for predictive models. Using a radial kernel SVM classifier, they obtained 69% classification accuracy in predicting depression of 81 participants of 209 were collected using questionnaires.

D. Rees et al. extracted predictive features from users' tweets to measure affect, linguistic style, and context; built models using these features with supervised learning algorithms and successfully discriminated between depressed and healthy material. Their data were collected from 105 of 204 depressed users and CESD scores depend on the identification of depressed users. The best classifier performance was obtained using the 1200-tree random forest classifier, increasing the accuracy to 0.866 compared to other study results.

E. Nadeem etc. used a bag-of-words approach for better depression detection, which uses frequency of word occurrence to measure the content of tweets measured at the document level. They used four types of binary classifiers: linear SVM classifier, decision tree (DT), naive Bayes (NB) algorithm, and logistic regressive approach, and found that NB outperformed other classifiers with an accuracy of 81% and a precision of 0.86. They used a corpus of more than 2.5 M tweets collected online from CLPsych 2015 Shared Task organizers who indicated they were diagnosed with depression (326) or PTSD.

F. Jamil etc. concluded that using sentiment analysis with the percentage of depressed tweets increases the precision and recall of depression detection. The classifier was trained on 95 users who self-reported depression (which equaled 5% of the users who participated in the study, while the remaining 95% were healthy users), using SVM, which provided a recall of 0.875 and a precision of 0.775. was

III. PROBLEM STATEMENT

Detecting depression in people using a social media platform to analyze their sentiment through text data, emotions and other features using deep learning technology. We use different machine learning approaches to train data and evaluate the effectiveness of our proposed method. For more accurate results, we would use a combination of different feature extraction techniques with machine tilt classifiers. The use of social media for health care, and especially the detection of depression, is feasible.

The main objectives are:

✓ The main objective of our study is to examine the posts of social media users to find out any factors that may reveal the depressed attitude of the respective online users.
✓ The purpose of this systematic review is to summarize the findings of previous studies related to applying machine learning (ML) methods to data text from social media to detect depressive symptoms and suggest directions for future research in this area.
✓ The features are based on linguistic analysis of the text followed by various machine learning approaches to get the best performance.

IV. METHODOLOGY

A quantitative study is conducted to train and test different machine learning classifiers to determine Whether a Twitter account user is frustrated, with user-initiated tweets or his/her activities on Twitter. Data preparation, feature extraction and classification tasks are performed. Classifiers are trained using 10-fold cross validation to avoid overfitting, and then tested on the hold-out test set. First, all tweets for depressed and non-depressed accounts, as well as user account and activity information such as number of followers, number of followers, total number of 6 posts, post time, number of mentions, and number of retweets, are retrieved. Next, all tweets from the account are assembled into one document.
Text preprocessing is applied to all documents. First, a corpus is created and the tweets in each document are tokenized. Next, normalization is applied, where all characters are converted to lowercase and punctuation marks, retweets, mentions, links, unknown emojis and symbols are removed. Usually, generalization involves removing stop words, such

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁒ Impact Factor 8.066 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10594

as first-person pronouns like "I," "me," and "you," but when we remove stop words, we keep first-person pronouns. Later, stemming is applied and a Document Term Matrix (DTM) is created for each account. The matrix shows the frequency of words in each tweet, where each row indicates a document of tweets and each column indicates all words used in all accounts. TF-IDF is used to measure the weight of terms. Features applied to DTM are then merged with account criteria extracted from social Network and user activities. The results of the merge are then treated as independent variables in a classification algorithm to predict the dependent variable of the outcome of interest. Finally, we decide on DT, a linear and radial kernel support vector classifier.

Feature Engineering

Feature engineering is referred to in machine learning as "the process of using domain knowledge of data to create features that can be used by machine learning algorithms to find patterns." Features are generated to extract information understood by machine learning algorithms and can be useful for prediction. The number of types and features has a significant impact on the performance of machine learning algorithms.

The Twitter platform contains a large amount of information about the user, various features can be obtained from Twitter users' activity history and tweets. Features are extracted from the text after text pre-processing when the text is in the desired format, where these features are counted for both training and test sets as . The features used for the classification model and their possible values, where the possible values T (true) and F (false) indicate whether the feature is used or not. For example, when the probability value for TF-IDF is T it means TF-IDF is used for experiment and if it is F it means word frequency is used instead. Feature Extraction aims to reduce the number of features in a dataset by creating new features from the existing ones (and then discarding the original features). These new reduced set of features should then be able to summarize most of the information contained in the original set of features. In this way, a summarised version of the original features can be created from a combination of the original set.

Machine Learning Classifiers

Support vector machine algorithm

Support Vector Machine or SVM is one of the most popular supervised learning algorithms, used for classification as well as regression problems . However, mainly, it is used for classification problems in machine learning.

The goal of the SVM algorithm is to create an optimal line or decision boundary that can divide the n-dimensional space into classes so that we can easily place new data points into the correct category in the future. This optimal decision boundary is called a hyperplane. SVM selects the extreme points/vectors that help to construct the hyperplane. These extreme cases are called support vectors, and hence the algorithm is known as a support vector machine. Consider the following diagram in which there are two different categories classified using a decision boundary or hyperplane:

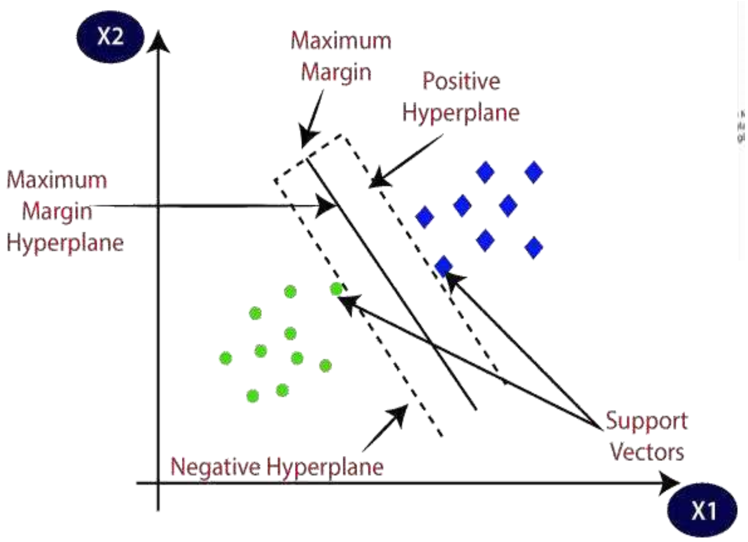


Fig 1: Diagram for Support Vector Machine with hyperplanes

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⁒ Impact Factor 8.066 ⁒ Peer-reviewed / Refereed journal ⁒ Vol. 10, Issue 5, May 2023

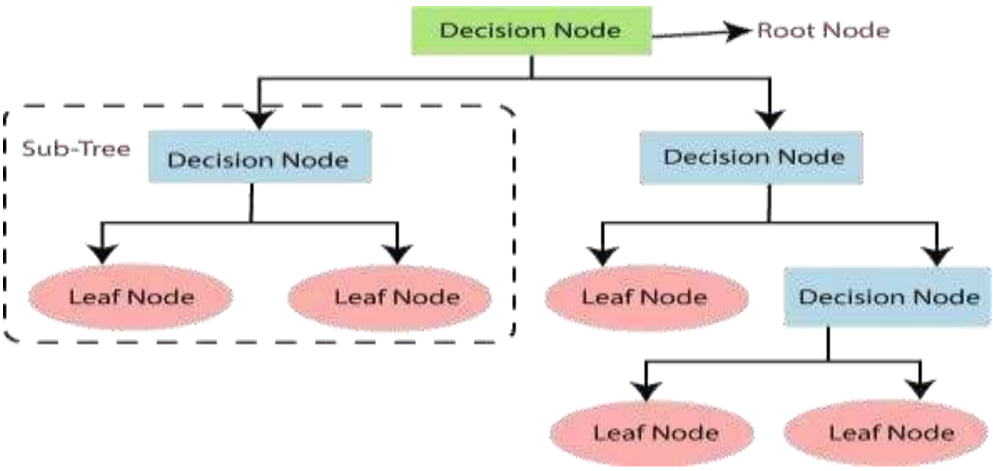
DOI: 10.17148/IARJSET.2023.10594

Decision Tree Classifier

Decision trees are a supervised learning technique that can be used for both classification and regression problems, but are mostly chosen for solving classification problems. It is a tree structured classifier, where internal nodes represent features of the dataset, branches represent decision rules, and each leaf node represents an outcome.

In a decision tree, there are two nodes, which are the decision node and the leaf node. Decision nodes are used to make any decision and have multiple branches, while leaf nodes are the output of those decisions and contain no further input. branches. Decisions or tests are made based on the attributes of a given dataset. It is a pictorial representation to derive all possible solutions to a problem/decision based on given conditions. It is called a decision tree because, like a tree, it starts from a root node, which extends to further branches and forms a tree-like structure.

To build the tree, we use the CART algorithm, which is used for classification and regression tree algorithms. A decision tree simply asks a question, and based on the answer (yes/no), it splits the tree into subtrees.



Naïve Bayes Classifier:

NB is based on "Bayes theorem" in probability. As a requirement of this theorem, NB can be applied only if the attributes are independent of each other.

$$P(X|Y) = \frac{P(Y|X)}{P(X)/P(Y)}$$

It is a prediction model that breaks down the probabilities behind each class and the possible circumstances of the class for each trait. It is commonly used in machine learning because of its ability to efficiently merge evidence from multiple features. Often, we know how often some evidence is observed, given a known outcome [17]. Certain evidence provides us with a conclusion with the knowledge that it is observed. The NB classifier is considered the simplest method in the machine learning field, although it is still competitive with SVM.

V. EXPERIMENT AND ANALYSIS

Dataset collection:

We use a dataset of depressed Twitter users. Self-reports are collected by searching Twitter using a regular expression ("..diagnosis of depression…"). Candidate users are manually filtered and then all their recent tweets are continuously searched using the Twitter Search API. To ensure that users are disclosing their own depression and not talking about a friend or family member, a human annotator reviews these tweets. For each user, up to 3,000 of their most recent public tweets are included in the dataset, and each user is isolated from the others. Note that this limit of 3000 tweets is derived from Twitter's archive policy . Non-depressed users are collected randomly and checked manually to ensure that they never post any tweet containing the character string "depressed". In an effort to minimize noisy and unreliable data, users with fewer than five Twitter posts are excluded.

DOI: 10.17148/IARJSET.2023.10594

Evaluation measures:

Critique and cross validation of the feasibility of these automatic predictions will be conducted by standardizing accuracy (Acc), precision (P), recall (R), and F1 scores, as well as confusion matrix (CM), and receiver operating classification curves. ROCs), which are defined as follows:

Accuracy: The simplest and most commonly used measure to evaluate a classifier. It is defined as the degree of correct prediction of the model (or conversely, the percentage of misclassification errors).

$$Acc = \frac{true\ positive + true\ negative}{true\ positive + true\ negative + false\ positive + false\ negative}$$

Precision: Defined as the fraction of correctly classified positives to the total predicted positives. Under our scenario, the objective is to find out how many of the users identified as depressed are actually depressed.

$$P = \frac{true\ positives}{true\ positives + false\ positives}$$

Recall: Defined as the fraction of correctly classified positives and total positives. In our scenario, it aims to determine how many of all frustrated users are correctly detected.

$$R = \frac{true\ positives}{true\ positives + false\ negatives}$$

A trade-off is made between recall (false negatives) and precision (false positives). Considering the F1-measure:

F1 score (F-measure): is the harmonic mean of precision and recall; It weights each metric equally, and is, therefore, commonly used as a classification evaluation metric.

$$F1 = \frac{2 * P * R}{P + R}$$

Therefore, it is important to achieve both high recall and high precision.

Confusion Matrix: is a form of contingency table that represents the differences between true and predicted classes for a set of labeled instances . It has four categories: true positive (TP), which refers to correctly identified positives; false positives (FP), which are falsely identified positives and assumed to be negatives; True negatives (TN), which refer to negatives correctly labeled as negatives, and false negatives (FN), correspond to positives that are incorrectly labeled as negatives.

CM can be used to generate a point in ROC space using a matrix defined as:

$$CM = \begin{bmatrix} TN & FP \\ FN & TP \end{bmatrix}$$

VI. RESULTS

Experiments are conducted on all possible combinations of feature values shown in the table, using different classification algorithms (SVM with different kernels, DT, and NB). The expected labels for any training/test sample are not depressed/depressed. From previous studies, "first person pronoun" and "TF-IDF" have been proven to be discriminative for the identification of depression. This finding has been proven during experiments conducted by measuring the correlation between features and class labels. In order to get the best classification results, different feature combinations are used. In addition, training and testing are conducted to arrive at the best feature selector.

After conducting several experiments with variations in the exploited features, the results emphasize that enriching the model with discriminative features gives better results. Moreover, the SVM-linear classifier shows superior results and invariant behavior despite its extensive operational complexity.

IARJSET ISSN (O) 2393-8021, ISSN (P) 2394-1588

International Advanced Research Journal in Science, Engineering and Technology

ISO 3297:2007 Certified ⪤ Impact Factor 8.066 ⪤ Peer-reviewed / Refereed journal ⪤ Vol. 10, Issue 5, May 2023

DOI: 10.17148/IARJSET.2023.10594

Figure 3 shows the overall effect of the feature set (SVM-linear), which increases the frequency of words used by depressed people and strengthens the words when fed to a classification model. It also reduces the number of words in the corpus, which reduces the computation time.

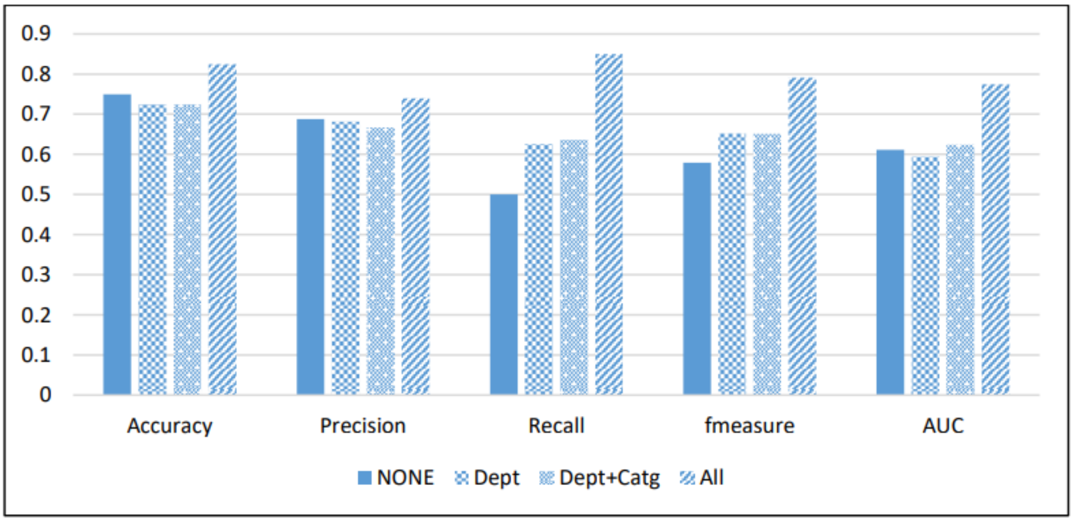


Figure 3: Results of Applying all features in SVM - L

VII. CONCLUSION AND FUTURE WORK

This paper defines the binary classification problem as identifying whether a person is depressed based on his tweets and Twitter profile activity. Different machine learning algorithms are used and different feature datasets are searched. Several preprocessing steps are performed, including data preparation and alignment, data labeling, and feature extraction and selection. The SVM model achieved the best accuracy metric combinations; It turns a highly non-linear classification problem into a linearly separable problem. Although the DT model is comprehensive and follows understandable steps, it can fail if it is exposed to brand new data. So this study can be considered as a step towards creating a complete social media-based platform

Analyzing and predicting mental and psychological problems and recommending solutions for these users. The main contribution of this study lies in exploiting a rich, diverse, and discriminative feature set that includes both tweet text and behavioral trends of different users. This study can be expanded in the future by considering more ML models that are highly unlikely to over-fit the data used and find a more reliable way to measure the effect of feature.

REFERENCES

[1] Shen, G., Jia, J., Nie, L., Feng, F., Zhang, C., Hu, T., Chua, T.-S., & Zhu, W. Depression detection via harvesting social media: A multimodal dictionary learning solution in IJCAI 3838-3844 (2017).
[2] Mowery, D., Bryan, C., & Conway, M. Feature studies to inform the classification of depressive symptoms from Twitter data for population health. Preprint at arXiv:1701.08229 (2017).
[3] Coppersmith, G., Dredze, M., & Harman, C. Quantifying mental health signals in Twitter in Book Quantifying mental health signals in Twitter 51-60 (2014).
[4] Stack, S. J. Mental illness and suicide. The Wiley Blackwell Encyclopedia of Health, Illness, Behavior, and Society. 1618-1623 (2014).
[5] De Choudhury, M., Kiciman, E., Dredze, M., Coppersmith, G., & Kumar, M. Discovering shifts to suicidal ideation from mental health content in social media in Discovering shifts to suicidal ideation from mental health content in social media 2098-2110 (ACM, 2016).
[6] Organization, W. H. Depression and other common mental disorders: global health estimates', 2017
[7] Yazdavar, A.H., Al-Olimat, H.S., Banerjee, T., Thirunarayan, K., & Sheth, A.P. Analyzing clinical depressive symptoms in twitter (2016).
[8] De Choudhury, M., Gamon, M., Counts, S., & Horvitz, E. Predicting depression via social media. ICWSM. 13, 1-10 (2013).
[9] Reece, A.G., et al. Forecasting the onset and course of mental illness with Twitter data. Sci. Rep. 7(1),13006 (2017).
[10] Nadeem, M. Identifying depression on Twitter. Preprint at arXiv:1607.07384 (2016).