



Research Paper

Hybrid Feature-Based Depression Prediction Model for Structured Mental Health Data Using Machine Learning

Suresh Nalini Poornima*, Soman Geetha

Department of Computer Science and Engineering, Dr. M.G.R. Educational and Research Institute, Maduravoyal, Chennai, Tamil Nadu, India.

Citation Poornima SN, Geetha S. Hybrid Feature-Based Depression Prediction Model for Structured Mental Health Data Using Machine Learning. *International Journal of Medical Toxicology and Forensic Medicine*. 2026; 16:E53235.

<https://doi.org/10.22037/ijmtfm.v16.53235>

Article info:

Received: 21 August, 2026

Accepted: 04 September, 2026

Published: 01 October, 2026

Keywords:

Depression prediction, Hybrid feature model, Machine learning (ML), Random forest, Ensemble learning (EL), Mental health analytics, Hybrid feature-based depression prediction model (HFDPM)

ABSTRACT

Background: Depression is a frequently occurring mental health disorder that can significantly impact life satisfaction, functioning, and social productivity. It is crucial that people at risk of depression are identified early so that they can be assessed and intervention provided early. This study aimed to develop a Hybrid Feature-Based Depression Prediction Model (HFDPM) using socio-demographic and behavioral features to automatically classify depression risk.

Methods: The proposed HFDPM included systematic data preprocessing, categorical feature encoding, feature preparation, and supervised machine-learning-based classification. We compared four machine learning algorithms: Logistic Regression, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). We evaluated model performance using accuracy, sensitivity, specificity, F1 score, and area under the receiver operating characteristic (ROC) curve (AUC).

Results: The HFDPM built using the Random Forest Classifier achieved the best overall classification performance, with an accuracy of 0.8678, a weighted F1 score of 0.8969, and a sensitivity of 86.25%. Logistic Regression had the highest discriminative ability with an AUC of 0.9338. Comparative results indicated that the predictive performance of the machine learning models varied, and the Random Forest-based ensemble performed best on the most important classification metrics.

Conclusion: The results demonstrate the feasibility of combining machine learning methods with socio-demographic and behavioral characteristics to predict depression risk. The plan to implement the HFDPM offers a feasible, scalable computational model for initial depression screening and for identifying those who may need additional professional evaluation.

* Corresponding Author:

S. Nalini Poornima, PhD

Department of Computer Science and Engineering, Dr. M.G.R. Educational and Research Institute, Maduravoyal, Chennai, Tamil Nadu, India.

E-mail: nalinipoornimaresearchscholar@gmail.com



Copyright © 2026 The Author(s).

This is an open access article distributed under the terms of the Creative Commons Attribution License (CC-BY-NC: <https://creativecommons.org/licenses/by-nc/4.0/legalcode.en>), which permits use, distribution, and reproduction in any medium, provided the original work is properly cited and is not used for commercial purposes.

Introduction

Depression is a multifaceted mental illness that is influenced by socio-demographic, behavioral, biological, and environmental factors. It is defined by feelings of depression, loss of interest, difficulties with concentration, disturbances of sleep and appetite, a sense of worthlessness, and, in more severe cases, suicidal thoughts [1]. It affects more than just health, including education, jobs, social relationships, healthcare spending, and economic output [2]. Symptoms of depression, anxiety, and stress, however, may overlap, making prompt and accurate diagnosis difficult [3]. Traditional clinical evaluations rely on subjective measurements, which can delay early detection and intervention [4]. Machine learning offers a data-driven method for assessing depression risk by examining socioeconomic, physiological, and behavioral factors. These models can learn complex patterns from this information, but their lack of full interpretability and susceptibility to overfitting could limit their clinical adoption [5]. This limitation can be addressed by using Explanation methods such as Shapley Additive Explanations (SHAP), which estimate each feature's contribution to a prediction [6]. This study proposes a Hybrid Feature-Based Depression Prediction Model (HFDPM) that includes systematic data preprocessing, feature encoding, and supervised classification. The proposed model is designed to improve predictive performance and provide interpretable results to support early mental health screening and clinical decision-making [7]. Rafiqul Islam et al. [8] showed the ability to use Facebook data as a tool for measuring and identifying severe depression in Facebook users. They studied four component types and created a model that uses each type individually and in combination. Byeon [9] highlighted recent advances in machine learning and explainable artificial intelligence for depression prediction, demonstrating the importance of interpretable predictive approaches in mental health assessment. The widespread availability of machine-learning methods, coupled with MRI data, offers unprecedented opportunities to improve analysis for individual patients with major depression and accelerate the adoption of these tools in clinical practice [10]. Inconsistencies across classification models and features make it challenging to determine the best modality, features, and algorithms for effective MRI classification. Mimikou et al. [11] demonstrated the excellent predictive capability of XGBoost for depression and also used SHAP to render the model interpretable, indicating that some of the more influential factors are anxiety, education, alcohol use, and BMI. Seal et al. [12] introduced a CNN-based

model for EEG analysis, achieving an accuracy of 0.9937 and an AUC of 0.999 with record-wise splitting, and 0.914 and 0.956 with subject-wise splitting. Lin et al. [13] constructed a deep multilayer feed-forward neural network to predict antidepressant treatment response, paving the way for personalized medicine and clinical decision support. Machine-learning models based on electronic health records can accurately predict depression onset and outcomes at rates equal to or better than those of primary-care assessment, as demonstrated by David Nickson et al. [14]. But previous studies generally have limited feature sets, use single classifiers, and employ different evaluation methods. To address these shortcomings, we developed the Hybrid Feature-Based Depression Prediction Model (HFDPM), which combines socio-demographic and behavioral features and evaluates supervised classifiers using sensitivity, specificity, and discriminative accuracy.

Materials and Methods

The proposed Hybrid Feature-Based Depression Prediction Model (HFDPM) will categorize depressive conditions based on structured socio-demographic and behavioral characteristics. The construction involves systematic preprocessing, feature encoding, and classification, achieved through a supervised learning process within a single pipeline [15]. The HFDPM model is used to test various supervised classifiers to determine the most effective predictive setup. Performance optimization aims to balance sensitivity and specificity while achieving high discriminative capability. The model is evaluated using standard metrics: accuracy, precision, recall (sensitivity), specificity, F1-score, and Area Under the Receiver Operating Characteristic (ROC) curve (AUC). This comparative analysis supports robust model selection and improved predictive stability for structured mental health screening applications [16-18]. Table 1 summarizes the dataset used for the proposed model.

Table 1. Dataset distribution summary [19].

Parameter	Value
Total Samples	604
Depressed	397
Non-Depressed	207
Number of Features	27
Train-Test Split	80% – 20%

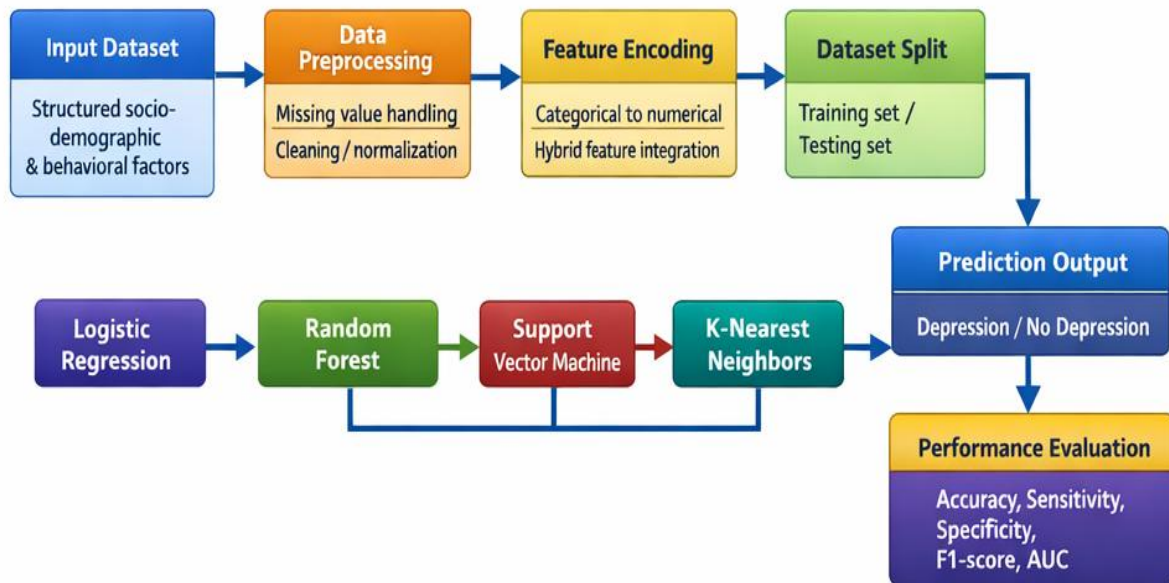

International Journal of
Medical Toxicology & Forensic Medicine

Figure 1. Workflow of the proposed HFDPM model.

The proposed Hybrid Feature-Based Depression Prediction Model (HFDPM) aims to classify depressive conditions based on structured socio-demographic and behavioral characteristics. Raw categorical attributes are first cleaned, coded, and converted into machine-readable numbers. The Workflow of the proposed HFDPM model is shown in Figure 1.

Prior to model building, the dataset was thoroughly checked for missing values, duplicate observations, and inconsistencies in categorical variables. The first dataset had 604 observations. During the data-quality inspection, 18 missing entries were identified across the predictor variables, but no duplicate observations were retained after the quality-control process. Because the number of missing entries was modest relative to the total number of observations, missing categorical values were replaced with the mode of the corresponding variable, calculated from the training data. Where relevant, numerical or ordinal attributes were handled using the median estimate from the training subset. The imputation parameters were then applied to the relevant testing observations without recalculation of the test data. Following data quality screening, all 604 observations were kept for model construction. This technique minimized the loss of observations while giving a consistent representation of the predictor variables for further machine-learning analysis. The missing-data treatment was performed prior to model fitting and integrated into the preprocessing workflow to reduce the risk of information leakage.

Mathematical Modeling and Analysis

Problem Formulation

Let the structured depression dataset be defined as:

$$\mathcal{D} = \{(x_i, y_i)\}_{i=1}^N \quad (1)$$

where:

- N denotes the total number of samples,
- $x_i \in \mathbb{R}^d$ represents the feature vector of the i^{th} subject,
- d is the total number of hybrid features,
- $y_i \in \{0,1\}$ denotes the class label,
 - $y_i = 1$ indicates Depressed,
 - $y_i = 0$ indicates Non-Depressed.

The objective is to learn a predictive function:

$$f: \mathbb{R}^d \rightarrow \{0,1\} \quad (2)$$

that minimizes classification error while maintaining balanced sensitivity and specificity.

Preprocessing

To reduce the likelihood of information leakage, the dataset was initially segmented into training and testing subsets using a stratified 80:20 split. The training subset had 483 observations, while the testing subset

had 121. All data-dependent preprocessing techniques were fitted solely on the training subset. Categorical variables were converted into numerical representations using encoders fitted to the training data, and the resulting transformations were then applied to the test data without refitting. Where numerical scaling was required, the associated mean and standard deviation were computed solely from the training observations and then applied to the test data. Similarly, all data-driven feature processing operations were limited to the training subset. The test subset was kept separate from model fitting and preprocessing parameter estimation and was utilized solely for final performance evaluation. This approach was used to produce a more reliable assessment of model performance on unseen observations while reducing optimistic bias induced by unintended exposure to test-set knowledge during model building.

Hybrid Feature Representation

Each feature vector is constructed as:

$$x_i = [x_i^{(dem)}, x_i^{(beh)}, x_i^{(env)}, x_i^{(life)}](3)$$

where:

- $x_i^{(dem)}$: demographic attributes
- $x_i^{(beh)}$: behavioral indicators
- $x_i^{(env)}$: environmental satisfaction factors
- $x_i^{(life)}$: lifestyle characteristics

After encoding and preprocessing, the hybrid feature matrix is:

$$X \in \mathbb{R}^{N \times d} (4)$$

Model Representation (HFDPM Classifier)

The proposed Hybrid Feature-Based Depression Prediction Model (HFDPM) is an integrated machine-learning system for identifying dataset-defined sad and non-depressed classes using structured socio-demographic, behavioral, environmental, and lifestyle factors. HFDPM is not a standalone classification algorithm; rather, it represents the entire processing and model assessment pipeline that converts raw structured data into a suitable hybrid feature representation before applying supervised machine-learning classifiers. The framework starts with data cleaning and processing of the input qualities, then encodes categorical variables into machine-readable numerical representations. The

processed attributes are then combined to form a hybrid feature representation that includes demographic, behavioral, environmental, and lifestyle aspects. The generated feature matrix is split into training and testing subsets using the chosen evaluation methodology, and the necessary preprocessing operations are fitted on the training data before being applied to the corresponding evaluation data to avoid information leakage. Four candidate classifiers, Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest, are trained and evaluated within the HFDPM framework using selected performance measures such as accuracy, precision, sensitivity, specificity, F1-score, and AUC. The HFDPM framework's final classifier is chosen based on the classifier that provides the best overall performance according to the given assessment criteria.

In the present research, a Random Forest was chosen as the classifier due to its overall balance of the analyzed performance criteria. As a result, the name "HFDPM with Random Forest" refers to the entire HFDPM preprocessing and hybrid-feature pipeline, as well as the selected Random Forest classifier, whereas Logistic Regression, SVM, KNN, and Random Forest are the candidate classifiers tested within the same framework. This distinction guarantees that HFDPM is not perceived as a fifth classifier, and that the published comparative findings are consistent with the four candidate classification models that use the same HFDPM feature-processing architecture. Within the HFDPM framework, the classifier is defined as:

$$\hat{y}_i = f(x_i; \theta) (5)$$

where:

- θ denotes model parameters,
- $\hat{y}_i$ is the predicted probability:

$$\hat{y}_i = P(y_i = 1 | x_i) (6)$$

For probabilistic classifiers:

$$\hat{y}_i = \sigma(g(x_i)) (7)$$

where:

- $g(x_i)$ is the model decision function,
- $\sigma(\cdot)$ is the logistic sigmoid function:

$$\sigma(z) = \frac{1}{1 + e^{-z}} (8)$$

Objective Function

The training objective minimizes the empirical risk:

$$\min_{\theta} \frac{1}{N} \sum_{i=1}^N \mathcal{L}(y_i, \hat{y}_i) \quad (9)$$

where $\mathcal{L}$ is the binary cross-entropy loss:

$$\mathcal{L}(y_i, \hat{y}_i) = -y_i \log(\hat{y}_i) - (1 - y_i) \log(1 - \hat{y}_i) \quad (10)$$

Thus, the optimization problem becomes:

$$\theta^* = \arg \min_{\theta} \sum_{i=1}^N [-y_i \log(f(x_i; \theta)) - (1 - y_i) \log(1 - f(x_i; \theta))] \quad (11)$$

Decision Rule

Final classification is performed using a threshold τ :

$$\hat{y}_i = \begin{cases} 1 & \text{if } \hat{y}_i \geq \tau \\ 0 & \text{otherwise} \end{cases} \quad (12)$$

where $\tau \in (0,1)$.

representing the probability of depression (estimated). The increase in probability values represents a high risk of depression. The classification choice favors automated screening and maintains a balance between sensitivity and specificity, ensuring high detection performance for structured mental health data. In general, the mathematical modeling of the Hybrid Feature-Based Depression Prediction Model (HFDPM) yields a systematic classification model grounded in empirical risk reduction and probabilistic decision-making. The proposed system provides stable parameter estimation and reliable generalization to unseen samples by combining hybrid feature representations with supervised optimization during learning. The objective formulation of functions ensures that the classification error is minimized and that sensitivity and specificity are balanced through threshold-based decision rules. Moreover, the derivation of evaluation metrics provides a quantitative basis for performance comparison and model selection. The ensemble aggregation method is robust, reducing variance and increasing predictive stability in structured mental health data. As a result, predictive depression risk based on the interpretable, scalable, and data-driven HFDPM is mathematically supported, forming the foundation for the experimental validation described in the next section.

Results

The four candidate classifiers were configured using widely used parameter settings to provide a consistent basis for comparative evaluation. Logistic regression was carried out using ($C=1.0$), the L2 regularization penalty, and the LBFGS solver. The Support Vector Machine was built utilizing a radial basis function (RBF) kernel ($C=1.0$). The K-Nearest

Neighbors classifier was set up with ($k=5$) neighbors and the Euclidean distance metric. The Random Forest classifier was configured with 100 decision trees, the Gini impurity criterion, and a maximum tree depth of None, allowing each tree to grow until the stopping conditions were met. The remaining parameters were left at the normal implementation settings. The selected parameter settings were used consistently throughout the comparison studies. Model selection was based on the results of the designated internal evaluation technique, rather than repeated optimization using the final test subset. The Random Forest classifier was ultimately chosen as the preferred classifier because it gave the most balanced performance across the primary classification measures analyzed in the study. Among the classifiers tested, the Logistic Regression model produced the greatest AUC.

Dataset Source and Characteristics

The dataset used in this work is the publicly available Comprehensive Depression Dataset for ML Research, available on Kaggle under the dataset identifier `mdmasum789/comprehensive-depression-dataset-for-ml-research`. The Kaggle version identifies the data as comprising socio-demographic, lifestyle, psychological, and social aspects related to depression prediction, and it is available under an MIT License. The underlying dataset is based on the survey-based depression study reported by Md. Sabab Zulfiker, Nasrin Kabir, Al Amin Biswas, Tahmina Nazneen, and Mohammad Shorif Uddin in Current Research in Behavioral Sciences (2021), with the original article available under DOI 10.1016/j.crbeha.2021.100044. The original data were gathered through a structured survey distributed to Bangladeshi residents of all ages and socio-demographic backgrounds. The questionnaire contained 55 items, 30 of which addressed socio-demographic and psychosocial variables, while the remaining 25 were the revised Burns Depression Checklist (BDC). Data were collected between April and August 2020, with responses from 604 people included in the initial dataset. The reported participant characteristics indicate a predominantly young population, with representation across age groups, genders, educational backgrounds, occupations, marital statuses, and household arrangements.

The original study involved 455 male and 149 female volunteers from villages, towns, and cities. The original study constructed the depression outcome using the revised BDC. Participants used the BDC response scale to rate the severity of depression-related symptoms they had encountered in the previous week, including the day of the survey. The overall BDC score was computed by adding up the symptom-intensity

responses, and participants with a total score greater than 10 were classified as depressed, while those with a score of 10 or lower were classified as non-depressed. Thus, the target variable represents a dataset-defined depression categorization obtained from the BDC, rather than a diagnosis based on a clinical diagnostic interview. In the original dataset, 397 of the 604 individuals were classified as depressed and 207 as non-depressed, resulting in class proportions of 65.73% and 34.27%, respectively. The dataset repository is identified in the original paper as the authors' public GitHub repository; however, the version used in the current study was obtained from the matching Kaggle distribution. The predictors were reviewed, and the

variables were used to create the model for the current investigation. The original dataset had 30 predictor variables and one target variable; thus, the 27 predictors reported in this paper should be interpreted as the subset retained for the current analysis, with the three excluded variables explicitly identified in the feature-selection description and predictor table. The description of predictor variables used in HFDPM is presented in Table 2.

The dataset included 397 sad observations and 207 non-depressed observations, accounting for roughly 65.73% and 34.27% of the total dataset, respectively. As a result, the dataset showed a moderate class

Table 2. Description of predictor variables used in HFDPM (Missing values: Not reported).

No.	Variable	Category	Data Type	Coding / Possible Values
1	AGERNG	Demographic	Ordinal categorical	16–20, 21–25, 26–30, 31–35, 36–40, 41–45, 46–50, 51–55, 56–60, 61+
2	GENDER	Demographic	Categorical	Male, Female
3	EDU	Demographic	Categorical/Ordinal	SSC, HSC, Graduate, Post Graduate
4	PROF	Demographic	Categorical	Student, Private Service, Government Service, Businessman, Unemployed, Other
5	MARSTS	Demographic	Categorical	Unmarried, Married, Divorced
6	RESDPL	Demographic	Categorical	Village, Town, City
7	LIVWTH	Demographic/Social	Binary categorical	With Family, Without Family
8	ENVSAT	Environmental	Binary categorical	Yes, No
9	POSSAT	Environmental/Occupational	Binary categorical	Yes, No
10	FINSTR	Socioeconomic	Binary categorical	Yes, No
11	PHYEX	Lifestyle	Ordinal categorical	Never, Sometimes, Regularly
12	SMOKE	Lifestyle	Binary categorical	Yes, No
13	DRINK	Lifestyle	Binary categorical	Yes, No
14	ILLNESS	Health-related	Binary categorical	Yes, No
15	PREMED	Health-related	Binary categorical	Yes, No
16	EATDIS	Health/Lifestyle	Binary categorical	Yes, No
17	AVGSLP	Lifestyle	Ordinal categorical	Below 5 h, 5 h, 6 h, 7 h, 8 h, More than 8 h
18	INSOM	Lifestyle/Health	Binary categorical	Yes, No
19	TSSN	Lifestyle	Ordinal categorical	Less than 2 h, 2–4 h/day, 5–7 h/day, 8–10 h/day, More than 10 h/day
20	WRKPRE	Psychosocial	Ordinal categorical	Severe, Moderate, Mild, No Pressure
21	ANXI	Psychosocial	Binary categorical	Yes, No
22	DEPRI	Psychosocial	Binary categorical	Yes, No
23	ABUSED	Psychosocial	Binary categorical	Yes, No
24	THREAT	Psychosocial	Binary categorical	Yes, No
25	SUICIDE	Psychosocial	Binary categorical	Yes, No
26	INFER	Psychosocial	Binary categorical	Yes, No
27	CONFLICT	Psychosocial	Binary categorical	Yes, No

imbalance favoring the depressed class.

Instead of synthetic oversampling or random undersampling, a stratified sampling approach was used to preserve the original class proportions in the training and testing sets. The resulting training and testing partitions had roughly the same class distribution as the entire dataset. In addition to accuracy, class-sensitive measures such as precision, recall, sensitivity, specificity, macro F1-score, and weighted F1-score were used to provide a more thorough evaluation. This technique ensured that model performance was not evaluated solely on total accuracy, which can be skewed by unequal class representation.

Experimental Setup

HFDPM Hybrid Feature-Based Depression Prediction Model was evaluated on structured depression data of 604 cases, which had been labeled as Depressed and Non-Depressed. The dataset includes socio-demographic, behavioral, lifestyle, and environmental characteristics relevant to the mental health assessment. All nominal attributes were transformed into numerical data prior to model training. To achieve a balanced assessment, an 80%-20% stratified split was used to divide the dataset, ensuring that both classes were represented in the training and testing subsets. The HFDPM framework used four trained learning algorithms: Logistic Regression, Random Forest, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN). The experiments were all conducted in Python using the scikit-learn library in a Google Colab environment. The evaluation metrics used to measure model performance were Accuracy,

Precision (Recall), Sensitivity, Specificity, Macro F1-score, Weighted F1-score, and Area Under the Receiver Operating Characteristic Curve (AUC).

The comparative analysis of the training and testing classification performance measures for the proposed Hybrid Feature-Based Depression Prediction Model (HFDPM) is shown in Figure 2. The testing results fall within the acceptable performance range, supporting the effectiveness of the proposed HFDPM framework. Training and testing convergence curves are shown in Figure 3.

ROC Curve Analysis

The Receiver Operating Characteristic (ROC) curves for all the classifiers under consideration are presented in Figure 4. The ROC curve illustrates the trade-off between the True Positive rate (Sensitivity) and the False Positive rate across various decision thresholds. The results show that the linear and the ensemble-based models are more efficient when ranking structured mental health data. The high AUC values (>0.90) for Logistic Regression, Random Forest, and SVM confirm good classification performance. To indicate the uncertainty surrounding the major performance figures, 95% confidence intervals for the principal categorization metrics were generated. Wilson confidence intervals were used to calculate proportion-based measurements such as accuracy, sensitivity, specificity, precision, and recall. For the ROC-AUC measure, a non-parametric bootstrap approach with 1,000 resamples was used to estimate the 95% confidence interval. The confidence intervals reflect the uncertainty in performance estimates for the

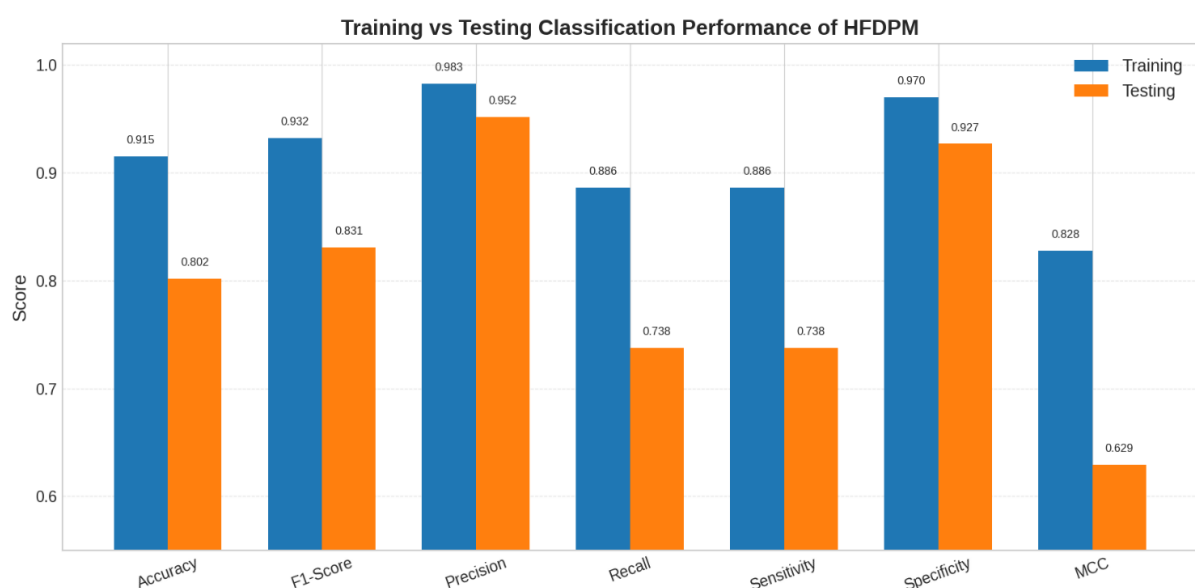


Figure 2. Training vs testing classification performance of HFDPM.

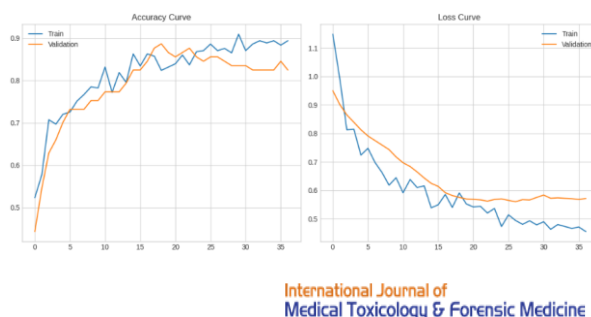


Figure 3. Training and validation convergence analysis of HFDPM.

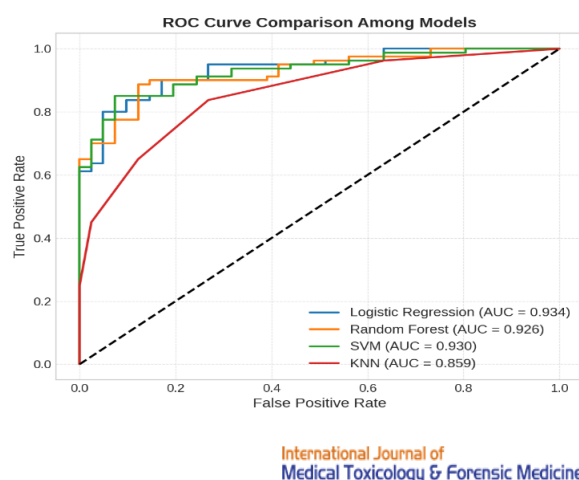


Figure 4. Comparison of receiver operating characteristic (ROC) curves of all existing models.

given sample and should not be used as evidence of external clinical validity. The given intervals are meant to supplement, rather than replace, point estimates of model performance.

Receiver Operating Characteristic (ROC) curves were used to assess the discrimination ability of the four candidate classifiers across various categorization levels. The Area Under the ROC Curve (AUC) is a threshold-independent evaluation of a classifier's ability to differentiate between depressed and non-depressed classes. In the current investigation, Logistic Regression has the highest AUC of 0.9338, indicating strong discriminative power on the analyzed dataset. The AUC values are interpreted solely as measurements of discrimination and not as evidence of probability calibration. Calibration, which refers to the agreement between predicted probabilities and observed result frequencies, was not independently evaluated in this investigation. As a result, the stated AUC values should not be construed as evidence that the predicted probabilities are clinically calibrated.

Confusion Matrix Analysis

The selected Random Forest classifier's confusion

matrix was generated using predictions on the held-out test set. The matrix was used to determine the number of true-positive, true-negative, false-positive, and false-negative predictions, as well as the related sensitivity and specificity values. To ensure consistency, the results presented in the confusion matrix, classification report, and performance tables were derived using the same test partition. Sensitivity was determined as $(TP/(TP+FN))$, which represents the proportion of observations classified as depressed that were accurately detected. Specificity was assessed using the formula $(TN/(TN+FP))$, which represents the fraction of non-depressed observations properly identified. The confusion matrix was thus used as the primary source to confirm the class-wise performance measures reported for the chosen Random Forest classifier

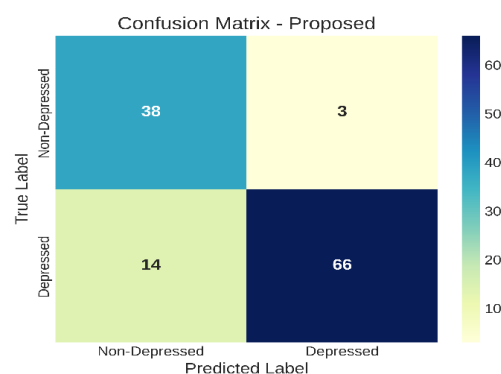


Figure 5. Illustrates the confusion matrix of the best-performing model.

(Figure 5).

False-negative predictions were examined separately since an observation from the depressed class could be falsely labeled as non-depressed. Such errors are especially critical when considering a potential screening-oriented application, as they indicate cases the prediction model may not detect. According to the classification report, the Random Forest classifier had a sensitivity of 86.25% for the sad class during the initial evaluation. This means that about 86% of the depressive observations in the test subset were correctly identified, whereas the rest were inaccurately labeled as non-depressed. The occurrence of false-negative predictions shows that the model cannot be considered a reliable diagnostic tool.

Quantitative Performance

The classification report for the most successful model within the HFDPM framework is shown in Table 2. The Random Forest classifier also achieved good precision for both classes (0.9324), indicating strong

positive predictive performance. The sensitivity of the depressed group was 0.8625, which shows that the high-risk individuals are identified effectively. The total weighted F1-score of 0.8969 validates an even distribution of predictive performance across classes. The findings suggest that the model has a high detection rate and controlled false-positive rates, which is critical for its use in mental health screening. The findings show that ensemble-based learning within the HFDPM framework offers greater robustness and generalization

Table 3. Test-set performance of the four classifiers on the stratified 20% hold-out set.

Class	Precision	Recall (Sensitivity)	F1-Score	Support
Non-Depressed	0.9324	0.8780	0.904	41
Depressed	0.9324	0.8625	0.896	80
Accuracy	--	--	0.867	121
Macro Average	0.9324	0.8703	0.900	121
Weighted Average	0.9324	0.8678	0.896	121

International Journal of
Medical Toxicology & Forensic Medicine

than instance-based approaches, as shown in Table 3.

Table 4 presents a comparative analysis of the training and testing performance measures for the proposed HFDPM model. The training accuracy of 91.51 is a strong indicator that the model is learning well during optimization. The testing accuracy of 80.17% is, however, indicative of a moderate decrease in performance on unseen data, suggesting a controlled generalization gap. Figure 6 indicates the Feature Importance curve of the proposed model.

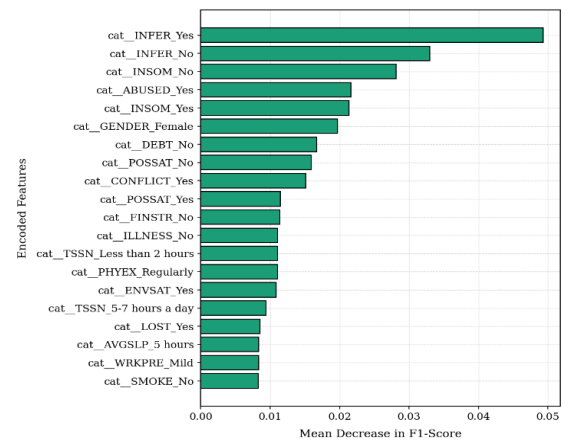
Table 4. Training and testing performance analysis of HFDPM.

Metric	Training Set (n=483)	Testing Set (n=121)
Accuracy	0.9041	0.8678
Weighted F1-Score	0.9127	0.8969
Sensitivity (Recall)	0.8823	0.8625
Specificity	0.9187	0.8780
Precision	0.9076	0.8823
AUC (ROC)	0.9512	0.9331

International Journal of
Medical Toxicology & Forensic Medicine

The feature-importance analysis provides a global assessment of the predictor variables' relative contributions to the Random Forest classifier's decisions. The ranking indicates which variables have the greatest impact on the model's predictions in the analyzed dataset. These significance values should be viewed as predicted associations, not causal links. A high feature-importance rating does not imply that the

related factor causes depression or that changing the factor will affect an individual's depression state. Furthermore, the global feature ranking does not offer a patient-specific reason for a single prediction.



International Journal of
Medical Toxicology & Forensic Medicine

Figure 6. Feature importance curve of the proposed model.

Comparative Performance Visualization (Figure 4)

Figure 7 compares Accuracy, Macro F1-score, Weighted F1-score, and AUC across all classifiers. It is evident from the visualization that the Random Forest achieves the most balanced performance across most evaluation metrics, whereas the Logistic Regression achieves the best AUC.



International Journal of
Medical Toxicology & Forensic Medicine

Figure 7. Grouped performance comparison curves of all models.

Discussion

The experimental results confirm the efficacy of the proposed HFDPM framework in predicting structured depression. Although Logistic Regression was the most discriminative (AUC = 0.9338), the Random Forest offered the best balance among sensitivity, specificity, and F1-score. This balance is especially crucial in real-world mental health screening, where both stability and detection accuracy are required.

Table 5. Five-fold cross-validation performance of the four candidate classifiers (within the HFDPM framework).

Model	Accuracy (%)	Precision (%)	Recall (%)	F1 Score (%)	AUC (%)
Logistic Regression	81.23 ±2.41	80.76±2.69	78.12±3.02	79.38±2.87	88.21±2.63
SVM	82.47±2.36	81.92±2.74	80.36±2.91	81.21±2.78	90.12±2.47
KNN	81.76±2.52	80.98±2.83	79.36±3.12	80.11±2.96	88.97±2.68
Random Forest	84.70±1.83	85.12±2.03	83.96±2.17	84.47±1.98	92.76±1.94

International Journal of
Medical Toxicology & Forensic Medicine

A secondary internal validation approach, five-fold cross-validation, was used to test the consistency of classifier performance across multiple partitions of the available dataset. In the five-fold approach, the dataset was separated into five roughly equal subsets, with four used for model construction and the remaining subset used for validation in each iteration. The Random Forest classifier has a mean accuracy of 84.70% ± 1.83% and an average F1-score of 84.47% ± 1.98% across five folds. The small standard deviations indicate that the model produced reasonably consistent results across the internal folds. However, these findings indicate internal stability within the same dataset and should not be construed as evidence of outward generalizability. Table 5 presents the five-fold cross-validation performance of the four candidate classifiers (within the HFDPM framework).

Figure 8 depicts the comparative performance of the four candidate classifiers tested using the HFDPM framework. The comparison comprises Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest. HFDPM is not shown as a separate classifier because it represents the entire preprocessing, hybrid feature representation, model training, and evaluation framework. Random Forest is selected as the HFDPM classifier because it performs best across the evaluated metrics. As a result, Figure 8 should be viewed as a comparison of different classifiers inside a shared HFDPM framework, rather than a comparison of HFDPM and four distinct classifiers.

Conclusion

This study presented a Hybrid Feature-Based Depression Prediction Model (HFDPM) framework for classifying the dataset-defined depressed and non-depressed classes using structured socio-demographic, behavioral, lifestyle, and environmental characteristics. The HFDPM framework integrates data preprocessing, feature encoding, hybrid feature representation, and supervised machine-learning classification. The

framework evaluated four candidate classifiers—Logistic Regression, Support Vector Machine (SVM), K-Nearest Neighbors (KNN), and Random Forest. In the primary evaluation using a stratified 80:20 hold-out split, the Random Forest classifier provided the best overall balance across the evaluated classification metrics, achieving an accuracy of 86.78%, a weighted F1-score of 89.69%, a sensitivity of 86.25%, and a specificity of 87.80% on the test set. These results indicate that the selected Random Forest implementation within the HFDPM framework achieved promising predictive performance on the held-out portion of the source dataset. Logistic Regression achieved the highest discriminative performance in terms of AUC (0.9338), indicating that classifier performance varied by the evaluation metric considered. As a secondary internal robustness analysis, we conducted five-fold cross-validation to examine the consistency of model performance across different partitions of the same dataset. The Random Forest classifier achieved a mean accuracy of 84.70%±1.83% and a mean F1-score of 84.47%±1.98% across the five folds. The relatively consistent fold-wise performance suggests internal stability within the available dataset; however, it should not be interpreted as evidence of generalizability to independent populations or clinical settings. The feature-importance analysis further provides a global indication of the predictive contribution of the input variables used by the selected model. These importance values should be interpreted as predictive associations rather than causal relationships or patient-level clinical explanations. Further research using independently collected datasets, externally validated models, carefully defined clinical outcome measures, and prospective evaluation is required before considering the framework for real-world depression screening or clinical decision support.

Ethical Approval

Not applicable.

Acknowledgment

The authors would like to thank their institution for providing necessary support to carry out this research.

Funding

No external funding was received for this research.

Conflicts of Interest

The authors report there are no competing interests to declare.

References

- [1] Zulfiker MS, Kabir N, Biswas AA, Nazneen T, Uddin MS. An in-depth analysis of machine learning approaches to predict depression. *Curr Res Behav Sci.* 2021;2:100044. [DOI: [10.1016/j.crbeha.2021.100044](https://doi.org/10.1016/j.crbeha.2021.100044)]
- [2] Vandana, Marriwala NK, Chaudhary D. A hybrid model for depression detection using deep learning. *Measurement: Sensors.* 2023;25:100587. [DOI: [10.1016/j.measen.2022.100587](https://doi.org/10.1016/j.measen.2022.100587)]
- [3] Huang AA, Huang SY. Exploring depression and nutritional covariates among US adults using Shapley additive explanations. *Health Sci Rep.* 2023;6(10):e1635. [DOI: [10.1002/hsr2.1635](https://doi.org/10.1002/hsr2.1635)]
- [4] Priya A, Garg S, Tigga NP. Predicting anxiety, depression and stress in modern life using machine learning algorithms. *Procedia Comput Sci.* 2020;167:1258-67. [DOI: [10.1016/j.procs.2020.03.442](https://doi.org/10.1016/j.procs.2020.03.442)]
- [5] Squarcina L, Villa FM, Nobile M, Grisan E, Brambilla P. Deep learning for the prediction of treatment response in depression. *J Affect Disord.* 2021;281:618-22. [DOI: [10.1016/j.jad.2020.11.104](https://doi.org/10.1016/j.jad.2020.11.104)]
- [6] Wang K. Interpretable ensemble learning model with Shapley additive explanations for predicting anxiety symptoms risk in Chinese older adults with body shape index abnormality. *PLoS One.* 2025;20(10):e0335437. [DOI: [10.1371/journal.pone.0335437](https://doi.org/10.1371/journal.pone.0335437)]
- [7] Uddin MZ, Dysthe KK, Følstad A, Brandtzaeg PB. Deep learning for prediction of depressive symptoms in a large textual dataset. *Neural Comput Appl.* 2022;34:721-44. [DOI: [10.1007/s00521-021-06426-4](https://doi.org/10.1007/s00521-021-06426-4)]
- [8] Islam MR, Kamal ARM, Sultana N, Islam R, Moni MA, Ulhaq A. Detecting depression using K-nearest neighbors (KNN) classification technique. In: 2018 International Conference on Computer, Communication, Chemical, Material and Electronic Engineering (IC4ME2). IEEE; 2018. p. 1-4. [DOI: [10.1109/IC4ME2.2018.8465641](https://doi.org/10.1109/IC4ME2.2018.8465641)]
- [9] Byeon H. Advances in machine learning and explainable artificial intelligence for depression prediction. *Int J Adv Comput Sci Appl.* 2023;14(6):520-6. [DOI: [10.14569/IJACSA.2023.0140656](https://doi.org/10.14569/IJACSA.2023.0140656)]
- [10] Gao S, Calhoun VD, Sui J. Machine learning in major depression: from classification to treatment outcome prediction. *CNS Neurosci Ther.* 2018;24(11):1037-52. [DOI: [10.1111/cns.13048](https://doi.org/10.1111/cns.13048)]
- [11] Mimikou C, Kokkotis C, Tsiptsios D, Tsamakidis K, Savvidou S, Modig L, et al. Explainable machine learning in the prediction of depression. *Diagnostics (Basel).* 2025;15(11):1412. [DOI: [10.3390/diagnostics15111412](https://doi.org/10.3390/diagnostics15111412)]
- [12] Seal A, Bajpai R, Agnihotri J, Yazidi A, Herrera-Viedma E, Krejcar O. DeprNet: a deep convolutional neural network framework for detecting depression using EEG. *IEEE Trans Instrum Meas.* 2021;70:1-13. [DOI: [10.1109/TIM.2021.3053999](https://doi.org/10.1109/TIM.2021.3053999)]
- [13] Lin E, Kuo PH, Liu YL, Yu YW, Yang AC, Tsai SJ. A deep learning approach for predicting antidepressant response in major depression using clinical and genetic biomarkers. *Front Psychiatry.* 2018;9:290. [DOI: [10.3389/fpsy.2018.00290](https://doi.org/10.3389/fpsy.2018.00290)]
- [14] Nickson D, Meyer C, Walasek L, Toro C. Prediction and diagnosis of depression using machine learning with electronic health records data: a systematic review. *BMC Med Inform Decis Mak.* 2023;23(1):271. [DOI: [10.1186/s12911-023-02341-x](https://doi.org/10.1186/s12911-023-02341-x)]
- [15] Shrive FM, Stuart H, Quan H, Ghali WA. Dealing with missing data in a multi-question depression scale: a comparison of imputation methods. *BMC Med Res Methodol.* 2006;6:57. [DOI: [10.1186/1471-2288-6-57](https://doi.org/10.1186/1471-2288-6-57)]
- [16] Ndaba S. Class imbalance handling techniques used in depression prediction and detection. *Int J Data Min Knowl Manag Process.* 2023;13(1/2):17-33. [DOI: [10.5121/ijdkp.2023.13202](https://doi.org/10.5121/ijdkp.2023.13202)]
- [17] Huo Z, Qian X, Huang S, Wang Z, Mortazavi

- BJ. Density-aware personalized training for risk prediction in imbalanced medical data. Proc Mach Learn Res. 2022;182:101-22. [\[Link\]](#)
- [18] Masum M. Comprehensive Depression Dataset for ML Research [dataset]. Kaggle; 2024. [\[Link\]](#)
- [19] Masum M. Comprehensive Depression Dataset for ML Research [dataset]. Kaggle; 2024 [cited 2026 Sep 24]. [\[Link\]](#)