**An integrated approach to mental health management with focus on depression detection**

Shiju George1, Asha Joseph2, Sona Sanal3

1,2,3Department of Information Technology, Amal Jyothi College of Engineering, Kanjirappally, Kottayam, 686518, Kerala, India.

[shijugeorge@amaljyothi.ac.in](mailto:shijugeorge@amaljyothi.ac.in), ashajoseph@amaljyothi.ac.in, sonasanal2023@it.ajce.in

Mental well-being is a crucial factor in today's competitive world. People experience a wide range of mental issues, including anxiety, bipolar disorder, depression, insomnia, and panic attacks, due to various reasons such as work pressure, peer pressure, relationship issues, and difficulties in maintaining work-life balance, among others. In this study, our main focus is on the mental issue of depression, and we analyze different methods for detecting it using advanced machine learning techniques. The algorithms are compared to identify the most appropriate model in terms of accuracy, with the support vector showing the highest accuracy at 99.7%.

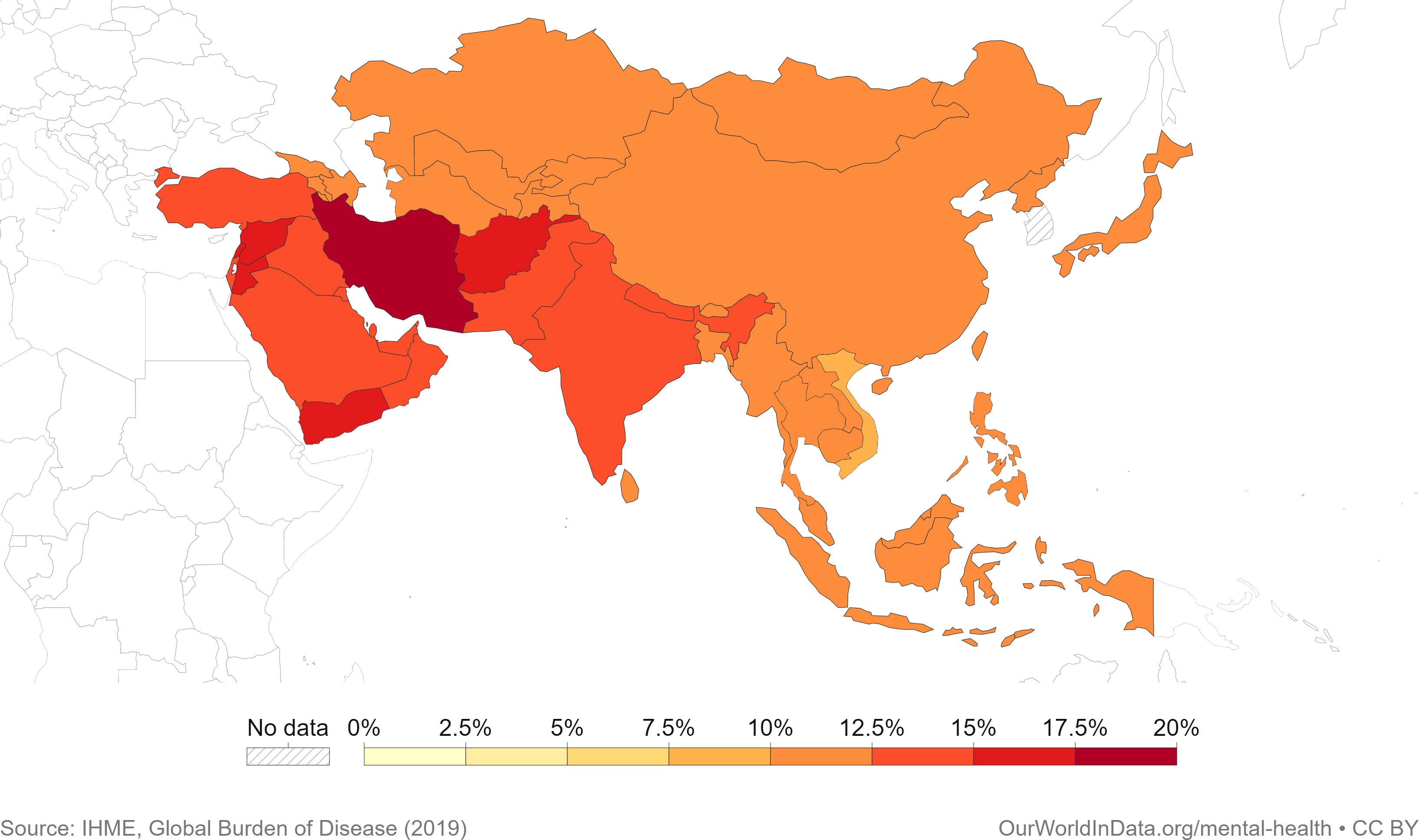
**Keywords:** Depression, Mental health, Anxiety, Machine Learning, Accuracy

# I INTRODUCTION

Major depressive disorder (MDD), also known simply as depression, is one of the most prevalent psychiatric disorders globally. According to the World Health Organization's Comprehensive Mental Health Action Plan 2013-2020, depression affects over 300 million people worldwide and ranks as one of the largest single causes of disability, especially for women. Currently, depression accounts for 4.3% of the global burden of disease and is projected to become the leading cause of disease burden in high-income countries by 2030. Depression, or major depressive disorder, is a common and severe medical illness that negatively impacts emotions, thoughts, and behaviors. Anxiety is the body's response to it. Both depression and anxiety are widespread mental health conditions affecting people worldwide and can significantly affect daily life, relationships, and overall well-being. Unfortunately, many individuals with depression remain undiagnosed, and even those seeking help may experience delays in receiving an accurate diagnosis.

Estimates from 2019 suggest that around 280 million people worldwide experience depression, comprising 5 percent of the global adult population and 5.7 percent of adults aged 60 and above. The National Institute of Mental Health (NIMH) indicates that in the U.S., 19.4 million adults had at least one major depressive episode in 2019, accounting for 7.8 percent of the U.S. adult population. In the past century, the medical community did not acknowledge the existence of depressive disorders in children, believing that children lacked the psychological and cognitive maturity to experience depressive symptoms. However, mounting evidence has shown that children can indeed experience the full spectrum of mood disorders and suffer from significant associated morbidity and mortality. Recent studies have confirmed a prevalence of depression in adolescents ranging from 10% to 60%. It is essential to note that due to underreporting and lack of help-seeking behavior, the actual number of individuals living with depression may be higher:

1. Research from 2021 suggests that nearly 60 percent of individuals with depression do not seek professional support, primarily due to the stigma associated with depression.
2. The World Health Organization (WHO) points out that more than 75 percent of individuals residing in low-income or middle-income countries do not receive treatment for depression because of various treatment barriers.



**Figure 1: Share of population with mental health disorders, 2019**

In recent years, there has been a growing interest in utilizing artificial intelligence (AI) and machine learning (ML) algorithms for the detection of mental health disorders, including anxiety and depression. These technologies hold the promise of enhancing diagnostic precision and speed, thereby facilitating more effective treatment and improved outcomes for patients. The objective of this project is to create a system capable of identifying early signs of depression in individuals. By identifying these indicators at an early stage, our aim is to offer timely access to mental health resources, support, and treatment for individuals in need.

**II LITERATURE REVIEW**

|  |  |  |  |
| --- | --- | --- | --- |
| **Author** | **Year** | **Key Concepts/Methods** | **Challenges/Issues** |
| Akshi  Kumar et al. | 2019 | Tweets of first 100 followers of MS India student forum are analyzed using various linguistic, semantic and activity features to detect anxious depression dis-  order. | Fine grain emotion analysis  can be done to detect anxiety indicators instead of using SentiWordNet which categorizes the words into three polarities. |
| Jana M.  Havigeo-a et al. | 2019 | Find models predicting the  depressivity of the writer based on the computational linguistic markers of his/her written text. | An unexpected limit of this  study is the higher percent- age of depressive men in the research sample. Due to the relatively small size of the research sample, we did not  further verify the results. |
| S. Smys,  Jennifer  S. Raj | 2021 | Proposes a machine-  learning algorithm to develop an early prediction from their depression mode, which can be protected from mental illness and suicide state of affairs. The combination of sup- port vector machine and Naive Bayes algorithm will be used to provide a good  accuracy level. | More datasets should be  used to verify the effective- ness of the proposed system. More number of attributes will be found from the emotional process fac- tor that should be included in future work. |
| Kim  Martinez et al. | 2021 | Aims to gather the most  current serious games, published from 2015 to 2020, with a new approach focusing on their applications: awareness, prevention, detection, and therapy. | The games should always  offer support while playing, in addition to collecting data on participant behavior during the game to better analyze their learning. |
| Astha  Singh, Divya Kumar | 2022 | The objective is to demodulate audio signals from a video clip if video data is in- appropriate to distinguish between stress, depression, and anxiety. | The face recognition task  can be performed using A. Singh and D. Kumar Micro- processors and Microsystems 95 (2022) 104681 15  thermal images from video input to minimize the noise  effect. |

**TABLE 1: Literature Review**

**III RELATED WORK**

There has been a significant amount of research and development in the field of comprehensive mental health detection. Some relevant examples of related work in the field of comprehensive mental health using machine learning and deep learning include:

1. Analysis of Deep Learning Techniques for Early Detection of Depression on Social Media Network [5]

This research article proposes a machine-learning algorithm to develop an early prediction from their depression mode, which can be protected from mental ill- ness and suicide state of affairs. The combination of support vector machine and Naive Bayes algorithm will be used to provide a good accuracy level. The classification model contains many cumulative distribution parameters, which should be classified and identified dynamically. This identification or detection is the features obtained from textual, semantic, and writing content.

1. Deep Learning for Depression Detection from Textual Data [11]

This paper proposes a productive model by implementing the Long-Short Term Memory (LSTM) model, consisting of two hidden layers and large bias with Recurrent Neural Network (RNN) with two dense layers, to predict depression from text, which can be beneficial in protecting individuals from mental disorders and suicidal affairs. They train RNN on textual data to identify depression from text, semantics, and written content. The proposed framework achieves 99.0% accuracy, higher than its counterpart, frequency-based deep learning models, whereas the false positive rate is reduced.

1. A Machine Learning based Depression Analysis and Suicidal Ideation Detection System using Questionnaires and Twitter [10]

In this paper, they have proposed a depression analysis and suicidal ideation detection system, for predicting the suicidal acts based on the level of depres- sion. They collected real time data from students and parents by making them fill questionnaires similar to PHQ-9 (Parent health questionnaire) consisting of questions like What’s your age? or Are you regular in school/college? and pro- cessed it into meaningful data with related features like age, sex, regularity in the school, etc. Then, classification machine algorithms are used to train and classify it in five stages of depression depending on severity - Minimal or none, mild, moderate, moderately severe and severe. Maximum accuracy i.e. 83.87% was achieved by using XGBoost classifier in this dataset. Also, data was col- lected in the form of tweets and were classified into whether the person who tweeted is in depression or not using classification algorithms. Logistic Regres- sion classifier gave the maximum accuracy i.e. 86.45 % for the same.

1. Depression detection from social network data using machine learning

techniques [9]

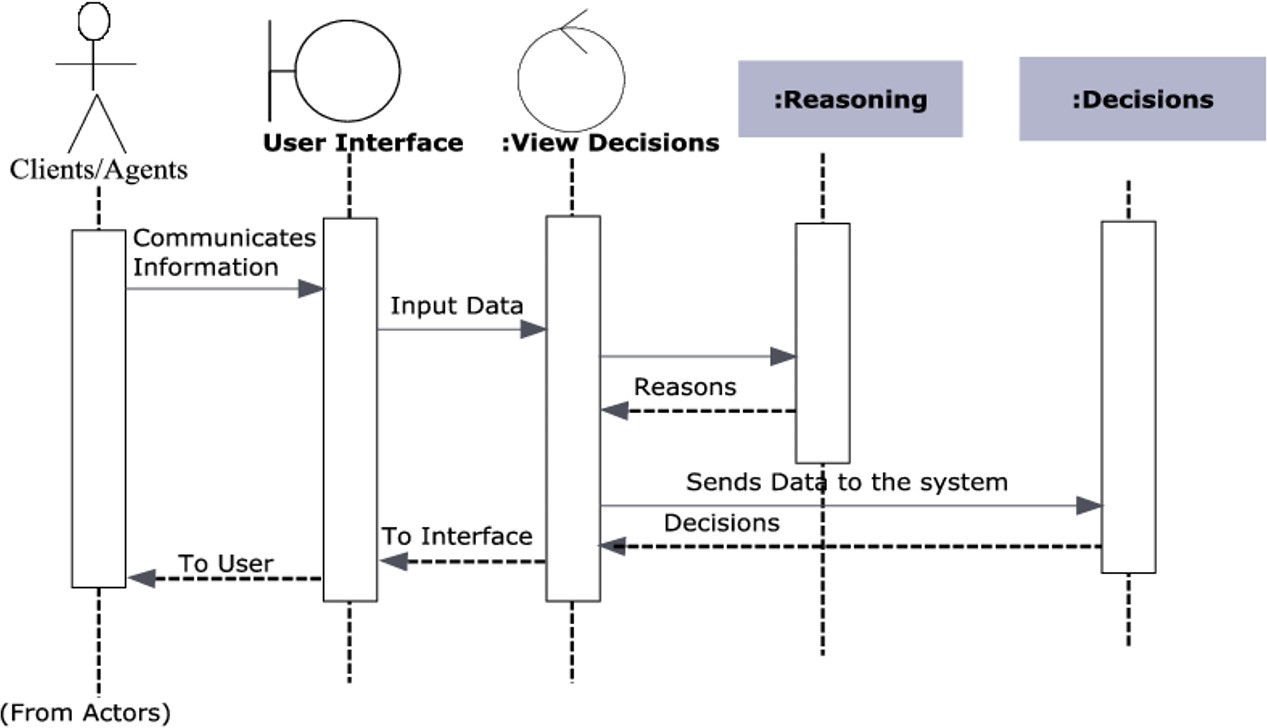
In this paper they have exhibited the capability of using Facebook as a tool for measuring and detecting major depression among its users. Facebook data to detect any factors that may reflect the depression of relevant Facebook’s users. Various machine learning techniques are employed for such purposes.

1. Machine Learning-based Approach for Depression Detection in Twitter

Using Content and Activity Features [12]

The present study aims to exploit machine learning techniques for detecting a probable depressed Twitter user based on both his/her network behavior and tweets. For this purpose, they trained and tested classifiers to distinguish whether a user is depressed or not using features extracted from his/her activities in the network and tweets. The results showed that the more features are used, the higher are the accuracy and F-measure scores in detecting depressed users. This method is a data-driven, predictive approach for early detection of depression or other mental illnesses. This study’s main contribution is the exploration part of the features and its impact on detecting the depression level.

**IV PROPOSED MODEL WORKFLOW**



**Figure 2. Work Flow**

Training a depression detection model involves several steps: data collection, data preprocessing, feature extraction, model selection, and model evaluation. Here is an overview of these steps:

1. **Data Collection:**

Collect a diverse dataset of labeled depression/non-depression data, which can include self-reported assessments, diagnostic interviews, or clinician ratings. Classic datasets such as the Iris flower data set, MNIST database, and others have been widely used in the statistical literature.

1. **Data Preprocessing:**

Clean the data by removing missing values and converting text data into numerical features. Techniques like tokenization, stemming, lemmatization, and stop-word removal are commonly used to transform the data into a format suitable for data mining and machine learning tasks.

1. **Feature Extraction:**

Extract features from the preprocessed text data, which can be done manually or automatically. Manual feature extraction involves identifying and describing relevant features, while automated feature extraction utilizes specialized algorithms or deep networks to extract features automatically.

1. **Model Selection:**

Choose an appropriate model for depression detection, such as traditional machine learning models (logistic regression, support vector machines, decision trees) or deep learning models (convolutional neural networks, recurrent neural networks). The statistical analysis helps select the best model based on goodness of fit and simplicity.

1. **Model Training:**

Train the selected model using the preprocessed and feature-extracted data. This involves splitting the data into training and validation sets and optimizing the model parameters using techniques like backpropagation and gradient descent.

1. **Model Evaluation:**

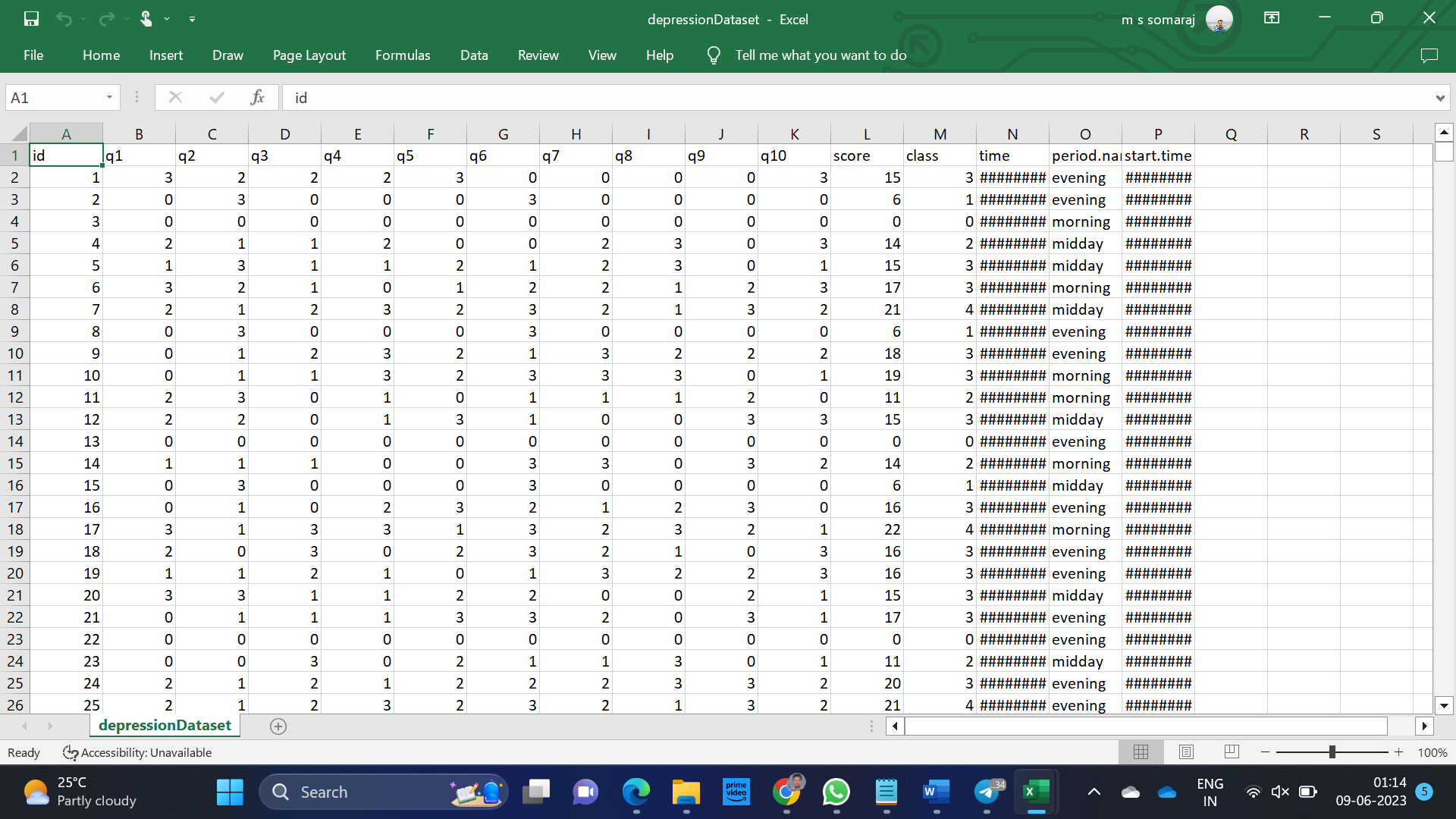
Evaluate the trained model on an unseen test set of data. Calculate metrics like accuracy, precision, and recall to assess the model's performance. Common evaluation metrics include accuracy, precision, confusion matrix, log-loss, and AUC (area under the ROC curve).

Model evaluation is essential to assess the model's efficacy during research phases and plays a role in model monitoring. Accuracy measures correct predictions, while precision measures the proportion of true positives among predicted positives. The confusion matrix provides a detailed breakdown of correct and incorrect classifications for each class, making it useful when distinguishing between classes with different misclassification costs or imbalanced data.

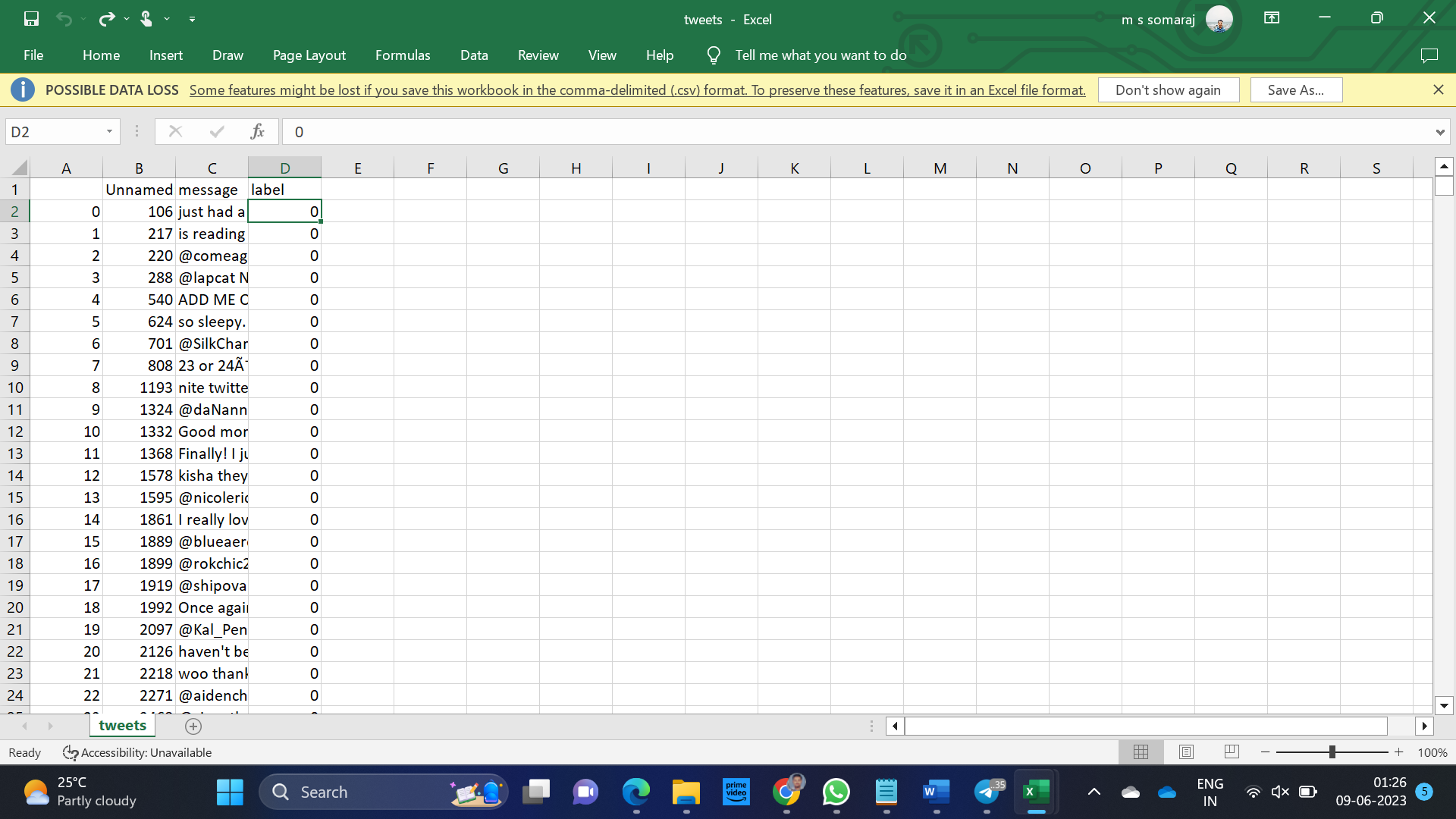
**V IMPLEMENTATION**

* 1. **Dataset Description**

In the proposed method we have used the PHQ-9 dataset for preparing the questionnaire for the quiz-based detection, another dataset having different scores of the 9 questions containing 15923 entries and a labelled tweet dataset containing 4629 tweets of different people for the detecting depression from the textual data entered by the user, by doing sentiment analysis on it. The 9-question Patient Health Questionnaire (PHQ-9) is a diagnostic tool introduced in 2001 to screen adult patients in a primary care setting for the presence and severity of depression. It rates depression based on the self-administered Patient Health Questionnaire (PHQ). The PHQ is part of Pfizer’s larger suite of trademarked products, called the Primary Care Evaluation of Mental Disorders (PRIME- MD). The PHQ-9 takes less than 3 minutes to complete and simply scores each of the 9 DSM-IV criteria for depression based on the mood module from the original PRIME-MD. Primary care providers frequently use the PHQ-9 to screen for depression in patients.



**Figure 3 : PHQ-9 based quiz dataset**



**Figure 4 : Twitter tweet dataset**

* 1. **Methodology**

The technique used for implementing an anxiety and depression detection app can vary depending on the specific approach used for detecting these conditions. The specific implementation of the app will depend on the chosen technique and the specific tools and technologies used for development.

In this analysis, we mainly utilized Quiz-based analysis and sentiment-based analysis. As mentioned earlier, we used the PHQ-9 dataset to prepare the questionnaire for the online quiz-based detection. Additionally, we employed another depression dataset with 15,923 entries containing scores for 9 questions to train the quiz-based model. For detecting depression from the user's textual data, we used a labeled tweet dataset containing 4,629 tweets from various individuals and performed sentiment analysis on it. Python was the language of choice for training, testing data, and result prediction. The depression dataset underwent classification using various machine learning algorithms to evaluate the results. The dataset was read from a CSV file, preprocessed by handling missing data, and then split into training and testing sets using the train-test split function from the sklearn.model selection module. We included several methods defining different classification algorithms using scikit-learn's implementations, including SVM, decision trees, random forests, naive Bayes, and KNN. Each method returns a trained classifier object. The accuracy method takes a trained classifier object as input, evaluates its accuracy on the test set using the confusion matrix, and then prints the accuracy percentage for the given classifier. Finally, we called the different classification methods and used the accuracy method to evaluate their performance.

For tweet sentiment analysis, we employed the Naive Bayes classifier. We loaded and preprocessed a dataset of tweets to train two classifiers: one based on Bag-of-Words (BOW) and another based on Term Frequency-Inverse Document Frequency (TF-IDF) methods. The classifiers were then tested on the dataset, and we computed metrics such as precision, recall, F-score, and accuracy to evaluate their performance.

The model takes two sample messages as input and classifies them as either positive or depressive tweets using both classifiers. The results are printed for each message. The code relies on several libraries, including pandas, numpy, and pickle. We also used the nltk library to download the Punkt tokenizer, which is used for tokenization in the preprocessing step. The Depression Detection class has an init () method that loads the tweet dataset and splits it into training and testing sets. It also includes a classify () method that takes a processed message and a method (either 'bow' or 'tf-idf') as input and returns a binary classification label (0 for positive, 1 for depressive tweet). Lastly, it has a metrics () method that takes true labels and predicted labels as input and computes several evaluation metrics.

The classifier takes in a set of training data, calculates the term frequency-inverse document frequency (TF-IDF) score for each word in the training data, and uses these scores to classify new tweets as either positive or depressive. The train method calculates either the TF-IDF scores or the probability scores for each word in the training set, depending on the value of the method parameter. The probability of a tweet being depressive or positive is calculated based on the frequency of words in the training set. The term frequency and inverse document frequency for each word in the training set are calculated, which is used to compute the TF-IDF score. The classify method takes in a processed tweet message and uses the scores calculated in the train method to classify the tweet as either positive or depressive. The method loads the scores from the pickle files created in the train method and calculates the probability of a tweet being depressive or positive based on the presence of words in the tweet.

The main block of the code instantiates a Depression Detection object and two Tweet Classifier objects (one for BOW and one for TF-IDF). It trains and tests the classifiers on the dataset using the metrics () method to evaluate their performance. Then, it classifies two sample messages using both classifiers and prints the results.

1. **COMPARISON OF ML TECHNIQUES**

There are several existing techniques for anxiety and depression detection, and each has its advantages and limitations. Here is a comparison of some of the most common techniques:

1. Audio analysis: If the app is using audio recordings as input for anxiety and depression detection, techniques such as speech signal processing and acoustic analysis can be used to extract features from the audio data. Machine learning models such as decision trees, support vector machines, and deep neural net- works can be used to classify the audio data as indicative of anxiety, depression, or other emotional states.
2. Self-reported assessments: Self-reported assessments, such as the Patient Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD- 7), are commonly used to assess anxiety and depression. These assessments are easy to administer and can be completed by the patient without the need for specialized equipment or expertise. However, self-reported assessments are subject to social desirability bias and may not accurately reflect the patient’s true emotional state.
3. Physiological signal analysis: Physiological signals, such as heart rate vari- ability and skin conductance, have been used to assess anxiety and depression. These signals can be measured using wearable sensors or other monitoring de- vices. Physiological signal analysis has the advantage of providing objective measures of emotional state, but it requires specialized equipment and may not be practical for everyday use.
4. Speech analysis: Speech analysis has been used to detect anxiety and depression based on features such as speech rate, pitch, and tone. This approach has the advantage of using easily accessible data, such as recordings of phone con- versations or speech samples, but it may be affected by factors such as language proficiency, dialect, and accent.
5. Natural language processing: Natural language processing techniques have been used to analyze text data, such as social media posts and emails, to detect anxiety and depression. This approach has the advantage of using readily available data and can be easily integrated into existing platforms. However, it requires sophisticated language models and may be subject to variations in language use and context.
6. Hybrid approaches: Hybrid approaches that combine multiple data sources, such as physiological signals and speech or text data, have been proposed for anxiety and depression detection. These approaches have the potential to pro- vide more comprehensive and accurate assessments of emotional state but may require more complex processing and analysis methods.
7. **RESULTS AND OBSERVATIONS**

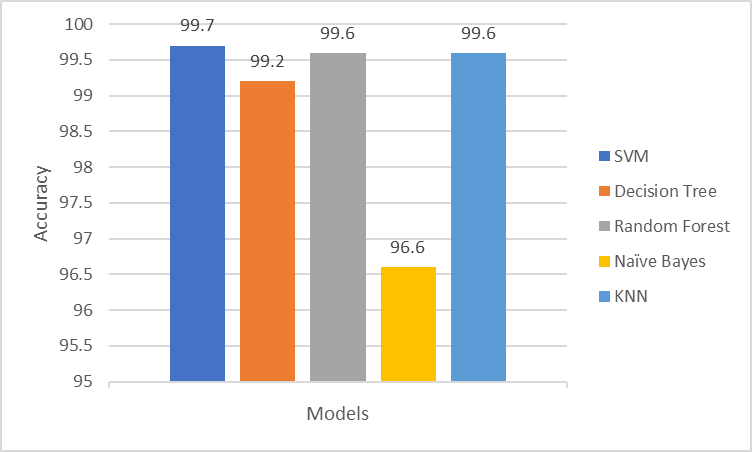
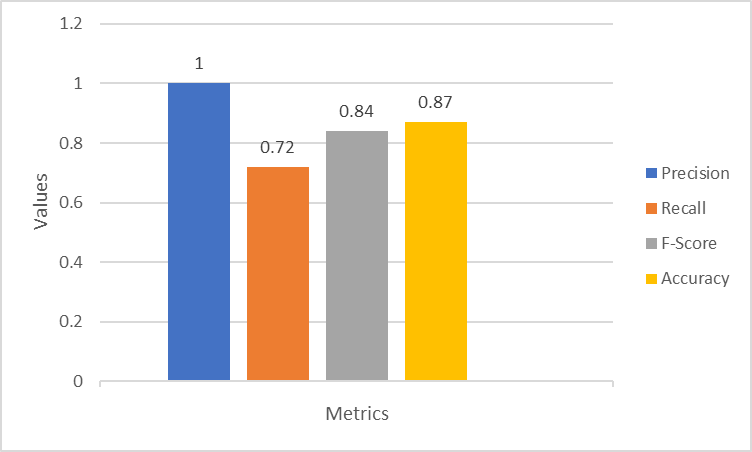


Figure 5. Shows the comparison of accuracies of different machine learning models in

**Figure 5 Accuracy of the Classifiers**

predicting the depression of the users from the quiz-based analysis. Svm, Decision tree, Random Forest, Naive Bayes and KNN classifiers obtained an ac- curacy of 99.7%, 99.2%, 99.6%, 96.6% and 99.6% respectively.



**Figure 6 Values of different evaluation metrics**

Figure 6, shows the values of different evaluation metrics such as precision, recall, F-score, and accuracy calculated based on the sentiment analysis, which gave the values 1.0 0.72, 0.84 and 0.87 respectively.

**VIII CONCLUSION**

Depression is a debilitating psychiatric disorder that exerts a profound impact on individuals' well-being. As technology continues to advance, there has been a burgeoning interest in developing applications and tools for detecting both anxiety and depression. Numerous methodologies have been proposed for this purpose, encompassing self-reported assessments, analysis of physiological signals, speech analysis, natural language processing, and hybrid approaches. Each technique possesses distinct strengths and limitations, and the most appropriate choice hinges on the specific application's requirements and the trade-offs between accuracy, convenience, and practicability.

Within the framework of this project, we leveraged the PHQ-9 dataset to develop a quiz-based depression detection system. Additionally, we utilized sentiment analysis techniques to detect depression in users' textual inputs, employing a dataset comprising diverse tweets. The quiz model was subjected to training using a range of classifiers, including SVM, Decision Tree, Random Forest, Naive Bayes, and KNN, yielding impressive accuracies of 99.7%, 99.2%, 99.6%, 96.6%, and 99.6%, respectively. For tweet sentiment analysis, a Naive Bayes classifier was employed to analyze and predict depression based on textual data provided by users. The Confusion matrix was utilized to determine the Precision, Recall, F-score, and Accuracy of the sentiment analysis, resulting in values of 1.0, 0.72, 0.84, and 0.87, respectively.

**IX FUTURE SCOPE**

In the realm of future prospects, numerous domains warrant further exploration and advancement in the domain of anxiety and depression detection. One of these domains pertains to the advancement of intricate machine learning models capable of accommodating individual variances and context-specific influences. Another crucial area involves amalgamating diverse data sources, encompassing audio, physiological signals, and textual data, to furnish comprehensive and precise evaluations of emotional states. Moreover, it is imperative to delve into the ethical and privacy ramifications of collecting and analyzing personal data for mental health assessment, as well as to scrutinize the efficacy of anxiety and depression detection applications and tools in real-world scenarios. On the whole, technology possesses significant potential to augment mental health outcomes, and sustained innovation and collaboration will be pivotal in realizing this potential

**REFERENCES**

[1] Martinez, K., Menéndez-Menéndez, M., & Bustillo, A. (2021). Awareness, Prevention, Detection, and Therapy Applications for Depression and Anxiety in Serious Games for Children and Adolescents: A Systematic Review. JMIR Serious Games, 9(4), 1-15.

[2] Singh, A., & Kumar, D. (2022). Detection of Stress, Anxiety, and Depression (SAD) in Video Surveillance using ResNet-101. Microprocessors and Microsystems, 9(4), 1-15.

[3] Kumar, A., Sharma, A., & Arora, A. (2019). Anxious Depression Prediction in Real-time Social Data. In International Conference on Advances in Engineering Science Management and Technology (ICAESMT), 1-6.

[4] Havigerová, J. M., Haviger, J., Kucera, D., & Hoffmannová, P. (2019). Text-Based Detection of the Risk of Depression. Frontiers in Psychology, 10, 1-9.

[5] Smys, S., & Raj, J. S. (2021). Analysis of Deep Learning Techniques for Early Detection of Depression on Social Media Network - A Comparative Study. Journal of Trends in Computer Science and Smart Technology (TCSST), 3(1), 24-39.

[6] Barua, P. D., Vicnesh, J., & Lih, O. S. (2022). Artificial Intelligence Assisted Tools for the Detection of Anxiety and Depression Leading to Suicidal Ideation in Adolescents: A Review. Cognitive Neurodynamics, 1-19.

[7] Nemesure, M. D., Heinz, M. V., & Huang, R. (2021). Predictive Modeling of Depression and Anxiety using Electronic Health Records and a Novel Machine Learning Approach with Artificial Intelligence. Scientific Reports, 11, 1-7.

[8] Sarkar, A., Singh, A., & Chakraborty, R. (2022). A Deep Learning-based Comparative Study to Track Mental Depression from EEG Data. Neuroscience Informatics, 2(4), 1-5.

[9] Islam, M. R., Kabir, M. A., Ahmed, A., Kamal, A. R. M., Wang, H., & Ulhaq, A. (2018). Depression Detection from Social Network Data using Machine Learning Techniques. Health Information Science and Systems, 6(1), 1-11.

[10] Sajja, G. (2021). Machine Learning-based Detection of Depression and Anxiety. International Journal of Computer Applications, 183(45), 20-23.

[11] Amanat, A., Rizwan, M., Javed, A. R., Abdelhaq, M., Alsaqour, R., Pandya, S., & Uddin, M. X. (2021). Deep Learning for Depression Detection from Textual Data. Electronics, 11(5), 1-12.

[12] Alsagri, H., & Ykhlef, M. (2020). Machine Learning-based Approach for Depression Detection in Twitter Using Content and Activity Features. IEICE Transactions on Information and Systems, E103-D(8), 1825-1832.

[13] Kroenke, K., Spitzer, R. L., & Williams, J. B. W. (2001). The PHQ-9. Journal of General Internal Medicine, 16, 606-613.

[14] Fisher, R. A. (1936). The Use of Multiple Measurements in Taxonomic Problems. Annals of Eugenics, 7, 179-188.