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

Shiju George1, Asha Joseph2, Sona Sanal3 , Anila Susan Philip4

1,2,3,4Department 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](mailto:sonasanal2023@it.ajce.in), anilasusanphilip2023@it.ajce.in

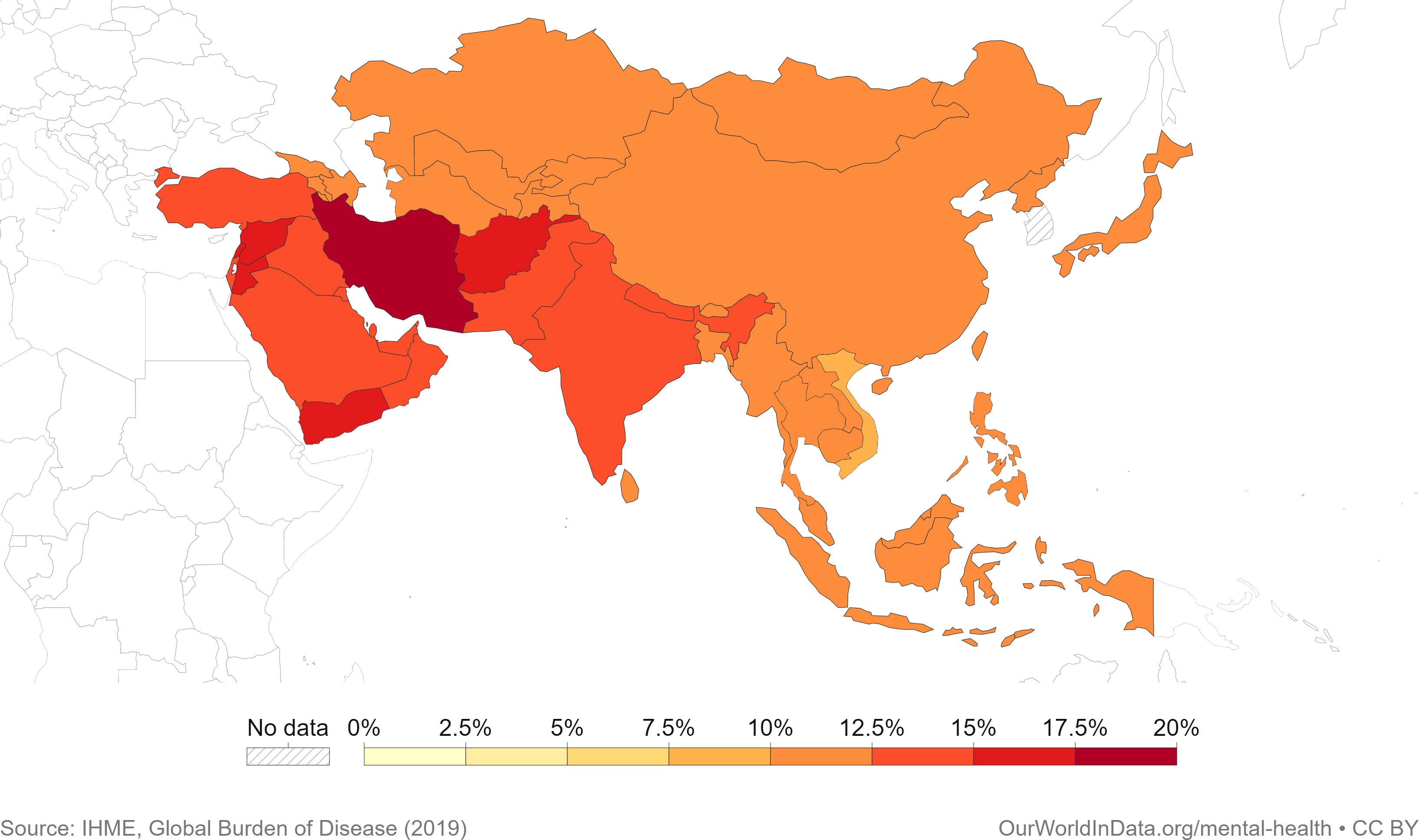
**Abstract:** Mental well-being is a crucial factor in today's competitive world. People experience a wide range of mental issues, including anxiety, bipolar disorder, depression, insomnia, and panic attacks, due to various reasons such as work pressure, peer pressure, relationship issues, and difficulties 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

One of the most common mental conditions worldwide is major depressive disorder (MDD), usually referred to as depression. According to the Comprehensive Mental Health Action Plan 2013-2020 of the World Health Organization, depression affects more than 300 million people around the globe and is one of the biggest causes of disability, particularly for women. By 2030, depression is expected to overtake diabetes as the top cause of illness burden in high-income nations, where it currently accounts for 4.3% of the global disease burden. Major depressive disorder, sometimes known as depression, is a serious medical condition that frequently affects feelings, thoughts, and actions. 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. Since children lacked the psychological and cognitive development to feel depressive symptoms, the medical establishment has refused to recognize the presence of pediatric depression disorders. Nevertheless, accumulating research has demonstrated that kids can certainly experience the whole range of mental disorders and have serious related morbidity and mortality. The prevalence of depression among teenagers, which ranges from 10% to 60%, has been validated by recent studies. It is important to remember that the real number of people living with depression may be larger due to underreporting and a lack of help-seeking behavior.

1. According to research from 2021, about 60% of depressed people do not seek professional help, largely because of the stigma attached to the illness.
2. The World Health Organization (WHO) notes that due to different treatment hurdles, more than 75% of people living in low- or middle-income countries do not receive treatment for depression.

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

The use of artificial intelligence (AI) and machine learning (ML) algorithms for the diagnosis of mental health issues, such as anxiety and depression, has gained popularity in recent years. These technologies have the potential to increase diagnostic accuracy and speed, enabling more efficient patient care and better outcomes. The goal of this project is to develop a system that can spot a person's first indications of depression. Our goal is to provide persons in need with timely access to mental health services, support, and therapy by identifying these signs at an early stage.

**II LITERATURE REVIEW**

|  |  |  |  |
| --- | --- | --- | --- |
| **Author** | **Year** | **Key Concepts/Methods** | **Challenges/Issues** |
| Akshi  Kumar et al. | 2019 | With the aim of identifying anxious depression disorder, the first hundred followers of the MS India student forum's Twitter account were subjected to linguistic, semantic, and activity analysis. | SentiWordNet, which divides words into three polarities, cannot be used to identify anxiety indicators because it is too coarse-grained. |
| Jana M.  Havigeo-a et al. | 2019 | On the basis of the computational linguistic indicators of the writer's written text, develop models for forecasting the writer's depressivity. | This study's increased percentage of older depressed men in the research sample presents an unanticipated limitation. We did not do any additional testing of the findings because the research sample was so small. |
| S. Smys,  Jennifer  S. Raj | 2021 | Proposes a machine-learning system to create an early depression prediction from their state, protecting them from mental illness and suicide circumstances. A good level of accuracy will be provided by combining the support vector machine and the Naive Bayes algorithm. | To confirm that the suggested system is effective, more datasets should be employed. The emotional process factor will reveal a greater quantity of qualities that ought to be taken into account in upcoming tasks. |
| Kim  Martinez et al. | 2021 | Aims to compile the most recent serious games, released from 2015 to 2020, with a fresh perspective concentrating on their applications: awareness, prevention, detection, and therapy. | In addition to gathering information on participant behavior during the game to better analyze their learning, the games should always provide assistance while being played.. |
| Astha  Singh, Divya Kumar | 2022 | If visual data is inadequate to distinguish between stress, depression, and anxiety, the goal is to demodulate auditory signals from a video clip. | A. Singh and D. Kumar Micro-processors and Microsystems 95 (2022) 104681 15 thermal images from video input can be used to conduct the face recognition task while reducing the noise effect. |

**TABLE 1: Literature Review**

**III RELATED WORK**

In the area of thorough mental health detection, there has been a lot of research and advancement. In the area of comprehensive mental health, some pertinent examples of related work employing machine learning and deep learning include:

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

In order to prevent mental illness and suicide, this study piece suggests a machine-learning algorithm to generate a predicted outcome from their depressive mood. To offer a high level of accuracy, the support vector machine and Naive Bayes method will be combined. There are numerous cumulative distribution parameters in the classification model that need to be dynamically categorized and identified. The features used for this identification or detection come from the textual, semantic, and writing content.

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

In order to predict depression from text, the Long-Short Term Memory (LSTM) model, which consists of two hidden layers, large bias, and recurrent neural network (RNN) with two dense layers, is implemented in this paper. This model can be useful in preventing people from mental illnesses and suicidal thoughts. To diagnose depression from language, semantics, and written content, they train RNN on textual data. In comparison to frequency-based deep learning models, the suggested framework achieves 99.0% accuracy while having a lower false positive rate.

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

For forecasting suicidal behaviors based on the amount of depression, they have presented a depression analysis and suicidal ideation detection system in this work. Parents and students were asked to complete PHQ-9-style surveys, which included questions like "What's your age?" to gather data in real time. or Do you attend classes regularly? and transformed it into useful data with correlated characteristics like age, sex, frequent attendance at school, etc. Then, classification machine algorithms are utilized to train and categorize the data into five severity-based phases of depression: minimal or none, mild, moderate, fairly severe, and severe. Using the XGBoost classifier, the maximum accuracy of 83.87% was attained in this dataset. Additionally, information was gathered in the form of tweets, and using classification algorithms, it was determined whether the tweeter was depressed or not. The highest accuracy, or 86.45%, was provided by the Logistic Regression classifier for the same.

1. Depression detection from social network data using machine learning

techniques [9]

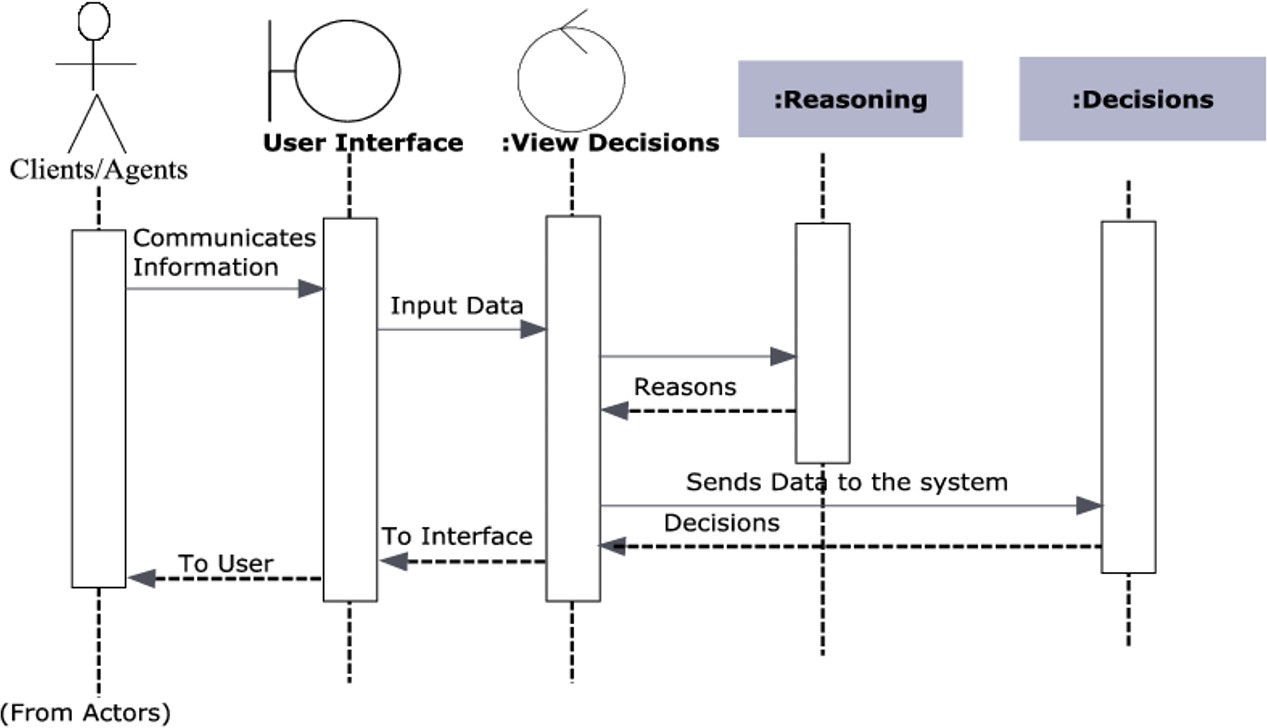
The possibility to use Facebook as a tool for assessing and detecting serious depression among its users has been demonstrated in this paper. Using Facebook data, researchers can look for any patterns that would indicate that certain Facebook users are depressed. For these aims, a variety of machine learning approaches are used..

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

Using Content and Activity Features [12]

The goal of the current study is to employ machine learning approaches to identify potential depressed Twitter users by analyzing their network activity and tweets. They developed and tested classifiers to do this, using information gleaned from a person's network activities and tweets to determine if the user is depressed or not. The findings indicated that the accuracy and F-measure scores in identifying depressed users increased with the number of features included. A data-driven, predictive strategy is used in this technique to identify depression or other mental diseases early on. The exploration of the traits and their impact on determining the degree of depression is the primary contribution of this work.

**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:**

Select a suitable machine learning model for detecting depression, such as deep learning models (convolutional neural networks, recurrent neural networks), or more conventional models (logistic regression, support vector machines, decision trees). The statistical analysis aids in choosing the optimum model based on simplicity and goodness of fit.

1. **Model Training:**

The preprocessed and feature-extracted data should be used to train the chosen model. This entails dividing the data into training and validation sets and utilizing tools like backpropagation and gradient descent to optimize the model parameters.

1. **Model Evaluation:**

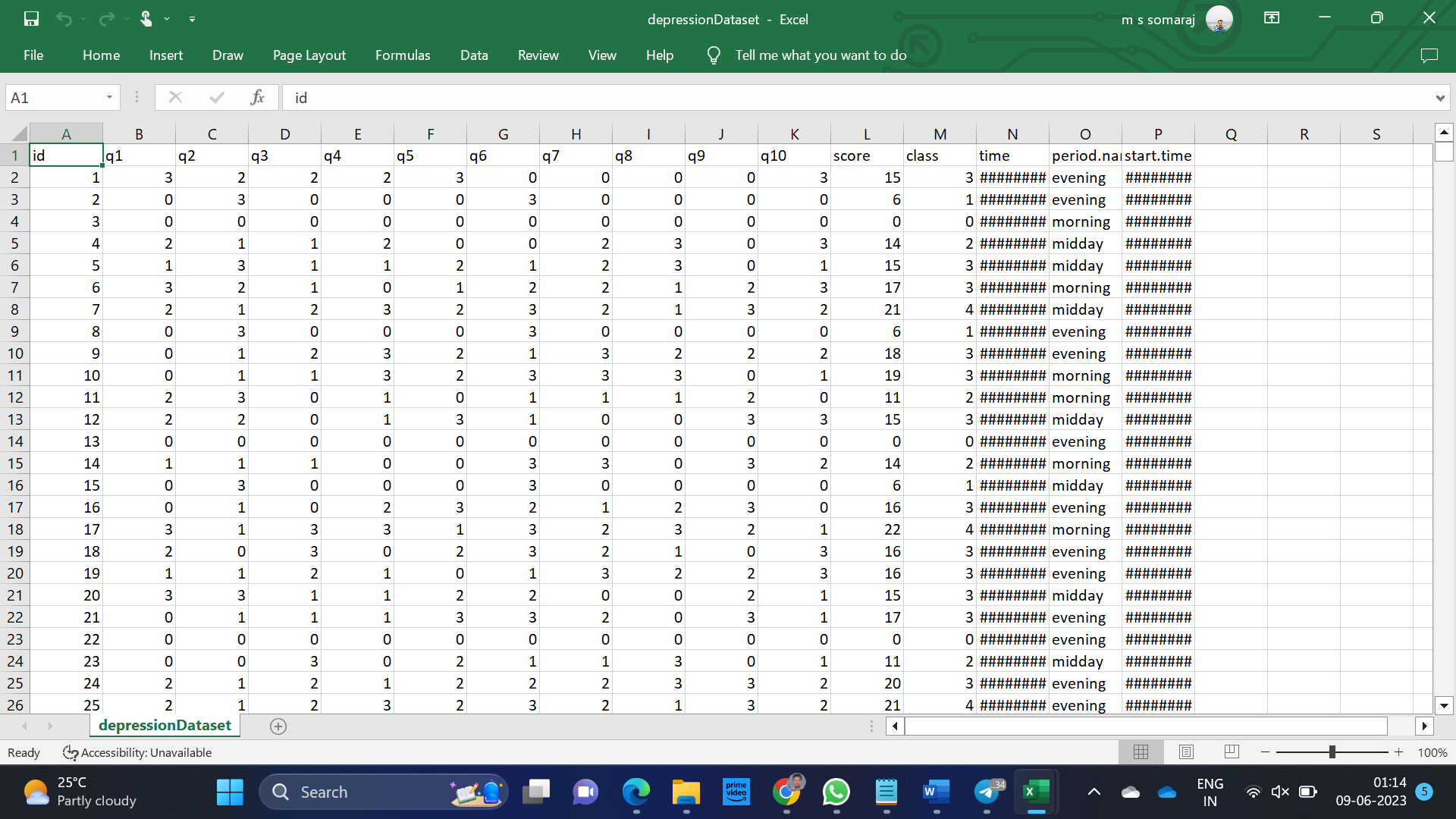
Evaluate the trained model on an unseen test set of data. Calculate metrics like accuracy, precision, and recall to assess the model's performance. Common evaluation metrics include accuracy, precision, confusion matrix, log-loss, and AUC (area under the ROC curve).

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.

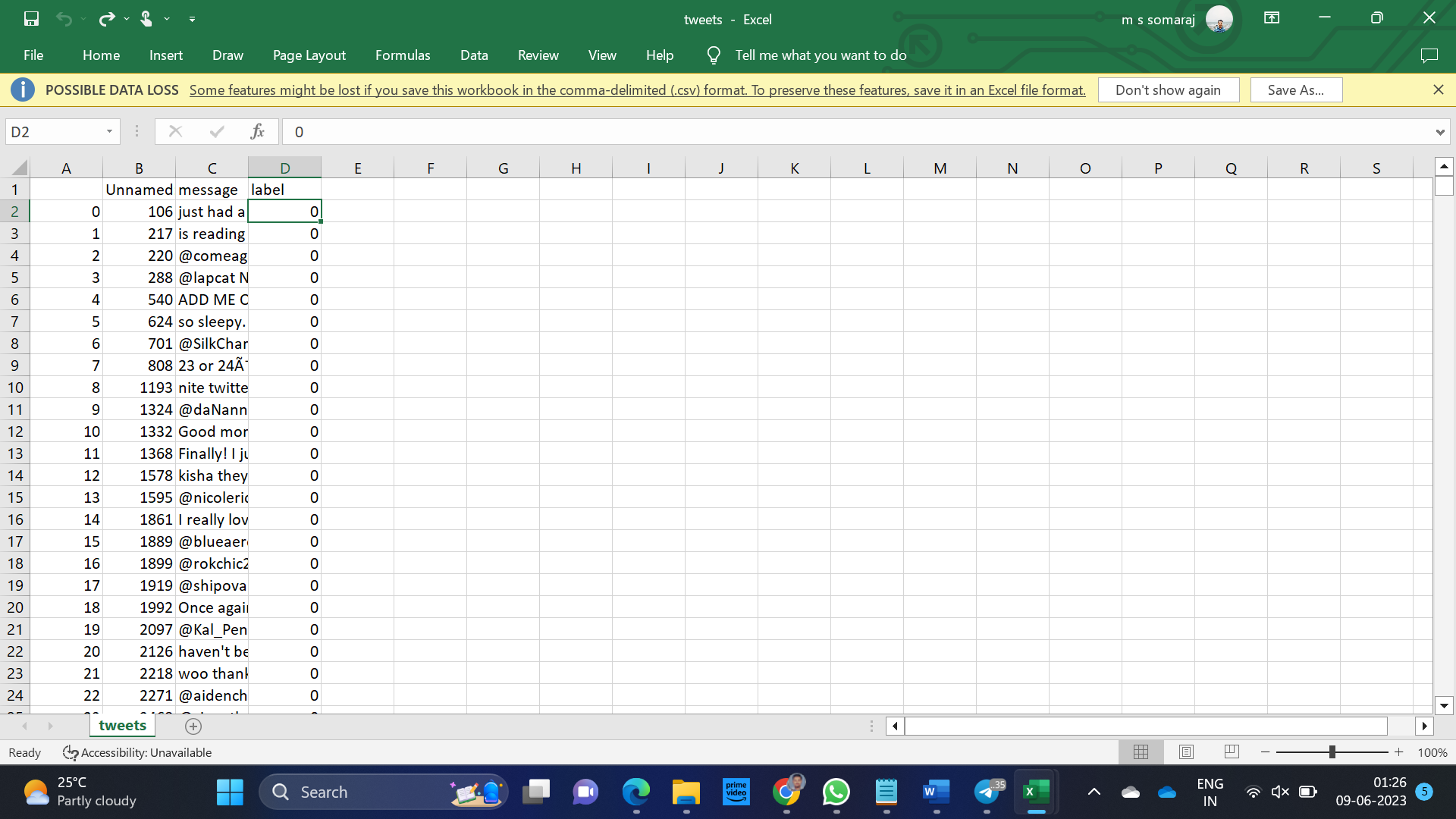
**V IMPLEMENTATION**

* 1. **Dataset Description**

The PHQ-9 dataset, another dataset with different scores for the nine questions, and a labelled tweet dataset, each containing 4629 tweets from various people, have all been used in the proposed method to prepare the questionnaire for quiz-based depression detection. Additionally, sentiment analysis has been applied to the text input by the user to identify depression from the inputted text. In order to assess the presence and severity of depression in adult patients receiving primary care, the 9-question Patient Health Questionnaire (PHQ-9) was developed in 2001. Based on the results of the self-administered Patient Health Questionnaire (PHQ), depression is rated. The Primary Care Evaluation of Mental Disorders (PRIME- MD) is a wider collection of trademarked goods from Pfizer that includes the PHQ. The PHQ-9 is based on the mood module from the original PRIME-MD and takes less than 3 minutes to complete. It simply scores each of the 9 DSM-IV criteria for depression. The PHQ-9 is often used by primary care doctors to check their patients for depression.



**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. Using the train-test split function from the sklearn.model selection module, the dataset was split into training and testing sets after being read from a CSV file and preprocessed by handling missing data. 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.

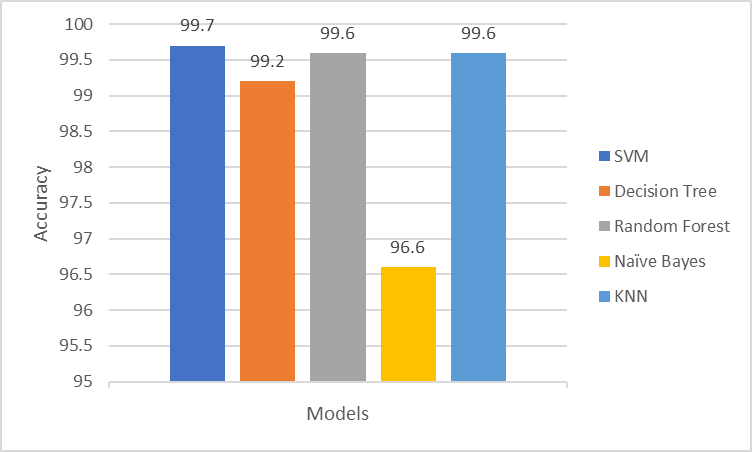
Fresh tweets are classified as either positive or depressing by the classifier using a set of training data and the term frequency-inverse document frequency (TF-IDF) scores it produces for each word. For each word in the training set, the train method determines either the TF-IDF scores or the probability scores depending on the method parameter's value. Based on how often terms appear in the training set, it is possible to determine if a tweet is depressing or upbeat. The TF-IDF score is generated by calculating the term frequency and inverse document frequency for each word in the training set. 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**

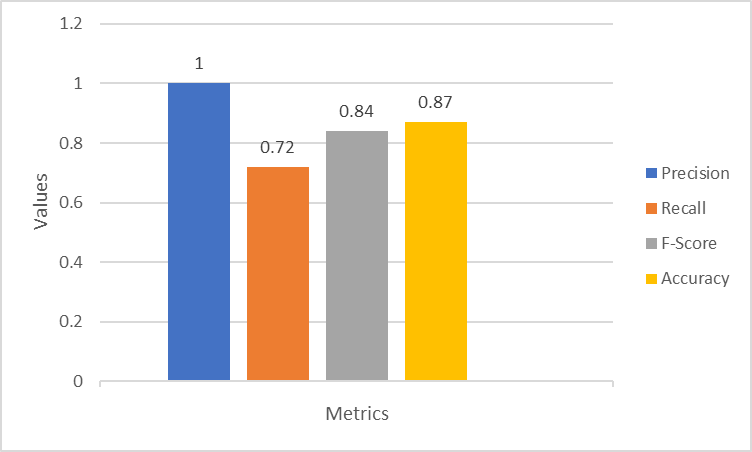
There are several existing techniques for anxiety and depression detection, and each has its advantages and limitations. Here is a comparison of some of the most common techniques:

1. Audio analysis: Speech signal processing and acoustic analysis techniques can be utilized to extract features from audio data if the app uses audio recordings as input for anxiety and depression detection. To classify the audio data as suggestive of anxiety, depression, or other emotional states, machine learning methods can be utilized, including decision trees, support vector machines, and deep neural networks.
2. Self-reported assessments: Self-reported tests are frequently used to evaluate anxiety and depression. Examples include the Patient Health Questionnaire-9 (PHQ-9) and the Generalized Anxiety Disorder-7 (GAD- 7). These examinations can be carried out by the patient without the aid of specialist tools or training because they are simple to administer. Self-reported evaluations, however, are susceptible to social desirability bias and could not be a real reflection of the patient's emotional condition.
3. Physiological signal analysis: Anxiety and depression have been measured using physiological signals including skin conductance and heart rate variability. Wearable sensors or other monitoring technologies can be used to measure these signals. Although objective assessments of emotional state can be obtained by physiological signal analysis, this method needs specialist equipment and might not be appropriate for everyday use.
4. Speech analysis: Based on characteristics including speech pace, pitch, and tone, speech analysis has been used to identify signs of anxiety and sadness. Although this method has the benefit of using readily available data, such as recordings of phone conversations or speech samples, it may be impacted by elements like language ability, dialect, and accent.
5. Natural language processing: In order to identify anxiety and depression in text data, such as emails and social media posts, natural language processing techniques have been applied. This strategy has the benefit of utilizing easily accessible data and is simple to integrate into current platforms. However, it necessitates complex linguistic models and might be affected by changes in linguistic usage and contextual factors.
6. Hybrid approaches: For the purpose of detecting anxiety and depression, hybrid techniques that mix various data sources, such as physiological signals and voice or text data, have been developed. These methods may necessitate more involved processing and analysis techniques but have the ability to provide assessments of emotional state that are more thorough and accurate.
7. **RESULTS AND OBSERVATIONS**



**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 accuracy 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 classification 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**

Debilitating psychiatric disorders like depression have a significant negative impact on people's wellbeing. An increasing number of people are interested in creating tools and applications for identifying anxiety and depression as technology develops. 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**

Future prospects show that there are many areas that merit more research and development in the area of detecting anxiety and depression. The development of complex machine learning models that can account for individual variations and context-specific variables is one of these areas Another crucial area involves amalgamating diverse data sources, encompassing audio, physiological signals, and textual data, to furnish comprehensive and precise evaluations of emotional states. Additionally, it is crucial to investigate the ethical and privacy implications of gathering and analyzing personal data for mental health assessment. Additionally, it is important to evaluate the effectiveness of applications and tools for anxiety and depression detection in real-world situations. Overall, technology has enormous promise to improve mental health outcomes, and maximizing this potential will need continued innovation and collaboration.

**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.