Ensemble Learning Approaches for Cardiovascular Disease Prediction

Ahmad Muzammil Halim
Faculty of Computer and Mathematical Sciences, University Teknologi MARA, Shah Alam, Malaysia
halimmuzammil614@gmail.com

Marshima Mohd Rosli
Faculty of Computer and Mathematical Sciences, University Teknologi MARA, Shah Alam, Malaysia
marshima@uitm.edu.my (corresponding author)

Norzilah Musa
Faculty of Computer and Mathematical Sciences, University Teknologi MARA, Shah Alam, Malaysia
norzi105@uitm.edu.my

Nurain Ibrahim
Faculty of Computer and Mathematical Sciences, University Teknologi MARA, Shah Alam, Malaysia
nurainibrahim@uitm.edu.my

Received: 18 September 2025 | Revised: 4 November 2025 and 20 November 2025 | Accepted: 3 December 2025

DOI: https://doi.org/10.48084/etasr.14877

ABSTRACT

Cardiovascular Disease (CVD) is a primary cause of mortality worldwide, which requires precise and timely predictions to implement effective prevention and intervention strategies. Existing prediction methods, dependent on a single classifier, frequently prove inadequate due to data complexity and class imbalance, leading to incorrect classification. This study aimed to compare the performance of ensemble learning approaches for predicting CVD using clinical datasets. Six ensemble classifiers were evaluated: XGBoost and Gradient Boosting (Boosting), Random Forest and Decision Tree (Bagging), and K-Nearest-Neighbor and Logistic Regression (Stacking). Boosting classifiers achieved the highest accuracy (73.0%) and G-mean (72.0%) for CVD prediction, while the bagging classifier for Random Forest achieved similar precision in identifying high-risk patients with CVD. Boosting classifiers demonstrate strong discriminative power in differentiating between risk, as confirmed by the ROC AUC (0.80), indicating their effectiveness for reliable CVD prediction.

Keywords-component; cardiovascular; ensemble learning; prediction; classifiers

I. INTRODUCTION

Cardiovascular Disease (CVD) has been classified as an epidemic and has been reported as the leading cause of mortality globally [1, 2]. The prevalence of CVD is increasing due to the modern lifestyle, urbanization, and increased life expectancy. CVD remains a critical challenge in healthcare, necessitating the advancement of predictive methods to address the prevailing concerns. Despite the existence of various diagnostic methods, including traditional procedures such as ECG and echocardiography, these tests often lack precision and do not detect the disease in its early stages. As awareness of early screening and clinical risk assessment of CVD increases, various ML methods are used for prediction. These methods aim to improve accuracy in detecting the risk of CVD and its advancement, thus facilitating effective early intervention and management [3-5].

Previous research on CVD classification has utilized ML classifiers, such as Logistic Regression (LR), Decision Trees (DTs), Support Vector Machines (SVM), and Neural Networks (NN) [6, 7]. Although these models allow early diagnosis, they are restricted by numerous limitations, including class imbalance, inadequate generalization, and lack of interpretability [8, 9]. Many classifiers demonstrate overfitting to certain datasets, making them unsuitable for diverse populations, while methods such as deep neural networks, despite their precision, provide no practical insight [10-12].

ML has progressed, with ensemble learning techniques demonstrating considerable potential to enhance predictive accuracy by integrating the strengths of many algorithms. Ensemble learning offers an improved approach by combining various classifiers to enhance predictive accuracy, reduce bias, and increase interpretability, which is particularly useful in the context of CVD. Ensemble learning methods have been shown to outperform traditional classifiers in terms of accuracy and robustness in the detection of CVD [13-15]. For example, boosting techniques such as XGBoost enhance recall for minority classes, but bagging methods such as RF mitigate overfitting and enhance generalization. Ensemble learning classifiers have shown significant improvements in the accuracy and robustness of CVD prediction models, offering important tools in clinical decision-making and early detection. [16-19].

Ensemble learning is a technique that combines multiple basic classifiers to enhance predictive performance relative to single models [20]. The primary objective is to improve accuracy and robustness by using the strengths of many models while mitigating their weaknesses [21]. Base classifiers are independent models whose predictions are combined to create a more robust model. The efficacy of an ensemble depends on the accuracy of the base learners and the diversity among them. Diverse models exhibit reduced susceptibility to the same errors, which enhances overall performance [22-24].

Common types of ensemble learning are bagging, boosting, and stacking. Bagging involves independent training of different models and the aggregation of their predictions by voting to improve accuracy. RF is a popular classifier in this category. Bagging emphasizes the reduction of variation and the prevention of overfitting [21, 25]. In boosting, base learners are trained consecutively, with each learner concentrating on cases that previous models struggled with, thus enhancing overall performance. Boosting is frequently employed to improve the efficacy of a base classifier. Although boosting might be computationally intensive and susceptible to noise, it often yields more accurate classifications than bagging. The most prominent boosting classifiers are XGBoost and Gradient Boosting (GB) [25, 26]. Stacking integrates the predictions of various models through a meta-learner to produce a final prediction, facilitating the capture of more complex correlations. Stacking is more effective when the basic classifiers exhibit strong learning performance and increased variance. Common models used as stacking meta-learner classifiers are K-Nearest Neighbors (KNN) and LR [25, 26].

The integration of feature selection with ensemble learning can improve the accuracy and interpretability of CVD risk models. Feature selection is the process of selecting important attributes in the dataset to improve model performance. In the context of classification, feature selection methods are commonly categorized into three main approaches: filter methods, wrapper methods, and embedded methods. Techniques such as Pearson Correlation (filter), RF Importance (embedded), and Recursive Feature (RFE) (wrapper) have been used to reduce high dimensionality, improve interpretability, and reduce overfitting. These methods ensure that only the most relevant features affect the final prediction model, hence improving classification accuracy and computational efficiency [26-28].

Previous research findings on CVD classification have consistently demonstrated the integration of data balancing strategies and ensemble learning to address data imbalance without altering class structures [29, 30]. These strategies involve adjusting the class distribution of the dataset to achieve a balanced representation, using techniques such as Random Undersampling (RUS), Random Oversampling (ROS), and the Synthetic Minority Oversampling Technique (SMOTE). Data rebalancing strategies allow for a holistic modeling approach by preserving the original class structure and analyzing the class distribution globally. Previous research has demonstrated that these methods have shown significant promise in mitigating class imbalance issues [31, 32]. By focusing on maintaining the natural class structure, rebalancing strategies offer a more suitable approach for medical data, as they minimize information loss and help maintain the relevance of features critical for accurately predicting target classes [31, 32]. However, most disease predictions applying ensemble learning with these strategies are often restricted to specific domains and tasks, limiting their generalizability.

This study expands on previous research by incorporating SMOTE for data balance, utilizing ensemble learning to improve classification accuracy, and implementing optimal feature selection to increase CVD prediction and patient support. This study compares the performance of ensemble learning classifiers: XGBoost and GB (Boosting), RF and DT (Bagging), and KNN, LR (Stacking). The purpose is to utilize these predictions to improve the identification of CVD in the context of health screening patients. SMOTE was used for data balancing, and RFE was used to discover relevant CVD predictors. The experimental results reveal that all classifiers exhibit adequate precision in predicting CVD among patients.

II. METHODOLOGY

Figure 1 shows the workflow followed in this study, having three phases: (i) Data Preprocessing, (ii) CVD Prediction, (iii) Performance Evaluation.

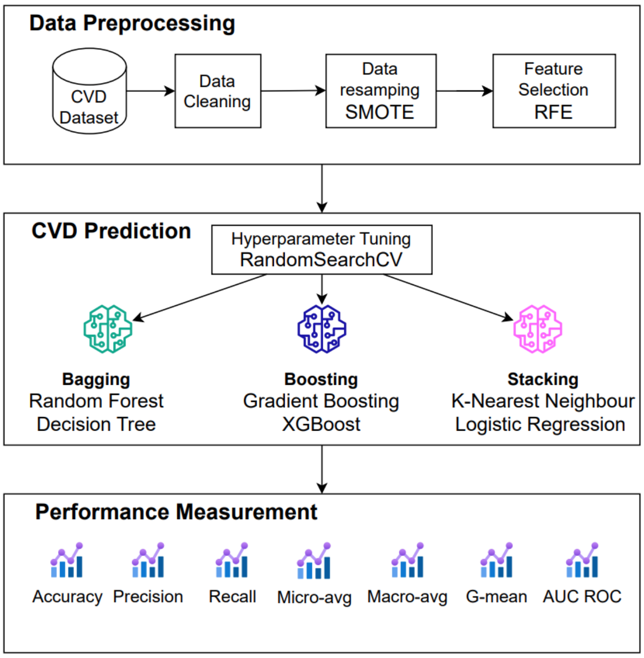

Fig. 1. Workflow of this study.

In Data Preprocessing, the raw dataset was processed using data cleaning, data resampling with SMOTE, and feature selection with RFE. In CVD Prediction, three types of ensemble classifiers, Boosting (GB, XGBoost), Bagging (RF, and RF meta-learner), and KNN, LR (as stacking meta-learners), were used to classify CVD. Finally, the effectiveness of each classifier was assessed.

A. Data Preprocessing

This study used a secondary dataset including 69,966 records with 11 independent variables [33]. The data were gathered during a public health screening program and represented demographic, behavioral, and clinical measurements associated with CVD risk. This dataset was selected due to its large sample size, balanced representation of risk factors, and frequent use in predictive modeling research. Table I shows the dataset's variables. The target variable indicates the presence or absence of CVD.

Redundant records were removed to ensure data integrity. Missing values were identified, and data imputation techniques were applied, replacing null values with the corresponding column averages. Outlier detection and removal were performed using the Z-score method, which effectively filtered out extreme values that impact model performance. The cleaned datasets consisted of 55,972 records.

TABLE I. CARDIOVASCULAR DATASET FEATURES

Variable	**Description**	**Value range**
Age	Years	(21,99)
Height	Height	Continuous numerical
Weight	Weight	Continuous numerical
Gender	Gender	0 = Female, 1 = Male
ap_hi	Systolic Blood Pressure	Continuous numerical
ap_lo	Diastolic Blood Pressure	Continuous numerical
Chol	Cholesterol	1 = Normal 2 = Above normal 3 = Well above normal
Gluc	Glucose	1 = Normal 2 = Above normal 3 = Well above normal
Smoke	Smoking	0 = Non-smoker 1 = Smoker
Alco	Alcohol Intake	0 = Non-consumption 1 = Consumes alcohol
Active	Physical Activity	0 = Inactive 1 = Active
Cardio	Presence or Absence of CVD	0 = No, 1 = Yes

Rebalancing procedures are commonly employed to address imbalances in class distribution. This study used SMOTE to rebalance the data before training. SMOTE oversamples the minority class by duplicating and generating new instances through existing samples, as shown in Figure 2. This study also employed RFE to identify the most relevant features for CVD classification.

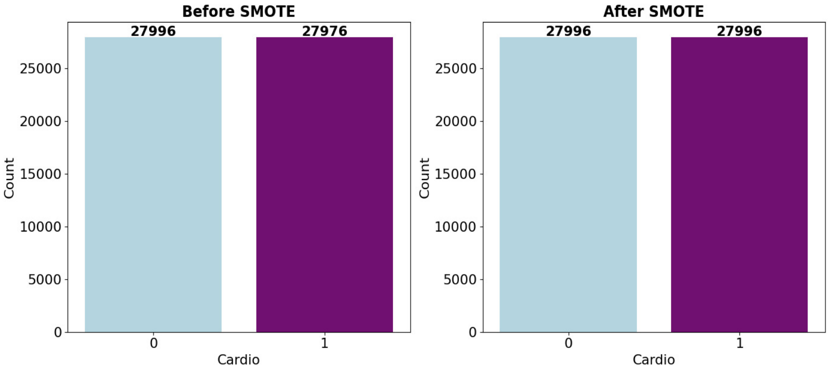

Fig. 2. Comparison before SMOTE and after SMOTE

B. CVD Classification

Experiments were conducted using three types of ensemble models described above. The boosting models, XGBoost and GB, were configured with 100 estimators, a learning rate of 0.1, and a maximum tree depth of 5, achieving a balance between accuracy and computational cost. RF was trained with 100 trees, each having a maximum depth of 10, and the DT classifier employed the same criterion for interpretability comparison. KNN utilised $k = 5$ neighbors with Euclidean distance, whereas LR served as the meta-learner with L2 regularization.

The dataset was partitioned into the training set, which included 80% of each class's data, and the test set, containing the remaining 20% (80:20). The partition was performed randomly using stratified sampling, ensuring that the class distribution of CVD and non-CVD remained consistent in both the training and test sets. The 80:20 ratio is commonly adopted in research, since it offers a balanced trade-off between having sufficient data for accurate measurement.

C. Performance Evaluation

Various performance metrics were considered to evaluate the performance of the ensemble learning approaches. The formulas for each metric, including accuracy, precision, recall, and the Area Under the Receiver Operating Characteristic (AUROC) curve, were computed as follows.

Accuracy is the overall percentage of correctly predicted instances across all the classes, where TP denotes true positives, TN is the true negatives, FP is false positives, and FN is false negatives.

$$\text{Accuracy} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{FP} + \text{TN} + \text{FN}} \tag{1}$$

Precision is the percentage of values that a model predicts as positive and are, in fact, positive (TP):

$$\text{Precision} = \frac{\text{TP}}{\text{TP + FP}} \quad (2)$$

Recall evaluates the model's ability to accurately predict the positive class:

$$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}} \quad (3)$$

A higher F1-score macro-averaging value indicates better overall classification performances in all the classes:

$$\text{MacroAverageF1} = 2 * \left(\frac{\text{MAP}*\text{MAR}}{\text{MAP+MAR}}\right) \quad (4)$$

where MAP is the Macro Average Precision and MAR is Macro Average Recall. For F1-score microaveraging, the global average calculations were made without taking into account the class differences:

$$\text{Micro AverageF1} = \frac{\sum_{k=1}^{K} TPk}{\text{Grand Total}} \quad (5)$$

G-mean was used to measure the balance between classes,

$$\text{G} - \text{mean} = \left(\prod_{i1}^{k} Ri\right) 1/k \quad (6)$$

where R is Recall.

III. RESULTS

The performance of each classifier in CVD prediction was assessed using the average of 5-fold cross-validation, and the results are presented in Table II. The boosting classifiers demonstrated the strongest performance among all ensemble learning approaches. XGBoost exhibited the highest accuracy at 73.04% and a G-mean of 71.95%. Similarly, GB showed comparable results with an accuracy of 72.92% and a G-mean of 71.7%. These classifiers also achieved higher precision and recall, signifying their capacity to distinguish between classes while preserving a balance between FP and FN.

The RF classifier (bagging) demonstrated moderate performance, achieving an accuracy of 72.40 and a G-mean of 70.9, indicating balanced performance across classes. The DT bagging meta-learner obtained an accuracy of 69.0 and a G-mean of 68.9. The stacking meta-learners exhibited varied performances. LR achieved an accuracy of 72.7% and a G-mean of 71.5, close to the performance of XGBoost. However, KNN demonstrated the lowest performance, achieving 69.2% accuracy and a G-mean of 69.0%, while maintaining essential sensitivity to the positive class. KNN consistently showed inadequate performance in all cases.

TABLE II. PERFORMANCE METRICS OF DIFFERENT ENSEMBLE LEARNING APPROACHES

	Boosting		Bagging		Stacking	
Metric	**XGBoost**	**GB**	**RF**	**DT**	**KNN**	**LR**
Accuracy	**73.0**	**72.9**	72.4	69.0	69.2	72.7
Precision	**75.1**	**74.9**	74.9	69.0	69.5	74.7
Recall	**68.9**	**68.7**	67.2	68.9	68.5	68.4
Micro-avgF1	**73.0**	**72.9**	72.4	69.0	69.2	72.7
Macro-avgF1	**72.9**	**72.8**	72.3	69.0	69.2	72.6
G-mean	**71.9**	**71.7**	70.9	68.9	69.0	71.5

Overall, boosting classifiers, specifically XGBoost and GB, outperformed other classifiers and established themselves as the most effective ensemble learning approaches for predicting cardiovascular disease. RF demonstrated similar performance to the boosting classifiers. The LR meta-learner offered competitive performance, although less effective.

Figure 3 illustrates the True Positive Rate (TPR) plotted against the False Positive Rate (FPR) on the ROC curves for the classifiers. All ensemble learning classifiers exhibited high predictive capacity, with points located near or within the upper region (0,1) signifying faultless classification. The boosting classifiers achieved the highest AUC-ROC performance with 0.80. RF (bagging) also achieved an AUC-ROC of 0.80, which is comparable with boosting classifiers, whereas the stacking LR meta-learner also achieved strong performance with an AUC-ROC of 0.79. Overall, ROC curves indicate that boosting and RF classifiers performed better than stacking classifiers, offering robust predictive accuracy for CVD prediction.

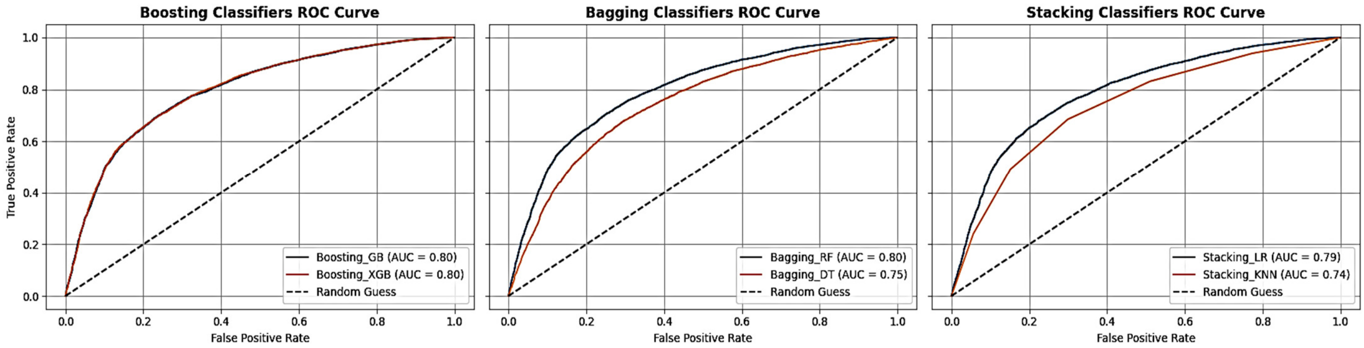

Fig. 3. Comparison of ROC curves among ensemble learning approaches.

IV. DISCUSSION

Comparative analysis of ensemble learning approaches can shed light on the advantages and drawbacks of boosting, bagging, and stacking classifiers in CVD prediction. Boosting and bagging surpassed stacking in all performance criteria, albeit with slight differences. XGBoost and GB demonstrated superior performance, achieving accuracies of 73.0% and 72.9%, respectively, along with the highest precisions of 75.1% and 74.9%, respectively. Both classifiers' ROC curves confirmed this performance, reaching AUC scores of 0.80, which demonstrates their competence in discriminating positive and negative classifications. However, their recall values were marginally superior to those of the DT bagging meta-learner, suggesting that while boosting enhances overall accuracy and precision, it may sacrifice some sensitivity in the classification of positive cases.

The RF approach (bagging) achieved competitive performance, with RF outperforming DT, achieving an accuracy of 72.7% and an AUC of 0.80, demonstrating comparability to boosting approaches. This aligns with the resilience of RF as a formidable baseline ensemble approach, capable of balancing and reducing variation. In contrast, the DT bagging meta-learner exhibited much inferior performance (accuracy = 69.0%, AUC = 0.75), despite its competitive recall.

LR, used as a stacking classifier, achieved balanced performance results, equal to the leading boosting and bagging approaches. In contrast, the KNN stacking meta-learner had an inferior prediction accuracy, which can be attributed to its data sensitivity and lack of resilience.

These results have practical relevance for the implementation of predictive models in decision support systems. The effectiveness of boosting and bagging classifiers in CVD prediction demonstrates their appropriateness for applications requiring stability and generalization, including medical diagnosis, financial risk assessment, or fraud detection. High precision values indicate a reduction in FP, which is essential in clinical and operational contexts, where unwarranted interventions or misclassifications might incur high costs. Ensemble learning approaches, especially boosting and RF (bagging), suggest a solid basis for developing reliable and interpretable predictive classifiers in practical settings.

V. CONCLUSION

This study evaluated the efficacy of three ensemble learning approaches (bagging, boosting, and stacking) using various performance metrics. The results indicate that boosting and bagging (RF) classifiers continuously achieve superior accuracy, precision, and AUC values, which affirms their robustness and dependability for classification applications. The results of this study enhance the understanding of how hybrid feature-ensemble models might optimize performance while mitigating overfitting and dimensionality concerns by systematically evaluating various ensemble paradigms within a unified experimental framework. The findings indicate that boosting methods achieved the best performance, showcasing their robustness and suitability for practical applications.

However, the analysis was restricted to a singular dataset, hence constraining generalizability. Further research into automated hyperparameter optimization, Explainable AI (XAI) techniques, and empirical clinical validation would improve both efficacy and comprehensibility. Moreover, further studies need to integrate additional performance metrics, including specificity, Matthews Correlation Coefficient (MCC), and calibration analysis, to provide a more thorough assessment of model reliability. The use of temporal or longitudinal health data may enhance early detection and individualized CVD risk evaluation.

ACKNOWLEDGMENT

The authors acknowledge the Ministry of Higher Education Malaysia and Universiti Teknologi MARA for their financial support to this project under the Fundamental Research Grant Scheme (FRGS) Grant No. FRGS/1/2023/ICT01/UITM/02/4. They also thank the Faculty of Computer and Mathematical Sciences, Universiti Teknologi MARA, Selangor, Malaysia, for all the support.

REFERENCES

[1] D. Adhikary, S. Barman, R. Ranjan, and H. Stone, "A Systematic Review of Major Cardiovascular Risk Factors: A Growing Global Health Concern," *Cureus*, Oct. 2022, https://doi.org/10.7759/cureus.30119.

[2] F. Rezaei *et al.*, "Comparison of laboratory-based and non-laboratory-based WHO cardiovascular disease risk charts: a population-based study," *Journal of Translational Medicine*, vol. 20, no. 1, Mar. 2022, Art. no. 133, https://doi.org/10.1186/s12967-022-03336-4.

[3] C. Bernand, E. Mirand, and M. Aryun, "Coronary heart disease prediction models using machine learning and deep learning algorithms," *AIP Conference Proceedings*, vol. 2838, no. 1, Surakarta, Indonesia, Feb. 2024, Art. no. 070009, https://doi.org/10.1063/5.0179929.

[4] C. Venkatesh, L. Sivayamini, T. Supriya, J. V. Kumar, B. V. K. Reddy, and N. Sujaritha, "A Hybrid Model for the Detection and Classification of Cardiovascular Diseases Based on Deep Learning and Optimization Techniques," in *Proceedings of the 6th International Conference on Communications and Cyber Physical Engineering*, Hyderabad, India, 2024, pp. 747–762, https://doi.org/10.1007/978-981-99-7137-4_73.

[5] K. Gunasekaran, V. D. A. Kumar, and K. Jayashree, "Prediction and risk analysis of Cardio Vascular diseases in IoHT by enhanced CHIO-based Residual and dilated gated network with Attention Mechanism," *Biomedical Signal Processing and Control*, vol. 94, Aug. 2024, Art. no. 106222, https://doi.org/10.1016/j.bspc.2024.106222.

[6] P. Dhaka, R. Sehrawat, and P. Bhutani, "An Innovative Approach to Cardiovascular Disease Prediction: A Hybrid Deep Learning Model," *Engineering, Technology & Applied Science Research*, vol. 13, no. 6, pp. 12396–12403, Dec. 2023, https://doi.org/10.48084/etasr.6503.

[7] N. K. Al-Shammari *et al.*, "Cardiac Stroke Prediction Framework using Hybrid Optimization Algorithm under DNN," *Engineering, Technology & Applied Science Research*, vol. 11, no. 4, pp. 7436–7441, Aug. 2021, https://doi.org/10.48084/etasr.4277.

[8] J. Azmi, M. Arif, M. T. Nafis, M. A. Alam, S. Tanweer, and G. Wang, "A systematic review on machine learning approaches for cardiovascular disease prediction using medical big data," *Medical Engineering & Physics*, vol. 105, July 2022, Art. no. 103825, https://doi.org/10.1016/j.medengphy.2022.103825.

[9] N. Biswas *et al.*, "Machine Learning-Based Model to Predict Heart Disease in Early Stage Employing Different Feature Selection Techniques," *BioMed Research International*, vol. 2023, no. 1, 2023, Art. no. 6864343, https://doi.org/10.1155/2023/6864343.

[10] S. Subramani *et al.*, "Cardiovascular diseases prediction by machine learning incorporation with deep learning," *Frontiers in Medicine*, vol.

10, Apr. 2023, Art. no. 1150933, https://doi.org/10.3389/fmed.2023.1150933.

[11] D. Sharma, N. Gotlieb, M. E. Farkouh, K. Patel, W. Xu, and M. Bhat, "Machine Learning Approach to Classify Cardiovascular Disease in Patients With Nonalcoholic Fatty Liver Disease in the UK Biobank Cohort," *Journal of the American Heart Association*, vol. 11, no. 1, Jan. 2022, Art. no. e022576, https://doi.org/10.1161/JAHA.121.022576.

[12] A. Khan, M. Qureshi, M. Daniyal, and K. Tawiah, "A Novel Study on Machine Learning Algorithm-Based Cardiovascular Disease Prediction," *Health & Social Care in the Community*, vol. 2023, pp. 1–10, Feb. 2023, https://doi.org/10.1155/2023/1406060.

[13] B. R. Devi, U. Sivaji, T. Swetha, J. Avanija, A. Suresh, and K. R. Madhavi, "Advanced Cardiovascular Disease Prediction: A Comparative Analysis of Ensemble Stacking and Deep Neural Networks," *International Journal of Intelligent Systems and Applications in Engineering*, vol. 12, no. 6, pp. 46–55, 2024.

[14] I. Yekkala, S. Dixit, and M. A. Jabbar, "Prediction of heart disease using ensemble learning and Particle Swarm Optimization," in *2017 International Conference On Smart Technologies For Smart Nation (SmartTechCon)*, Bengaluru, India, Dec. 2017, pp. 691–698, https://doi.org/10.1109/SmartTechCon.2017.8358460.

[15] X. Y. Gao, A. A. Ali, H. S. Hassan, and E. M. Anwar, "Improving the Accuracy for Analyzing Heart Diseases Prediction Based on the Ensemble Method," *Complexity*, vol. 2021, no. 1, Jan. 2021, Art. no. 6663455, https://doi.org/10.1155/2021/6663455.

[16] D. Zhou, H. Qiu, L. Wang, and M. Shen, "Risk prediction of heart failure in patients with ischemic heart disease using network analytics and stacking ensemble learning," *BMC Medical Informatics and Decision Making*, vol. 23, no. 1, May 2023, Art. no. 99, https://doi.org/10.1186/s12911-023-02196-2.

[17] P. Gupta and D.D. Seth, "Improving the Prediction of Heart Disease Using Ensemble Learning and Feature Selection," *International Journal of Advances in Soft Computing and its Applications*, vol. 14, no. 2, pp. 37–40, Aug. 2022, https://doi.org/10.15849/IJASCA.220720.03.

[18] K. Shilpa and T. Adilakshmi, "Applying Ensemble Techniques of Machine Learning to Predict Heart Disease," in *Proceedings of the International Conference on Cognitive and Intelligent Computing*, A. Kumar, G. Ghinea, S. Merugu, and T. Hashimoto, Springer Nature Singapore, 2022, pp. 775–783.

[19] E. Akkur, "Prediction of Cardiovascular Disease Based on Voting Ensemble Model and SHAP Analysis," *Sakarya University Journal of Computer and Information Sciences*, vol. 6, no. 3, pp. 226–238, Dec. 2023, https://doi.org/10.35377/saucis...1367326.

[20] N. Tabassum and T. Ahmed, "A theoretical study on classifier ensemble methods and its applications," in *2016 3rd International Conference on Computing for Sustainable Global Development (INDIACom)*, New Delhi, India, Mar. 2016, pp. 374–378.

[21] T. N. Rincy and R. Gupta, "Ensemble Learning Techniques and its Efficiency in Machine Learning: A Survey," in *2nd International Conference on Data, Engineering and Applications (IDEA)*, Bhopal, India, Feb. 2020, pp. 1–6, https://doi.org/10.1109/IDEA49133.2020.9170675.

[22] V. Chang, V. R. Bhavani, A. Q. Xu, and M. Hossain, "An artificial intelligence model for heart disease detection using machine learning algorithms," *Healthcare Analytics*, vol. 2, Nov. 2022, Art. no. 100016, https://doi.org/10.1016/j.health.2022.100016.

[23] A. Ogunpola, F. Saeed, S. Basurra, A. M. Albarrak, and S. N. Qasem, "Machine Learning-Based Predictive Models for Detection of Cardiovascular Diseases," *Diagnostics*, vol. 14, no. 2, Jan. 2024, Art. no. 144, https://doi.org/10.3390/diagnostics14020144.

[24] M. Pal, S. Parija, G. Panda, K. Dhama, and R. K. Mohapatra, "Risk prediction of cardiovascular disease using machine learning classifiers," *Open Medicine*, vol. 17, no. 1, pp. 1100–1113, June 2022, https://doi.org/10.1515/med-2022-0508.

[25] M. Shah, K. Gandhi, K. A. Patel, H. Kantawala, R. Patel, and A. Kothari, "Theoretical Evaluation of Ensemble Machine Learning Techniques," in *2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT)*, Tirunelveli, India, Jan. 2023, pp. 829–837, https://doi.org/10.1109/ICSSIT55814.2023.10061139.

[26] V. J. Prakash and N. K. Karthikeyan, "Enhanced Evolutionary Feature Selection and Ensemble Method for Cardiovascular Disease Prediction," *Interdisciplinary Sciences: Computational Life Sciences*, vol. 13, no. 3, pp. 389–412, Sept. 2021, https://doi.org/10.1007/s12539-021-00430-x.

[27] N. V. MahaLakshmi and R. K. Rout, "An intelligence method for heart disease prediction using integrated filter-evolutionary search based feature selection and optimized ensemble classifier," *Multimedia Tools and Applications*, vol. 83, no. 13, pp. 39841–39865, Oct. 2023, https://doi.org/10.1007/s11042-023-16924-3.

[28] M. R. Islam, Md. Durul Hoda, Md. A. Rashid, S. Alam Suha, and M. T. Islam Miya, "Data-Driven Heart Disease Prediction by Ensemble Feature Selection and Machine Learning Techniques," in *2022 25th International Conference on Computer and Information Technology (ICCIT)*, Cox's Bazar, Bangladesh, Dec. 2022, pp. 575–580, https://doi.org/10.1109/ICCIT57492.2022.10054998.

[29] J. Chen, H. Huang, A. G. Cohn, D. Zhang, and M. Zhou, "Machine learning-based classification of rock discontinuity trace: SMOTE oversampling integrated with GBT ensemble learning," *International Journal of Mining Science and Technology*, vol. 32, no. 2, pp. 309–322, Mar. 2022, https://doi.org/10.1016/j.ijmst.2021.08.004.

[30] J. Zhang, T. Wang, W. W. Y. Ng, and W. Pedrycz, "Ensembling perturbation-based oversamplers for imbalanced datasets," *Neurocomputing*, vol. 479, pp. 1–11, Mar. 2022, https://doi.org/10.1016/j.neucom.2022.01.049.

[31] A. M. Sowjanya and O. Mrudula, "Effective treatment of imbalanced datasets in health care using modified SMOTE coupled with stacked deep learning algorithms," *Applied Nanoscience*, vol. 13, no. 3, pp. 1829–1840, Mar. 2023, https://doi.org/10.1007/s13204-021-02063-4.

[32] C. S. Tarimo, S. S. Bhuyan, Q. Li, W. Ren, M. J. Mahande, and J. Wu, "Combining Resampling Strategies and Ensemble Machine Learning Methods to Enhance Prediction of Neonates with a Low Apgar Score After Induction of Labor in Northern Tanzania," *Risk Management and Healthcare Policy*, vol. 14, pp. 3711–3720, Sept. 2021, https://doi.org/10.2147/RMHP.S331077.

[33] S. Ulianova, "Cardiovascular Disease dataset." Kaggle, Nov. 04, 2025, Available: https://www.kaggle.com/datasets/sulianova/cardiovascular-disease-dataset.