International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-6, April 2020

Relative Perusal of ML Classifiers for Depression Detection in Twitter Feeds

Piyusha Sahni, **C.N. Subalalitha**

Abstract**: *Depression is viewed as a significant cause of suicidal inclination. It affects the style of writing manifested in the text. Analyzing linguistic markers in social media posts help in recognizing and classifying whether thoughts or sentiments expressed in source text correspond to a depressed user. A large amount of emotion-rich data generated by social media platforms is in the form of tweets, feeds, blog posts, etc. Analysis of this user-generated data helps in understanding an individual's state of mind. The main focus is to scrutinize the posts of users of twitter to analyze the depression attitudes of users. Natural Language Processing and ML techniques like MNB, TF-IDF, SVC, SGD, and LR have been utilized for training the data set and estimating the efficacy of our proffered approach. Firstly, the words are reduced into their morphological form during pre-processing. Then, a depression analysis model is built based on the suggested method and various features of depressed users derived from psychological research. Tweets with the hashtags #depression are classified based on their content and their relative tendencies towards depression. Tweets related to social distance, workplace stress, peer pressure, family problems, personal weakness, failure, mocking, and self-stigma helped in depression detection. The results have been rendered using the key evaluation measures, which include accuracy, precision, and F1-score. The results of the study may be beneficial in assisting mental health awareness and supporting organizations to provide data about resources and counter common depression.***

Keywords**: *depression analysis, ML, natural language processing, twitter.***

I. INTRODUCTION

Depression is a condition that affects the general state of mental wellness impairing many aspects of our lives. Despite its inconceivable symptoms, it has an acute impact. The pressures of everyday incidents have increased the odds of depression. It often coexists with stress or other psychosomatic disorders; and affects the state of mind and behavior of the affected individuals. According to the World Health Organization (WHO), One in five people suffers from depression. Worldwide, depression is still under-analyzed and left with no satisfactory treatment, which can usher to a genuine self-discernment and, at the very least, to suicide. Moreover, the social shame encompassing discouragement keeps many affected people from looking for fitting proficient help.

Revised Manuscript Received on April 30, 2020.
***** Correspondence Author
Piyusha Sahni*, Computer Science Engineering, SRM institute of Science and Technology, Chennai, India. Email: piyusha14.sahni@gmail.com
Dr. C.N. Subalalitha, Computer Science Engineering, SRM institute of Science and Technology, Chennai, India. Email: subalaln@srmist.edu.in

Subsequently, they go to less conventional assets, for example, social media. In this day and age, the worries of day by day life occasions may build odds of melancholy. With the increase in the use of the internet, individuals have begun to impart their experiences to psychological wellness issues through posts on social media platforms like twitter. Their online exercises enlivened numerous researchers to aid novel health care measures and strategies for rathe identification of depression. Various text analysis and NLP methods have been incorporated to analyze depression attitude in the tweets.

II. LITERATURE SURVEY

It employed the lexicon-based approach to identify the words mostly used by depressed users. MLP classifier gave the best results for finding the signs of depression in the combined feature set. Besides, bigram showed maximum accuracy among single feature sets [1].

In this work, SVM, MNB, and RF were used to extract the emotion of users in conversations related to football matches in their tweets. SVM showed the best results with the BOW model(unigram). The negative responses can also be used to convey the authorities about possible riots after the match [2].

The paper identifies the presence of sarcasm in the tweet using various classification techniques like RF, KNN, SVM, and Maximum Entropy. POS tags have been used to identify tenors, indicating the degree of sarcasm in them [3].

Emotions have been classified on the basis of the lexicon. BOW has been used for feature extraction. Besides, the data set also includes weather statistics to estimate its impact on crowd sentiment [4].

The researchers have classified tweets as negative or neutral based on an organised list of words that relate to depression. MNB has shown maximum accuracy of 83% as compared to that of SVM, which is 79% [5].

In this project, the emoticons in the tweets are considered along with text analysis. The positive and negative emoticons are replaced with EPOS and ENEG in the text, which further helps in condensing the correct sentiment of the tweet. MNB and SVR have been used to classify the mixed tweets on the grounds of depression and anxiety. SVR shows a higher accuracy of 79.7% in comparison to MNB [6].

This work focuses on classifying the overall sentiment polarity of text on twitter based on five sentiment classes and generating a polarity score. Amongst different classifying techniques like RF, SVR, DT, and SoftMax. Decision Trees shows the highest accuracy at 91.81% [7].

This research paper lays emphasis on the numerous kinds of pre-processing methods which can effectively bring vast changes in the accuracy of the classification.

Replacing negation and expansion of acronyms has proved more productive in comparison to commonly used noise removing techniques like elimination of stop words, punctuations, numbers, and URL's [8].

Identification of "hate speech" in the text is implemented using SVM, RF, and J48graft. J48graft has been proved better than the other algorithms, with an accuracy of 87.4% in terms of text classification on the basis of their polarities, i.e., offensive and clean text. It also arrives at an accuracy of 78.4% when classified in terms of hateful, offensive and clean [9].

The study focuses on finding the relation between the usage of expressions and the actual emotion of the person writing the text. It evaluates dependability and certainty of various expressions used in the digital world with respect to the actual sentiment of the user using Naive Bayes classifier and accentuate the basis for assigning polarity to each expression [10].

III. METHODOLOGY

The earlier projects displayed a primary study of the existing methods which can be incorporated together to accomplish better outcomes from feature extraction of text. Our data set consists of 1.6 million tweets consisting of catchphrases, for example, depression, stress, anxiety, and mental illness, which relate to a gloomy attitude of the user towards life. Tweet pre- processing is the elementary step in sentiment analysis. Also, feature extraction is requisite to achieve the desired results. Different classifying techniques have been employed in order to compare the polarity of the text for each classifier.

Table- I: Data Set Summary.

Dataset	Total
Training	1200000
Testing	400000

A. *Pre-processing*

The pre-processing of the data acts as a foundation for the structure of our work. The aim behind pre-processing is to remove all the noise and redundancy, causing words or characters from the text that has no relation to sentiment classification. This initial step makes further working more efficient. Thus, it's an unavoidable step for any research work. The following steps are involved in data pre-processing.

- Case conversion: All the words in the data set are converted to lowercase.
- Removal of punctuation, unnecessary blank spaces, and hashtags.
- Usernames: Users often include Twitter usernames of people in their feeds by inserting the @ symbol before the username. All such usernames are to be taken care of by replacing the text following @ symbol with a token.
- Repeated letters: Tweets contain very candid usage of the writing style. To count that set of language in our work, we will use pre-processing in order to replace the letter occurring more than two times in a row is with two occurrences. For instance, if you search "depression" with a random count of ss in it (e.g., depresssssion, depressssion, depression), then all these words would be reworked into the token depression.
- Links in a tweet: All URL's in the feed is replaced into a token. For instance, conversion of URL like "http://tinyurl.com/abc99" to the token 'URL'.
- Stemming: Reducing the terms into their morphological form to team up terms that are alike.
- Removal of stopwords: The words that do not provide any meaningful content to the data set. So, removing them cleans the data significantly.

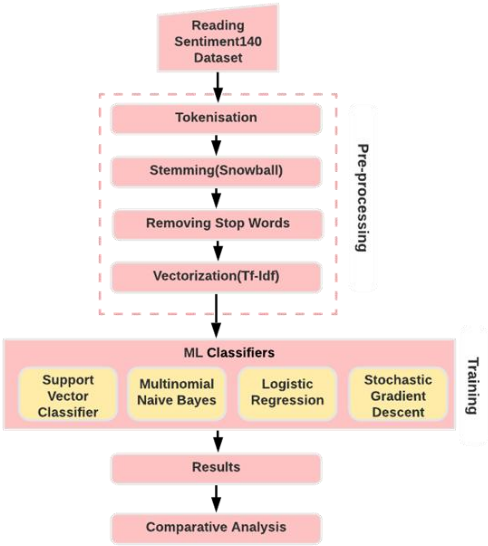

Fig. 1.Training and Testing Framework.

B. *Feature Extraction*

- Tokenization: An entire text document consisting of phrase, sentence, or paragraph into smaller units called tokens. Tokens can further be used as features that can be used as inputs for different classifiers.
- Stemming: Removing the suffix of the word to get its base form so that redundancy can be reduced is called stemming. This helps in improving natural language processing because different words are batched into one term.
- Lemmatization: Lemmatization does a morphological analysis of the words. For example, words like stressed, stresses, stressful, and stress can be grouped into a single word "stress".
- N-Gram: The sentence is converted into a sequence of tokens with n words in each parsing with inclusive use of porter stemmer class for word stemming. In the key-value pair, the key is the word generated after tokenization.
- TF-IDF: It counts the occurrence of a term in the given text and reflects it in the given document. We use it to generate n-grams. It is mainly plied to minimize the less informative tokens in comparison to those which occur less frequently but are highly edifying and thus are designated with a higher value of TF-IDF.
- Modeling of n-grams is done using the TF-IDF vectorizer.

Retrieval Number: F3721049620/2020©BEIESP
DOI: 10.35940/ijitee.F3721.049620
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

International Journal of Innovative Technology and Exploring Engineering (IJITEE)
ISSN: 2278-3075 (Online), Volume-9 Issue-6, April 2020

Fig. 2.Tag Cloud of Depressive Words.

Fig. 3.Tag Cloud of Positive Words.

C. Classifiers

- Case conversion: All the words in the data set are converted to lowercase.
- Multinomial Naive Bayes (MNB): Given a class, it estimates the conditional probability of a particular word, term, or token as the relative frequency of a specific term belonging to that class in the document. It is employed mostly when numerous instances have a lot of significance in the classification of texts.
- Support Vector Machine (SVM): It analyzes the data and detects patterns used for classification. Data are represented as polar coordinates in a high dimensional region bestowed for division in which the parameters associated with distinct categories are widely divided. Test data is then depicted in the same area and depending on which region they fall into; they are assigned a class. In the context of this paper, the Support Vector Classifier helps in detecting whether the given twitter feed belongs to the set predefined categories using multiple optimal strategies.
- Stochastic Gradient Descent (SGD): NLP and text classification problems use SGD to deal will large ML challenges. Updating of the coefficients is performed for each training instance when the gradient descent procedure is executed and not at the end of the batch of instances.
- Logistic Regression (LR): Based on one or more predictors and features, LR works as a linear classifier algorithm that estimates the probabilistic occurrence of a binary response. While it may technically not qualify as a classification method, it represents a discrete choice model, and we, therefore, use it as such.

IV. RESULT AND DISCUSSION

To test the classification techniques mentioned above, test metrics such as the accuracy of estimations (Accuracy) and F- score consisting of precision and recall has been worked with. It is hinged on an error matrix, which helps in determining the performance of classifiers used. Accuracy of estimations measures the precise classification of text.

$$\text{Accuracy}=\frac{\text{Truepositive+Truenegative}}{\text{Total Predictions}} \quad (1)$$

$$\text{Precision}=\frac{\text{Truepositive}}{\text{Truepositive+Falsepositive}} \quad (2)$$

$$\text{Recall}=\frac{\text{Truepositive}}{\text{Truepositive+Falsenegative}} \quad (3)$$

$$\text{F1 Score}=2.\frac{\text{Precision.Recall}}{\text{Precision+Recall}} \quad (4)$$

F1 score assesses the accuracy of the experiment. Precision computes how many samples were correctly identified, remembering how well the samples were correctly identified. A higher F1 score is obtained when both the values are close enough. Several true-positive, true-negative, false-positive, and false-negative predictions have been found in the evaluation metrics. F1 score is considered the chief criteria for evaluation of performance and accuracy is the subsidiary one. It can be seen from the basis of the conclusion the Vector Classifier performed very well with 73% accuracy and the Logistic Regression achieved 72% accuracy. Fig.5, Fig.6, Fig.7 and Fig.8 represents the normalized confusion matrix for SVC, MNB, LR, and SGD, respectively. The classifiers are awarding slightly lower accuracy because the feeds include texts which is not in a definitive style. For example, "gn" is commonly used instead of Good Night. Therefore, training the classifier and obtaining significant results is a tad demanding. According to the results, the best approach for our model will be the Support Vector Classifier since it gives a better combination accuracy, precision, and F1 score. Besides, it also predicts the tweets accurately in which the sentiment is not clearly specified in the form of negation or affirmation.

This is due to the massive data set which we have used. This inference has been communicated through Fig.4.

```
[31] classify_with_models('I feel like overeating after all what has happened')

I feel like overeating after all what has happened
LR: Depressed   MNB: Depressed   SVC: Depressed   SGD: Depressed
```

Fig. 4.Execution of Tweet.

V. CONCLUSION

In this work, a relative analysis based on different classifiers has been performed for detecting depressive nature in the feeds on Twitter and performance measures have been enhanced. A nigh link between the emotional aspect of the text and style of writing has been determined using text classification and NLP techniques. According to our results, predictors of depressed language had words related to depression, sorrow, consternation, wrath, loneliness, or self-destruction. Various text classification methods have been employed to measure depressive symptoms and the achievements of single features, as well as mixed feature sets, have been scrutinized. The strength and efficacy of integrated features demonstrated using SVC have achieved 73% accuracy and 0.72 F1 score. Besides, unigram has come up as the best one during feature extraction. Although our experiment indicates that the performance of the methods used is competent, the figures in the metrics call for further grilling in this study. It is presumed that our work accentuates the framework for novel mechanisms utilized in distinct regions of wellness programs to identify depressive tendencies. The people suffering from psychological disorders can also be benefited by using it to drive them towards their speedy recuperation. We will also attempt to explore the bond between the user's demeanor and their behaviors related to the depressions expressed on social media in our future work.

Table-II: Comparison of Metrics of Classification Algorithms.

Name	Precision	Recall	F1 Score	Accuracy
LR	0.75	0.68	0.72	0.7275
MNB	0.71	0.73	0.72	0.7143
SVC	0.76	0.68	0.72	0.7306
SGD	0.76	0.67	0.71	0.7262

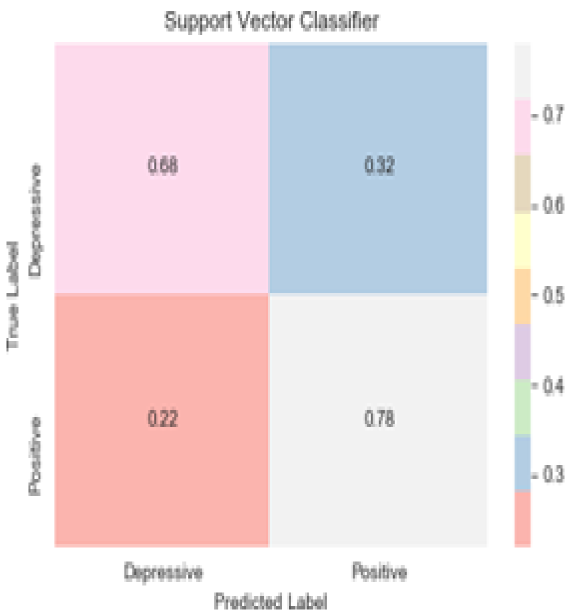

Fig. 5.Error Matrix for SVC.

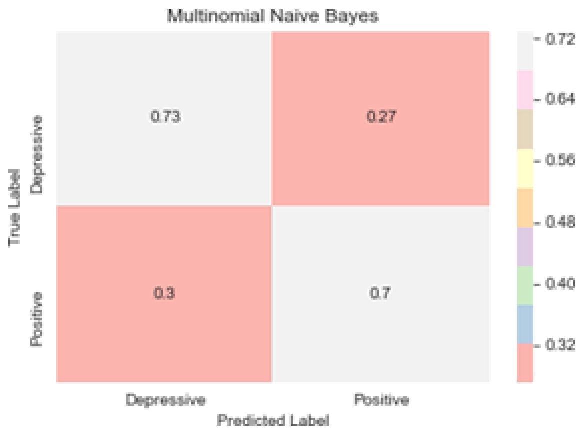

Fig. 6.Error Matrix for MNB.

Fig. 7.Error Matrix for LR.

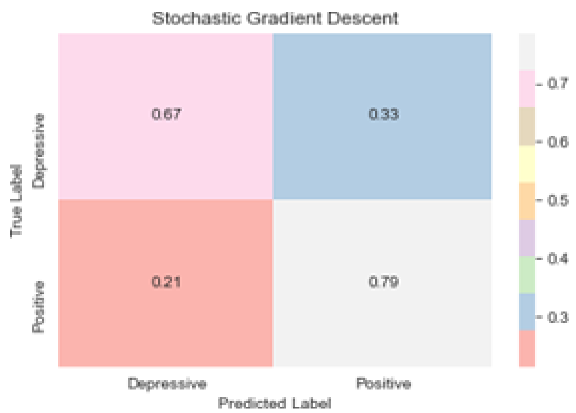

Fig. 8.Error Matrix for SGD.

REFERENCES

1. M. M. Tadesse, H. Lin, B. Xu and L. Yang, "Detection of Depression-Related Posts in Reddit Social Media Forum," in IEEE Access, vol. 7, pp. 44883-44893, 2019.
2. S. Aloufi and A. E. Saddik, "Sentiment Identification in Football-Specific Tweets," in IEEE Access, vol. 6, pp. 78609-78621, 2018.
3. M. Bouazizi and T. Otsuki Ohtsuki, "A Pattern-Based Approach for Sarcasm Detection on Twitter," in IEEE Access, vol. 4, pp. 5477-5488, 2016.
4. Kamal, Rashid & Shah, Munam & Maple, Carsten & Masood, Mohsin & Wahid, Abdul & Mehmood, Amjad. (2019). Emotion Classification and Crowd Source Sensing; A Lexicon based Approach. IEEE Access. PP. 1-1. 10.1109/ACCESS.2019.2892624.

Retrieval Number: F3721049620/2020©BEIESP
DOI: 10.35940/ijitee.F3721.049620
Journal Website: www.ijitee.org

Published By:
Blue Eyes Intelligence Engineering
& Sciences Publication

5. M. Deshpande and V. Rao, "Depression detection using emotion artificial intelligence," 2017 International Conference on Intelligent Sustainable Systems (ICISS), Palladam, 2017, pp. 858-862.
6. P. Arora and P. Arora, "Mining Twitter Data for Depression Detection," 2019 International Conference on Signal Processing and Communication (ICSC), NOIDA, India, 2019, pp. 186-189.
7.]S. E. Saad and J. Yang, "Twitter Sentiment Analysis Based on Ordinal Regression," in IEEE Access, vol. 7, pp. 163677-163685, 2019.
8.]Z. Jianqiang and G. Xiaolin, "Comparison Research on Text Pre-processing Methods on Twitter Sentiment Analysis," in IEEE Access, vol. 5, pp. 2870-2879, 2017.
9.]H. Watanabe, M. Bouazizi and T. Ohtsuki, "Hate Speech on Twitter: A Pragmatic Approach to Collect Hateful and Offensive Expressions and Perform Hate Speech Detection," in IEEE Access, vol. 6, pp. 13825-13835, 2018.
10.]F. T. Giuntini et al., "How Do I Feel? Identifying Emotional Expressions on Facebook Reactions Using Clustering Mechanism," in IEEE Access, vol. 7, pp. 53909-53921, 2019.

AUTHORS PROFILE

Piyusha Sahni is pursuing Bachelor of Technology in Computer Science Engineering from SRM Institute of Science and Technology, Kattankulathur Campus, Chennai. She is presently in her final year in SRM. She has completed her higher secondary education from Carmel Convent Girls Senior Secondary School, BHEL, Bhopal (2004 – 2016) and has been certified for overall general proficiency from 2012-2016. She was also the content writer for MOZILLA campus club in her college. Her research interests include Natural Language Processing, Machine Learning, Sentiment Analysis, Personality Prediction and Data Analytics. She has completed projects in the field of Artificial Intelligence, Natural Language Processing, Big Data and Cloud Computing.

Dr. C.N. Subalalitha, Ph. D., associate professor, Department of Computer Science and Engineering, Kattankulathur Campus, SRM Institute of Science and Technology. She graduated from SSN College of Engineering, Chennai, Anna University in 2006 with a Master of Engineering in Computer Science and Engineering. In 2014, she received a doctoral degree in Computer Science and Engineering from College of Engineering Guindy, Anna University. She has been a member of Indian Science Congress and IET and has received Gold Medal in the Research Day held at SRM IST in the years, 2016, 2017 and 2018 consecutively. Since 2015, she has been involved in the teaching with specialization in Natural Language Understanding, Natural Language Processing, Text Mining, Computational Linguistics, Machine Learning, Data Structures & Programming and Design in C. At present, her research content comprises of Discourse Analysis, Text Extraction, Image Extraction and Chatbots. She has authored many papers, and has published them in various International Conferences and journals. She has also delivered keynote lectures on topics such as Semi-Supervised Machine Learning Approach for NLP task, Tensor-Flow and Document Text Mining.