iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** V **Month of publication:** May 2024

DOI: https://doi.org/10.22214/ijraset.2024.62513

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 12 Issue V May 2024- Available at www.ijraset.com

Machine Learning Approach to Detect Depression through Social Media Posts

Aloukik Verma[1], Archana Verma[2], Er. Shubham Kumar[3]
[1, 2]*Undergraduate Students, Department of Computer Science and Engineering, SRMCEM, Lucknow, Uttar Pradesh, India,*
[3]*Assistant Professor, Department of Computer Science and Engineering, SRMCEM, Lucknow, Uttar Pradesh, India*

Abstract: ***Depression is one of the biggest issues which is increasing rapidly day by day. It has always been a topic for various researchers to discover new methods and techniques to overcome it. It is basically a kind of disorder in which people do not feel happy and joy. Since with the increase of the social networking sites people are expressing their feelings on the social media platforms which are helping the various researchers to analyze these posts and find an effective solution. Now a days there are various studies that aim to exploit machine learning techniques for detecting a probable depressed Twitter (Now X) user tweets. Textual data can be collected from the comments on the post, exchanging information etc. Using Facebook, Twitter, and Instagram. There is a huge amount of data which is present on these platforms which can be used for research. The aim of this study is that whether machine learning can be able to effectively detect the various signs of depression by analyzing the user post. This paper is basically on the sentimental analysis using NLP that identifies that the post is positive, negative or neutral.***
Keywords: Machine learning, NLP, social media, depression, sentimental analysis

I. INTRODUCTION

In our current society, depression is a common condition that many people have. Depression is one of the most frequent and devastating mental illnesses that has a significant social impact. It is common knowledge that mental health is an important aspect of public health. Depression is one of the most common causes of disability in the world. Millions of individuals suffer from depression around the world. So, using various techniques we can be able to identify depression. With the various algorithms we can also classify which type of tweet it is. The emotion and language used in social media postings may indicate feelings of worthlessness, guilt, helplessness, and self-hatred that characterize major depression. Additionally, depression sufferers often withdraw from social situations and activities. Such changes in activity might be salient with changes in activity on social media [4]. Globally, the proportion of the population with depression in 2015 was estimated to be 4.4% (more than 332 million people). Depressive disorders are ranked as the single largest contributor to non-fatal health loss. More than 80% of this non-fatal disease burden occurs in low- and middle-income countries. Furthermore, between 2005 and 2015 the total estimated number of people living with depression increased by 18.4% (World Health Organization, 2017) [1]. Users are categorized as risk or non-risk (of depression). Each user produced a sequence of reddit posts, written within a given period of time. The pilot task was organized in two stages: training and test, each having a different dataset divided into 10 chunks. During the training stage, a dataset containing a sequential set of posts per user was provided along with the user's category. All training chunks were made available, containing the complete user post sequence [2]. There are various users of social media which are also increasing day by day. There is an urgent need to tackle depression by identifying risks as soon as possible using social media platforms. Identifying depressed people online can pave the way for locating and helping people who need professional care but do not have access to them owing to the aforementioned barriers [3]. In the paper [5] the depression was predicted from Twitter data in a Japanese sample where he showed that the features based on a topic modeling are useful in the tasks for recognizing depressive and suicidal users. Bentoni et al. In this paper we want to examine the relationship between depression and user's language usage, and which type of language the user is posting on the social platforms.

II. LITERATURE REVIEW

In the paper [1] (Burdisso, S. G., Errecalde, M., & Montes-y-Gómez, M. (2019)) it is mentioned that there is a need for the early detection of the depression with the help of the machine learning techniques, and it is also able to understand that there is enough data available for the training of the model that may be able to identify the depression in the human being or not. This model is called SS3, whose goal is to provide support for incremental classification, early classification and explainability in a unified, simple and effective way.

©IJRASET: All Rights are Reserved

The paper [2] uses the reddit posts of a user to identifythe risk of mental health issues from user-generated content in social media. Several approaches based on supervised learning and information retrieval methods were used to estimate the risk of depression for a user. In this have designed a multipronged approach that combines results obtained from both Information Retrieval (IR) and Supervised Learning (SL) based systems. The combination is performed by a decision algorithm.

In the paper [3] it has been mentioned that they have used the Facebook and twitter APIs to collect the post of the user for identifying whether the person is depressed or not and to train the naïve bayes classifier, also computing the (Term Document Matrix) and convert the data set in frequency table using TF-IDF (term frequency–inverse document frequency and using the naïve bayes to identify that the comments are depressed or not depressed.

The paper [4] describes that the data is collected with crowdsourcing, then used the CES-D (Center for Epidemiologic Studies Depression Scale)2 questionnaire as the primary tool to determine the depression levels of the crowd workers. Also used the Beck Depression Inventory (BDI) for this purpose (Beck et al., 1976). Like CES-D, BDI is also used commonly as a test for depression by healthcare professionals to measure depression. For behavioral exploration and prediction, we collected data from the Twitter feeds of all users. We used the Twitter Firehose made available to us via our organization's contract with Twitter. Here they have used four measures of the emotional state of users in our dataset: positive affect (PA), negative affect (NA), activation, and dominance.

The paper [5] describes that as ground truth data, they have used the results of a web-based questionnaire for measuring degree of depression of Twitter (now X) users and extracted several features from the activity histories of Twitter users. By leveraging these features, they construct models for estimating the presence of active depression. Through experiments, it is shown that (1) features obtained from user activities can be used to predict depression of users with an accuracy of 69%, (2) topics of tweets estimated with a topic model are useful features, (3) approximately two months of observation data are necessary for recognizing depression, and longer observation periods do not contribute to improving the accuracy of estimation for current depression; sometimes, longer periods worsen the accuracy.

In the paper [6] (Kaur, 2015) describe regarding geographic area flood information set collected from twitter and realize the opinion of individuals. In this the Naive Bayes formula is used for the classification of information and result they got 67% accuracy. They need to collect several resolutions from the individuals that are useful for each government and non-government organization to handle such scenario in an exceedingly higher manner.

III. METHODOLOGY

In today's world, social media networking sites serve as a new knowledge gateway for people of all age groups. These platforms have become venues for expressing sentiments in the form of opinions, judgments, feelings, expressions, and reviews on a wide range of topics such as movies, brands, products, the clothing industry, social activities, and more. The reviews or expressions can be positive, negative, or neutral.

The automated process of analyzing these opinions or textual data is known as sentiment analysis. Sentiment analysis can be described as a standardized analysis of online expressions.

There is a growing array of methodologies and techniques for detecting depression levels from social media posts. Our study provides a technical description of techniques used for depression identification utilizing Natural Language Processing (NLP) and text classification methods. The framework includes steps for data preprocessing, feature extraction, application of machine learning classifiers, feature analysis of the data, and experimental results.

A. Dataset

The dataset consists of tweets collected using the Twitter dataset. We have gathered a total of 31962 datasets to develop both the training and testing modules for our model.

The training is done on 30000 datasets. The training dataset will include a curated list of words associated with mental illness and depression, such as 'depressed', 'sad', 'suicide', 'gloomy', 'unhappy', 'low', 'down' and others, and contains positive words such as happy, thank you, love etc. The testing dataset will comprise randomly collected tweets that contain a mix of positive, neutral and negative components.

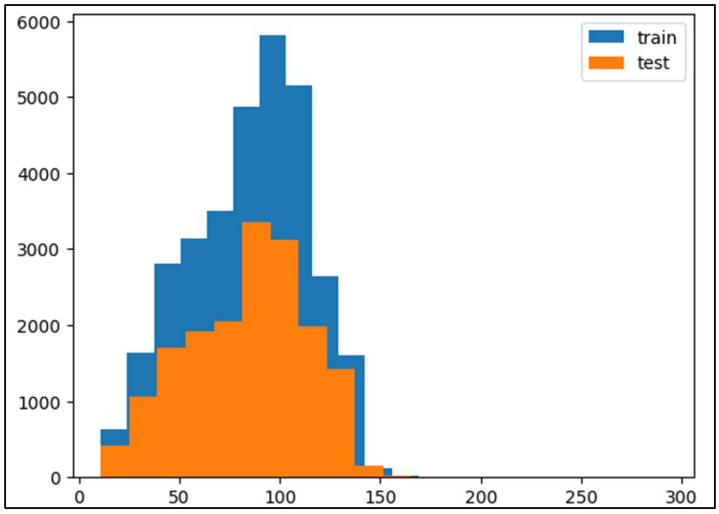


Fig. 1 Total number of trained datasets and number of datasets tested

B. Implementation

Initially, we gather all tweets using the datasets. Subsequently, text pre-processing techniques are applied to these documents. Firstly, a corpus is constructed, and the tweets within each document are tokenized. Following tokenization, normalization procedures are implemented, including converting all characters to lowercase, removing punctuation, retweets, mentions, links, unrecognized emojis, and symbols. While it's common to eliminate stop words like "I," "me," and "you" during normalization, we retain first-person pronouns in this process. Additionally, stemming is performed. The TF-IDF method is then employed to assess the weight of words. Features derived from the DTM are integrated with user activity data extracted from the social network. The resulting merged dataset, considered as independent variables, is inputted into a classification algorithm to predict the outcome of interest, serving as the dependent variable. After careful consideration, we opt for algorithms like Bernoulli Naive Bayes, SVM (Support Vector Machine), Logistic Regression models to find the accuracy, F1 score of the given datasets.

C. Data Pre-Processing

From the [8] we have found that how to perform the data pre-processing. Data pre-processing is an essential step in constructing a Machine Learning model, especially when dealing with unstructured information like text. Natural Language Processing (NLP) tools are utilized to preprocess the dataset before moving on to feature selection and training. The initial pre-processing step involves Tokenization, which breaks down Twitter posts into individual tokens. Subsequently, URLs, punctuations, and stop words are removed to prevent unpredictable outcomes. However, emojis or emoticons are retained during Sentiment Analysis as they often convey crucial sentiment-related information. Additionally, stemming is applied to reduce words to their root form and group similar words together.

D. Feature Extraction

Machine Learning algorithms operate by learning from a predetermined set of options within the training data to generate output for the test data. However, the primary challenge in working with language processing is that machine learning algorithms cannot directly process raw text. Therefore, feature extraction techniques are necessary to convert text into a matrix or vector of features. Some popular methods of feature extraction include:

1) Bag-of-Words

The bag-of-words model is a simplified representation used in natural language processing (NLP) and information retrieval (IR). In this model, a text (like a sentence or document) is treated as a collection of its words, disregarding grammar and word order while preserving multiplicity. In the context of sentiment analysis for depression detection, positive sentiments are indicated by words like 'joyful,' 'happy,' and 'amazing,' while negative sentiments are conveyed by words like 'sad,' 'depressed,' 'dejected,' 'sorrowful,' and 'miserable.'

Creating a Bag of Words model involves three key steps:

a) The first step in text preprocessing involves converting the entire text to lowercase characters and removing all punctuation and unnecessary symbols.

b) The second step includes creating a vocabulary of all unique words from the text collection.

c) The third step involves creating a feature matrix by assigning a separate column for each word, with each row corresponding to the sentiment. This process, known as Text Vectorization, indicates the presence or absence of a word in the sentiment.

2) TF-IDF (Term Frequency Inverse Document Frequency)

TF-IDF, short for Term Frequency-Inverse Document Frequency, is a method used to determine the significance of sentences composed of words while mitigating the limitations of the Bag of Words technique. This approach is beneficial for text classification and aids in helping machines understand words in context rather than just counting them.

3) Parts of Speech (POS) Tagging

Part of speech tagger is a piece of software that reads text in some language and assigns parts of speech to each word such as nouns, verbs and adjectives etc.

IV. ALGORITHMS

A. Naïve Bayes

Sentiment analysis is a specialized field focused on extracting subjective emotions and sentiments from text data. A common application of sentiment analysis is determining whether a text conveys negative or positive feelings. Written reviews provide excellent datasets for sentiment analysis due to their accompanying scores that can be used for classifier training. Naive Bayes classifiers are widely utilized for text classification and machine learning tasks in text analysis. While it is relatively straightforward, a Naive Bayes classifier often performs comparably to more complex solutions. It requires less training time and data. When presented with a feature matrix X and a target vector y, we can formulate our problem for a Naive Bayes classifier as follows:

$$P(y \mid X) = \frac{P(X \mid y)P(y)}{P(X)}$$

Where, y is category variable and X could be a dependent feature vector with dimension d i.e., X = (x1, x2, x2, xd), wherever d is that the range of variables/features of the sample.

B. Support Vector Machine

The Support Vector Machines (SVM) algorithm determines the most effective decision boundary between vectors that belong to a specific cluster or category and those that do not. It can be applied to any type of vectors representing various types of data. In SVM, we represent each data point as a point in an n-dimensional space, where n represents the number of features, with each feature's value corresponding to a specific coordinate. Support vectors are essentially the coordinates of individual observations. The SVM Classifier is a boundary that effectively separates the two categories, often represented as a hyperplane or line.

C. Logistic Regression

Logistic regression is a statistical method used for binary classification tasks, where the outcome variable has only two possible classes, such as "yes" or "no," "positive" or "negative," etc. It's a type of regression analysis that predicts the probability of occurrence of an event by fitting data to a logistic curve. In logistic regression, the output or dependent variable is binary, and the input or independent variables can be continuous or categorical. The model calculates the probability that an observation belongs to a particular class using the logistic function, also known as the sigmoid function. This function ensures that predicted probabilities fall between 0 and 1.

During training, logistic regression optimizes its parameters (coefficients) using techniques like maximum likelihood estimation or gradient descent to minimize the error between predicted probabilities and actual class labels. Once trained, the model can predict the probability of an input belonging to the positive class, and a threshold is typically applied to convert these probabilities into class labels.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |

V. RESULT

For all the three models we have created the confusion matrix and ROC-AUC curve whose result are as follows

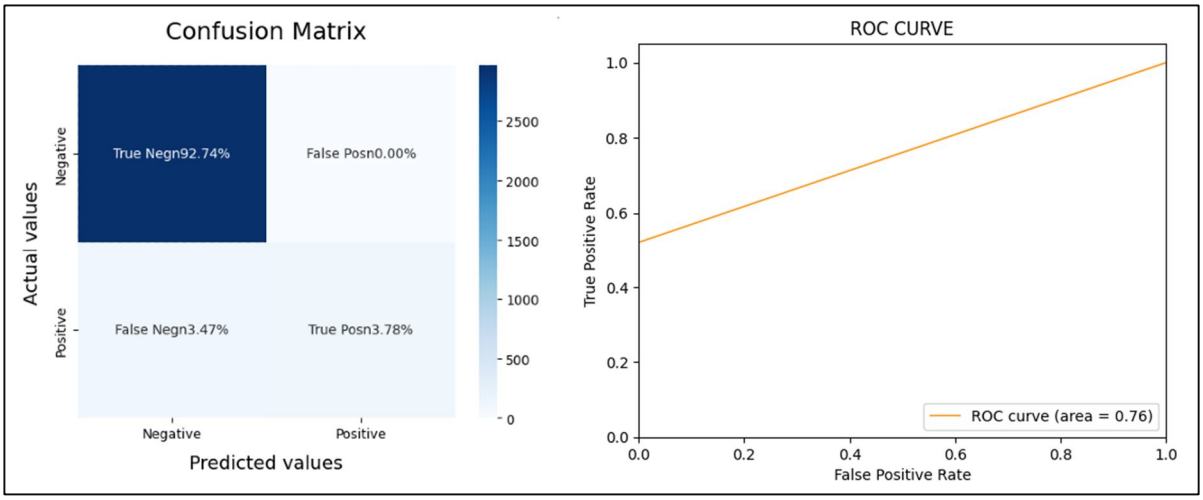


Fig. 2 Result of Naive Bayes model

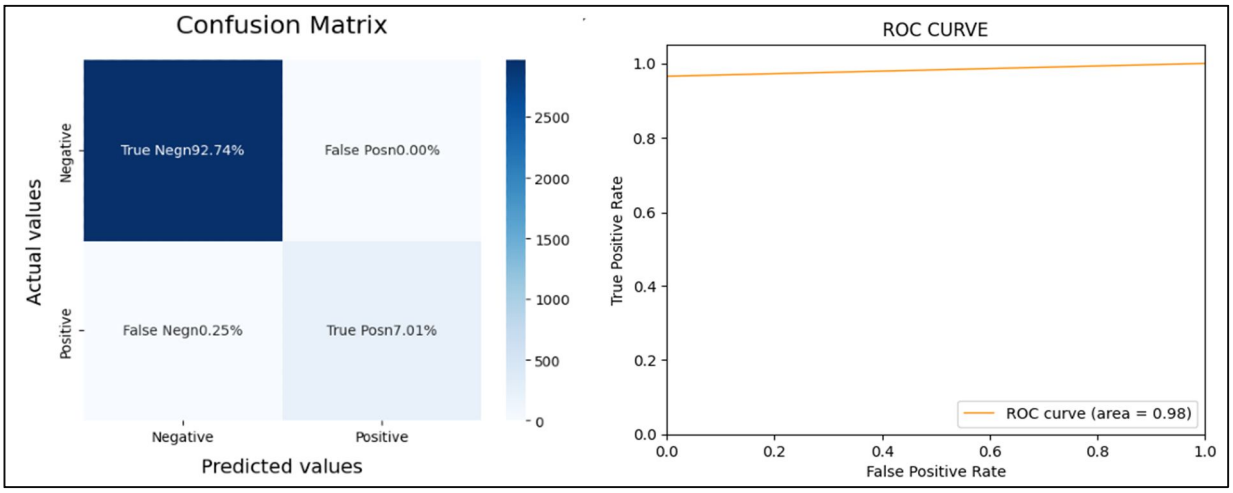


Fig. 3 Result of SVM (Support Vector Machine) model

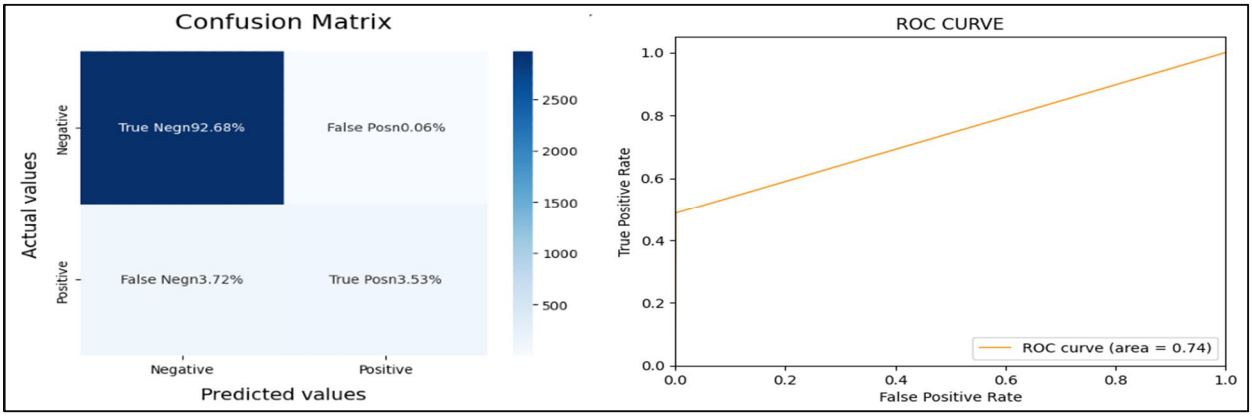


Fig. 4 Result of Logistic Regression model

Based on our evaluation of all three models we have concluded that

Accuracy: As far as the accuracy of the model is concerned Logistic Regression performs better than SVM which in turn performs better than Bernoulli Naive Bayes.

F1-score: The F1 Scores for class 0 and class 1 are:
(a) For class 0: Logistic Regression (accuracy = 0.98) < Bernoulli Naive Bayes (accuracy = 0.98) < SVM (accuracy =1.00)
(b) For class 1: Logistic Regression (accuracy = 0.65) < Bernoulli Naive Bayes (accuracy = 0.69) < SVM (accuracy = 0.98)
AUC Score: All three models have the same ROC-AUC score.

VI. FUTURE SCOPE

In future research, enhancing the model's performance could involve integrating more aspects of online user behavior, such as the timing of tweets or engagement with other users.

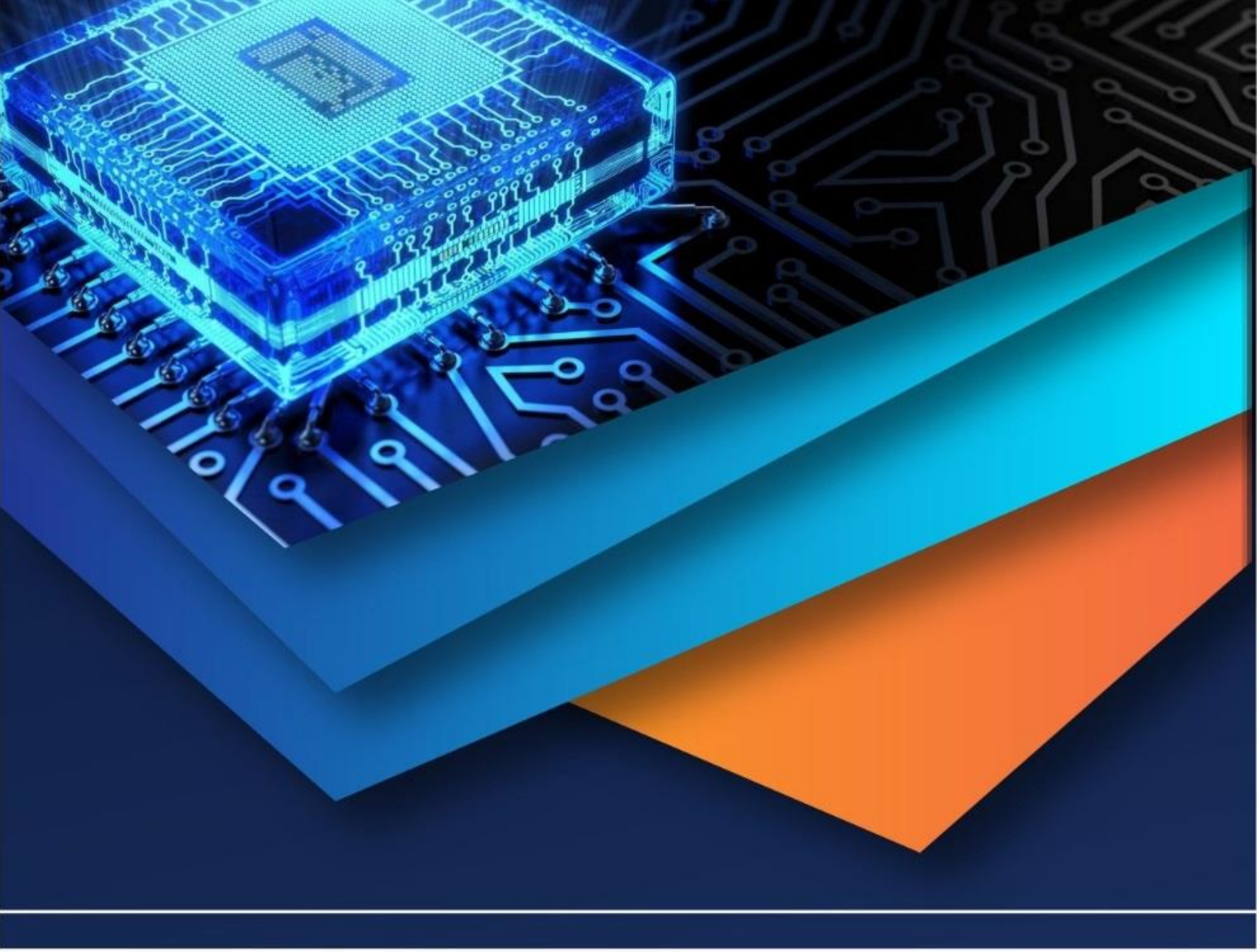
VII. CONCLUSION

In conclusion, we provide evidence regarding whether users express their depressive feelings or acknowledge their depression on popular platforms like Twitter. We developed a predictive model to identify if a user's tweet indicates depression, utilizing a supervised learning approach for sentiment analysis. We evaluated the performance of classifiers using a dataset from Twitter, based on a manually constructed corpus with labeled data (positive, negative). Our observations indicate that individuals suffering from depression tend to be more socially isolated, as shown by their interactions with trending hashtags they use in their tweets. The accuracy was around 90%.

REFERENCES

[1] Burdisso, S. G., Errecalde, M., & Montes-y-Gómez, M. (2019). A text classification framework for simple and effective early depression detection over social media streams. Expert Systems with Applications, 133, 182-197.
[2] Hayda Almeida, Antoine Briand, and Marie-Jean Meurs. 2017. Detecting Early Risk of Depression from Social Media User-generated Content 8th International Conference of the CLEF Association, CLEF 2017, Dublin, Ireland, September 11-14, 2017.
[3] Chatterjee, R., Gupta, R. K., & Gupta, B. (2021). Depression detection from social media posts using multinomial naive theorem. In IOP Conference Series: Materials Science and Engineering (Vol. 1022, No. 1, p. 012095). IOP Publishing.
[4] Munmun De Choudhury, Michael Gamon, Scott Counts, and Eric Horvitz. 2013. Predicting Depression via social media. In ICWSM. 2.
[5] S. Tsugawa, Y. Kikuchi, F. Kishino, K. Nakajima, Y. Itoh and H. Ohsaki, "Recognizing depression from twitter activity", Proc. 33rd Annu. ACM Conf. Hum. Factors Comput. Syst., pp. 3187-3196, Apr. 2015.
[6] Kaur, H. J. (2015). Sentiment Analysis from Social Media in Crisis Situations. IEEE, (pp. 251-256).
[7] Khafaga, D. S., Auvdaiappan, M., Deepa, K., Abouhawwash, M., & Karim, F. K. (2023). Deep learning for depression detection using Twitter data. Intelligent Automation & Soft Computing, 36(2), 1301-1313.
[8] Pachouly, S. J., Raut, G., Bute, K., Tambe, R., & Bhavsar, S. (2021). Depression detection on social media network (Twitter) using sentiment analysis. Int. Res. J. Eng. Technol, 8, 1834-1839.
[9] Hassan, A. U., Hussain, J., Hussain, M., Sadiq, M., & Lee, S. (2017, October). Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression. In 2017 international conference on information and communication technology convergence (ICTC) (pp. 138-140). IEEE.
[10] Arora, P., & Arora, P. (2019, March). Mining twitter data for depression detection. In 2019 international conference on signal processing and communication (ICSC) (pp. 186-189). IEEE.
[11] Ng, A., & Jordan, M. (2001). On discriminative vs. generative classifiers: A comparison of logistic regression and naive bayes. Advances in neural information processing systems, 14.
[12] Amanat, A., Rizwan, M., Javed, A. R., Abdelhaq, M., Alsaqour, R., Pandya, S., & Uddin, M. (2022). Deep learning for depression detection from textual data. Electronics, 11(5), 676.
[13] Shahare, F. F. (2017, June). Sentiment analysis for the news data based on the social media. In 2017 International Conference on Intelligent Computing and Control Systems (ICICCS) (pp. 1365-1370). IEEE.
[14] Sun, C., Qiu, X., Xu, Y., & Huang, X. (2019). How to fine-tune bert for text classification?. In Chinese computational linguistics: 18th China national conference, CCL 2019, Kunming, China, October 18–20, 2019, proceedings 18 (pp. 194-206). Springer International Publishing
[15] Feuston, J. L., & Piper, A. M. (2018). Beyond the coded gaze: Analyzing expression of mental health and illness on instagram. Proceedings of the ACM on Human-Computer Interaction, 2(CSCW), 1-21.
[16] Vij, A., & Pruthi, J. (2018). An automated psychometric analyzer based on sentiment analysis and emotion recognition for healthcare. Procedia computer science, 132, 1184-1191.
[17] Andrade, L., Caraveo-Anduaga, J. J., Berglund, P., Bijl, R. V., Graaf, R. D., Vollebergh, W., ... & Wittchen, H. U. (2003). The epidemiology of major depressive episodes: results from the International Consortium of Psychiatric Epidemiology (ICPE) Surveys. International journal of methods in psychiatric research, 12(1), 3-21
[18] Oxman, T. E., Rosenberg, S. D., & Tucker, G. J. (1982). The language of paranoia. The American journal of psychiatry, 139(3), 275-282.
[19] Robinson, M. S., & Alloy, L. B. (2003). Negative cognitive styles and stress-reactive rumination interact to predict depression: A prospective study. Cognitive Therapy and research, 27, 275-291.
[20] Moreno, M. A., Jelenchick, L. A., Egan, K. G., Cox, E., Young, H., Gannon, K. E., & Becker, T. (2011). Feeling bad on Facebook: Depression disclosures by college students on a social networking site. Depression and anxiety, 28(6), 447-455.

©IJRASET: All Rights are Reserved | SJ Impact Factor 7.538 | ISRA Journal Impact Factor 7.894 |



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)