INDIAN JOURNAL OF SCIENCE AND TECHNOLOGY

RESEARCH ARTICLE

Depression Prediction Model based on Ensemble Learning Classifier

Pratiksha Deshmukh[1]*, **Harshali Patil**[2]

1 Assistant Professor, Computer Engineering Department, Thakur College of Engineering and Technology, Mumbai, India
2 Professor, Computer Engineering Department, Thakur College of Engineering and Technology, Mumbai, India

OPEN ACCESS

Received: 22-01-2024

Accepted: 25-04-2024

Published: 16-10-2024

Citation: Deshmukh P, Patil H (2024) Depression Prediction Model based on Ensemble Learning Classifier. Indian Journal of Science and Technology 17(39): 4084-4093. https://doi.org/10.17485/IJST/v17i39.159

* **Corresponding author**.

pratiksha.deshmukh@thakureducation.org

Funding: None

Competing Interests: None

Published By Indian Society for Education and Environment (iSee)

ISSN
Print: 0974-6846
Electronic: 0974-5645

Abstract

Objective: The main objective of this research was to develop a suitable prediction model to classify the symptoms of depression experienced by people. **Methodology:** This research incorporates the dataset of the "Centres for Disease Control and Prevention National Health and Nutrition Examination Survey," which was available on GitHub. After that, pre-processing of the dataset was done using the infinite latent feature selection (ILFS) algorithm to extract the appropriate features from the dataset. After that, the dataset was split into 70:30 ratios. About 70% of the data is employed for training the machine learning algorithm, while the remaining 30% is used for testing. Further, for classification purposes, multiple machine learning algorithms were taken into consideration using the ensemble learning classifier, and the final prediction was made based on the voting classifier. The voting classifier helps determine which machine learning algorithm has superior performance over other machine learning algorithms. Finally, the proposed approach is evaluated via simulation on the dataset, with a number of metrics for performance being obtained. **Finding:** The result demonstrates that the suggested model achieves a high-value accuracy of 0.9166, a prediction value of 0.9177, a sensitivity value of 0.9984, a specificity value of 0.9984, and an F1-Score value of 0.9564. Finally, the comparison shows that the proposed mode outperforms machine learning algorithms, namely, ANN, XGBoost, Adaboost, random forest, stochastic gradient boosting, and SVM. **Novelty:** The proposed model was classified using various machine learning algorithms in place of one. Therefore, the accuracy of the proposed model was high.

Keywords: Anxiety; Classification; Depression; Ensemble Learning; ILFS; Linear Regression; ML; Prediction; Stress

1 Introduction

1.1 Motivation

In the present scenario, the modern lifestyle of the people has a psychological impact on their minds and leads to emotional distress and depression [1].

Depression is a widespread mental disorder that limits a person's capacity for new cognitive growth and thought. Over 300 million individuals worldwide experience depression [2], while according to the WHO, mental health diseases impact one billion individuals worldwide [3]. Suicidal ideation is mostly a symptom of depression. Annual suicide rates are estimated at over 800,000. Therefore, a comprehensive response is necessary to address the growing number of mental health problems [4,5]. An individual's socioeconomic position may cause them to suffer from depression. People with depression are more hesitant to engage in social activities. Depression may be treated with counseling and other psychological treatments.

1.2 Machine Learning for Depression Classification

In this first stage, conventional methods of clinical diagnosis are used. These treatments for depression do not adequately capture depression's nuanced nature. Therefore, at the current moment, it is simple to identify and forecast the composition of symptoms connected to mental diseases such as depression by using machine learning techniques. Machine learning (ML) is an approach to computer science that tries to teach computers to learn how to recognize complicated patterns without any human input. This skill is useful for solving new challenges by drawing on past experience. Most ML algorithms are designed to provide predictable results [6,7]. Algorithms in machine learning (ML) may be broken down into many broad categories which include unsupervised learning, semi-supervised learning, supervised learning, and reinforcement learning. Unsupervised ML techniques [8] reveal hidden patterns and clusters in the provided data, whereas supervised machine learning algorithms [9] use primary inputs to forecast recognized values. Between unsupervised and supervised learning is semi-supervised learning [10], which focuses on understanding system behavior using a combination of unlabeled and labeled data. One of the main focuses of reinforcement learning [11] is learning how to respond in response to feedback about how well those actions went. In the literature, machine learning for depression classification is done using the following way.

Initially, the standard dataset of depression is read and pre-processing is done on it to fill in the null attributes/select the appropriate features from it. Further, machine learning algorithm performance depends on the dataset, feature selection pre-processing technique, and training and testing size of the dataset. These considerations are included in the development of depression prediction models in this study. In the literature, filtering and wrapping are the two types of feature selection [12,13]. The filtering-based approach selects the subset of the dataset independently, whereas the wrapper-based approach uses a recursive approach to select the subset. In this research, the wrapper-based approach is taken into consideration. Further, the machine learning algorithm is utilized for classification purposes. In the literature, numerous researchers have designed machine learning-based depression prediction models [14–19]. In the literature, decision trees [20,21], Naïve Bayes [22], random forests [23], support vector machines, KNN, logistic regression, and XGBoost are successfully used for depression models. The decision tree is very sensitive to input data. On the other hand, random forest is slow due to very few branches. Further, SVM is superior for small datasets, and its performance is degraded for large datasets. Besides that, in the previous work, no hybridization of the machine learning algorithm is done. Finally, the performance evaluation of the depression model is done using the various performance metrics.

1.3 Main Contribution of the Paper

The main contribution of this study is to design a depression prediction model based on algorithms of machine learning to classify whether the user is depressed or not. The key highlights of this research are as follows:

- In this research, a feature selection method is designed using the infinite latent feature selection (ILFS) algorithm to extract the appropriate features from the dataset. The benefit of this algorithm is that it comes under the wrapper method of feature selection. Therefore, it uses the recursive approach to select the appropriate features from the dataset.
- In this research, an ensemble learning-based classifier is designed using multiple machine learning algorithms to enhance the performance of the proposed method. Further, the benefit of an ensemble learning-based classifier is that it helps reduce the overfitting effect during prediction. In this work, we have used bagging, subspace, and ridge regression in the ensemble learning classifier. Finally, the best classifier is determined using the mode function.
- The simulation evaluation and comparative analysis are done depending on several performance metrics. As can be seen from the results, the suggested model outperforms conventional techniques of classification.

1.4 Organisation of the Paper

The rest part of the paper is structured as follows. Section 1 provides an overview introduction of the research topic. Section 2 shows the related work. This section explains the feature selection and machine learning algorithms. Section 3 explains the proposed depression prediction model. In Section 4, we see how the suggested model was tested in simulation using different measures of performance. Section 5 concludes with a discussion of implications and future directions.

2 Methodology

In order to understand the proposed method, in this section, a detailed description of the dataset, feature selection algorithm, and machine learning algorithms are given which are utilized to build an ensemble learning classifier for classification purposes.

2.1 Dataset:

In this research, a publicly available dataset of depression is taken into consideration. This dataset is available on Github[24]. A detailed description of the dataset is given below.

The dataset was prepared by examining US adults through the Centres for Disease Control and Prevention national health and nutrition examination survey. The dataset contains a total of 36259 entries, and this data was prepared by examining the 2005–2018 data.

2.2 Feature Selection using the ILFS Algorithm

The ILFS method has three fundamental steps. The method begins with a preprocessing phase, then moves on to a weighted graph, and finally, a ranking phase[13]. The pre-processing phase's objective is to quantitatively characterize the distribution of characteristics xi in a matrix representation. Then, determine the worth of a token t such that it suppresses characteristic xi. Discriminative quantization is the technical term for this procedure. In order to get a vector from a feature, the Fisher criteria approach is used. Graph weighting is the next stage. The goal of the weighting procedure is to build a graph at each node that links all of the features to one another. By applying the PLSA method, we can determine the weights based on the probabilities of feature and token co-occurrences. The Expectation Maximization (EM) method is used to fine-tune the weight towards the end. Specifically, the researcher lays forth the steps involved in the ILFS method.

- Discriminative Quantization Process: Before doing many-to-few mapping, it is necessary to generate a representation matrix that indicates how well the supplied feature reflects the class. As a first step, we may use Equation (1) to derive a modified Fisher criteria that is capable of dealing with the score vector in a multi-class context. Where K shows the total number of features, s depicts the size of the sample, σ is the standard deviation, μ is the mean, Z represents the normalization factor, and ϕ represents the Score vector.

$$\phi = \frac{1}{Z}\left[\frac{(s-\mu_1)^2}{\sum_{k=1}^{K}\sigma_k^2}\cdots\cdot\frac{(s-\mu_k)^2}{\sum_{k=1}^{K}\sigma_k^2}\right], \forall k \in K \tag{1}$$

The next stage is quantization, which determines the value of the discriminative feature in a certain "token" interval (lower values reflect less well-represented samples, and vice versa).

- From correlation to graph weighting Weights assigned to nodes in the graph depending on their relative significance, as determined by the training data. Probabilities P(token-feature) and P(factor feature) are determined under the condition that both token and factor have the same feature. In addition, an Expectation-Maximization (EM): maximum likelihood is carried out to fine-tune this parameter.
- Probabilistic Infinite Feature: The geometric series used to extend the route of the Selection Matrix acquired in the previous step into infinity is then computed. Gelfand's formula is used in this study[18].

$$C = ((1-rA))^{-1} - I \tag{2}$$

where A is the initial matrix produced in step 2, I is the matrix identity of A, and r is the result of step 3.

$$r = \frac{1}{max(eig(A))} \tag{3}$$

The route length's energy score is then ranked according to its correlation with the feature by adding up the matrix's dimensions.

2.3 Machine Learning Algorithms are Utilized for Ensemble Learning Classifier

In the ensemble learning classifier, three algorithms are used. Out of these algorithms, two algorithms are based on the bagging of a decision tree known as the ensemble algorithm whereas the third algorithm is based on linear regression.

2.3.1 Ensemble Algorithm

By integrating several models, ensemble learning may assist machine learning and provide better results. Better prediction performance may be achieved with this method than with a single model[25,26]. Stacking, bagging, and boosting are the major classes of ensemble learning algorithms. Bagging is a technique where numerous decision trees are trained on different subsamples of a dataset and their predictions are then aggregated. Additionally, stacking involves fitting many models to the same data set and then using a different approach to optimize prediction possibilities. Furthermore, boosting adds members to the ensemble one at a time to correct the forecasts of the previous models, ultimately producing a weighted average of the predictions.

2.3.2 Linear Regression Algorithm

An example of a supervised machine learning technique, linear regression estimates the linear connection between a target variable and a set of explanatory variables. When there is just one independent feature, we call it univariate linear regression; when there are many features, we call it multivariate linear regression. The purpose of the method is to identify the most accurate linear relationship between the dependent and independent variables from which a prediction can be made. Using this equation, we may see the connection between the independent and dependent variables as a straight line. According to[27], the slope of a line indicates the ratio of the change in the dependent variable to the change in the independent variables, represented as a proportion.

2.4 Performance Evaluation Metrics

Table 1 shows the performance metrics determined for the proposed model.

Table 1. Performance Metrics

Parameter	Equation
Accuracy	$\frac{TP+TN}{TP+FP+TN+FN}$
Precision	$\frac{TP}{TP+FP}$
Sensitivity/Recall	$\frac{TP}{TP+FN}$
Specificity	$\frac{TP}{TP+FN}$
F1-Score	$\frac{2*(Recall*Precision)}{(Recall+Precision)}$

Note: TP: True Positive, FP: False Positive, TN: True Negative, FN: False Negative

2.5 Proposed Depression Prediction Model

The purpose of the proposed model for depression classification is put forth in this section. Figure 1 depicts the suggested model flowchart. Next, we have explained how the proposed model works to predict the depression.

Initially, in the proposed model, the standard dataset is read. The dataset is publicly available on the GitHub website. After that, the dataset is pre-processed to determine the input and output attributes. Next, feature selection is performed using the infinite latent feature selection algorithm to select the most relevant features. Furthermore, the data set is partitioned into a 70:30 ratio. For training, we utilize 70% of the data, and for testing 30% is used. Next, the ensemble learning approach with a voting classifier is trained using the training dataset and then tested on the testing dataset. In the ensemble learning approach, multiple machine learning algorithms are utilized for predicting a single output. Further, in this research, a heterogeneous ensemble learning algorithm is taken into consideration over homogeneous. In this algorithm, multiple machine learning algorithm is trained and tested using the same dataset for classification purposes. Finally, the proposed model's performance is assessed over a number of performance indicators, and a comparison is made to the state-of-the-art in machine learning.

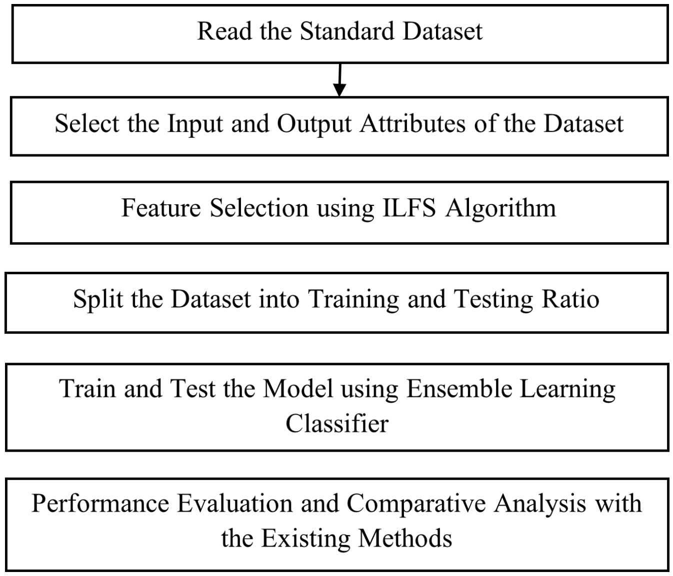

Fig 1. Flowchart of the Proposed Prediction Model

2.6 Statistical Analysis

A dataset which is used in this research, statistical analysis is carried out. On dataset, first, Cronbach's alpha is calculated to check the reliability and correlation, by doing feature selection. Table 2 shows the top 16 Parameters (features) selected for reliability testing with Split-Half Reliability and Test-Retest Reliability Test and Correlation values are calculated using Cronbach alpha in Table 3. Data profiling is carried out after that feature selection to prioritize features based on their correlation with the depression column using random forest and the accuracy achieved is 92.93%.

Table 2. Split-Half Reliabilityand Test-Retest Reliability Test

Parameters (Features) considered for Reliability Test	Split-Half Reliability values	Test-Retest Reliability values
age	-0.025771	-0.001364
lymphocyte_ percent	0.021257	-0.018396
mo0cyte_percent	-0.034257	-0.010538
white_BCC	0.001149	-0.000010
SEQN	0.002843	1.000000
platelet_count	0.001368	0.031570
Rx_days_CLONAZEPAM	-0.004613	-0.005108
asthma_onset	-0.026324	-0.001364
neutrophils_percent	0.024199	0.004310
Rx_days_ALPRAZOLAM	-0.006937	-0.007681
basophils_percent	0.027929	-0.008915
mo0cyte_count	0.003730	-0.012531
cocaine_use	-0.016621	-0.006290
time_in_current_job	-0.002145	0.026433
sleep_hours	0.048053	-0.007740
health_problem_Back or Neck	-0.021393	0.024773

Note: Features that have Split-Half Reliability and Test-Retest Reliability between 0 and 1.

Table 3. Correlation values using Cronbach alpha

Parameters (Features) considered for Inter Item Correlation Test	Inter Item Correlation Test using Cronbach alpha values
age	0.0376168220497651
Lumphocyte_percent	0.07781387309005054
mo0cyte_percent	0.06285074646870391
white_BCC	0.07148790893433908
SEQN	0.1062787300234007
platelet_count	0.07587632172354551
Rx_days_CLONAZEPAM	0.08384725214481026
asthma_onset	0
neutrophils_percent	0.09889543296087955
Rx_days_ALPRAZOLAM	0.07500043260200658
basophils_percent	0.049472959559578845
mo0cyte_count	0.08109774715545821
cocaine_use	0
time_in_current_job	0.06457940041285881
sleep_hours	0.0758238876431175
health_problem_Back or Neck	0

Three reliability tests that are Cronbach, split-half reliability, and test-retest reliability testing carried out, so further Statistical data is computed Minimum, Maximum, Mean (SD), Skewness, and Kurtosis on selected features based on priority.

3 Result and Discussion

The suggested model's simulation is compared using the various performance metrics. The prediction model was implemented and tested on MATLAB 2018a software with system configurations are i7 processor, 16GB RAM, 64-bit operating system, and Windows 10. Further, Table 5 shows the simulation setup configuration of the machine learning algorithms that are initialized for the design of the proposed depression prediction model.

Table 4. Minimum, Maximum, Mean (SD), Skewness, Kurtosis for selected features based on priority

Parameters (Features) considered for Statistical Test	Minimum	Maximum	Mean (SD)	Skewness	Kurtosis
age	18	85	44.51 (392.65)	0.371152	-1.006448
lymphocyte_ percent	0.0	86.6	28.83 (115.85)	-0.203112	1.531308
mo0cyte_percent	0.0	34.1	7.52 (8.02)	0.265651	4.62154
white_BCC	0.0	55.9	7.21 (8.24)	3.246755	52.799426
SEQN	31131	37129	34155.53 (3008380.55)	-0.039146	-1.197975
platelet_count	0	702	267.81 (7605.13)	-0.409216	2.358265
Rx_days_CLONAZEPAM	0	5475	5.19 (17349.42)	34.493751	1297.333274
asthma_onset	0	85	2.55 (89.91)	4.78895	25.076558
neutrophils_percent	0.0	90.2	56.25 (235.35)	-1.880843	5.024325
Rx_days_ALPRAZOLAM	0	7300	15.73 (75115.34)	22.922528	572.288175
basophils_percent	0.0	10.8	0.64 (0.33)	5.615104	70.198013
mo0cyte_count	0.0	2.7	0.54 (0.05)	0.799741	5.650843
cocaine_use	0	1	0.12 (0.10)	2.35855	3.562758
time_in_current_job	0	720	46.11 (8236.85)	3.20768	12.57133
sleep_hours	0	12	6.89 (2.21)	-0.237194	1.604742
health_problem_Back or Neck	0	1	0.10 (0.09)	2.754012	5.584583

Five performance metrics—specificity, precision, sensitivity, accuracy, and F1-score are utilized to assess the efficiency of the proposed model. To determine these performance metrics, initially, a confusion matrix is determined. Table 6 shows the obtained results of the proposed model. According to the findings, the suggested methodology can achieve high sensitivity (0.9984) and specificity (0.9984), as shown in Figure 2.

Table 5. Simulation Setup Configuration of the Proposed Depression Prediction Model

Parameter	Values
Alpha	0.82
Data Split Ratio	70:30

Table 6. Performance Metrics for the Proposed Model

Parameter	Proposed Model
Accuracy	0.9166
Precision	0.9177
Sensitivity	0.9984
Specificity	0.9984
F1-Score	0.9564

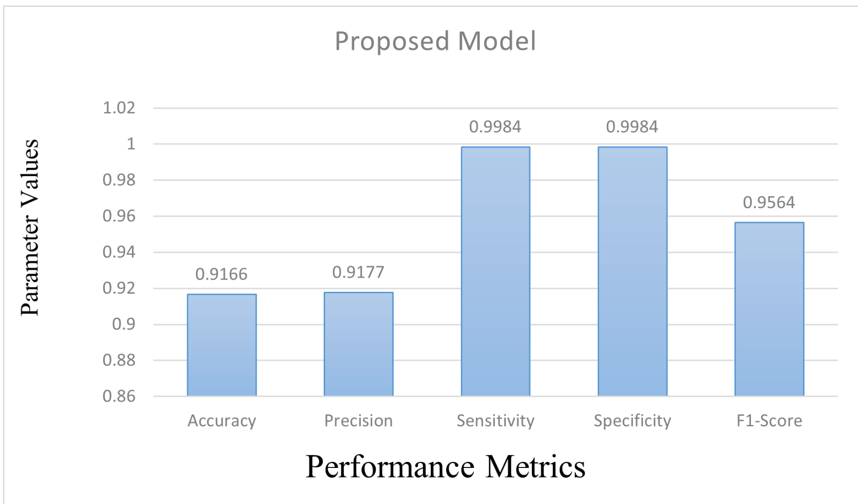

Fig 2. Evaluation of the Proposed Model Using Various Performance Metrics

Figure 3 represents the results of the comparison with the current models that use machine learning methods. The machine learning methods that are already in use are ANN, Adaboost, random forest, XGBoost, Stochastic Gradient Boosting, and SVM. Further, the performance metrics taken into consideration for comparison purposes are precision, accuracy, specificity, sensitivity, and F1-Score.

The result shows that the proposed model achieves the highest accuracy of 0.9166 over previous methods such as ANN of 0.706, random forest of 0.686, Adaboost of 0.673, and stochastic gradient boosting of 0.707, as shown in Figure 3**(a)**. Further, the result shows that the proposed model achieves a lower precision of 0.9177 over previous methods such as ANN of 0.961, random forest of 0.958, Adaboost of 0.956, and stochastic gradient boosting of 0.956, as shown in Figure 3**(b)**. Moreover, the result shows that the proposed model achieves the highest sensitivity of 0.9984 over previous methods such as ANN of 0.697, random forest of 0.676, Adaboost of 0.662, and stochastic gradient boosting of 0.701, as shown in Figure 3**(c)**. Finally, the proposed model also achieves the highest specificity and F1-score over the existing model, as shown in Figure 3 **(d–e)**.

3.1 Discussion

The deployment of the ILFS algorithm selects the appropriate features from the dataset. The model robustness is checked by training 70% data randomly and 30% data for testing purposes. The ensemble learning classifier helps to predict a single output by considering the multiple machine learning algorithms. In this research, the maximum voting method is utilized for the final classification.

By calculating statistical measures like Split-Half Reliability and Test-Retest Reliability Test and Correlation values using Cronbach alpha, it has been observed that values are between 0 and 1 and items are closely related in a group.

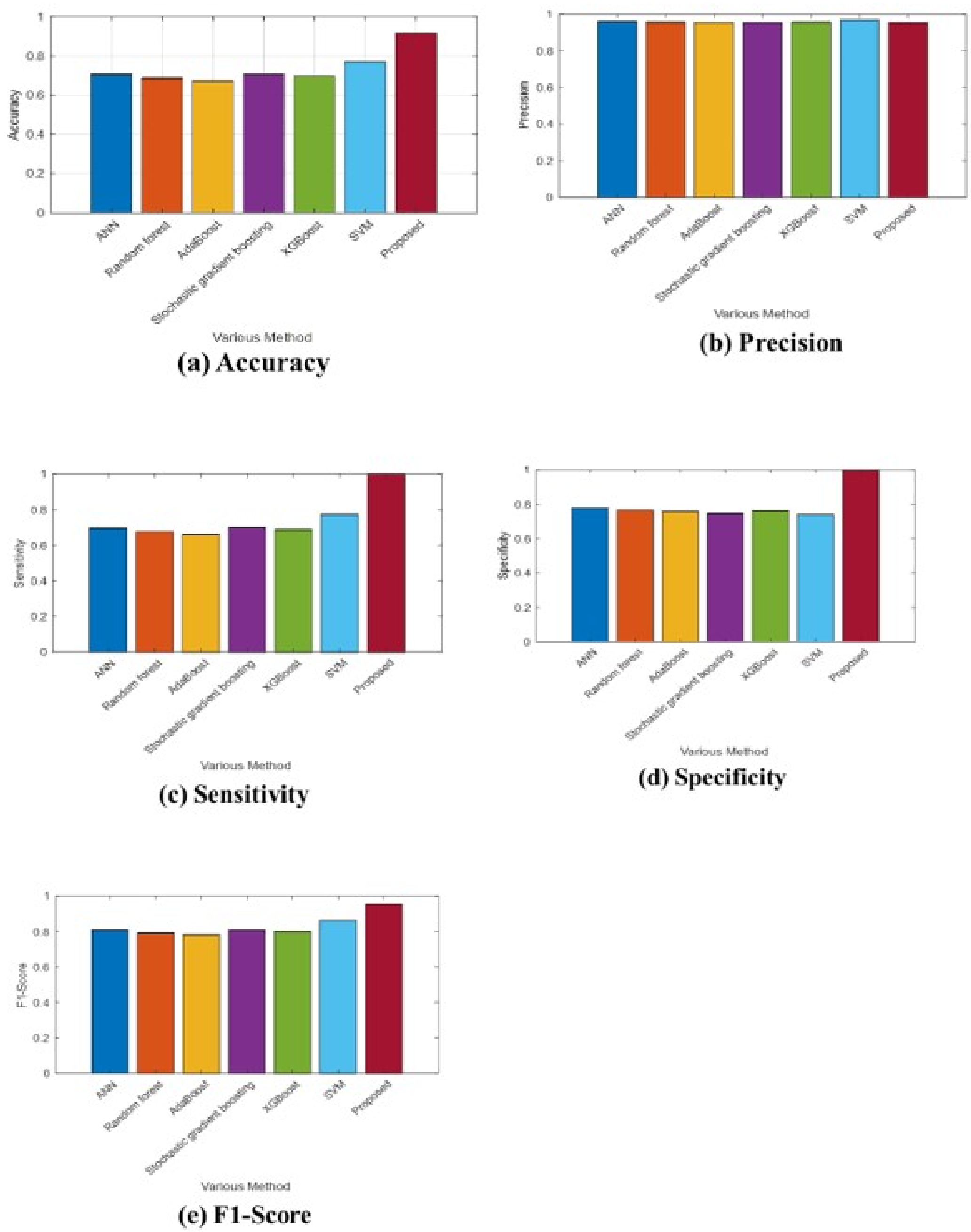

Fig 3. Comparative Analysis with the Existing Machine Learning Algorithms based on (a) Accuracy (b) Precision (c) Sensitivity (d) Specificity (e) F1-Score

4 Conclusion

This study presents a depression prediction model based on an ensemble learning classifier. In this model, initially, input and output attributes are determined, and appropriate features from the dataset are selected using the ILFS algorithm. After that, the dataset is split into 70:30 ratios to assess the machine learning algorithms used in the proposed model. Further, two ensemble learning models and one linear regression model are used in the ensemble learning classifier for classification purposes. The simulation evaluation based on five parameters is done. The findings indicate that the model provided in this study shows a high level of accuracy, with a value of 0.9166. Additionally, the model exhibits strong predictive capabilities, as seen by a value of 0.9177. Furthermore, the model demonstrates excellent sensitivity and specificity, with values of 0.9984 for both. Lastly, the F1-Score value, which measures the model's overall performance, is determined to be 0.9564. In the future, hyper tuning of the machine learning algorithm will be done, which is taken into consideration in the ensemble learning classifier. The hyper tuning process of the machine learning algorithm finds the most optimal parameter value.

References

1) Aleem S, Huda N, Amin R, Khalid S, Alshamrani SS, Alshehri A. Machine Learning Algorithms for Depression: Diagnosis, Insights, and Research Directions," . *Electronics*. 2022;11(7):1111–1111. Available from: http://dx.doi.org/10.3390/electronics11071111.
2) World Mental Health Day: an opportunity to kick-start a massive scale-up in investment in mental health. 2020. Available from: https://www.who.int/news/item/27-08-2020-world-mental-health-day-an-opportunity-to-kick-start-a-massive-scale-up-in-investment-in-mental-health.
3) Squires M. Deep learning and machine learning in psychiatry: a survey of current progress in depression detection, diagnosis and treatment. *Brain Informatics*. 2023;10(1). Available from: https://doi.org/10.1186/s40708-023-00188-6.
4) Safdarian M, Trinka E, Rahimi-Movaghar V, Thomschewski A, Aali A. Gdiom Gebreheat Abady, Semagn Mekonnen Abate et al, "Global, regional, and national burden of spinal cord injury, 1990-2019: a systematic analysis for the Global Burden of Disease Study. *The Lancet Neurology*. 2019;22(11):287–296. Available from: https://doi.org/10.1016/S1474-4422(23)00287-9.
5) Dobson KS, Dozois DJA. Handbook of Cognitive-Behavioral Therapies. Guilford Publications. 2019. Available from: https://www.guilford.com/books/Handbook-of-Cognitive-Behavioral-Therapies/Dobson-Dozois/9781462547722.
6) World Health Organization. Depression and Other Common Mental Disorders: Global Health Estimates. . Available from: https://apps.who.int/iris/bitstream/handle/10665/254610/WHO-MSD-MER-2017.2-eng.pdf.
7) Dhall D, Kaur R, Juneja M. Machine learning: a review of the algorithms and its applications. *Proceedings of ICRIC 2019: Recent Innovations in Computing*. 2020;p. 47–63. Available from: https://doi.org/10.1007/978-3-030-29407-6_5.
8) Kadhim AI. Survey on supervised machine learning techniques for automatic text classification. *Artificial Intelligence Review*. 2019;52(1):273–292. Available from: https://doi.org/10.1007/s10462-018-09677-1.
9) Aichele S, Handing E, Strobl C, Jiao Y, Feliciano L, "predictors Of Depression Among Middle-Aged, et al. PREDICTORS OF DEPRESSION AMONG MIDDLE-AGED AND OLDER MEN AND WOMEN IN EUROPE: A MACHINE LEARNING APPROACH,. *Innovation in Aging*. 2022;6(Supplement_1):179–179. Available from: http://dx.doi.org/10.1093/geroni/igac059.716.
10) Aissaoui OE, Madani YEAE, Oughdir L, Allioui YE. Combining supervised and unsupervised machine learning algorithms to predict the learners' learning styles. *Procedia Computer Science*. 2019;148:87–96. Available from: http://dx.doi.org/10.1016/j.procs.2019.01.012.
11) Engelen V, Jesper E, Hoos HH. A survey on semi-supervised learning. *Machine learning*. 2020;109(2):373–440. Available from: https://doi.org/10.1007/s10994-019-05855-6.
12) Liu J, Shi M. A Hybrid Feature Selection and Ensemble Approach to Identify Depressed Users in Online Social Media. *Frontiers in Psychology*. 2021;18(12):802821–802821. Available from: https://doi.org/10.3389/fpsyg.2021.802821.
13) Miftahushudur T, Wael CBA, Praludi T. Infinite Latent Feature Selection Technique for Hyperspectral Image Classification. *Jurnal Elektronika dan Telekomunikasi*. 2019;19(1):32–32. Available from: https://dx.doi.org/10.14203/jet.v19.32-37.
14) Prudencio RF, Maximo MROA, Colombini EL. A Survey on Offline Reinforcement Learning: Taxonomy, Review, and Open Problems. *IEEE Transactions on Neural Networks and Learning Systems*. 2023;p. 1–1. Available from: https://doi.org/10.1109/tnnls.2023.3250269.
15) Iyortsuun NK, Kim SH, Jhon M, Yang HJ, Pant S. A Review of Machine Learning and Deep Learning Approaches on Mental Health Diagnosis. *Healthcare*. 2023;11(3):285–285. Available from: https://doi.org/10.3390/healthcare11030285.
16) Haque UM, Kabir E, Khanam R. Detection of child depression using machine learning methods. *PLOS ONE*. 2021;16(12):261131–261131. Available from: http://dx.doi.org/10.1371/journal.pone.0261131.
17) Sajjadian M, et al. Machine learning in the prediction of depression treatment outcomes: a systematic review and meta-analysis. *Psychological Medicine*. 2021;51(16):2742–2751. Available from: http://dx.doi.org/10.1017/s0033291721003871.
18) Priya A, Garg S, Tigga NP. Predicting Anxiety, Depression and Stress in Modern Life using Machine Learning Algorithms. *Procedia Computer Science*. 2020;167:1258–1267. Available from: https://doi.org/10.1016/j.procs.2020.03.442.
19) Kumar P, Garg S, Garg A. Assessment of Anxiety, Depression and Stress using Machine Learning Models. *Procedia Computer Science*. 2020;171:1989–1998. Available from: https://doi.org/10.1016/j.procs.2020.04.213.
20) Costa VG, Pedreira CE. Recent advances in decision trees: an updated survey. *Artificial Intelligence Review*. 2022;56(5):4765–4800. Available from: https://doi.org/10.1007/s10462-022-10275-5.
21) Kumar P, Garg S, Garg A. Assessment of Anxiety, Depression and Stress using Machine Learning Models. *Procedia Computer Science*. 1989;171. Available from: http://dx.doi.org/10.1016/j.procs.2020.04.213.
22) M SR, L AR, A A. Detection of Depression among Social Media Users with Machine Learning. *Webology*. 2022;19(1):250–257. Available from: http://dx.doi.org/10.14704/web/v19i1/web19019.
23) Liu D, Feng XL, Ahmed F, Shahid M, Guo J. Detecting and Measuring Depression on Social Media Using a Machine Learning Approach: Systematic Review. *JMIR Mental Health*. 2022;9(3):27244–27244. Available from: http://dx.doi.org/10.2196/27244.

24) Heythatsviv. HeyThatsViv/Predicting-Depression: Project using machine learning to predict depression using health care data from the CDC NHANES website. A companion dashboard for users to explore the data in this project was created using Streamlit. Written with python using jupyter notebook for the main project flow/analysis and visual studio code for writing custom functions and creating the dashboard. . Available from: https://github.com/HeyThatsViv/Predicting-Depression/tree/main.
25) Ensemble Classifier Data Mining. *GeeksforGeeks*. 2022. Available from: https://www.geeksforgeeks.org/ensemble-classifier-data-mining/.
26) Brownlee J. A Gentle Introduction to Ensemble Learning Algorithms. *MachineLearningMasterycom*. 2021. Available from: https://machinelearningmastery.com/tour-of-ensemble-learning-algorithms/.
27) Linear Regression in Machine learning. *GeeksforGeeks*. 2023. Available from: https://www.geeksforgeeks.org/ml-linear-regression/.