DEPRESSION DETECTION BY ANALYZING SOCIAL MEDIA POST OF USER: A REVIEW

Prof. Dhanashree Patil, Yogesh Palve, Ketaki Todkar, Akash Shinde, Anisha Biradar, Tejaswini Patil

Abstract— **In contemporary times, identifying depression at an early stage has become a critical issue within the field of psychology. Mental health problems, notably depression, afflict a significant portion of the global population, with an estimated 300 million individuals currently affected. Leveraging the vast amount of user-generated content on social media platforms, researchers are exploring ways to utilize this data to identify potential mental health issues in users. Depression remains a pressing concern in society and a topic of ongoing research globally. Despite numerous studies attempting to understand various mood disorders such as depression, anxiety, and stress using activity logs from devices like smartphones, accurately predicting depressive episodes remains challenging. Social media analysis has emerged as a popular approach to tackle this challenge. In this paper, we propose a system for assessing depression and detecting suicidal ideation by predicting suicidal tendencies based on the severity of depression. Our study aims to develop effective machine learning algorithms for identifying potential signs of depression in individuals through their social media posts. To achieve this, we trained and tested classifiers to discern whether a person exhibits signs of depression using features extracted from their social media activities. Various machine learning algorithms were employed to train and classify users into different levels of depression on a scale ranging from 0 to 100%. Data in the form of posts were collected and categorized into depressed or non-depressed using machine learning algorithms. Our primary contribution lies in exploring a range of features and their effectiveness in detecting the degree of depression. We aim to build a deep learning model capable of categorizing users with depression by learning from individual-level labels to infer post-level labels. By amalgamating the probabilities of post label categories, we can create temporal posting profiles to classify users with depression. Our findings reveal distinct variations in posting behaviors between depressed and non-depressed users, as indicated by the combined probabilities of post label categories. In our study, we utilized natural language processing (NLP) techniques, employing the BERT algorithm, to efficiently identify potential depression indicators in social media content. This approach offers a more accessible and effective means of detecting depression.**

Keywords— Machine Learning, NLP, BERT Algorithm, Depression, Classification, Social Media Post.

I. INTRODUCTION (*HEADING 1*)

In today's psychological landscape, the early identification of depression has become an imperative concern. Depression, a pervasive mental health issue, is exacerbated by the pressures and challenges of modern life, potentially impacting individuals worldwide. With over 350 million people globally grappling with depression, it affects approximately 5% of the global population. Alarmingly, suicide claims nearly 800,000 lives annually, making it the second leading cause of death among individuals aged 15–29, often closely associated with depression. Recent studies underscore depression as a primary cause of disability and severe somatic illnesses. The proliferation of internet and communication technologies, particularly through online social networks like Facebook, Twitter, and Instagram, has revolutionized human interaction and expression. These platforms serve as not only repositories for written and multimedia content but also as arenas for users to articulate their emotions, sentiments, and feelings on various subjects or issues. While this facilitates open discourse and engagement, it also presents an opportunity for health professionals to gauge the mental states of individuals based on their online interactions. To gain insights into these mental states, machine learning techniques offer specific features for analyzing patterns within online communication, thereby revealing emotional states such as happiness, sadness, anger, anxiety, and depression among social media users. Additionally, a growing body of literature highlights the role of social networks in influencing various aspects of social relationships, mental health, substance use, and suicidal ideation, particularly among vulnerable populations such as young adults, racial/ethnic minorities, essential workers, and unpaid caregivers. Youth, defined as individuals aged fifteen to twenty-four, undergo significant physical, psychological, and social changes during this developmental stage. While this period is characterized by a quest for happiness, love, action, and independence, it also presents challenges that can lead to either normalcy or mental health disorders. Depression, in particular, can profoundly impact an individual's functioning across various domains, including work, school, and social activities. The ubiquitous nature of social media platforms like Facebook, Twitter, and Instagram has led to a culture where individuals express their emotions, opinions, and daily experiences through posts, photos, and videos. This study aims to analyse social media posts to identify indicators of depression among relevant users. Various machine learning techniques and algorithms will be employed for this purpose, with the ultimate goal of developing effective tools for detecting depression based on social media content. By harnessing the power of machine learning, researchers can delve into the intricacies of online

communication patterns to uncover subtle cues indicative of mental health states. This approach holds promise for early intervention and support for individuals at risk of depression, potentially reducing the burden of the condition on both individuals and society as a whole. Furthermore, understanding the role of social networks in shaping mental health outcomes can inform targeted interventions and policies aimed at promoting well-being and preventing adverse mental health outcomes among vulnerable populations. However, it's important to consider the ethical implications of mining personal data from social media platforms for mental health assessment. Privacy concerns, data security, and the potential for algorithmic bias must be carefully addressed to ensure that such endeavors uphold ethical standards and protect the rights and dignity of individuals. By navigating these challenges thoughtfully, researchers can harness the potential of machine learning and social media data to advance our understanding of depression and improve mental health outcomes for individuals worldwide.

II. NLP(NATURAL LANGUAGE PROCESSING)

The artwork discussed in this paper falls within the domain of Natural Language Processing (NLP), particularly focusing on text classification. The roots of text classification tasks can be traced back to early research endeavours aimed at automatically categorizing documents based on a statistical analysis of key terms, dating as far back as 1961. Subsequent research efforts led to the development of rule-based text classification systems like CONSTRUE in 1990, marking a significant milestone. However, the field gradually transitioned towards the utilization of machine learning algorithms around the year 2000. In addition to text categorization, machine learning has been increasingly applied to various text-based tasks, including sentiment analysis. Sentiment analysis involves extracting opinions and sentiments from text documents and was initially employed alongside machine learning to discern positive or negative reviews in movie critiques. Over time, it expanded into diverse domains such as social media monitoring and general assessments of consumer attitudes. More recently, deep learning techniques have been integrated into text classification, extending its applications beyond traditional areas such as image classification. State-of-the-art results in numerous text-based tasks are now achieved through advanced methods like transfer learning, exemplified by techniques like Universal Language Model Fine-tuning (ULM Fit) and Google's Bidirectional Encoder Representations from Transformers (BERT). These approaches facilitate the training of language representations, enabling enhanced performance across various text classification tasks. The availability of pre-trained models, including BERT and others, has further accelerated progress in this field, providing researchers with powerful tools to tackle complex text analysis challenges.

III. LITERATURE REVIEW

Opportunities for analysing client behaviour in social networks are always evolving. Specifically, computational linguistics techniques are effectively used to the analysis of social media messages.

1) Within the paper, a records-analytic-based approach is suggested to identify an individual's depression. The data is collected from the posts made by clients on Facebook and Twitter, two well-known social media platforms. Device analysis is employed in this study to handle the scraped data collected from users of social networking sites. Natural Language Processing (NLP) refers to the application of Support Vector Machine (SVM) and Naïve Bayes algorithms to identify sadness potentially in a more convenient and environmentally friendly way. [1]

2) The research employs Natural Language Processing (NLP) strategies to increase a melancholy detection set of rules for the Thai language on Facebook in which human beings use it as a device for sharing opinions, feelings, and existence events. [2]

3) The fitness tweets are analysed for Depression, Anxiety from the blended tweets via way of means of the usage of Multinomial Naive Bayes and Support Vector Regression (SVR) Algorithm as a classifier in paper. [3]

4) In the study, researchers provide a method for determining an individual's level of melancholy by analysing and deriving emotions from the text, applying emotion theories, device learning techniques, and natural language processing techniques on distinct social media platforms. [4]

5) The study aims to use natural language processing to Twitter streams in order to conduct a melancholic emotion assessment. Tweets are classified as neutral or negative based on a carefully chosen phrase list that evokes feelings of melancholy. A Naive-Bayes classifier and guide vector device had been employed in the sophistication prediction method. The use of the top category metrics, such as the confusion matrix, accuracy, and F1-score, had been made available to the consequences. [5]

6) The study suggests a method for measuring sadness and identifying suicidal thoughts in order to forecast suicide behavior entirely based on the degree of melancholy. Data collected in real time took the form of surveys and tweets. Next, based on severity, categorization device algorithms are employed to teach and categorize it into five levels of sadness. [6]

7) Yates et al. demonstrated the risks of self-harm and depression using a neural community model based solely on postings from Reddit and Twitter, and they verified the high diagnostic accuracy of this method. The authors think that the methodologies they have suggested might be applied not just to scientific therapy but also to large-scale research on intellectual fitness. [8]

8) O'Dea et al. conducted a study to explore the hypothesis that Twitter may be used to progressively identify a person's mental health state, including suicidality and depression. According to their research, it is feasible to use both a human coder and a preprogrammed device classifier to determine the level of fear present in tweets related to suicide.[10]

There is a severe and developing variety of methodologies and strategies for detection of the melancholy degree from the posts on Social Media networks. In our study, we consolidate a technical description of strategies implemented for melancholy identity the usage of the Natural Language Processing approach labeled the usage of the BERT set of rules to hit upon melancholy. The framework is created from Data pre-processing steps.

IV. PROBLEM STATEMENT

Depression has been verified to have an impact at the language of individuals. To expand a software to investigate and hit upon depression of social media posts of clients thru tool learning techniques. This assignment pursuits to use natural language processing, tool learning techniques, and neural network architectures to build, tune, and have a look at models that classify social media Post of Users as "depressed" or "non-depressed.

V. OBJECTIVE

The targets are as follows:
1. System will constantly hold on tracking the posts and chats of users. And if it detects the bad notion sort of conduct then the device will mechanically put up the high-quality put up on his/her wall primarily based totally on the extent of depression.
2. Assist the individual in overcoming depression.

VI. METHODOLOGY

Machine Learning Classification Techniques used for the mode

1. BERT Algorithm:

"BERT stands for Bidirectional Encoder Representations from Transformers. It is designed to pre-educate deep bidirectional representations from the unlabeled textual content with the aid of using collectively conditioning on each left and proper context. As a result, the pre-skilled BERT version may be fine-tuned with simply one extra output layer to create today's fashions for a huge variety of NLP tasks." We advanced a Bidirectional Encoder Representations from Transformers (BERT)-primarily based totally version, that's a brand-new language illustration version as defined in. As the call suggests, it changed into designed to pre-educate deep bidirectional representations that may be fine-tuned with an extra output layer. For this project, this outputs layer - a pooled output - changed into used for the binary type of the comments. From the various pre-skilled fashions available, we selected the English-language uncased (all lowercase earlier than tokenization) version of BERT, as case records isn't always specially essential to the venture of social media remark type
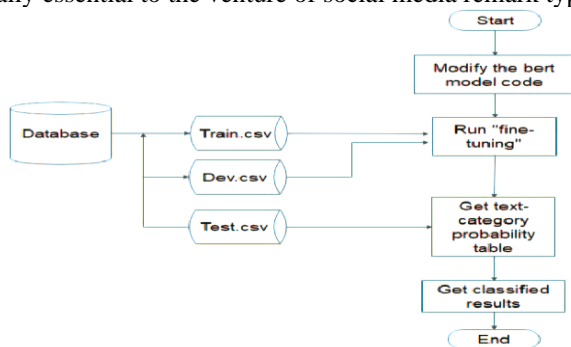

Fig.BERT Algorithm

VII. MATHEMATICAL MODEL

Relevant Mathematics Associated with The Project:

System Description:
S= I, O, F, DD, NDD, Failure, Success

Where,
S=System
I= Input
O=Output
F=Failure
S=Success

I is Input of system
Input I = set of Inputs

Where,
I= {Users Social media posts}
F is Function of system
F = set of Function

Where,
F1= {Input Dataset}
F2= {Json to CSV Conversion}
F3={Pre-processing}
F4={Cleaning}
F5= {Train test split}
F6= {Sentiment Dictionary}
F7= {Classifier (BERT Algorithm)}
F8={Tokenization}
             is Output of system
Output O1= {Depression detection}
Success Conditions: Product working Smoothly. depression detection successfully.

Failure Conditions: if internet connection Unavailable.
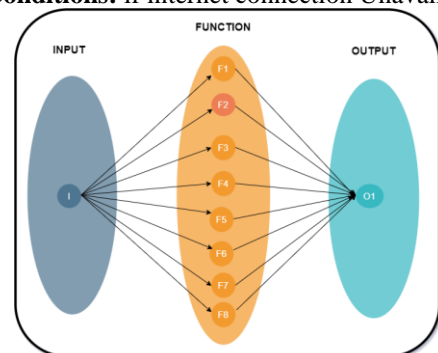

Fig.Venn Diagram

Where,
    I = {Users Social media posts}
    F1= {Input Dataset}
    F2= {Json to CSV Conversion}
    F3={Pre-processing}
    F4={Cleaning}
    F5= {Train test split}
    F6= {Sentiment Dictionary}

F7= {Classifier (BERT Algorithm)}
F8={Tokenization}
Output O1 = {Depression detection}

VIII. SYSTEM ARCHITECTURE

Depression is a severe task in private and public fitness. One of the predominant answers to this hassle is an in-depth look at an individual's conduct attributes. These attributes are to be had on diverse social networking websites together with Facebook, Twitter, Instagram, etc. Social networking platform is the pleasant manner to recognize a person's conduct, wondering style, mood, egoistic networks, evaluations, etc. The use of social networking websites is growing in particular the various younger generation. The humans on social media-specific their feelings, day by day sports, evaluations approximately diverse topics, etc. So social networking platforms serve as screening tools for predicting levels of depression by accessing users' experiences, reviews, social interactions, and personality traits. The advanced technique of analysis of the affected person isn't always so applicable however the use of consumer-generated content material on social media posts enables one to are expecting the intellectual fitness stages and melancholy of a specific individual. Our challenge goal is to extract facts from social media posts and via way of means of having a clean knowledge of a person's behavioral attributes and tried questionnaires, melancholy stages of the consumer are predicted. A quantitative look is carried out to teach and check diverse device mastering classifiers to decide whether or not a social media submit of the consumer is depressed, from posts initiated via way of means of the consumer or his/her sports on social media. The following parent illustrates the melancholy detection of the use of the pastime and content material capabilities type model. First, all tweets for depressed and non-depressed accounts, in addition to facts of consumer accounts and sports together with the number of followers, the quantity of following, time of posts, the number of mentions, and quantity of reposts, are retrieved. Subsequently, all the posts from an account are compiled into a unified file. Text pre-processing is carried out on all documents. First, a corpus is created and posts in every file are tokenized. BERT Classification Algorithm may be used.
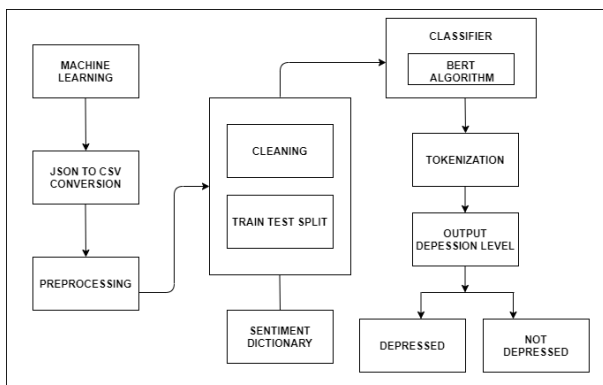

Fig.System Architecture

IX. EXISTING SYSTEM

The already present device affords a far smooth flowing manner to decide the melancholy stage of customers the use of the Naïve Bayes algorithm. The extraction of textual information is finished through the extraction elegance from Facebook with the assist of the Facebook Graph API. After extraction, the information is pre-processed. Pre-processing addresses missing or redundant attributes. Techniques like tokenization, decrease case conversion, and phrase stemming and phrases elimination are used for Pre-processing of information. In the proposed device, consistent with customers' Facebook submit version can discover whether or not he/she is depressed or not. But simplest reading posts won't deliver correct outcomes so we additionally examine the feedback through the consumer and his buddies and his chats also are analysed because the consumer will sincerely proportion his melancholy together along with his friend. On foundation of those analyses, the customers may be labelled as burdened and non-burdened.
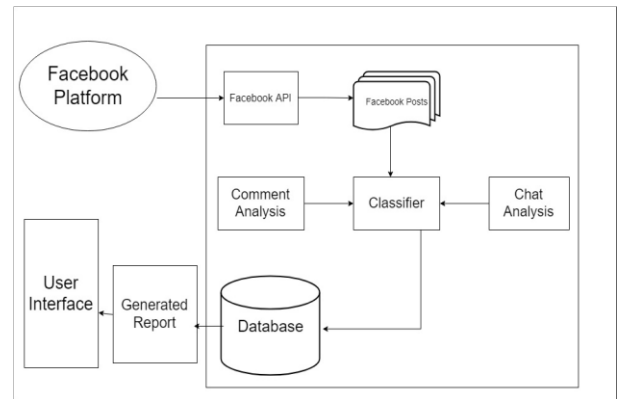

Fig.Exisiting System

X. CONCLUSION

The proposed system may help the suspected purchaser to save his/her life, through manner of approach of know-how in advance whether or not or now no longer the customer is depressed or maybe the system will deliver some motivational posts to the customer based mostly on the volume of his depression. We end the system can be very useful in today's global wherein most humans don't have time to fulfill our friends, percent their thoughts and feelings as we applied in older days due to busy schedules. So, our system plays a important role proper right here to avoid any unwanted human loss. The system will inform their family individuals or spouses and kids approximately the situation of a depressed man or woman. So that very own family or buddy circle will help the man or woman to come out of depression.

REFERENCES

[1] N. A. Asad, M. A. Mahmud Pranto, S. Afreen and M. M. Islam, "Depression Detection by Analyzing Social Media Posts of User," 2019 IEEE International Conference on Signal Processing, Information, Communication & Systems (SPICSCON), Dhaka, Bangladesh, 2019, pp. 13-17, doi: 10.1109/SPICSCON48833.2019.9065101.

[2] K. Katchapakirin, K. Wongpatikaseree, P. Yomaboot and Y. Kaewpitakkun, "Facebook Social Media for Depression Detection in the Thai Community," 2018 15th International Joint Conference on Computer Science and Software Engineering (JCSSE), Nakhonpathom, 2018, pp. 1-6, doi: 10.1109/JCSSE.2018.8457362.

[3] 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, doi: 10.1109/ICSC45622.2019.8938353.

[4] A. U. Hassan, J. Hussain, M. Hussain, M. Sadiq and S. Lee, "Sentiment analysis of social networking sites (SNS) data using machine learning approach for the measurement of depression," 2017 International Conference on Information and Communication Technology Convergence (ICTC), Jeju, 2017, pp. 138-140, doi: 10.1109/ICTC.2017.8190959.

[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, doi: 10.1109/ISS1.2017.8389299.

[6] S. Jain, S. P. Narayan, R. K. Dewang, U. Bhartiya, N. Meena and V. Kumar, "A Machine Learning based Depression Analysis and Suicidal Ideation Detection System using Questionnaires and Twitter," 2019 IEEE Students Conference on Engineering and Systems (SCES), Allahabad, India, 2019, pp. 1-6, doi: 10.1109/SCES46477.2019.8977211.

[7] B. Yalamanchili, N. S. Kota, M. S. Abbaraju, V. S. S. Nadella and S. V. Alluri, "Real-time Acoustic based Depression Detection using Machine Learning Techniques," 2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE), Vellore, India, 2020, pp. 1-6, doi: 10.1109/icETITE47903.2020.394.

[8] Yates, A., Cohan, A., and Goharian, N.: Depression and self-harm risk assessment in online forums. arXiv preprint arXiv:1709.01848 (2017).

[9] Seabrook, E.M., Kern, M.L., Fulcher, B.D., and Rickard, N.S.: Predicting depression from language-based emotion dynamics: longitudinal analysis of Facebook and Twitter status updates. Journal of Medical Internet Research 20 (5), e168 (2018).

[10] O'Dea B, et al. Detecting suicidality on Twitter. Internet Interv. 2015;2(2):183–188. doi: 10.1016/j.invent.2015.03.005.

[11] S. Grover, A. Dutt, and A. Avasthi, "An overview of indian research in depression," Indian journal of psychiatry, vol. 52, no. Suppl1, p. S178, 2010.

[12] M. Reddy, "Depression: the disorder and the burden," Indian journal of psychological medicine, vol. 32, no. 1, p. 1, 2010.

[13] "Depression." [Online]. Available: https://www.who.int/news-room/fact-sheets/detail/depression

[14] M. M. Aldarwish and H. F. Ahmad, "Predicting depression levels using social media posts," in 2017 IEEE 13th international Symposium on Autonomous decentralized system (ISADS). IEEE, 2017, pp. 277–280.

[15] P. Tzirakis, G. Trigeorgis, M. A. Nicolaou, B. W. Schuller, and S. Zafeiriou, "End-to-end multimodal emotion recognition using deep neural networks," IEEE Journal of Selected Topics in Signal Processing, vol. 11, no. 8, pp. 1301–1309, 2017.

[16] "Depression and suicide: Overview, etiology of depression and suicidality, epidemiology of depression and suicide," Feb 2019. [Online]. Available: https://emedicine.medscape.com/article/805459-overview