Mental Health Evaluation for The Social Media Platform Using Machine Learning

Ayushi Yadav
AMITY INSTITUTE OF INFORMATION TECHNOLOGY

Abstract— **This system focuses on developing a comprehensive system for mental health evaluation within social media platforms utilizing Convolutional Neural Networks (CNN). The system provides an in-depth understanding of mental health, leveraging the vast amount of data generated on social media platforms for prediction and analysis through machine learning and deep learning algorithms. By monitoring the social media activities of individuals, the system aims to predict various mental health factors such as depression, anxiety, and stress. Utilizing the online social media data, the system explores the correlation between users' mental health and the content they post on their social media handles. Through this approach, mentors including teachers and doctors can access weekly analyses of individuals' stress levels, facilitating tailored consultation and support. The CNN model's advanced capabilities offer in analyzing complex textual and visual data, enabling more accurate and nuanced evaluations of users' mental health states. Ethical considerations, including user privacy and data protection, are prioritized throughout the development and deployment of the system. By integrating this system into social media platforms, it holds significant potential to provide timely support and resources for individuals struggling with mental health issues, contributing to the well-being of users in the digital age.**

Index Terms— **Mental health evaluation, K-Means clustering, Python, Behavioural patterns, Digital mental health, Social media analytics**

I. INTRODUCTION

Social media has become an integral part of contemporary society, enabling individuals to share their lives and express themselves with unprecedented ease. The ubiquitous nature of social media platforms facilitates the observation of users' emotional and mental states through their posts and interactions. Despite the accessibility to vast amounts of usergenerated content, accurately assessing an individual's mental well-being remains a challenge. However, leveraging realworld social media data presents an opportunity to discern patterns indicative of mental health issues. The importance of detecting and addressing mental health concerns on social media platforms cannot be overstated. With a significant portion of users expressing their thoughts, emotions, and experiences online, there lies a potential to identify symptoms of psychological distress such as depression, anxiety, and stress. This capability holds promise for providing timely interventions and medical assistance to individuals in need,

potentially mitigating the adverse effects of declining mental health.

Machine learning and deep learning techniques, particularly convolutional neural networks (CNNs), offer powerful tools for analyzing social media data and detecting subtle indicators of mental health issues. By mining textual and emotional cues from users' posts and interactions, these algorithms can identify patterns associated with various mental health conditions. Furthermore, the integration of artificial intelligence (AI) methods, including deep learning, into mental health evaluation signifies a paradigm shift in healthcare management. These technologies enable healthcare professionals to make informed decisions based on comprehensive analyses of individuals' historical data, including social media usage patterns. This research aims to explore the application of CNNs in evaluating mental health on social media platforms. By harnessing the capabilities of deep learning and data analysis, this study seeks to contribute to the advancement of mental health awareness and intervention strategies, ultimately fostering well-being in the digital age.

II. LITERATURE SURVEY

In today's digital era, social media has become an integral part of daily life for people worldwide. It serves not only as a means to connect with others but also as a platform for sharing personal experiences and insights. Users engage on social media in multifaceted ways, extending beyond mere conversations. While it undeniably facilitates connections between individuals separated by geographical barriers and fosters relationships based on shared interests and opinions, excessive usage can pose significant risks to mental wellbeing. The COVID-19 pandemic has impacted over 200 nations, inducing worldwide apprehension and consequently leading to a decline in individuals' mental well-being and heightened stress levels [5]. Furthermore, in the era of social media dominance, numerous individuals find themselves inundated with rumours and misinformation pertaining to the virus, much of which is unverified and untrue [6]. This news also triggers unwarranted panic, fear, and anxiety among the general population of India.

Twitter stands as a ubiquitous social media platform, freely accessible and extensively utilized globally. The default setting renders tweets public, unless users opt for privacy settings. Moreover, users can employ hashtags to categorize their tweets. Consequently, owing to its inherently public nature,

Twitter serves as a valuable resource for gauging public opinions on diverse subjects and topics. Additionally, it offers insights into users' emotional states and mental well- being, thus presenting an invaluable avenue for mental health evaluation, treatment, and prevention initiatives [7].

Research findings from a study [8] outline eight distinct gratifications that drive individuals to use social media. These motives encompass self-documentation, information sharing, social interaction, entertainment, passing time, self-expression, medium appeal, and convenience. Among these, users primarily motivated by self-documentation, self-expression, and social interaction are more inclined to utilize social media platforms as a means to discuss their personal experiences and mental health. The relationship between users and social media manifests in various forms, including conditions like social media anxiety disorder, social anxiety, and anxious depression social media expression [1]. Detecting these disorders involves employing deep learning models to assess users' mental health based on their social media activity. Notably, a research paper [9] focuses on identifying depression among youth in the Middle East and North Africa (MENA) region by analyzing self-reported messages on social media. This study utilizes natural language processing (NLP) techniques and machine learning (ML) algorithms such as Support Vector Machines (SVM). Data collection from Twitter involves leveraging the Twitter API, which provides additional contextual information like location and profile descriptions. Sentiment analysis of tweets is conducted using Ekman's model [10], which categorizes emotions into six categories: joy, sadness, anger, fear, disgust, and surprise.

In a separate study, an analysis of sentiment was carried out at the sentence level on tweets, employing a three-step process. This involved pre-processing the data, constructing a feature vector with pertinent features, and subsequently categorizing the tweets into positive and negative sentiments utilizing various classifiers [11]. Positive sentiments were assigned a weight of "1," while negative ones were assigned a weight of "-1." Twitter-specific elements like hashtags and emoticons were eliminated, followed by additional feature extraction to simplify the tweet into plain text. The tweets were then represented as a compilation of words using a unigram method. Several classifiers, including Multinomial Naive Bayes, Support Vector Machine (SVM), Random Forest, KNearest Neighbors (KNN), Maximum Entropy, and Ensemble classifiers, were examined and exhibited similar levels of accuracy.

III. PROPOSED SYSTEM ARCHITECTURE

This system is designed to provide mental health evaluation for users of social media platforms, with a primary focus on Twitter, accessible through a user-friendly mobile application. Leveraging Convolutional Neural Networks (CNNs), the system aims to analyze users' tweets for detecting and assessing their mental well-being. Convolutional neural networks (CNNs), often referred to as CoVNet in this context, have been chosen as the primary algorithm for their high

accuracy in mental health analysis, as demonstrated in previous research [12]. While originally developed for image analysis, CNNs have proven effective for processing various data types, including textual data. The CNN architecture comprises convolutional, pooling, and fully connected layers, facilitating feature extraction and classification [12]. The system operates as a single-party application, requiring users to log in using their Twitter handle authenticated via Twitter OAuth 2.0. Through the Twitter Developer system, the application securely accesses public tweets for analysis, utilizing the Twitter API to fetch real-time data. A dataset of over 40,000 tweets is used for training the CNN model, which is preprocessed and trained using categorical cross-entropy classification. To enhance the accuracy of analysis, the system incorporates word embeddings, specifically Word2Vec, trained on a vast corpus of text data from Google News [13]. Once trained, the model categorizes incoming tweets as positive, negative, or neutral, providing users with graphical representations of their mental health status on a weekly and monthly basis. An important feature of the system is its proactive alert mechanism. If the number of negative tweets surpasses a predetermined threshold, the user receives a notification via WhatsApp, integrated using the Twilio API. Additionally, the system offers recommendations for medical assistance based on the user's location, utilizing the Google Maps API to identify nearby healthcare facilities. The mobile application is developed using Flutter, ensuring a seamless and intuitive user experience. Privacy and data security are paramount, with users' data analysis conducted securely within the application. It's important to note that the system's performance relies on the data obtained from users' Twitter accounts, emphasizing the need for continuous monitoring and feedback to improve accuracy and relevance. The proposed system offers a comprehensive solution for mental health evaluation on social media platforms, empowering users with insights into their mental well-being and facilitating timely interventions when necessary.

IV. DESIGN AND METHODOLOGY

The methodology for developing a mental health evaluation system for social media platforms, particularly Twitter, revolves around several key stages, each meticulously crafted to ensure the system's effectiveness and accuracy. To access it, users must possess a Twitter account with a sufficient tweet history for accurate evaluation. Upon logging into the application using their Twitter handle, users undergo authentication via Twitter OAuth 2.0, ensuring legal authorization and permissions for data extraction. The system's architecture comprises distinct layers, including the API layer, application layer, and database layer. The API layer interacts with the Twitter API to retrieve real-time data, which is then processed and stored in a Firestore database. Within the application layer, Flask and Python libraries handle data cleaning, preprocessing, and model integration tasks.

Upon fetching data from the Twitter API, an exploration and cleaning process ensues. Public and secured Twitter data is

stored in CSV format and formatted using Pandas data frames. Data cleaning involves employing regular expressions (RegEx), tokenization, lowercase conversion, and stop word removal using the Natural Language Processing Toolkit (NLTK). The cleaned data undergoes division into training and testing sets, with a 90-10 ratio allocation, respectively. Vocabulary building determines maximum lengths for training and testing, followed by tokenization and sequence padding to ensure uniform input data. Preprocessed data is fed into the CNN model, comprising convolutional, pooling, and fully connected layers. Multiple filter sizes are applied to each text, with categorical cross-entropy classification employed. The model undergoes training with a specified number of epochs to iteratively optimize performance. Backend functionalities are segregated into a Flask-based server and a real-time Firestore database, facilitating data management and tweet classification requests driven by the CNN model. The frontend encompasses a user-friendly Flutter-based mobile application interface. Upon login, users grant data access permissions, enabling the system to fetch and analyze their tweets. Graphical representations display weekly tweet analyses, with WhatsApp alerts issued if negative trends surpass predefined thresholds. Optional medical consultation recommendations are available via Google Maps integration.

The CNN model demonstrates impressive performance, achieving an initial test accuracy of 85.3%, which further improves to 89.21% post additional training and data augmentation. This performance superiority over alternative algorithms underscores the effectiveness of CNNs for mental health evaluation on social media platforms, particularly Twitter. The outlined encompasses comprehensive data processing, model training, and system integration steps, culminating in a robust mental health evaluation system tailored for social media platforms.

V. RESULT AND ANALYSIS

The decision to adopt the Convolutional Neural Network (CNN) model for mental health evaluation on social media platforms was made following a rigorous comparative analysis with various other algorithms, including Multinomial Naive Bayes (MNB), Support Vector Machine (SVM), Random Forest (RF), and K-nearest Neighbors (KNN) models. Extensive planning and meticulous analysis were conducted to select the most suitable model, leading to enhanced clarity in the overall design and implementation process. The CNN model showcased remarkable capabilities in maintaining its embeddings while effectively training and testing data, thereby yielding superior accuracy compared to alternative algorithms. The resultant accuracy achieved post-experimentation stands at an impressive 89.21%, surpassing the standard benchmark for accuracy in mental health evaluation. This noteworthy performance underscores the effectiveness of the CNN model in detecting and analyzing social media data for mental health assessment, offering promising avenues for further research and development in the field. Given the escalating challenges to mental balance and stability in contemporary society, the

analysis of Twitter data for detection and analysis, coupled with subsequent assistance, assumes paramount importance. This system employs a tweet extraction method, which is subsequently utilized to assess the individual's mental state, yielding commendable accuracy results. The utilization of the CNN model for implementation further enhances the system's efficacy, offering users a reliable tool for mental health evaluation on social media platforms. The successful completion of the detection and analysis stages of social media data, coupled with the superior accuracy attained by the CNN model, underscores the significance of leveraging advanced computational techniques for mental health assessment. As research endeavors continue to evolve, the integration of innovative methodologies such as CNNs holds immense promise in advancing our understanding and management of mental health issues in the digital age.
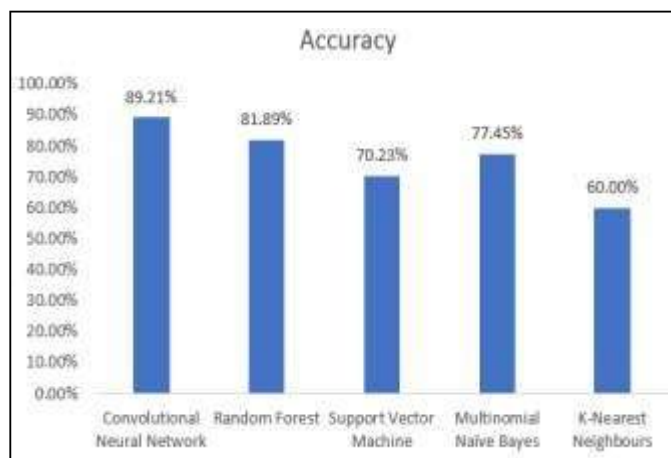
Sno	Algorithms	Accuracy
1	Convolutional Neural Network	89.21%
2	Random Forest	81.89%
3	Support Vector Machine	70.23%
4	Multinomial Naïve Bayes	77.45%
5	K-Nearest Neighbours	60.00%

VI. CONCLUSIONS

In conclusion, the application efficiently analyzes social media content, providing graphical representations, and offering recommendations for nearby medical assistance if negativity exceeds a set threshold. It serves as a supportive tool for individuals grappling with mental health issues like stress and depression, offering insights into their social media behavior and suggesting potential avenues for improvement. Additionally, the application delivers crucial solutions postdetection to aid individuals in enhancing their mental wellbeing. This holistic approach not only provides valuable assistance to users but also fosters greater awareness and understanding of mental health challenges in the digital age.

REFERENCES

[1] Akshi Kumar, Aditi Sharma, Anshika Arora, Anxious Depression Prediction in Real-time Social Data, International Conference on Advanced Engineering, Science, Management and Technology . 2019 (ICAESMT19), March 2019

[2] Ahmed Husseini Orabi, Prasadith Buddhitha, Mahmoud Husseini Orabi, Diana Inkpen, Deep Learning for Depression Detection of Twitter Users, University of Ottawa, ONKIN 6N5 Canada, June 2018

[3] Huijie Lin, Jia Jia, Jiezhon Qiu, Yongfeng Zhang, Lexing Xie, Jie Tang, Ling Feng, and Tat-Seng Chua, Detecting Stress Based on Social Interaction in Social Networks, Journal of Latex Class Files, Vol. 13, No. 9, September 2014

[4] Nhan Cach Dang, María N. Moreno-García, Fernando De la Prieta, Sentiment Analysis Based on Deep Learning: A Comparative Study, MDPI, Biotechnology, Intelligent Systems and Educational Technology (BISITE) Research Group, University of Salamanca, 37007 Salamanca, Spain, 14 March 2020

[5] Golechha M, COVID-19, India, lockdown and psychosocial challenges: What next?, International Journal of Social Psychiatry. 2020;66(8):830-832. doi:10.1177/0020764020935922

[6] Anant Kumar, K. Rajasekharan Nayar, COVID 19 and its mental health consequences, Journal of Mental Health, DOI: 10.1080/09638237.2020.1757052

[7] Chandler McClellan, Mir M Ali, Ryan Mutter, Larry Kroutil, Justin Landwehr, Using social media to monitor mental health discussions − evidence from Twitter, Journal of the American Medical Informatics Association, Volume 24, Issue 3, May 2017, Pages 496–502

[8] Liu, Ivy L.B.; Cheung, Christy M.K.; and Lee, Matthew K.O., Understanding Twitter Usage: What Drive People Continue to Tweet, PACIS 2010 Proceedings. Paper 92. http://aisel.aisnet.org/pacis2010/92

[9] Wajdi Zaghouani, A Large-Scale Social Media Corpus for the Detection of Youth Depression (Project Note), Procedia Computer Science, Volume 142, 2018, Pages 347-351, ISSN 1877-0509, https://doi.org/10.1016/j.procs.2018.10.483.

[10] Paul Ekman, An argument for basic emotions, Cognition and Emotion, 6:3-4, 169-200, DOI: 10.1080/02699939208411068, 1992

[11] M. S. Neethu and R. Rajasree, Sentiment analysis in twitter using machine learning techniques, 2013 Fourth International Conference on Computing, Communications and Networking Technologies (ICCCNT), Tiruchengode, India, 2013, pp. 1-5, doi: 10.1109/ICCCNT.2013.6726818.

[12] Shuichi Hashida, Keiichi Tamura, Tatsuhiro Sakai, Classifying Tweets using Convolutional Neural Networks with Multi-Channel Distributed Representation, IAENG International Journal of Computer Science, 46:1, IJCS_46_1_08, 1 February 2019

[13] Y. Kim, Convolutional neural networks for sentence classification, in Proceedings of the 2014 Conference on Empirical Methods in Natural Language Processing, EMNLP 2014, October 25-29, 2014, Doha, Qatar, A meeting of SIGDAT, a Special Interest Group of the ACL, A. Moschitti, B. Pang, and W. Daelemans, Eds. ACL, 2014, pp. 1746– 1751.

[14] Munir Ahmad, Shabib Aftab, Syed Shah Muhammad, Sarfraz Ahmad, Machine Learning Techniques for Sentiment Analysis: A Review, International Journal of Multidisciplinary Sciences and Engineering, Vol. 8, No. 3, April 2017

[15] Abhilasha Tyagi, Naresh Sharma, Sentiments Analysis of Twitter Data using K- Nearest Neighbour Classifier, International Journal of Engineering Science and Computing, April 2018.