www.nature.com/scientificreports

scientific reports

OPEN

Ensemble of hybrid model based technique for early detecting of depression based on SVM and neural networks

Dip Kumar Saha[1], Tuhin Hossain[2], Mejdl Safran[3]✉, Sultan Alfarhood[3], M. F. Mridha[1]✉ & Dunren Che[4]

The prevalence of depression has increased dramatically over the last several decades: it is frequently overlooked and can have a significant impact on both physical and mental health. Therefore, it is crucial to develop an automated detection system that can instantly identify whether a person is depressed. Currently, machine learning (ML) and artificial neural networks (ANNs) are among the most promising approaches for developing automated computer-based systems to predict several mental health issues, such as depression. This study propose an ensemble of hybrid model-based techniques that aims to build a strong detection model that considers many psychological and sociodemographic characteristics of an individual to detect whether a person is depressed. Support vector machines (SVM) and multilayer perceptrons (MLP) are the two fundamental methods used to construct the suggested ensemble approach. The hybrid DeprMVM served as a meta-learner. In this study, the hybrid DeprMVM is a level-1 learner, whereas the SVM and MLP networks are level-0 learners. After the classifiers are trained and tested at level 0, their outputs are based on both the independent and dependent variables in the new data set that was used to train the meta-classifier. The training data class imbalance was reduced by applying the synthetic minority oversampling technique (SMOTE) and cluster sampling together, which improved the accuracy for detecting depression. Additionally, it can effectively reduce the risk of over-fitting from simply duplicating data points. To further confirm the effectiveness of the proposed method, various performance evaluation metrics were calculated and compared with previous studies conducted on this specific dataset. In conclusion, among all the techniques for identifying depression, the suggested ensemble approach had the best accuracy, at 99.39%, and an F1-score of 99.51%.

The present way of life has given rise to many psychological health issues in a significant number of individuals. Guo et al.[1] considered several Psycho-pathological conditions, such as anxiety, depression, stress, schizophrenia, bipolar illness, and personality disorders exhibit certain shared characteristics that cause feelings of depression and isolation. Depression is currently recognized as a medical disorder and is considered to be one of the most prevalent mental illnesses. It ranks third in terms of its impact on the global disease burden and is the primary cause of disability in medium-income and high-income countries . It is treated as a severe illness that affects an individual's emotional well-being and also inflicts bodily damage on the patient. Le et al.[2] presented the most common mental health disorders (MHD) include anxiety, nervousness, sleep deprivation, eating, addiction, depression, trauma, and stress-related disturbances . It affects an individual's thinking abilities, leads to mood instability, and frequently diminishes productivity. Over 300 million people worldwide[3], or over 4.4% of the population, suffer from depression. Thus, early detection and treatment of depression symptoms can significantly increase the likelihood of successfully managing depression and lessen its bad effects on a person's general well-being, physical health, and socioeconomic standing. Alada et al.[4] identified the depressed symptoms of persons relies on analyzing their online social media information, posts, and comments to determine their emotional state is a significant activity . Nevertheless, their emotional state is robust and may result in confusion regarding

[1]Department of Computer Science, American International University-Bangladesh, 1229 Dhaka, Bangladesh. [2]Department of Computer Science and Engineering, Daffodil International University, Dhaka, Bangladesh. [3]Department of Computer Science, College of Computer and Information Sciences, King Saud University, 11543 Riyadh, Saudi Arabia. [4]Department of Electrical Engineering and Computer Science, Texas A\M University-Kingsville, 78363 Kingsville, TX, USA. ✉email: mejdl@ksu.edu.sa; firoz.mridha@aiub.edu

their identification with depression. Wang et al.[5] presented the current most commonly used methods are clinical interviews and questionnaire surveys conducted by hospitals and organizations. These assessments utilize psychiatric evaluation tools to determine probable outcomes of mental disorders. In recent times, significant progress in the scientific and medical domains has yielded very efficient technological advancements that have enabled the early detection and anticipation of diseases.

Detecting depression in sociology and public health is challenging. The prediction of depression is primarily dependent on an individual's psychological and sociodemographic traits. Rapid progress in information and technology has led to an increased demand for the utilization of ML algorithms. Robles-Palazón et al.[6] used several ML models to extract meaningful patterns from data across various sectors. Although ML algorithms have been extensively used in the healthcare and medical fields, their utilization in the psychological sector remains limited. In general, there are multiple ML and deep learning (DL) algorithms, and continued research is aimed at producing optimal results. It is important to acknowledge that no single learning algorithm consistently outperforms others in all areas therefore, it is still crucial to determine which type of algorithm is most effective for a specific task. To address this difficulty, we developed an automated detection system that relies on a sophisticated ensemble learning model, known as ensemvble of DeprMVM. This model enabled us to quickly and accurately identify depression symptoms in individuals.

The motivation behind this research was to detect the depression of a person at an early stages, which would enable medical practitioners to instantly recognize the signs of depression and provide proper treatment without any delay. Therefore, there can be a notable decline in the number of spontaneous, unexpected suicide instances among depressed people. Thus, our proposed methodology has a noteworthy influence on the field of public health, particularly in medical diagnosis and health consulting. Another aspect of this work is to compare it to the existing work done on this dataset, considering the model accuracy, and finally, demonstrate the effectiveness of the proposed model in terms of several performance parameters. Hence, to construct our desired system, we investigated several ML and ANN classifiers, such as the Random Forest classifier (RFC), K-nearest neighbor (KNN), extreme gradient boosting (XGB), and MLP in a written inquiry dataset for immediate guidance of individuals regarding their self-perceived signs of depression.

This work makes significant contributions in the following areas:

- We propose a cutting-edge ensemble approach of a hybrid model for the early detection of depression that employs a combination of ML and ANN.
- Our automated approach can play a vital role in detecting depression by using the psychological and socioeconomic conditions of a person.
- We integrate data manipulation methods, such as cluster sampling and SMOTE, to enhance the accuracy of the ensemble model performance.
- This research approach improves the predictive accuracy of depression in comparison to previous studies.The remainder of this paper is organized as follows: The relevant literature is presented in Section 2, and the entire methodology of the proposed model is detailed in Section 3. The methods used in this study are described in Section 4. Section 5 presents the experimental outcomes along with a comparison of the suggested work with other pertinent research. Section 6 concludes the report, provides a summary of the outcomes, and suggests potential directions for future studies.

Related works

This research problem requires an in-depth investigation because of its complexity. This section examines numerous related research works to learn about the tools and methods utilized in previous studies and to identify research gaps.

Traditional ML techniques

Nemesure et al.[7] used several ML algorithms based on the patient's electronic health records to predict the presence of major depression disorder (MDD) and generalized anxiety disorder (GAD) in patients. 59 undergraduate people with unique characteristics participated in that study. The ML models achieved 73% accuracy for GAD and 67% for MDD. The second aim of their research was to determine the main features of patients with GAD or MDD based on the Shapley additive explanation score. Chung et al.[8] identified pre-disease signs of a mental health patient using an automated computer system. The system was developed by applying different supervised ML classification algorithms and ensemble ML approaches on the "OSMI Mental Health in the Tech Survey" dataset. Several distinct ML models have been developed with different capabilities for detecting these symptoms. Among all the applied supervised classification algorithms, the Gradient Boosting algorithm generated the highest accuracy of 88.80%, whereas the ensemble models generated only 85.60%. In addition, the Neural Network achieved 88% accuracy. Susanty et al.[9] used a geriatric depression scale to assess whether adults experience depression. The normal range of this scale is from 0 to 4, mild depression ranges from 5 to 8, moderate depression ranges from 9 to 11, and severe depression ranges from 12 to 15. They applied four ML models in their work. Thus, depressive symptoms among community-dwelling older adults, especially those who are illiterate can be automatically predicted without any preliminary screening for depressive symptoms. According to internal validation in their study, the RF and deep-insight visible neural network were the top models. For the external validation set containing a non-local ethnic group, the RF model achieved an AUROC score of 0.619 (95% CI 0.610 to 0.627). Additionally, Rois et al.[10] suggested an ML approach to identify stress among university students in Bangladesh. This approach can efficiently detect key predictive factors and accurately forecast psychological issues. A survey dataset comprising 355 students from 28 different universities in Bangladesh was used in that research. The dataset includes various questions covering anthropometric measurements, academic performance, lifestyle choices, and health-related information. These questions aimed to assess respondents' perceived stress levels. Therefore, multiple ML models, including DT, RF, SVM, and logistic regression(LR), have been applied to predict the current mental state of university students. Compared with other ML techniques, the RF model achieved a maximum accuracy of 89.3% and showed superior predictive ability for stress detection. Zhaozhao et al.[11] developed a classification technique that anticipates the presence of depression among university students by identifying important risk factors associated with depression. Six distinct ML classifier models were used to construct a classification model to predict the existence of depression. Among the six classifiers, the stacking classifier with 24 characteristics demonstrated the highest accuracy of 77.45%. In addition, it provided a precision of 0.77, recall of 0.64, and 0.80 AUC, which is better for psychological diagnosis. Nayan et al.[12] focused on the effectiveness of ML algorithms in predicting depression and anxiety among university students in Bangladesh during the initial stages of the COVID-19 pandemic. It also considered demographic factors, the patient health questionnaire(PHQ-9), and the GAD assessment-7 scale. Several widely recognized ML techniques were employed here, and linear discriminant analysis to forecast the occurrence of mental illness among university students residing in Dhaka city. Based on several performance criteria, the RF model performed better than the other models at predicting depression, with an accuracy of 89%. On the other hand, SVM yielded the highest accuracy of 91.49% than other models in predicting anxiety. Ali et al.[13] created an entirely new survey dataset that implemented several trained ML approaches to assist in predicting depression in postmenopausal women. Among all the applied classifiers, the RF and XGB models are the top performers, achieving an accuracy of 99.04% by utilizing the 14 most significant features. Priya et al.[14], utilized different ML methods to ascertain five distinct levels of severity of anxiety, despair, and stress. Data for this study were gathered from employed and unemployed persons from various cultures and groups using a depression, anxiety, and stress scale questionnaire. After exploring various ML techniques, it was determined that there was an imbalance in the distribution of the classes within the confusion matrix.

Xin et al.[15] employed the RF classifier to predict depression in older adults with disabilities based on six variables: social relationships, family dynamics, health behavior, health condition, and demographics. 30% of the test set and 70% of the training set were used to develop the prediction model. The percentage of elderly people with disabilities living in rural areas was 57.67%, while that in urban areas was 44.59%. Urban disabled seniors with depression, obtained 78.1% specificity and 64.2% sensitivity. AlSagri et al.[16] introduced here for binary classification, which involves determining whether a person is experiencing depression based on information from their tweets and behavior on Twitter. Considering several significant features, various ML classifiers such as SVM, NB, RF, and linear SVM have been utilized to anticipate mental and psychological issues and provide recommendations for these to specific individuals. Compared with other classifiers, the linear SVM classifier achieved the highest accuracy of 82.5%. Kyoung et al.[17] built an ML-based predictive model based on survey data collected from community-dwelling adults of the Republic in Korea to identify the future symptoms of depression. In their research, SMOTE was used to overcome the issue of class-imbalanced data. Lastly, the predictive model was developed based on the RF classifier, which generated a maximum accuracy of 86.20%. Zulfiker et al.[18] suggested two main factors that influence the onset of depression. One is satisfaction with socio-familial relationships, and the other is satisfaction with health. Another thing attempted to identify the most common signs of depression, particularly concerning socio-demographic and psychosocial factors. These factors can be employed to determine whether an individual is depressed. The most crucial psychosocial and socio-demographic characteristics of depressed persons are extracted using a variety of ML feature selection strategies. Finally, the AdaBoost classifier relies on the SelectKBest factor-picking strategy was the ideal model for forecasting depression, with an accuracy of 92.56%. Sau et al.[19] developed an automated system that utilizes various ML algorithms to identify mental health disorders in seafarers as early as possible. The dataset was compiled with targeted inquiries that encoded the socio-demographic, occupational, and health-related data of 470 seafarers in their research. They employed multiple ML classifiers to predict the existence of anxiety and depression, as evaluated using the Hospital Anxiety and Depression Scale. Overall the classifiers, the CatBoost classifier achieved the highest accuracy rate of 89.3%.

Deep learning techniques

Kour et al.[20] predicted one's mental health condition using Twitter data separating depressed and non-depressed users. This is accomplished by building a hybrid DL model known as the CNN-biLSTM model, which combines two DL architectures: bi-directional long short-term memory (biLSTM) and convolutional neural network(CNN). The implementation of this model on the Twitter Depression dataset yielded the highest accuracy of 94.28% compared with all previous studies conducted on this dataset. Uddin et al.[21] introduced a promising approach using LSTM-based Recurrent Neural Networks (RNN) to detect text that reflected self-perceived signs of depression in two distinct depression datasets, including the Norwegian depression dataset. Subsequently, the suggested DL methodology is used to train the time-sequential factors that differentiate texts that depict signs of depression from posts that have no such elaboration, or non-depression posts. In addition, because the features are predicated on probable signs of depression, the system may produce insightful justifications from ML models utilizing the explainable artificial intelligence (XAI) method known as Local Interpretable Model-Agnostic Explanations (LIME). However, the suggested feature-based strategy for depression symptoms performs better than conventional general word frequency-based techniques, which prioritize the frequency of characteristics rather than the particular symptoms of depression. Therefore, our trained suggested model can automatically predict depression posts with a maximum accuracy of 99% for these datasets when compared to other traditional techniques. Amanat et al.[22] suggested an efficient approach for predicting depression from text to prevent suicidal thoughts and MHD through the implementation of the LSTM-based model, which has two hidden layers and a large bias with RNN, including two dense layers. In comparison to its counterpart in the emotions of many social media users, this suggested technique shows the viability of RNN and LSTM by achieving remarkable results of 99% with a reduction in the false positive rate for early detection of depression.

Ghosh et al.[23] implemented an automated system to detect the existence of depression among Bengali individuals who use social media, by analyzing the content of their Bengali posts on different social platforms, to prevent unexpected suicidal cases. This system was validated using both the ML and DL models. The LSTM-GRU-CNN hybrid DL model outperformed all other models with an accuracy of 92.25%, sensitivity of 94.46%, and specificity of 91.15%. Depression can arise due to various factors, including anxiety disorders, bipolar disorder, and sleep disturbances. Marriwala et. al[24] presented a strategy for an automated depression detection system utilizing a DL approach. Three models were constructed in this study. Textual CNN was the first model in which CNN was trained using only text-based features. The second is a CNN model trained using only audio features. The last is a hybrid model that integrates audio and textual features and incorporates LSTM algorithms. The experimental results demonstrate that the audio CNN approach is superior to the text CNN model for detecting depression using the DL approach. The audio CNN model achieved a high accuracy of 98% and a low loss of 0.1%, enabling it to accurately predict the early symptoms of depression. In contrast, with only a 0.2% loss, the text CNN model achieved an accuracy of 92%. Kumar et al.[25] forecasts five levels of severity for anxiety, sadness, and stress by employing eight distinct ML models. These models are classified into four distinct categories: Bayes, neural networks, lazy, and trees. These approaches were used on two distinct datasets, and the findings indicated that the NN model, specifically the radial basis function network (RBFN), outperformed all other models in terms of depression prediction in both datasets.

Considering what we have covered so far, it seems that the majority of the existing research has been conducted to detect signs of mental illness among a particular group of populations, considering only specific factors, such as certain age ranges and medical health status. However, we suggest a more precise system that considers a wide variety of factors such as sociodemographic status, specific age ranges, medical health status, and socioeconomic background of an individual, to detect the early symptoms of depression. The methodology section will now reveal the techniques employed in this study to correctly identify depression in its early phases.

Research methodology

In this section, we discuss our recommended approach for extracting the best results from a given dataset as shown in Figure 1. Many ML classifiers and an ANN were employed in our study. To accomplish this and for our study, we selected some of the best training methods. These consisted of the RFC, KNN, SVM, XGB, and MLP. Because we have employed tabular data to detect depression, the proposed methodology will benefit from this ML methods. To better diagnose depression, we suggest a hybrid model DeprMVM, that combines ML and ANN. The proposed ensemble method combines an SVM with two neural networks: MLP and DeptrMVM. The depression framework consists of two layers: Level 0 and Level 1. The outputs of the classifiers were based on both the dependent and independent variables in the new data set that was used to train the meta-classifier.

Data analysis

In our study, we utilised the dataset[18], which included information on 604 individuals in total. A total of 397 out of 604 individuals in the sample were diagnosed with depression. 65.7% Of the survey respondents, 65.7% reported having depression, which was significantly higher than the actual prevalence of depression in the general population[26]. Given that the COVID-19 pandemic and the period in which the dataset was gathered coincided with a global emotional and socioeconomic crisis, it is possible that this contributed to an increase in the prevalence of depression among the participants. It was discovered that 34.3% of the respondents did not

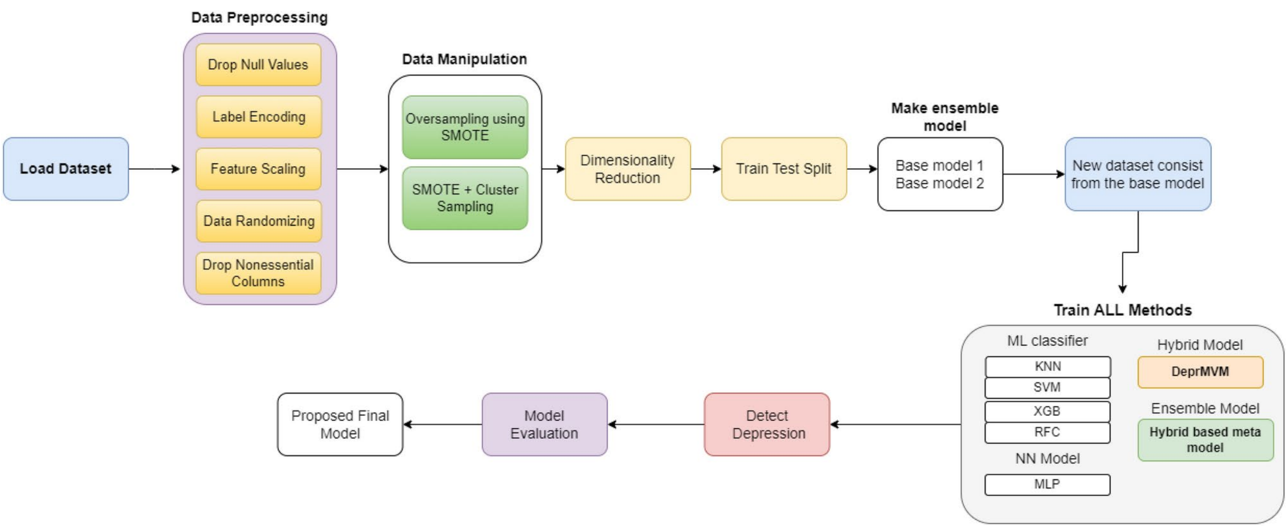

Fig. 1. The proposed methodology of this research (i) Load the dataset (ii) Dataset preprocessing (iii) Data manipulation (iv) Spiliting of the dataset (v) Make ensemble method using two base model (vi) Generate new dataset from the base model (vii) Train models applying ML and hybrid algorithms (viii) Detect depression (ix) Evaluate best model

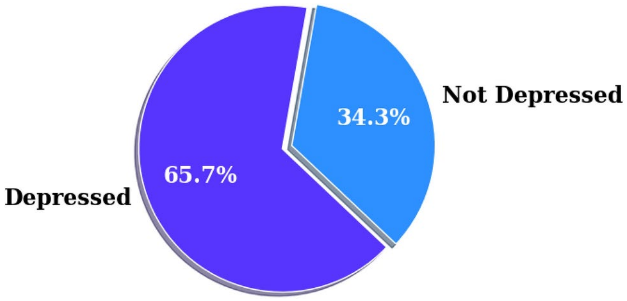

Fig. 2. Distribution of the dataset's participants-those with and without depression.

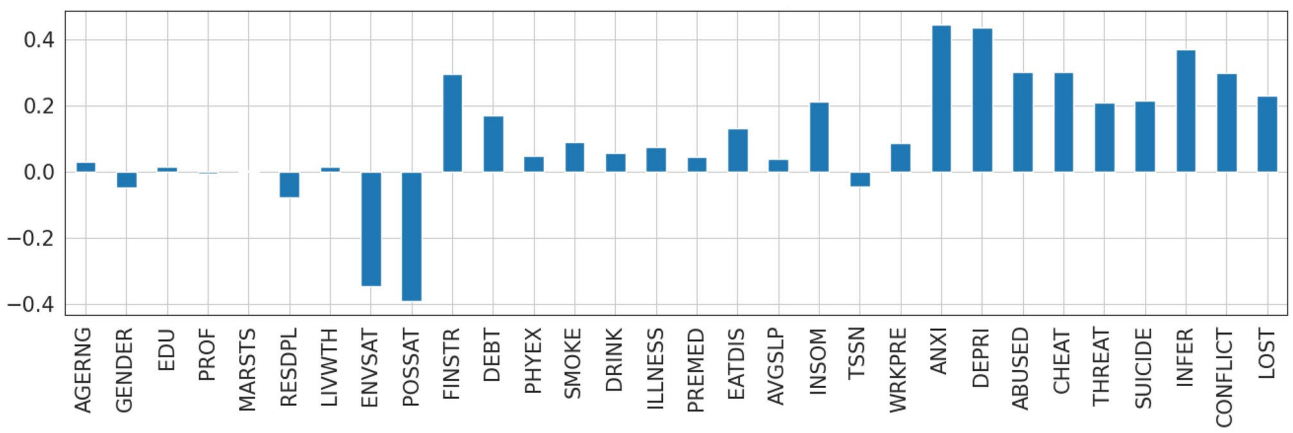

Fig. 3. Correlation of features after standard scaling.

experience depression. The distribution of participants in the dataset who were and were not depressed is shown in Figure 2.

Dataset preprocessing

Data pre-treatment is crucial for data analysis and ML. It is essential to maintain the quality of data because it cleans, fixes errors, and handles missing numbers. It also makes feature engineering easier, which by converting data, choosing pertinent features, and producing new useful qualities, can greatly improve the performance of ML models. Data pretreatment opens the door to more precise insights and reliable predictive models by guaranteeing that the data are in the best possible form for analysis. Preprocessing in this study comprised feature scaling, label encoding, and null value tracing. These steps are essential to achieving improved accuracy and less data loss.

Null Value Tracing and Label Encoding To guarantee data quality, completeness, and statistical validity all of which have an impact on the efficacy of ML models and the precision of analytical insights it is imperative that null values are checked in datasets. Accurate results depend on properly treating null values, and knowing why data are missing might yield insightful information about the domain. The null values were tracked, but none were discovered. Most ML techniques rely on numerical data to obtain better results. During the variable encoding stage, the categorical variables from the training and test datasets were converted into their equivalent numeric form using the Label Encoder function of the scikit-learn module. The numerical values 0 (NO) and 1 (YES) are already present in the target variable "DEPRESSED".

Feature Scaling Next, we moved on to the data transformation step, and a well-liked data preparation method called a standard scaler was used in ML to scale or standardize the characteristics in our dataset. This changes the data such that the mean is equal to zero and the standard deviation was equal to one. This is performed to guarantee that every feature has the same size, which might be crucial for some ML algorithms, especially those that rely on gradients and distances, such as k-means clustering and support vector machines (SVM). Certainly, here's the equation for standard scaling is:

$$z = \frac{x - \mu}{\sigma} \tag{1}$$

where z represents the standardized value, x represents the original data point, μ represents the mean (average) of the feature values, and σ is the standard deviation of the feature values. In Figure 3 The difference when using the standard scaler is shown.

Parameter name	Parameter value	Parameter description	Possible values
AGERNG	Predictor	Age range for the individual taking part (in years)	16-20, 21-25, 26-30, 31-35,36-40, 41-45, 46-50, 51-55, 56-60, 61+
GENDER	Predictor	participant's gender	Male, Female
EDU	Predictor	The individual's educational background	SSC, HSC, Graduate, Post Graduate
PROF	Predictor	The participant's occupation	Student, Service holder (Private), Service holder (Government), Businessman, Unemployed, Other
MARSTS	Predictor	The state of the participant's marriage	Unmarried, Married, Divorced
RESDPL	Predictor	Kind of residence for the concerned parties	Village, Town, City
LIVWTH	Predictor	It shows whether or not the individual resides with his family	With Family, Without Family
ENVSAT	Predictor	Whether or not the individual is content with his living situation	Yes, No
POSSAT	Predictor	If or not the participant is happy with his/her academic standing or current position	Yes, No
FINSTR	Predictor	Whether or whether the individual is experiencing financial strain	Yes, No
DEBT	Predictor	Whether or if the participant is in trouble	Yes, No
PHYEX	Predictor	how frequently the participant engages in physical activity	Never, Sometimes, Regularly
SMOKE	Predictor	Regardless of the participant's smoking status	Yes, No
DRINK	Predictor	Whether or not the person taking part consumes alcohol	Yes, No
ILLNESS	Predictor	Whether or not the person is afflicted with a serious sickness	Yes, No
PREMED	Predictor	Whether or not the individual takes any prescribed medication	Yes, No
EATDIS	Predictor	Whether or not the individual is experiencing eating disorders such as binge eating or desire loss	Yes, No
AVGSLP	Predictor	How long does the participant sleep at night on average	Below 5 h, 5 h, 6 h, 7 h, 8 h, More than 8 h
INSOM	Predictor	Whether or not the individual has sleeplessness	Yes, No
TSSN	Predictor	The participant's average number of hours on social media every day	Less than 2 h, 2-4 h a day, 5-7 h a day, 8-10 h a day, more than 10 h a day
WRKPRE	Predictor	The participant's present job or academic stress	Severe, Moderate, Mild, No Pressure
ANXI	Predictor	If the participant has experienced anxiety recently	Yes, No
DEPRI	Predictor	Whether or not the individual has experienced a recent sense of being denied something they are entitled to	Yes, No
ABUSED	Predictor	Whether or not the individual has experienced physical, sexual, or emotional abuse in the recent past	Yes, No
CHEAT	Predictor	Whether or not the individual has recently felt deceived by someone	Yes, No
THREAT	Predictor	Whether or not the person has recently experienced anything potentially fatal	Yes, No
SUICIDE	Predictor	Whether or not the participant has recently entertained suicide thoughts	Yes, No
INFER	Predictor	Whether or if the participant has experienced an inferiority complex recently	Yes, No
CONFLICT	Predictor	Whether or whether the individual has recently gotten into any sort of arguments with friends or family	Yes, No
LOST	Predictor	Whether the individual has experienced a recent loss of a close relative	Yes, No
DEPRESSED	Target	The goal variable indicates whether or not the subject is depressed	0 (Not depressed), 1 (Depressed)

Table 1. Dataset description for all parameters

Feature encoding was not required. Once the scaling was complete, we separated the features from the target variable, "DEPRESSED".

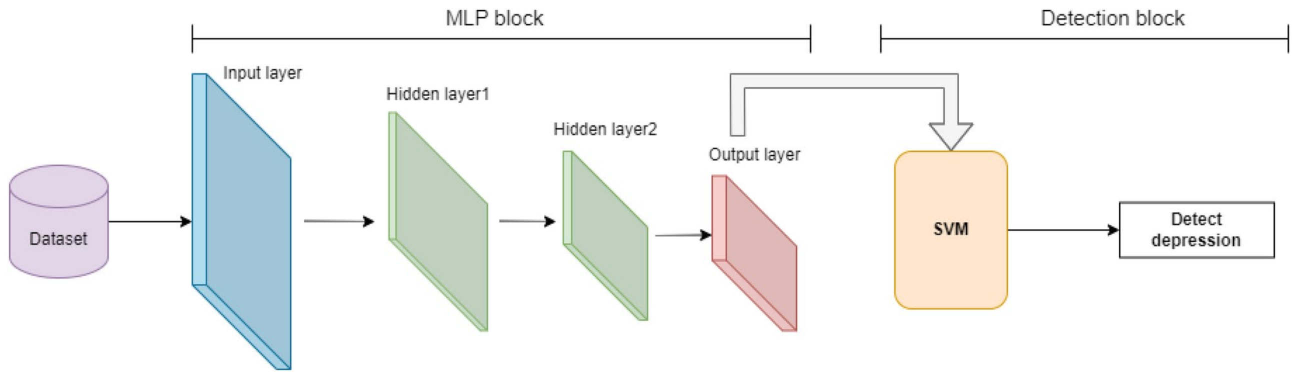

Fig. 4. The architecture of the hybrid model DeprMVM, combines the MLP and SVM classifiers. The MLP block and detection block have shown here.

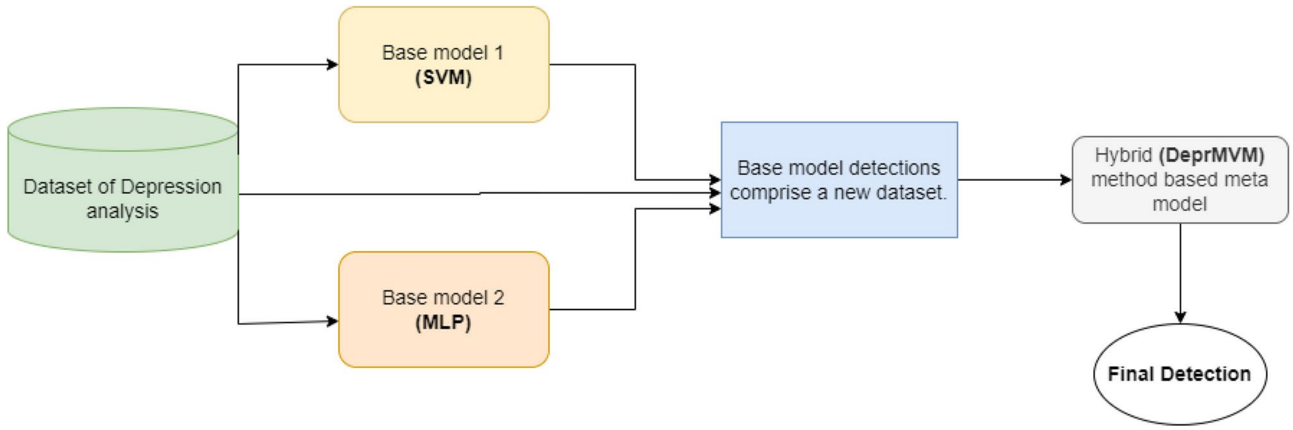

Fig. 5. In the proposed ensemble of hybrid model architecture, Two base classifiers combine to create a new dataset, and the DeprMVM method serves as a meta-model to produce the ultimate detection.

Technique	Classifier	Specificity%	Recall %	Precision %	F1-score %
Applying only SMOTE	KNN	97.47	53.75	95.56	68.80
	SVM	56.96	99.90	70.18	82.47
	XGB	89.87	88.75	89.87	89.31
	RFC	91.14	87.50	90.91	89.17
	MLP	93.67	91.25	93.59	92.41
	Ensemble of DeprMVM	93.67	92.50	93.67	93.08
Applying Smote and Cluster Sampling	KNN	93.92	85.00	92.17	85.37
	SVM	60.38	90.00	74.39	87.42
	XGB	95.27	94.40	95.27	94.67
	RFC	96.61	94.50	96.36	94.52
	MLP	99.29	99.00	99.12	97.22
	Ensemble of DeprMVM (Proposed)	99.48	99.26	99.64	99.51

Table 2. The effectiveness of the classifiers using various methods of data manipulation.

Data manipulation

ML approaches, such as oversampling and under-sampling data, are employed to rectify class imbalance in datasets. Oversampling helps the model better comprehend the minority class by increasing the number of cases in the minority class to balance the distribution, sometimes by duplication or the creation of synthetic data. Under-sampling, on the other hand, decreases the instances of the majority class to prevent model bias but may result in information loss. To successfully address class imbalance, other strategies such as cost-sensitive learning and ensemble techniques need to be considered. The decision between these approaches relies on the datasets

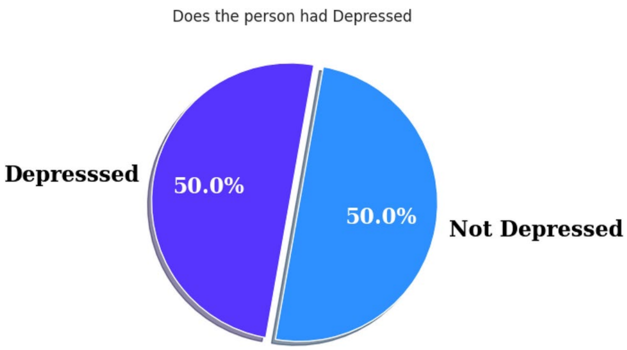

Fig. 6. Representation of the dataset after applying SMOTE.

and algorithms. We decided to use data modification techniques to address the imbalance in this dataset and to evaluate the performance of ML and DL.

SMOTE: SMOTE stands for the "Synthetic Minority Over-sampling Technique", which is a popular method for oversampling imbalanced datasets in ML[27]. To balance the distribution of classes, the SMOTE function generates synthetic instances of the minority class. Using this method, a data point from the minority class was chosen, and its k-nearest neighbors were then determined. Next, by interpolating between the chosen data point and its neighbors, SMOTE creates additional data points. Consequently, a set of synthetic data points is produced that, although introducing a variety, is comparable to the minority class examples that already exist. SMOTE effectively reduces the danger of over-fitting that can arise from simple duplication of data points while reducing bias towards the majority class. Our proposed methodology uses SMOTE to create synthetic instances of the minority class while balancing the distribution of classes. It minimizes bias towards the majority class and the risk of over-fitting that might impact on simple data point duplication. It's a useful technique for enhancing the classifier performance on unbalanced datasets. The techniques used were as follows:

$$\mathbf{x}_{\text{new}} = \mathbf{x}_i + (\mathbf{x}_{\text{nn}} - \mathbf{x}_i) \times \text{random\_value} \tag{2}$$

where the $\mathbf{x}_{\text{new}}$ represents the new synthetic sample, $\mathbf{x}_{\text{i}}$ is the original minority class instance, $\mathbf{x}_{\text{nn}}$ is one of the k-nearest neighbors of $\mathbf{x}_{\text{i}}$, and the random_value value is a random number between 0 and 1 used to control the interpolation.

Cluster sampling: Cluster sampling is a statistical sampling technique that is used in surveys and research. In cluster sampling, the population is divided into clusters or groups, and a random selection of these clusters is made for inclusion in the sample[11] After that, every cluster that has been identified is examined in depth, which frequently entails gathering information from every person or entity inside the clusters that have been selected. Because cluster sampling may lower the cost and logistical constraints of data collection, it is particularly helpful when it is impracticable or too expensive to directly sample people or units. To guarantee that the sample is representative of the population of interest, clusters should be chosen at random, as this might introduce heterogeneity within the clusters. In this study, it was used to evaluate several ML and DL models by undersampling an unbalanced dataset. Direct sampling of every feature in the depression dataset is helpful, as cluster sampling is utilized in our suggested methodology to reduce the dataset's cost and practical constraints.

Dataset splitting

The train-test split is a key method in ML and model assessment. In this process, a dataset is split into two distinct subsets: a training set and a testing set. An ML model can learn various patterns and correlations in the data by training on the training set. The testing set, on the other hand, was kept apart and employed to assess the generalization and performance of the model. The capacity of the model to provide accurate predictions of fresh, unknown cases can be evaluated by evaluating how well it performs on unseen data from the testing set. To avoid overfitting, a situation in which a model retains the training data but is unable to generalize to new data in which the train-test split is crucial. First, the training and test data were separated from the acquired dataset. 85% of the dataset was used as training data. The testing was performed on the remaining 15% of the dataset. Finally, a new set of data that is applied to the meta-model is derived from the two base models.

Methods for detecting depression

We selected various ML classifiers, one DL model, a hybrid model, and an ensemble meta-model for the training procedure. We selected the RFC, KNN, SVM, and XGB classifiers as the ML classifiers. We also propose a state-of-the-art hybrid model, DeprMVM. This model incorporates the DL and ML architectural components. With the help of the ML and DL models, we created an ensemble model that achieved greater accuracy in detecting depression.

Support vector machine

Support Vector Machine is an ML method based on the theory of statistical learning. It is a popular classifier used in many practical applications, particularly in classification problems. The basic design philosophy of SVM is to maximize the classification boundaries and its purpose is to maximize the hyper-plane[28]. SVM can handle both linearly separable and non-linearly separable problems by mapping data into a high-dimensional feature space. It uses a boundary-detection technique to retain the potential support vector and improve the learning generalization ability. SVM aims to minimize the empirical risk and achieve the minimization of existential risk and confidence range. This is a powerful tool for classification purposes. The mathematical equation representing the SVM classifier is expressed as follows:

$$f(x) = \mathrm{sign}(W \cdot X + b) \tag{3}$$

where $f(x)$ is the classification function, X is the input data, W is a weight vector. b is the bias term. In the case of a linear SVM, the objective is to determine the best W and b values to maximize the margin between the two classes[29].

DL model (multi-layer perceptron)

The MLP model consisted of three types of layers in this study. An input layer, one or more hidden layers, and an output layer are added to the many layers of linked nodes or neurons that constitute this network. Weighted connections and activation functions allow each neuron in the network to process information and forward it to the subsequent layer. MLPs[30] are ideal for a variety of applications, such as pattern recognition, regression, and classification because they may represent intricate and non-linear connections in data. When training an MLP, weights, and biases are often adjusted using optimization methods such as back-propagation to decrease the error between the expected outputs and the real objectives.

The mathematical equation of the MLP classifier consists of separate equations for each layer of the network. The forward pass of an MLP with a single hidden layer is represented by the following equations:

1. Input to the first hidden layer:

$$z_1 = W_1 \cdot X + b_1 \tag{4}$$

Here, z_1 is the weighted sum of input features X for the first hidden layer. The weight matrix for the links between the input layer and the first hidden layer is denoted as W_1. For the initial hidden layer, the bias vector is denoted as b_1.

2. Activation function for the first hidden layer:

$$a_1 = \sigma(z_1) \tag{5}$$

is typically a non-linear activation function, such as a sigmoid, Rectified Linear Unit(Relu), or hyperbolic tangent.

3. Input to the output layer:

$$z_2 = W_2 \cdot a_1 + b_2 \tag{6}$$

Here: z_2 is the weighted sum of activations from the first hidden layer to the output layer. The weight matrix W_2 represents the relationship between the output layer and first hidden layer. The bias vector of the output layer is denoted as b_2 .

4. Output of the MLP:

$$a_2 = \mathrm{softmax}(z_2) \tag{7}$$

In multiclass classification, the softmax function is often used to convert raw scores z_2 into class probabilities.

The above equations describe a simple feed-forward neural network with one hidden layer. This would keep building hidden layers on top of one another, each with the same processing pattern, to develop deeper networks. The goal of training an MLP is to determine the optimal weight matrices (W_1 and W_2) and bias vectors (b_1 and b_2) by adjusting them during the training process to minimize a chosen loss function, which is typically related to the difference between the predicted and actual target values. Training is often performed using techniques such as backpropagation and gradient descent[31].

Hybrid method (DeprMVM)

After using the models above, we combined the best-performing models to create an ensemble-based hybrid model to determine whether the hybrid model outperformed the other models.

DeprMVM is a hybrid model that combines the MLP and SVM models to accurately detect depression. This offers a hybrid strategy that leverages the benefits of both ANN[32] and conventional ML methods. The MLP performs feature extraction and complex data representation learning in this combination. At the same time, the

SVM performs the role of classifier, using the characteristics learned by the MLP to inform its conclusions. This combination is quite advantageous, especially when dealing with high-dimensional or non-linear data because the SVM is good at successfully differentiating different classes, whereas the MLP is good at capturing intricate patterns.

The hybrid DeprMVM is shown in Figure 4. The technique uses three hidden layers of the NN, with the output of the last hidden layer serving as the input for the SVM, which more precisely detects the depression. The method is slightly more complex than using SVM and MLP classifiers independently, but it typically improves several attributes, such as accuracy, precision, recall, and the f1-score. The mathematical equation of DeprMVM comprises separate equations for each layer of the network. An MLP with an input layer and two hidden layer forward passes is represented by the following equation:

Step 1: Input to the output layer:

$$Z_1 = W_1 \cdot (a_1 + a_2) + b_1 \tag{8}$$

Here: Z_1 is the weighted sum of the activations from the input layer to the hidden output layer. W_1 is the weight matrix for the connections between the first hidden layer and output layer. In addition, a_1, and $a_2$are hidden layers. b_1 denotes the bias vector of the output layer.

Step 2: Output equation of the DeprMVM:

$$A_1 = \mathrm{softmax}(Z_1) + \mathrm{sign}(W_2 \cdot X_0 + b_2) \tag{9}$$

In multiclass classification, the softmax function is often used to convert raw scores Z_2 into class probabilities. Here, X_0 is the connection between the MLP output and SVM input layers. Then W_2 is the weight vector and b_2 is the bias term. We obtained more accurate results when we utilized the hybrid model because the final hidden layer of the MLP had more crucial characteristics for every training data set, which was then further classified using SVM.

Following a thorough evaluation of the various ML, Hybrid, and DL models, we determined which model performed the best and detected the greatest impact on this type of dataset.

Input: Sociodemographic dataset, D
$(c_1, d_1), (c_2, d_2), \ldots\ldots, (c_m, d_m)$
The two selected base learners SVM and MLP.
The Ensemble meta learner hybrid model (DeprMVM)
Methodology
Step1: Preparing the models at level-0.
while $t = 1, \ldots\ldots, T$ **do**
| Utilizing D, train a base classifier P_t
end
Step2: Comprising a new dataset D_N
while $i = 1, \ldots\ldots, m$ **do**
| Create a new sample $\hat{c}_i$, $\hat{d}_i$ using Equation 10
end
Step3: Adapt the meta-model $\hat{P}$ that is hybrid-based using D_N.
return $M(c) = \hat{P}(P_1(c), P_2(c), \ldots\ldots, P_T(c))$
Output: Hybrid model based ensemble meta classifier M

Algorithm 1. Proposed hybrid model ensemble.

Proposed ensemble model

The proposed ensemble[33] approach combines two neural networks, MLP and DeptrMVM, with an SVM. There are two layers in the depression framework: Level 0 and Level 1. Both the dependent and independent variables in the fresh data set utilized for training the meta-classifier are the foundation for the classifiers' outputs after they have been trained and tested at level 0 using out-of-sample examples.

In this research, the hybrid DeprMVM is a level-1 learner, whereas the SVM and MLP networks are level-0 learners. The primary rationale for choosing MLP and SVM as foundational learners is their exceptional prediction performance in sequential data modeling[34] and their resilience in such scenarios. Additionally, because different base models are likely to generate various types of errors, their differences guarantee a variety in the ensemble, which is crucial. Figure 5 shows the flowchart of the proposed process. The proposed model approach is described in Algorithm 1.

First, the SVM and MLP networks were used to train the basis models. To ensure that there was no data leakage, an ensemble was developed using the two-fold cross-validation (CV) approach. The second stage creates a new dataset by transforming the out-of-fold predictions produced by the two base models and adding actual labels. Specifically, in the new dataset, the expected target labels were utilized as attributes, whereas the initial class labels comprised the response variable. Furthermore, the base learners used identical CV indices to train separately on the training data, whereas the hybrid model was trained using out-of-fold detections. As they were not utilized to train the level-0 models, actual labels were attached to these occurrences. Depressed and not depressed are thus represented by the base learners' outputs of 1 and 0, respectively.

For instance, assuming that each sample in a sociodemographic dataset D is c_i, d_i , a new sample $\hat{c}_i$, $\hat{d}_i$ is created,

$$\hat{c}_i = P_1(c_i), P_2(c_2), ..., P_T(c_i) \quad (10)$$

where the base models were $\{P_1, P_2, ..., PT\}$ and $\hat{c}_i$ are sample detections. Finally, the hybrid-based meta-model combines the base models trained on the resulting dataset in the third stage. The final ensemble detection for c is $\hat{P}$ $\{P_1(c), P_2(c), ..., P_T(c)\}$, where $\hat{P}$ represents the meta model.

Results and discussion

Dataset description

The dataset[18] was collected from Bangladeshis across various age groups and sociodemographic characteristics. To generate a survey for the dataset, 55 questions were designed. The purpose of the first 30 questions was to collect detailed psychosocial and demographic information about the individuals. And the final 25 questions were taken from the Burns Depression Checklist (BDC). The survey was conducted between April and August 2020. The responses of 604 participants were included in the dataset.

The cumulative severity of all the symptoms reported by a person was used to calculate the overall BDC score. If a person's BDC total score was more than 10, they were considered sad. Otherwise, it was assumed that the participant did not have depression. Based on their BDC scores, each study participant was categorized as depressed or non-depressed. The survey dataset consisted of 30 predictor variables and one target variable. Applying the BDC to all participants yielded the target variable. Table 1displays every parameter's potential value, parameter type, and parameter explanation.[1]

Evaluation metrics

In ML, accuracy is determined by dividing the total number of predictions by the number of accurate predictions made from the model. True positive (TP) and true negative (TN) are represented by a few abbreviations in equations (11) through (15), which demonstrate the accuracy of each class detection. On the other hand, there can be an alternative situation where the detection is accurate conceptually but inaccurate practically. These two circumstances are referred to as false positive (FP) and false negative (FN) scenarios. This provides a clear indication of the overall accuracy of the model, but it may not be appropriate for imbalanced datasets, where other metrics such as precision and recall are more useful.[35].

$$\text{Accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \times 100 \quad (11)$$

Precision Precision is an ML metric that calculates the proportion of correctly identified instances- or true-positive predictions- of all the positive predictions generated by the model. This ratio provides information on the model's ability to accurately identify positive examples while reducing false positives.

$$\text{Precision} = \frac{TP}{TP + FP} \times 100 \quad (12)$$

Recall: Recall, sometimes referred to as sensitivity or the rate of true positive predictions, is an ML metric that measures the percentage of correctly detected cases, or true positive predictions, among all of the dataset's real positive examples. It evaluates a model's capacity to detect and not overlook pertinent cases, making it especially important in situations where false negatives can be expensive or detrimental.

$$\text{Recall} = \frac{TP}{TP + FN} \times 100 \quad (13)$$

F1-Score: An ML metric called F1 score strikes a compromise between recall and precision by combining them into a single number. This is especially useful when trying to find a trade-off between accurately predicting positive cases and documenting all positive cases, as it accounts for both false positives and negatives. The harmonic mean of the recall and precision was used to quantify it.

$$\text{F1 Score} = \frac{2 \cdot \text{Precision} \cdot \text{Recall}}{\text{Precision} + \text{Recall}} \times 100 \quad (14)$$

Evaluating prediction models sometimes requires more than just accuracy and F1-score. Two further statistics were employed in the evaluation process: the receiver operating characteristics curve (ROC). The area under the curve (AUC) was computed using the ROC curve.[36]. The ROC was obtained by plotting the true positive rate (TPR) against the false positive rate (FPR). The recall is the only component of the TPR, and the FPR is calculated using the following equation.

$$\text{FPR} = \frac{FP}{FP + TN} \quad (15)$$

[1] https://github.com/Sabab31/Depression-Repository

Classifier	Accuracy %	Log Loss %	AUC %
KNN	75.47	8.84	86.00
XGB	89.31	3.85	89.00
RFC	89.31	3.85	94.00
SVM	78.62	7.71	96.00
MLP	92.45	2.72	97.00
Ensemble of DeprMVM	93.08	2.49	97.00

Table 3. Applying only SMOTE the accuracy, loss and AUC score of all evaluated methods.

Classifier	Accuracy %	Log Loss %	AUC %
KNN	86.18	6.57	89.00
XGB	94.67	3.85	97.00
RFC	94.67	3.85	96.00
SVM	83.33	7.71	94.00
MLP	97.33	2.95	97.00
Ensemble of DeprMVM (Proposed)	99.39	2.04	98.00

Table 4. Applying SMOTE and cluster sampling together, the accuracy, loss and AUC score of all evaluated methods.

The effectiveness of the proposed model in detecting depression is evaluated in this section. The two possible outcomes are as follows: I) preprocessed data using SMOTE II) combined with preprocessed data using SMOTE and Cluster Sampling.

Implementation results

Scikit-learn library and Python were used for the implementation of this study. The dataset contained worth of data of 604 participants. Thirty predictor variables and one target variable comprised this dataset. The target variable indicated whether the participant was depressed. The dataset was loaded into the proposed system after preparation. This section assesses how our proposed model improves the ability to detect depression.

The scenarios are as follows: I) Preprocessed data with SMOTE II) Balanced data with SMOTE and with Cluster Sampling combined.

Experiment with only SMOTE applied

The issue with this dataset is- that it contains imbalanced data. 65.7% Of the data showed depression, whereas 34.3% did not. SMOTE was used to address the issue of the class imbalance in the training datasets owing to its extreme imbalance. The percentages of individuals in the training datasets who were depressed and those who were not, both before and after SMOTE, are displayed in Figure 6.

Table 2 presents the training outcomes after balancing the dataset. Table 2 shows a significant improvement in the outcomes after using SMOTE and cluster sampling. The proposed ensemble model achieved a higher accuracy when we applied it. The ensemble of the DeprMVM model achieved a precision of 93.67%, which was the best accuracy when only SMOTE was applied. In addition, the hybrid model performed well on the balanced datasets. Table 3 describes the accuracy, loss, and AUC scores of all evaluated methods when applying only SMOTE. Here also, we can see that the proposed ensemble model achieved an accuracy of 93.08% and AUC score of 97%. Now we can make a difference by applying SMOTE and cluster sampling together.

Experiment with the combination of smote and cluster sampling

Table 2 shows that following cluster sampling and SMOTE together, the ensemble model maintained the best result, with a precision of 99.64%, which is promising in this research. In addition, the MLP has the greatest precision at 99.12%. Since the new dataset contains actual data, we can state with certainty that ensemble meta model and the hybrid DeprMVM perform remarkably well on this balanced dataset. Additionally, we can observe that our model's performance has improved when combining cluster sampling and smote together. Table 4 described the accuracy, loss, and AUC scores of all methods applying SMOTE and cluster sampling together. Here also the proposed ensemble of DeprMVM method has achieved remarkable results with an accuracy of 99.39% and a lower loss of 2.04%. From these two tables above, we can mention that the proposed ensemble architecture applying cluster sampling and SMOTE is a clear contribution in this research.

Classifiers performance matrix

This information can be found in the confusion matrix for each classifier. The model's classifier performance is displayed in a table-like manner. In either case, it shows the detection outcomes using test set data. The predicted

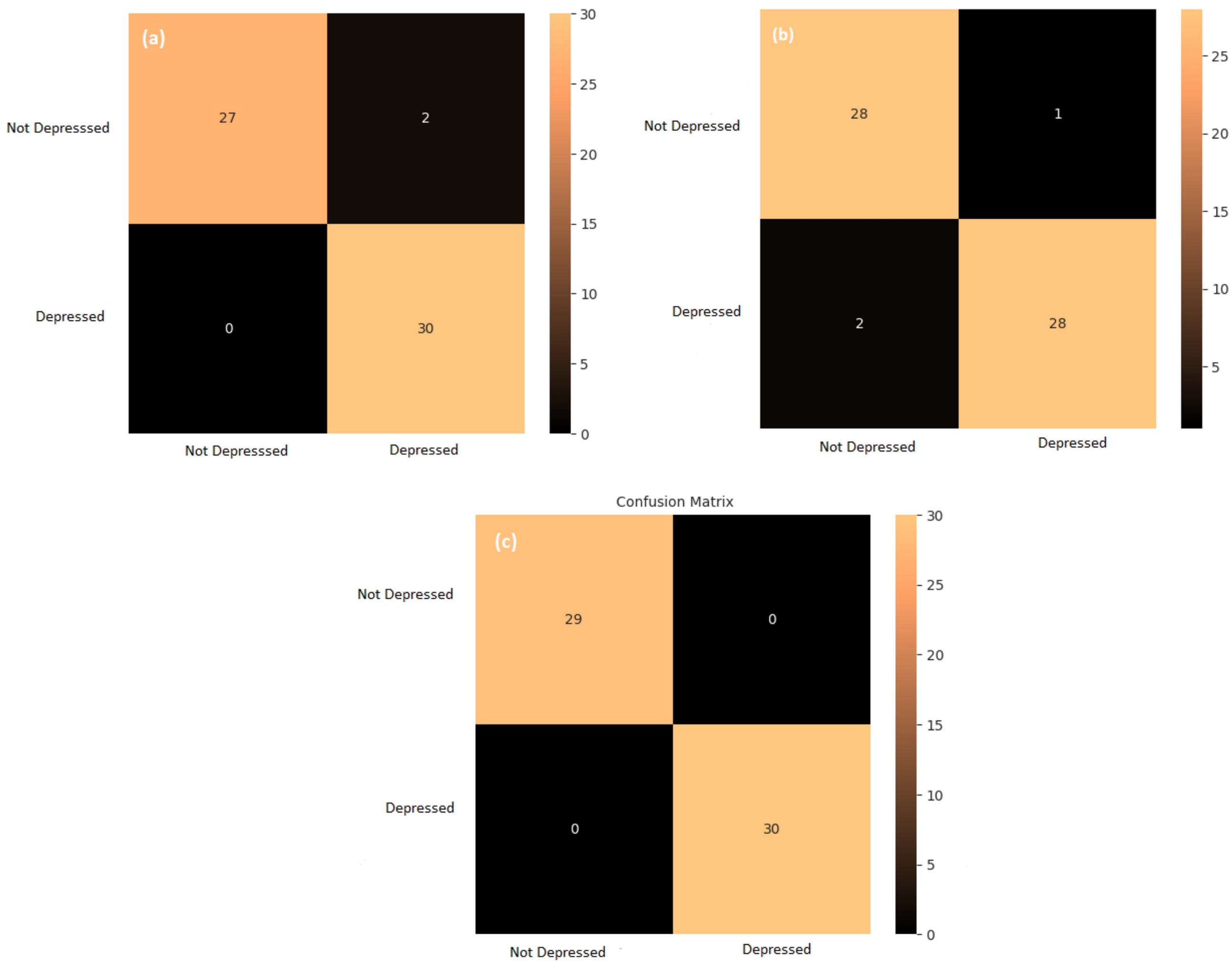

Fig. 7. Confusion Matrix of best three models applying SMOTE and cluster sampling together, (**a**) RFC, (**b**) MLP, and (**c**) Proposed Ensemble of hybrid model.

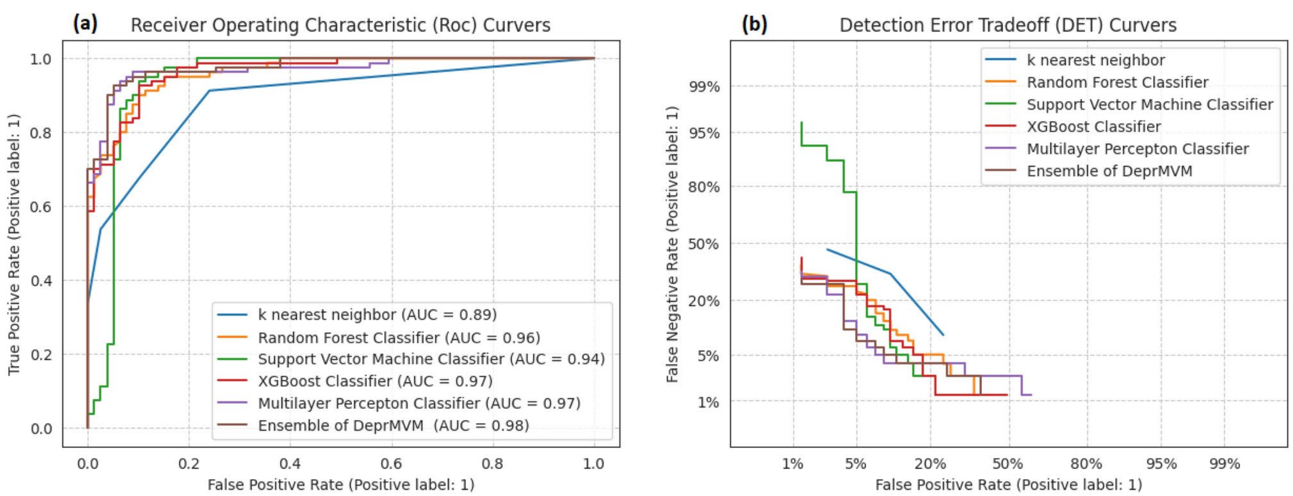

Fig. 8. The (**a**) ROC and (**b**) DET curve applying cluster sampling and SMOTE together are displayed.

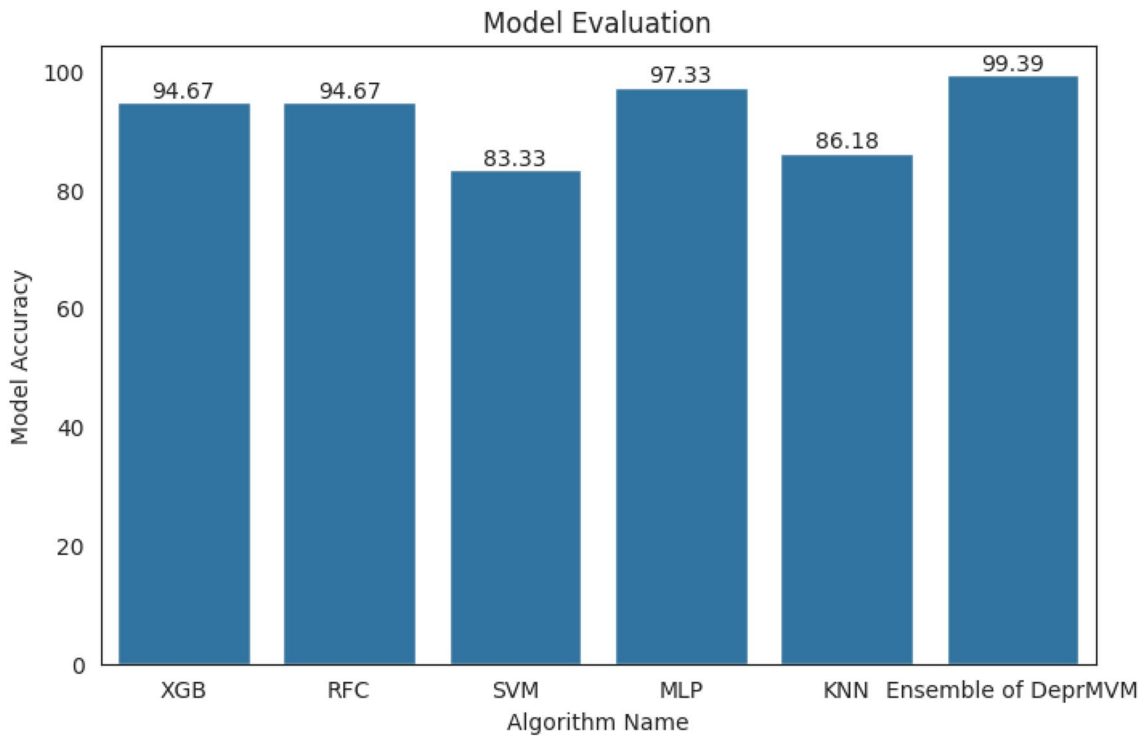

Fig. 9. Evaluation of best model applying SMOTE and cluster sampling together.

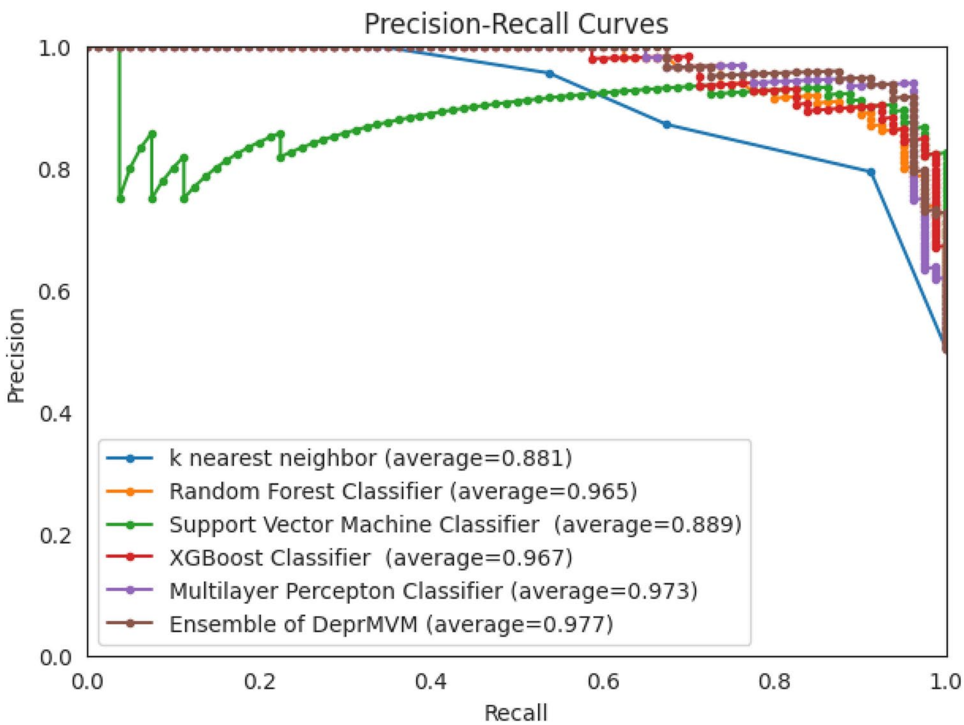

Fig. 10. The precision-recall Curves of all models applying SMOTE and cluster sampling together.

Works Done	Number of participants	Types of gathered information	discovered what causes depression in initial situation	Best Model	Accuracy (%)
Rois et al.[10]	355	Socio-demographic	No	Random Forest	89.3%
Munir et al.[37]	520	Socio-demographic, Job-Seeking stress-related Questions, and (PHQ-9)	Yes	Stacking classifier	77.45%
Priya et al.[14]	348	Socio-demographic, occupational information	Yes	Random Forest	85.5%
Nayan et al.[12]	2121	Socio-demograpchic	Yes	Random Forest	89%
AlSagri et al.[16]	111	Twitter posts, comments	Not applicable	SVM	82.5%
Na et al.[17]	6588	Socio-demographic, economic, clinical information	Yes	Random Forest	86.20%
Sau et al.[19]	470	Socio-demographic, occupational information	Yes	CatBoost	89.30%
Zulfiker et al.[18]	604	Socio-demographic and psychosocial information	Yes	AdaBoost	92.56%
Proposed work	604	Socio-demographic, and psychosocial information	Yes	Ensemble of DeprMVM	99.39%

Table 5. Comparing the results of our work with that of others.

outcome for each class is displayed in the columns, whereas the actual class is displayed in the rows. On the other hand, the matrix's inclination indicates the number of correctly detected photos.

The confusion matrix for the four best models used in this architecture is shown in Figure 7. The confusion matrix generated by the ensemble of the hybrid model is shown in Figure 7(c). The confusion matrix shows that the proposed model performed extremely well using cluster sampling and SMOTE. As a result, the research we propose creates a high level of categorization accuracy. Comparatively, MLP and RFC outperformed all ML classifiers.

Classifier parameter results

In the previous section, we mentioned that the ensemble model outperformed the other models in this study. When we applied SMOTE and cluster sampling together, the results outperformed those of previous studies. Figure 9 shows that the proposed ensemble of DeprMVM and MLP models achieves the highest accuracy in detecting depression.

ROC Curve: A model's classification efficiency is graphically represented by the ROC curve. Plotting the ROC curve involves dividing the true positive rate (TPR) by the false positive rate (FPR). Figure 8 (a) shows that the ensemble of DeprMVM, MLP, and XGB was the highest-performing model in our evaluation. Compared with the other ML models employed in these models, these were better at detecting depression. In addition, Figure 8 (b) shows the detection error tradeoff (DET) curves for all evaluated models.

figure 10 shows a precision-recall curve after applying cluster sampling and SMOTE. In addition, the proposed ensemble model achieved an average of higher precision. In addition, MLP achieved the greatest precision. From the above scenario, we can see that the proposed ensemble of the DeprMVM model is a significant contribution to this study. To increase the importance of this research, we need to compare it with a few recent publications.

Comparative study and findings

Comparing the findings and impact of this research with those of other published articles is essential for appreciating its excellence. We conduct a comparative analysis in which the most significant psycho-social and socio-demographic factors contributing to depression have been identified. However, this study aimed to detect depression in individuals from various age groups, occupations, and socioeconomic backgrounds. The most important psychosocial and sociodemographic factors that contributed to depression were identified in this study. Table 5 presents a comparative analysis of this study with previous studies. Here we can see that using a novel strategy of the ensemble model for generating the new dataset, produced a higher accuracy of 99.39%, which is remarkable in this field.

Discussion

Promoting remission, avoiding recurrence, and lessening the emotional burden of the illness all depend on early identification and treatment of depression. The majority of current diagnoses are arbitrary, inconsistent between medical professionals, and costly for those who could need assistance immediately. This study presents a novel method for automated questionnaire-based instantaneous depression identification using a revolutionary SVM and neural network-based ensemble strategy.

Interpretation of the main findings

For testing purposes in this study, 91 individuals' replies were utilized. Of these, 34.03% of participants were not depressed, and 65.67% were depressed from the beginning. Numerous criteria, including model precision, recall, specificity, accuracy, and F1 score, were already demonstrated in this study. Furthermore, these models' AUC values have been computed, and Figure 7 shows the confusion matrix of several applied models. This study uses a variety of data manipulation methods, including cluster sampling and SMOTE. After using SMOTE as a data manipulation approach, Table 2 displays the specificity, recall, precision, and f1-score as 93.67%, 92.50%, and 93.67%, respectively. However, our suggested ensemble DeprMVM model's specificity, recall, accuracy, and f1-score all significantly improved when we used both SMOTE and cluster sampling strategies; they are now

99.48%, 99.26%, 99.64%, and 99.51%, respectively. Furthermore, we observed a noteworthy improvement in accuracy when we only used SMOTE for our suggested model, with an overall 92.50%, accuracy of 93.08%, and an AUC of 97.00%, as shown in Table 3. Table 4 shows that there was a significant improvement in the model accuracy and AUC values, which were 99.39% and 98.00% when using both SMOTE and cluster sampling strategies. By comparing the outcomes of several models, we can determine that our suggested ensemble of the DeprMVM model, which uses the SMOTE and cluster sampling approach, performs better than the rest of the models in terms of accuracy, AUC values, and other performance metrics that are assessed. Figure 9 illustrates the greatest accuracy of 99.39% and the maximum AUC value of 97.00% produced by this ensemble of the DeprMVM model. In addition, Table 5 presents a comparative analysis of our study findings with those of previous studies conducted using the same dataset.

Implication
There are several mental health issues that have become increasingly frequent among the general public in today's fast-paced modern world. Some of these illnesses include stress, anxiety, and depression. However, when healthcare professionals manually identify depression based on patient background information, it may be challenging to provide assistance to depressed individuals in the healthcare sector. For the purpose of detective modelling in the field of health research, machine learning technology is the most cutting-edge technological advancement. Therefore, the fundamental objective of this study has been to develop an automated system that can determine whether or not a person is depressed by considering their demographic and psychological characteristics in an instant. This article examines and contrasts a variety of ML and ANN classifiers, including KNN, XGB, SVM, RF, MLP, and a hybrid model, namely DeprMVM, with the purpose of assisting in the identification of persons who may be suffering from mental health issues, such as depression. Taking into account a number of performance criteria, the model that we have proposed, DeprMVM, performs very well, having a detective accuracy of 99.39%. Finally, we can assert that this will be of great assistance to researchers and healthcare professionals in the prompt diagnosis of depression problems in clinical trials that are conducted in the real world, as well as in the provision of helpful advice for patients who are suffering from depression.

Limitations
However, most studies face particular challenges, and our current research is not free from limitations that should be carefully considered when interpreting data. One initial constraint is to treat the BDC as an absolute standard for diagnosing depression. Furthermore, biological variables that are associated with depression were not taken into consideration, and the quantity of the dataset that was used in this investigation was inadequate.

Future work
In the future, we have a plan to consider several biological factors and apply different dimensionality reduction algorithms during the data pre-processing stage on a large dataset. This enhanced the performance of our models and allowed us to compare their performance with the results obtained in the current study.

Conclusion and future work
Depression is a common mental disorder that can be associated with a variety of causes and is prevalent throughout the world. However, depending on the rating scale, it can be extremely challenging and time-consuming for humans to manually perform this task to ensure immediate care. To ensure rapid treatment, we developed an automated computer-based detection system that can identify key factors responsible for causing depression. This will assist researchers and medical practitioners in immediately diagnosing depression concerns in real-world clinical studies and provide valuable recommendations for patients with depression. The dataset used in this study contained sociodemographic and psychological variables to identify cases of depression. To effectively identify depression, this study involved several data preprocessing procedures. These steps include data labeling and the utilization of feature selection techniques to derive the most significant demographic and psychosocial factors that cause depression. Numerous ML and ANN classification techniques, including RF, KNN, SVM, XGB, MLP, a hybrid model called DeprMVM, and an ensemble method, were utilized on this processed dataset. SVM and MLP are the two primary methods used to construct the proposed ensemble approach. The hybrid DeprMVM has been utilized as a meta-learner. While the hybrid DeprMVM is a level-1 learner, the SVM and MLP networks are level-0 learners. Both independent and dependent factors are present in the new data set that was used to train the meta-classifier. These variables are considered in the classifiers' outputs after they have been trained and tested at level-0. Among the applied models, an automated system was developed based on the outperforming model. Experimental investigations have determined that the proposed ensemble of the DeprMVM model achieved the highest accuracy of 99.39%, recall of 99.26%, precision of 99.64%, AUC of 98.00%, and F1-score of 99.51% by applying SMOTE and cluster sampling together. In the future, we plan to pre-process the data by utilizing different dimensionality reduction algorithms and popular explainable artificial intelligence (XAI) techniques and consideration of several biological elements of a large dataset. These results provide a strong case for the importance and relevance of this research in the area of public health.

Data availability
Data is available in a publicly accessible link: https://github.com/Sabab31/Depression-Repository.

Received: 15 August 2024; Accepted: 21 October 2024
Published online: 26 October 2024

References

1. Guo, L.-N. et al. Psychometric assessment of the depressive cognition scale among older chinese people. *Archives of Psychiatric Nursing* **31**, 477–482 (2017).
2. Le, H.-N. & Boyd, R. C. Prevention of major depression: Early detection and early intervention in the general population. *Clin Neuropsychiatry* **3**, 6–22 (2006).
3. Depression, W. Other common mental disorders: global health estimates. *Geneva: World Health Organization* **24** (2017).
4. Aladağ, A. E., Muderrisoglu, S., Akbas, N. B., Zahmacioglu, O. & Bingol, H. O. Detecting suicidal ideation on forums: proof-of-concept study. *Journal of medical Internet research* **20**, e9840 (2018).
5. Wang, S., Peng, G., Zheng, Z. & Xu, Z. Capturing emotion distribution for multimedia emotion tagging. *IEEE Transactions on Affective Computing* **12**, 821–831 (2019).
6. Robles-Palazón, F. J. et al. Predicting injury risk using machine learning in male youth soccer players. *Chaos, Solitons & Fractals* **167**, 113079 (2023).
7. Nemesure, M. D., Heinz, M. V., Huang, R. & Jacobson, N. C. Predictive modeling of depression and anxiety using electronic health records and a novel machine learning approach with artificial intelligence. *Scientific reports* **11**, 1980 (2021).
8. Chung, J. & Teo, J. Single classifier vs. ensemble machine learning approaches for mental health prediction. *Brain informatics* **10**, 1 (2023).
9. Susanty, S., Sufriyana, H., Su, E.C.-Y. & Chuang, Y.-H. Questionnaire-free machine-learning method to predict depressive symptoms among community-dwelling older adults. *PLoS One* **18**, e0280330 (2023).
10. Rois, R., Ray, M., Rahman, A. & Roy, S. K. Prevalence and predicting factors of perceived stress among bangladeshi university students using machine learning algorithms. *Journal of Health, Population and Nutrition* **40**, 1–12 (2021).
11. Xu, Z. et al. A cluster-based oversampling algorithm combining smote and k-means for imbalanced medical data. *Information Sciences* **572**, 574–589 (2021).
12. Nayan, M. I. H. et al. Comparison of the performance of machine learning-based algorithms for predicting depression and anxiety among university students in bangladesh: A result of the first wave of the covid-19 pandemic. *Asian Journal of Social Health and Behavior* **5**, 75–84 (2022).
13. Ali, M. M. et al. Development and performance analysis of machine learning methods for predicting depression among menopausal women. *Healthcare Analytics* **3**, 100202 (2023).
14. Priya, A., Garg, S. & Tigga, N. P. Predicting anxiety, depression and stress in modern life using machine learning algorithms. *Procedia Computer Science* **167**, 1258–1267 (2020).
15. Xin, Y. & Ren, X. Predicting depression among rural and urban disabled elderly in china using a random forest classifier. *BMC psychiatry* **22**, 118 (2022).
16. AlSagri, H. S. & Ykhlef, M. Machine learning-based approach for depression detection in twitter using content and activity features. *IEICE Transactions on Information and Systems* **103**, 1825–1832 (2020).
17. Na, K.-S., Cho, S.-E., Geem, Z. W. & Kim, Y.-K. Predicting future onset of depression among community dwelling adults in the republic of korea using a machine learning algorithm. *Neuroscience Letters* **721**, 134804 (2020).
18. Zulfiker, M. S., Kabir, N., Biswas, A. A., Nazneen, T. & Uddin, M. S. An in-depth analysis of machine learning approaches to predict depression. *Current research in behavioral sciences* **2**, 100044 (2021).
19. Sau, A. & Bhakta, I. Screening of anxiety and depression among seafarers using machine learning technology. *Informatics in Medicine Unlocked* **16**, 100228 (2019).
20. Kour, H. & Gupta, M. K. An hybrid deep learning approach for depression prediction from user tweets using feature-rich cnn and bi-directional lstm. *Multimedia Tools and Applications* **81**, 23649–23685 (2022).
21. Uddin, M. Z., Dysthe, K. K., Følstad, A. & Brandtzaeg, P. B. Deep learning for prediction of depressive symptoms in a large textual dataset. *Neural Computing and Applications* **34**, 721–744 (2022).
22. Amanat, A. et al. Deep learning for depression detection from textual data. *Electronics* **11**, 676 (2022).
23. Ghosh, T. & Kaiser, M. S. Bangla depressive social media text detection using hybrid deep learning approach. In *Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering: TCCE 2021*, 111–120 (Springer, 2022).
24. Marriwala, N., Chaudhary, D. *et al.* A hybrid model for depression detection using deep learning. *Measurement: Sensors* **25**, 100587 (2023).
25. Kumar, P., Garg, S. & Garg, A. Assessment of anxiety, depression and stress using machine learning models. *Procedia Computer Science* **171**, 1989–1998 (2020).
26. Siddiqua, R., Islam, N., Bolaka, J. F., Khan, R. & Momen, S. Aida: Artificial intelligence based depression assessment applied to bangladeshi students. *Array* **18**, 100291 (2023).
27. Kumari, R., Singh, J. & Gosain, A. Sms: Smote-stacked hybrid model for diagnosis of polycystic ovary syndrome using feature selection method. *Expert Systems with Applications* **225**, 120102 (2023).
28. Kurani, A., Doshi, P., Vakharia, A. & Shah, M. A comprehensive comparative study of artificial neural network (ann) and support vector machines (svm) on stock forecasting. *Annals of Data Science* **10**, 183–208 (2023).
29. Han, K.-X., Chien, W., Chiu, C.-C. & Cheng, Y.-T. Application of support vector machine (svm) in the sentiment analysis of twitter dataset. *Applied Sciences* **10**, 1125 (2020).
30. Wang, M., Zhou, J., Cai, H. & Gong, M. Crowdmlp: Weakly-supervised crowd counting via multi-granularity mlp. *Pattern Recognition* **144**, 109830 (2023).
31. Maqbool, J., Aggarwal, P., Kaur, R., Mittal, A. & Ganaie, I. A. Stock prediction by integrating sentiment scores of financial news and mlp-regressor: a machine learning approach. *Procedia Computer Science* **218**, 1067–1078 (2023).
32. Panja, M. et al. An ensemble neural network approach to forecast dengue outbreak based on climatic condition. *Chaos, Solitons & Fractals* **167**, 113124 (2023).
33. Mienye, I. D. & Sun, Y. A deep learning ensemble with data resampling for credit card fraud detection. *IEEE Access* **11**, 30628–30638 (2023).
34. Liang, M. et al. A stacking ensemble learning framework for genomic prediction. *Frontiers in genetics* **12**, 600040 (2021).
35. Saha, D. K., Joy, A. M. & Majumder, A. Yotransvit: A transformer and cnn method for predicting and classifying skin diseases using segmentation techniques. *Informatics in Medicine Unlocked* 101495 (2024).
36. Saha, D. K. An extensive investigation of convolutional neural network designs for the diagnosis of lumpy skin disease in dairy cows. *Heliyon* (2024).
37. Munir, U. B., Kaiser, M. S., Islam, U. I. & Siddiqui, F. H. Machine learning classification algorithms for predicting depression among university students in bangladesh. In *Proceedings of the Third International Conference on Trends in Computational and Cognitive Engineering: TCCE 2021*, 69–80 (Springer, 2022).

Acknowledgements

The authors extend their appreciation to King Saud University for funding this research through Researchers Supporting Project Number (RSPD2024R1027), King Saud University, Riyadh, Saudi Arabia.

Author contributions

Methodology, D.K.S.; Software, D.K.S. and T.H; Validation, D.K.S., T.H., S.A.; Formal analysis, D.K.S.and T.H.; Investigation, D.K.S., T.H.; Resources, D.K.S., T.H.; Writing - original draft, D.K.S, T.H.; Writing - review & editing, D.K.S, T.H., M.F.M. and S.A.; Visualization, D.K.S, T.H.; Supervision, M.F.M. and D.C.; Project administration, M.S. and S.A.; Funding acquisition, M.S. and S.A. All authors have read and agreed to the published version of the manuscript.

Funding

This research was supported by the Researchers Supporting Project Number (RSPD2024R1027), King Saud University, Riyadh, Saudi Arabia.

Declarations

Competing interests

The authors declare no competing interests.

Additional information

Correspondence and requests for materials should be addressed to M.S. or M.F.M.