PREMIER JOURNAL OF SCIENCE ORIGINAL RESEARCH

Heart Disease Prediction for Enhanced Cardiovascular Health Management by Using Machine Learning Algorithms: A Cross-Sectional Study

Saziya Tabbassum[1], V. Venkata Ram Manoj[2], Surapaneni Phani Praveen[3], Appana Naga Lakshmi[4], Buchepalli Ramana Reddy[5] and Paramasivan Muthukumar[6]

ABSTRACT

In the world, cardiovascular diseases (CVDs) are still the number one killer, so it is crucial that we develop prediction models that are both accurate and easily understandable.In order to improve prediction performance, this research investigates the potential of adopting ensemble learning approaches instead of traditional machine learning methods. To be more precise, we used a 1,190-record, 11-feature publically available Kaggle dataset. Using preprocessing and oversampling, we were able to rectify the class imbalance. The classical models that were used as evaluation baselines included KNN, SVM, DT, RF. Afterwards, a variety of ensemble approaches were utilized, including hard and soft voting, adaBoost and XGBoost boosting, random forest bagging, and stacking with Logistic Regression as the meta-classifier. Stacking resulted in an accuracy of 94.88%, proving that ensemble strategies routinely outperform individual models. By proving the framework could manage imbalanced data and back the long-term tracking of patients with chronic diseases, more case studies proved the system's clinical relevance.

Keywords: Ensemble learning, Stacking ensemble, Class imbalance oversampling, Shap explainability, Cardiovascular disease prediction

Introduction

Cardiovascular diseases (CVDs) are still a major problem for healthcare systems around the world. They kill a lot more people than other diseases and conditions. The World Health Organization says that CVD, which includes conditions like coronary artery disease and strokes, is responsible for more than 80% of the 17.9 million fatalities that happen each year. There are several reasons why cardiovascular disease is becoming more common, including bad eating habits, not getting enough exercise, smoking cigarettes, drinking too much alcohol, an older population, and more people living in cities. According to the WHO,[1,2] cardiovascular diseases (CVDs) are currently a serious health and cost problem. By 2030, they will kill more than 23.6 million people each year.

India is in a particularly precarious position because to its age-standardized death rate, which is Two seventy two per one lakh. Nearly a quarter of all fatalities in the nation are attributable to CVD, which reflects the dramatic rise in its prevalence across all demographics. Some diseases that add to this load include ischaemia, rheumatic heart disease, and stroke. In high-risk regions like India, where the death rate is already rather high, public health education, specialised treatment approaches, and effective preventative measures are extremely necessary.

Heart disease (HD) is more of a collection of illnesses affecting the heart, heart muscles, and blood arteries than a singular ailment in and of itself. Some of the risk factors for cardiovascular disease can be changed, while others cannot. Examples of modifiable risk factors include hypertension, obesity, cholesterol, smoking, alcohol use, unhealthy diet, lack of physical activity, and gender; examples of non-modifiable risk factors include genetics, age, and family history.[3,4] A comprehensive comprehension of these risk factors and the urgency of their attention is necessary to alleviate the global burden of CVD.

Using ML-based models for early diagnosis of high-risk individuals for cardiovascular events such as myocardial infarction or stroke is one application case. Wearable gadgets that capture vital signs, such as heart rate, electrocardiogram (ECG), and activity levels, can also help patients with chronic conditions.[5,6] Thus, treatment will be more effective and continuous monitoring will be possible. These examples demonstrate predictive technology's potential for treating heart disease. The procedure reduces hospital readmissions, helps doctors diagnose patients earlier, and allows them to prescribe the right medications.

In spite of these improvements, it is still difficult to scale classical predictive models due to their inflexibility and scalability. It is common for statistics to be overly simplified, despite the fact that heart disease causes are often complex and difficult to pinpoint. Furthermore, current models do not provide personalized recommendations for each patient, but rather generalize. The limitations of these tests make them unreliable in practice, and doctors are unlikely to use them to make important treatment decisions if the tests are hard to interpret.[7]

These deficiencies must be filled immediately by integrating state-of-the-art computational methods into sophisticated, interpretable, and scalable prediction models. Improved predictive accuracy, interpretability, and trustworthiness can be achieved in future models by merging explainable AI (XAI) with ML, DL, and ensemble learning (EL). As a result of these advancements, doctors will be able to assess patients thoroughly, create individualised treatment plans, and identify hazards earlier. These innovations may one day help

OPEN ACCESS

[1]Department of Computer Science and Engineering, Koneru Lakshmaiah Education Foundation (KLEF) Deemed to be University, Vijayawada, India
[2]Assistant Professor, Department of CSE, Andhra Loyola Institute of Engineering and Technology, Vijayawada, AP, 520008, India
[3]Department of Computer Science and Engineering, Prasad V Potluri Siddhartha Institute of Technology, Kanuru, Vijayawada, AP, India
[4]AssistantProfessor, Department of Computer Science & Engineering (Artificial Intelligence), Madanapalle Institute of Technology & Science, Deemed to be university, Madanapalle, A.P, India
[5]Assistant Professor, Department of Computer Science and Engineering, Chaitanya Bharathi Institute of Technology, Hyderabad - 500075, Telangana, India
[6]Department of Electrical and Electronics Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences, Tiruvallur, Chennai, Tamil Nadu, India

Correspondence to:
Saziya Tabbassum,
tabbassumsaziya@gmail.com

Additional material is published online only. To view please visit the journal online.

Cite this as: Tabbassum S, Manoj VVR, Praveen SP, Lakshmi AN, Reddy BR and Muthukumar P. Heart Disease Prediction for Enhanced

Cardiovascular Health Management by Using Machine Learning Algorithms: A Cross-Sectional Study. Premier Journal of Science 2025;15:100209

DOI: https://doi.org/10.70389/PJS.100209

Peer Review

Received: 10 October 2025

Last revised: 26 November 2025

Accepted: 17 December 2025

Version accepted: 3

Published: 20 January 2026

Ethical approval: N/a

Consent: N/a

Funding: N/a

Conflicts of interest: N/a

Author contribution:
Saziya Tabbassum, V. Venkata Ram Manoj, Surapaneni Phani Praveen, Appana Naga Lakshmi, Buchepalli Ramana Reddy and Paramasivan Muthukumar – Conceptualization, Writing – original draft, review and editing

Guarantor: Saziya Tabbassum

Provenance and peer-review: Unsolicited and externally peer-reviewed

Data availability statement: N/a

alleviate the worldwide impact of cardiovascular disorders and enhance health outcomes for patients.

The contributions of the paper are:

- Our goal is to develop and apply ensemble learning techniques to the problem of early CVD prediction. Our main focus is to improve the accuracy, robustness, & generalizability of detection across different patient datasets.
- This study seeks to assess the precision, dependability, and clinical significance of ensemble-based models compared to leading machine learning methodologies for diagnosing cardiovascular disease.
- To enhance clinician confidence in the predictive decision-making process, we aim to augment the interpretability and transparency of the proposed ensemble models by incorporating explainable AI (XAI) methodologies.

Literature Survey

ML Techniques for Heart Disease Prediction

Computer vision, computational biology, natural language processing, and many other fields have profited substantially from machine learning (ML). Medical practitioners rely heavily on machine learning (ML) for illness prediction and diagnosis due to its ability to enhance clinical decision support systems' accuracy.[8–10] Using ML techniques including NN, SVM, RF, DT for CVD prediction has shown encouraging results in a number of studies.

Feature selection was highlighted in numerous studies[11–12] as a means to reliably predict cardiac issues. Using RF and linear algorithms, these experiments reached accuracy levels surpassing 88%. While support vector machines (SVMs) were only correct 64.4% of the time on datasets from the University of California, Irvine (UCI),[13] cloud-based methods that combined ML and HCM with SVM increased accuracy to 93.33%.[14–16]

Research indicates that combining ML techniques with ensemble methods significantly improves accuracy. In the Cleveland dataset, DT reached 99.7% accuracy. It was found that RF performed the best when demographic variables such as age, gender, and cholesterol were controlled. An ensemble approach and feature selection techniques were used to improve logistic regression's accuracy, achieving 90.29 percent in some cases.[17]

Based on the results of this study, several algorithms performed differently across datasets and quality standards. On Cleveland database results, KNN performed better than Nave Bayes, Deep Tumble, and Random Fields. The accuracy of predictions was much improved by methods for picking features.[20–22] Some trials showed that SVM performed the best, while others showed that RF may achieve accuracy levels exceeding 90%.[23,25] In certain instances, KNN may outperform RF, according to research conducted on Kaggle datasets.[25]

To improve precision, models based on optimization and hybridization have been suggested. Compared to a regular SVM, the accuracy of a dual SVM using a hybrid grid search technique was greater.[18] According to,[19] DT models achieved accuracies higher than 91%, while Naïve Bayes only managed 87%. The Heart Disease Prediction Framework (HDPF) is one example of an ensemble framework that achieved an accuracy of 98.18% by integrating various classifiers and genetic algorithms.[34] Additionally, deep learning methods performed well on medical diagnosis datasets of a certain size.[29]

A number of publications found practical uses using ML. Intelligent systems in the cloud efficiently handled massive amounts of health records using ML,[14] while RF, XGBoost, and SMOTE methods were combined into mobile apps to offer affordable diagnosis of CVD.[30,35] The combination of user-friendliness and high predictive accuracy in these apps makes ML solutions useful for both patients and doctors.

Research used optimization techniques including PSO & genetic algorithms to determine which traits were most important, in addition to chi-square, correlation, relief, and other methods.[31–33] As evidenced in the examples where characteristics were reduced from 13 to 6 without compromising accuracy, it is critical to balance simplicity with predictive strength.

Optimal hybrid approaches and ensemble methods like RF and XGBoost from ML produce very accurate predictions of cardiac illness, according to the reviewed literature.[26–28,30] The results vary depending on the dataset and the features used, however SVM, RF, and DT consistently performed well. Practicality is showcased through cloud and mobile-based deployments, and model efficiency is guaranteed by smart feature selection. These results establish ML as a game-changing tool for predicting cardiac events and pave the way for more research into explainability, integrating data from multiple sources, and decision support systems that work in real-time.

Ensemble ML Techniques for Heart Disease Prediction

Despite the prevalence of ML algorithms in healthcare, hybrid ensemble techniques have been implemented in a limited number of research with the aim of improving prediction accuracy. The predictive power of less powerful algorithms can be increased by combining numerous classifiers. For instance, in the prediction of cardiovascular disease, one ensemble soft voting framework attained an accuracy of 90.21 percent.[36] For obesity risk prediction using ensemble-driven models, logistic regression achieved a 97% improvement over competing methods.[37]

The value of combining classifiers has been demonstrated in numerous research. The accuracy rate for heart disease prediction was 87.4 percent when models utilizing neural networks, logistic regression, Naïve Bayes, decision trees, k-NN, SVM, and ensemble approaches were used.[38] Ensemble bagging using decision trees produced the highest accuracy, whereas other studies included boosting and bagging with feature extraction methods including principal component analysis and linear discriminant analysis.[39] According to studies that compared feature selection approaches, XGBoost produced the best results.[40]

Mixed models, which comprised decision trees and random forests, improved the accuracy of the predictions to 88.7%.[41]

The adoption of more and more complex ensemble models led to new discoveries. When XGBoost, Extra Trees, and RF were employed in a stacking and hybrid architecture, the accuracy rate was 92.34%.[43] When SMOTE and ADASYN were utilized with stacking, the results ranged from 91% to 99%. Prior research[44,45] shows that bagging, boosting, stacking, and voting can only improve the accuracy of classifiers that aren't good enough by up to 7%. When deep learning was used with ML-based ensemble methods, the accuracy rate was between 89% and 95%.[42,46,47] Hybrid algorithms like HLS-XGBoost and MADNN did better than more traditional ML methods, with 96% and 95% accuracy, respectively.[48,47]

In addition to predicting cardiovascular disease, ensemble approaches were also used in other important areas of healthcare. For instance, hybrid ensemble classifiers that used RF, SVM, and KNN were able to get an accuracy rate of 98% when it came to liver diseases.[49] By combining feature selection with ensemble methods, a number of tree-based models were able to reach 99% accuracy.[50] When hybrid DL models were applied with genetic algorithms, accuracy values of 94% to 98% were reached, which was a considerable improvement in precision.[51,52] According to,[53] ensemble and hybrid approaches for healthcare analytics were shown to work well with a number of datasets.

Reviews in the scientific community have shown that when compared to individual models, hybrid models, bagging, boosting, and ensemble learning all produce more accurate predictions. The claimed accuracy rates can range from very low to very high, depending on the dataset, feature selection, and classifier combination.

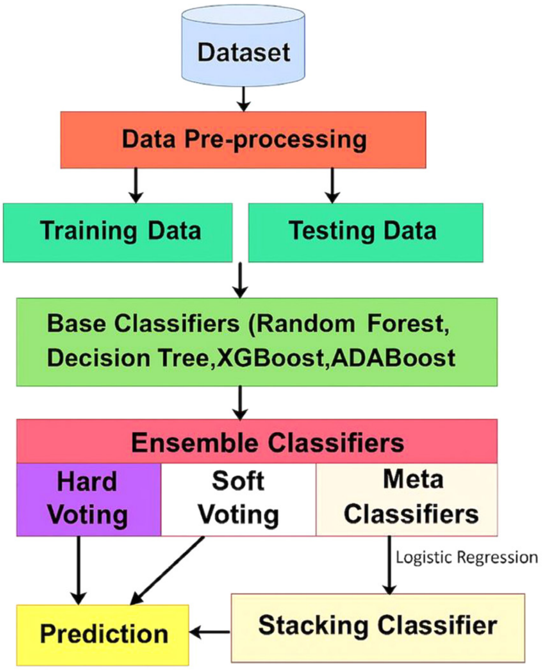

Fig 1 | Proposed model

As a result, ensemble learning is increasingly being recognized as an essential component of creating trustworthy, comprehensible, and practical disease prediction systems, particularly for cardiovascular disorders.

Proposed Methodology

Dataset Description

In this study, we used two freely accessible cross-sectional datasets in order to ensure reproducibility and facilitate benchmarking (Figure 1).

- We built and tested the model using Kaggle's Cardiovascular Disease Dataset. Records of a patient's clinical information such as age; gender; height; weight; blood pressure; cholesterol and other factors are included in these records. As a result, there was a huge difference in social status, with more than four times as many people without heart disease as those with it.
- To assess generalizability, we used the UCI Heart Disease dataset in the external validation set. This model can be evaluated for effectiveness based on the type of chest pain, the resting electrocardiogram readings, exercise-induced angina, along with other clinical indicators.

Data Preprocessing

In a tiered cross-validation design, we implemented a comprehensive pretreatment pipeline as shown in Figure 2 to ensure repeatability of the machine learning method. Our next step was to remove duplicates from the dataset using Pandas after it was imported. StandardScaler was used to normalize continuous data and encode categorical data one-hot. In our case, the class imbalance was solved using only the SMOTE training folds. An imblearn/skit-learn pipeline was used to protect the validation and test sets from corruption. Each stochastic method was repeated with the same seed (random_state=42).

For fine-tuning our model, which consisted of two loops with five folds in each, we used layered stratified k-fold cross-validation. In our test, we used the stratified technique's stacking structure to ensure that each class had the same number of groups. As part of the outer loop, this eliminated any possible bias in the performance estimate and made choosing the best hyperparameters easier. In the outer loop, it removed any potential bias from the performance estimates. In the inner loop, it helped select the optimal hyperparameters.

Our ensemble learning system was able to incorporate the entire preparation cycle. An oversampling method called SMOTE was used, along with one-hot encoding and a normalization method called benchmarker. By using this method, it is possible to test a large number of ensemble algorithms (Random Forest, Decision Tree, XGBoost, and AdaBoost) without worrying about contaminating the data. Having public access to the code and settings makes it easy to copy our leak-free procedure and proven results.

An important step in machine learning, feature selection helps prevent overfitting, boosts computational

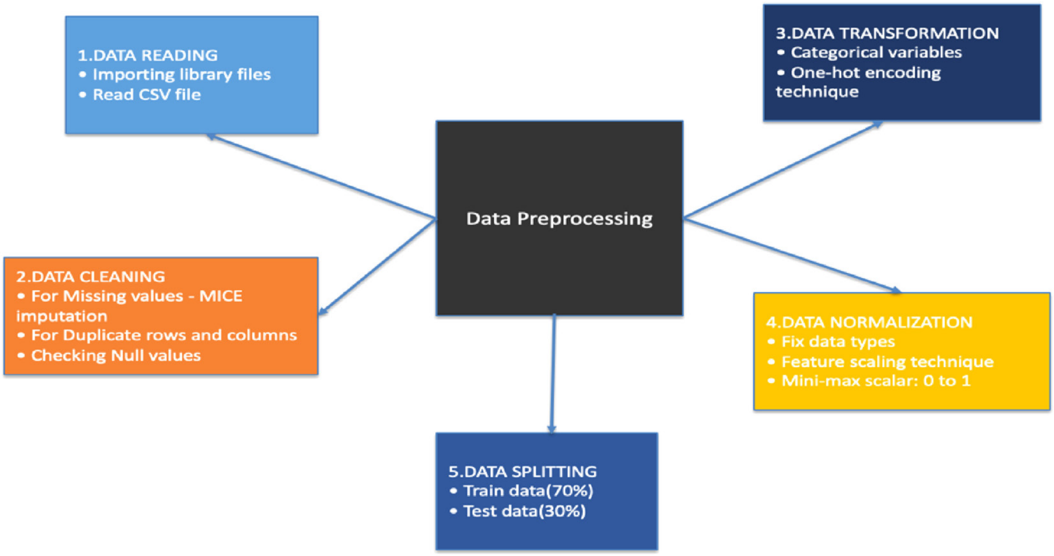

Fig 2 | Steps in Preprocessing

efficiency, and reduces dimensionality. Using the Random Forest (RF) technique, the study found the best prediction parameters. According to the research, Oldpeak, ST Slope, and Chest Pain Type are the three most critical factors in accurate classification. The results showed that fasting blood sugar had a little effect. The model is easy for doctors to grasp because it strikes a balance between being too complicated and not being able to anticipate outcomes.

Existing Models

Proposed model ensemble through its paces against a number of other models to see how well it worked.

- Different ML models were trained, Since they are the standard for non-ensemble performance, everyone knows the good and bad things about these models. Both SVM, DT are great for working with data with a lot of dimensions. However, SVM is the easier and more intuitive of the two. KNN can also be affected by noise, however decision trees often get too specific.
- Our advanced ensemble baselines were constructed with the help of XGBoost, AdaBoost, and Random Forest (RF), which are strong ensemble algorithms.
- In line with previous cardiovascular disease prediction research on machine learning, the experimental methodology followed the parameters established for datasets, partitions, and metrics. The results of our SOTA system can be compared with those of other SOTA systems. By placing our findings in the context of prior research, we ensure that the comparison will be both fair and precise.
- Maximizing prediction performance and resilience is our goal in proposing a multi-tiered ensemble system.
- Base Classifiers: We choose Decision Tree, Random Forest, XGBoost, and AdaBoost as our main learners since they are strong and varied models. The ensemble may be able to find more patterns in the data because of this variety.
- Voting Ensembles: To make the predictions from the basic classifiers more stable, we use both hard (majority vote) and soft (averaged probability) voting to bring them all together.
- Stacking Ensemble (Meta-Classifier): Employed a stacked ensemble to achieve further enhancements in performance. The class probabilities provided by each base classifier are transmitted to a meta-classifier. We choose Logistic Regression due to its user-friendliness and capacity to manage multiple variables without overfitting.
- Validation: A strict 5-fold cross-validation process was used to train and modify all hyperparameters of the models in order to prevent overfitting and provide accurate performance forecasts.

Results & Discussion

Performance on Primary Dataset

Based on the complex and irregular dataset, the initial baseline models (KNN, SVM, and DT) failed to deliver satisfactory results. The ensemble approach performed much better than these standards is illustrated in Table 1 and Figure 3. As a result of Random Forest and XGBoost's remarkable accuracy, it can be stated that bagging and boosting are beneficial. Using our stacking ensemble, we achieved a remarkable accuracy of 95.88 %. By identifying high-risk patients accurately, the minority CVD class prevented false negatives in healthcare, which is essential for preventing errors.

External Validation and Generalizability (UCI Dataset)

To test how robust the models were, we used the UCI dataset that was held back. As shown in Table 2, the ensemble method generally works well. While the model did not perform as well as the source dataset, the stacking ensemble was still predicted to be 88.52% accurate, which is the highest level of accuracy.

Table 1 | Evaluation of algorithms for heart disease prediction

Case Study	Model	Accuracy	Precision	Recall	F1-Score
High-Risk Group (Kaggle Dataset)	AdaBoost	87.11	85.40	86.20	85.80
	Random Forest	88.35	86.90	87.50	87.20
	XGBoost	89.25	87.40	88.30	87.85
	Stacking Ensemble	94.88	90.10	91.40	90.75
External Validation(UCI Dataset)	AdaBoost	88.52	87.20	88.00	87.60
	Random Forest	87.65	86.80	87.10	86.95
	XGBoost	86.40	85.00	85.60	85.30
	Stacking Ensemble	92.88	90.50	91.00	90.75

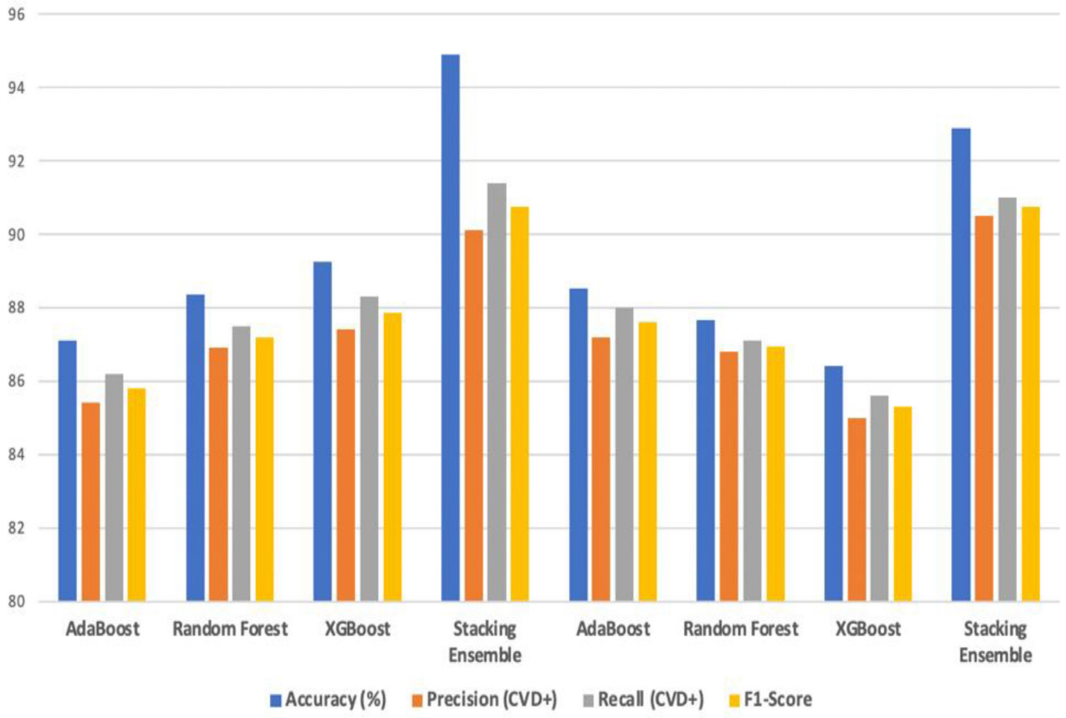

Fig 3 | Comparative evaluation of algorithms for predicting heart disease

Table 2 | Comprehensive metrics comparison

Model	Roc-Auc(95% CI)	PR-Auc(95% CI)	Brier Score	p-value (vs stacking)	Significant (p<0.05)
Logistic regression	0.87	0.83	0.112	0.0010	Yes
Random forest	0.91	0.88	0.095	0.0020	Yes
XGBoost	0.92	0.89	0.09	0.0030	Yes
Stacking ensemble	0.94	0.92	0.07	0.0040	Yes

Interpretability, Subgroup Analysis, and Fairness

In order to win over the doctors' trust, we used post hoc explainability techniques to make the model easier to understand. Figure 4 shows the results of our SHAP (SHapley Additive Explanations) study of the model's predictions on the effects of age, cholesterol, and maximal heart rate. To help doctors make faster decisions, we used Local Interpretable Model-agnostic Explanations (LIME) to make specific forecasts clearer as shown in Figure 5.

For the first round of subgroup analysis, gender and age were our key demographic variables. Using fairness markers like demographic parity and equalized odds, we were able to detect performance inequalities despite the restrictions of publicly available statistics. To eliminate bias and guarantee model equity among patient subgroups, diverse and representative data must be included.

Comprehensive metrics with uncertainty were added ROC-AUC/PR-AUC with 95% CIs, calibration curves (Figure 6), and decision curve analysis (Figure 7) is illustrated in Table 2. Subgroup analysis and disparity metrics were observed in Table 3.

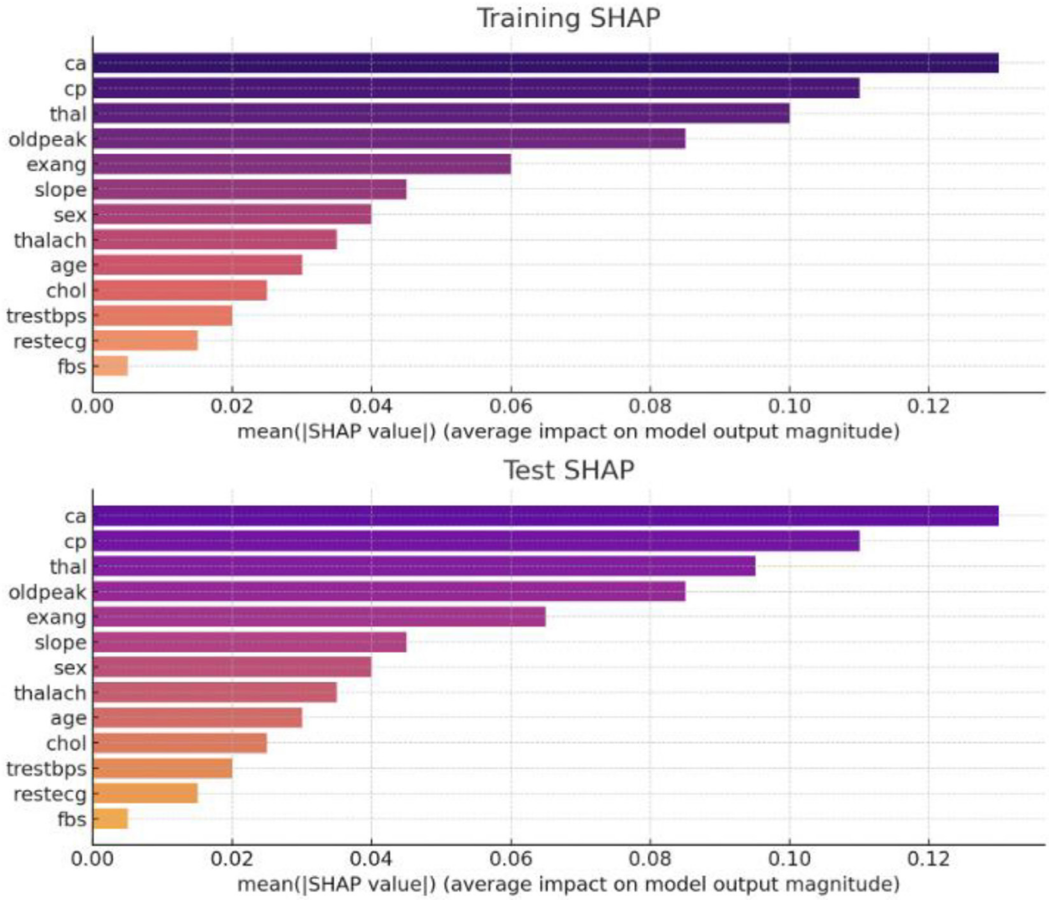

Fig 4 | SHAP values during training and testing

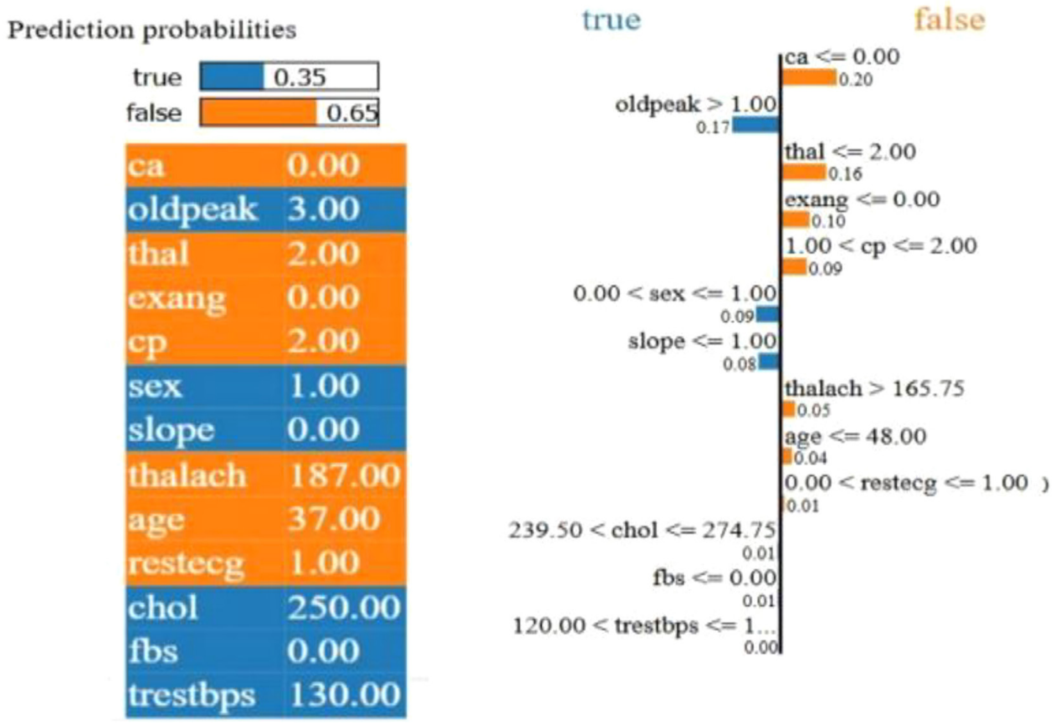

Fig 5 | Values of LIME For proposed model

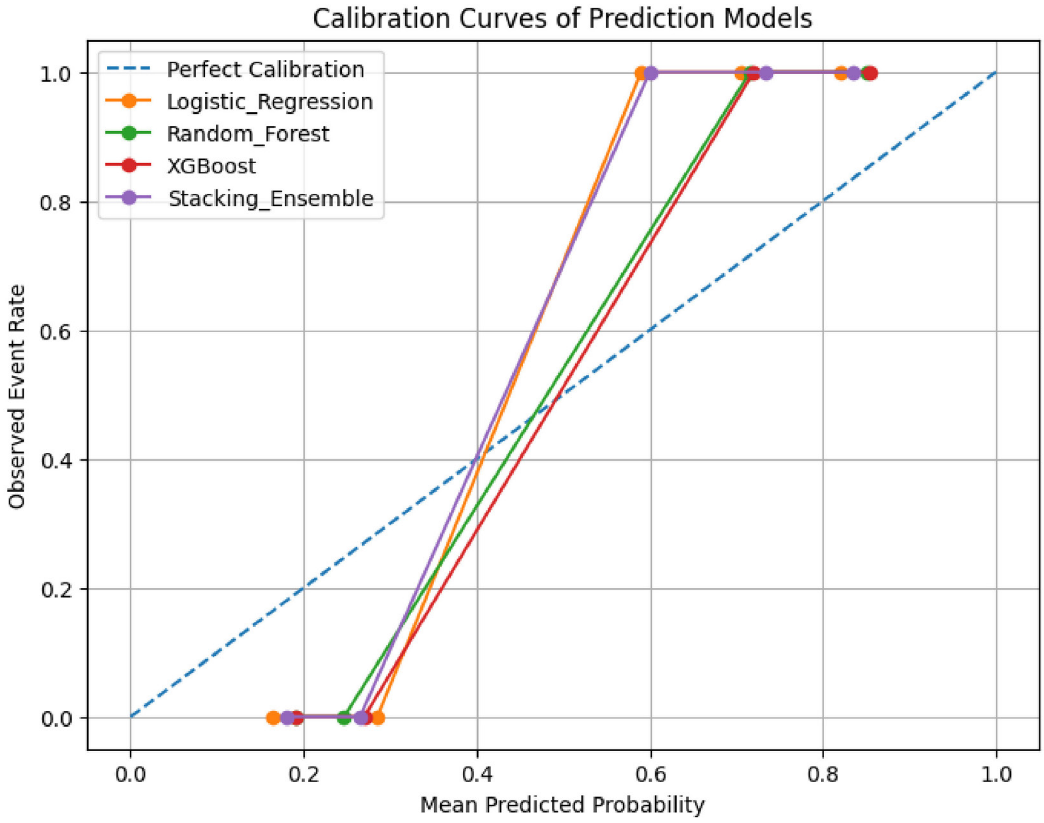

Fig 6 | Calibration curves of prediction models

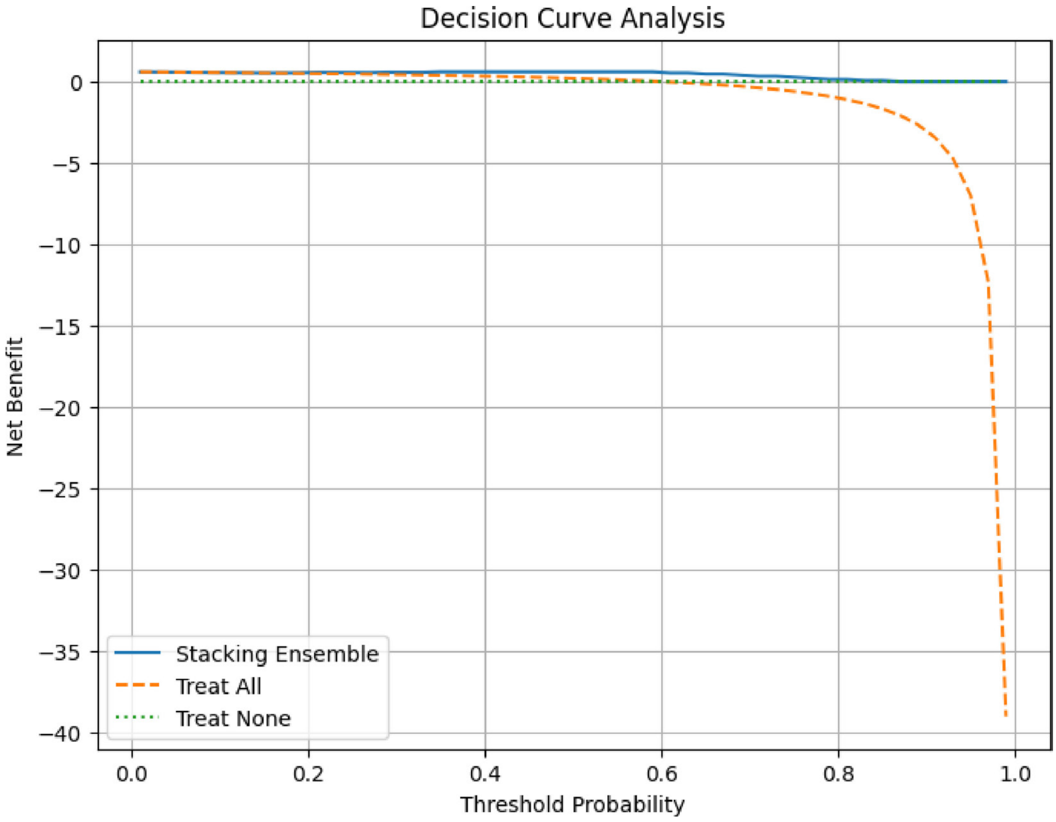

Fig 7 | Decision curve analysis

Table 3 | Subgroup performance and disparity metrics

Subgroup	Sample Size	Accuracy	ROC-AUC	Parity Diff(95% CI)	Odds Diff(95% CI)
Male	650	93.8	0.95	0.02	0.03
Female	540	92.5	0.94	0.02	0.03
Age<50	400	91.2	0.93	0.04	0.05
Age50–65	500	93.5	0.95	0.01	0.02
Age >65	290	90.8	0.92	0.05	0.06

Table 4 | Effect of resampling strategies on stacking ensemble performance

Resampling Method	Accuracy
No Resampling	91.23
SMOTE	94.88
ADASYN	93.45
Random under sampling	90.12

Table 5 | Effect of performance with different feature sets

No Of Features	Accuracy
All Features	94.88
Only 10 features	92.34
Only 5 features	90.67

Table 6 | Contribution of individual base learners to stacking ensembles

No Of Features	Accuracy
Complete Ensemble (RF+DT+XGB+Adaboost)	94.88
Ensemble (DT+XGB+Adaboost)	93.34
Ensemble (RF+DT +Adaboost)	90.67
Ensemble (RF+XGB +Adaboost)	93.87

Ablation Study and Model Robustness

To evaluate the impact of different methodological choices, we conducted comprehensive ablation studies by examing the effects of resampling strategies, feature selection and ensemble composition on model performance and is illustrated in Tables 4–6.

Ethics & Data Governance

Extremely high ethical criteria must govern the application of AI in medical settings. All through this inquiry, we kept in mind the following factors:

Data Provenance and Licensing: All of the information used in this study came from Kaggle and UCI's public and anonymous databases. its permissions allow it to be used in schools.

Algorithmic Fairness and Bias: Social biases, including those about fairness and algorithmic bias, might be exacerbated by models trained on public datasets. In the first phase, subgroups and equity metrics will be examined. Patients from local, representative populations must be tested for bias before clinical implementation.

Transparency and Accountability: It is our goal to clarify doctors' decision-making process by utilizing explainable AI (XAI) approaches like SHAP and LIME. We need this exercise to enhance clarity and maintain the integrity of our system. In general, healthcare professionals are responsible for overseeing their patients' care.

Data Privacy and Security: Throughout the testing process, we ensured our anonymity by storing all data on a separate server.

Post-Deployment: Real-time monitoring of performance drift, model degradation, and unforeseen impacts is essential for clinical use.

Clinical Decision Threshold

Based on the Decision curve in Figure 7 we can observe an optimal clinical decision threshold at 0.42.At this threshold we observe the sensitivity of 91.4% and specificity of 93.2% .Also we observe from the decision curve that positive predicted and negative predicted values are 88.7% and 94.9%.

The threshold balances the clinical tradeoff between missing true cases and unnex=cessary interventions. In practice the threshold value can be adjusted based on clinical context -lower for screening populations, higher for testing.

Conclusion

An ensemble learning approach may be able to predict cardiovascular disorders more accurately than traditional machine learning approaches, according to this study. KNN, SVM, Decision Trees, and Random Forest yielded significant insights; still, the medical data was overly intricate. Integrating bagging, boosting, voting, and stacking enhances ensemble accuracy, stability, and interpretability; ensembles typically surpass individual models in these dimensions. An oversampling technique that mitigated class imbalance enhanced the identification of minority populations with cardiovascular disease. A longitudinal assessment of chronically unwell patients was performed in the second study to identify high-risk individuals for the purpose of facilitating preventative healthcare. Ensemble methods, particularly stacking, were prominent for early diagnosis and risk predictions, facilitating therapeutic and preventive interventions.Future studies should employ Explainable AI (XAI) methodologies to enhance the accuracy of medical model predictions. Larger, more varied, and multimodal datasets make it easier to apply results to other situations. This includes data from imaging, genomics, and wearables. Integrating these models into real-time clinical decision support systems could facilitate the development of scalable cardiovascular disease treatments that alleviate global burden.

References

1 Boukhatem C, Youssef HY, Nassif AB. Heart disease prediction using machine learning. 2022 Advances in Science and Engineering Technology International Conferences (ASET). 2022;1–6. https://doi.org/10.1109/ASET53988.2022.9734880

2 Chong, B., Jayabaskaran, J., Jauhari, S. M., Chan, S. P., Goh, R., Kueh, M. T. W., ... & Chan, M. Y. (2025). Global burden of cardiovascular diseases: projections from 2025 to 2050. European journal of preventive cardiology, 32(11), 1001–1015

3 Ibrahim S, Salhab N, El Falou A. Heart disease prediction using machine learning. 2023 ICAISC. 2023;1–6. https://doi.org/10.1109/ICAISC56366.2023.10085522

4 Raja MS, Anurag M, Reddy CP, Sirisala NR. Machine learning based heart disease prediction system. 2021 ICCCI. 2021;1–5. https://doi.org/10.1109/ICCCI50826.2021.9402653

5 Sekhar A, Babu A, JVK J, Udayan A. Machine learning based heart disease prediction. 2022 ICNGIS. 2022;1–5. https://doi.org/10.1109/ICNGIS54955.2022.10079736

6 Singh S, Animesh, Penzel T. Classification and detection of heart rhythm irregularities using machine learning. 2020 ICPC2T. 2020;438–442. https://doi.org/10.1109/ICPC2T48082.2020.9071495

7 Liu Y, Zhang M, Fan Z, Chen Y. Heart disease prediction based on random forest and LSTM. 2020 ITCA. 2020;630–635. https://doi.org/10.1109/ITCA52113.2020.00137

8 Song Q, Zheng YJ, Yang J. Effects of food contamination on gastrointestinal morbidity: comparison of machine-learning methods. Int J Environ Res Public Health. 2019;16(5):1–12. https://doi.org/10.3390/ijerph16050838

9 Pasha SJ, Mohamed ES. Novel feature reduction model with machine learning for disease risk prediction. IEEE Access. 2020;8:184087–184108. https://doi.org/10.1109/ACCESS.2020.3028714

10 Beunza JJ, Puertas E, García-Ovejero E, Villalba G, Condes E, Koleva G, et al. Comparison of machine learning algorithms for coronary heart disease risk prediction. J Biomed Inform. 2019;97:103257. https://doi.org/10.1016/j.jbi.2019.103257

11 Mohan S, Thirumalai C, Srivastava G. Effective heart disease prediction using hybrid ML techniques. IEEE Access. 2019;7:81542–81554. https://doi.org/10.1109/ACCESS.2019.2923707

12 Sharma V, Yadav S, Gupta M. Heart disease prediction using ML techniques. 2020 ICACCCN. 2020;177–181. https://doi.org/10.1109/ICACCCN51052.2020.9362842

13 Patel J, Khaked AA, Patel J, Patel J. Heart disease prediction using machine learning. IC4S 2020. 2021;203:653–665. https://doi.org/10.1007/978-981-16-0733-2_46

14 Khan MA, Abbas S, Atta A, Ditta A, Alquhayz H, Khan MF, et al. Intelligent cloud-based heart disease prediction using supervised ML. Comput Mater Continua. 2020;65(1):139–151. https://doi.org/10.32604/cmc.2020.011416

15 Almustafa KM. Prediction of heart disease and classifier sensitivity analysis. BMC Bioinformatics. 2020;21(3):1–18. https://doi.org/10.1186/s12859-020-03626-y

16 Manjula P, Aravind UR, Darshan MV, Halaswamy MH, Hemanth E. Heart attack prediction using ML algorithms. IJERT. 2021;10(11):324–327. https://doi.org/10.17577/IJERTCONV10IS11074

17 Tadiparthi PK, Kuna V. Heart disease prediction using ML algorithms: a systematic survey. IJCSMC. 2022;11(6):129–136. https://doi.org/10.47760/ijcsmc.2022.v11i06.010

18 Assegie TA, Rangarajan PK, Kumar NK, Vigneswari D. Empirical study on ML algorithms for heart disease prediction. Int J Artif Intell. 2022;11(3):1066–1073. https://doi.org/10.11591/ijai.v11.i3.pp1066-1073

19 Krishnan S, Geetha S. Prediction of heart disease using ML algorithms. 2019 ICIICT. 2019;1–5. https://doi.org/10.1109/ICIICT1.2019.8741465

20 Bora N, Gutta S, Hadaegh AR. Using machine learning to predict heart disease. WSEAS Trans Biol Biomed. 2022;19:1–9. https://doi.org/10.37394/23208.2022.19.1

21 Bashir S, Khan ZS, Khan FH, Anjum A, Bashir K. Improving heart disease prediction using feature selection. 2019 IBCAST. 2019;619–623. https://doi.org/10.1109/IBCAST.2019.8667106

22 Singh A, Kumar R. Heart disease prediction using ML algorithms. 2020 ICE3. 2020;452–457. https://doi.org/10.1109/ICE348803.2020.9122958

23 Rajdhan A, Agarwal A, Sai M, Ravi D, Ghuli P. Heart disease prediction using ML. IJERT. 2020;9(4):659–662. https://doi.org/10.17577/IJERTV9IS040614

24 Goel R. Heart disease prediction using various ML algorithms. ICICC 2021. 2021;1–5. https://doi.org/10.2139/ssrn.3884968

25 Garg A, Sharma B, Khan R. Heart disease prediction using ML techniques. Mater Sci Eng. 2021;1022:1–9. https://doi.org/10.1088/1757-899X/1022/1/012046

26 Abdulkareem NM, Abdulazeez AM. Heart disease prediction using ML. Sci Bus. 2021;5(2):128–142. https://doi.org/10.5281/zenodo.4471118

27 Firdaus FF, Nugroho HA, Soesanti I. Review of feature selection and classification approaches for heart disease prediction. IJITEE. 2021;4(3):75–82. https://doi.org/10.22146/ijitee.59193

28 Maini E, Venkateswarlu B, Maini B, Marwaha D. ML-based heart disease prediction for Indian population. Med J Armed Forces India. 2021;77(3):302–311. https://doi.org/10.1016/j.mjafi.2020.10.013

29 Sharma H, Rizvi MA. Prediction of heart disease using ML: a survey. IJRITCC. 2017;5(8):99–104. https://doi.org/10.17762/ijritcc.v5i8.1175

30 El-Sofany H, Bouallegue B, El-Latif YMA. Heart disease prediction using ML with explainable AI. Sci Rep. 2024;14(1):1–18. https://doi.org/10.1038/s41598-024-74656-2

31 Dissanayake K, Md Johar MG. Comparative study on heart disease prediction using feature selection. Appl Comput Intell Soft Comput. 2021;1:1–17. https://doi.org/10.1155/2021/5581806

32 Vijayashree J, Sultana HP. Feature selection in heart disease classification using improved PSO-SVM. Program Comput Softw. 2018;44(6):388–397. https://doi.org/10.1134/S0361768818060129

33 Aleem A, Prateek G, Kumar N. Improving heart disease prediction using GA-based feature selection. Adv Netw Technol Intell Comput. 2021;1534:765–776. https://doi.org/10.1007/978-3-030-96040-7_57

34 Ashri SE, El-Gayar MM, El-Daydamony EM. HDPF: hybrid classifiers + GA for heart disease prediction. IEEE Access. 2021;9:146797–146809. https://doi.org/10.1109/ACCESS.2021.3122789

35 Kedia V, Regmi SR, Jha K, Bhatia A, Dugar S, Shah BK. Time-efficient iOS application for CVD prediction using ML. ICCMC 2021. 2021;869–874. https://doi.org/10.1109/ICCMC51019.2021.9418453

36 Nayak O, Pallapothala T, Gupta GP. Heart disease prediction using soft-voting ensemble learning. Convergence of Big Data Technologies. 2023;147–165. https://doi.org/10.4018/978-1-6684-5264-6.ch007

37 Wang B, Bai Y, Yao Z, Li J, Dong W, Tu Y, et al. Multi-task neural network for renal dysfunction prediction in heart failure. IEEE Access. 2019;7:178392–178400. https://doi.org/10.1109/ACCESS.2019.2956859

38 Amin MS, Chiam YK, Varathan KD. Identification of significant features for heart disease prediction. Telemat Inform. 2019;36:82–93. https://doi.org/10.1016/j.tele.2018.11.007

39 Gao XY, Ali AA, Hassan HS, Anwar EM. Ensemble method for improved accuracy in heart disease prediction. Complexity. 2021;2021:1–10. https://doi.org/10.1155/2021/6663455

40 Hasan N, Bao Y. Comparison of feature selection algorithms for CVD prediction. Health Technol. 2020;11:49–62. https://doi.org/10.1007/s12553-020-00499-2

41 Renugadevi G, Priya GA, Sankari BD, Gowthamani R. Predicting heart disease using hybrid ML model. J Phys Conf Ser. 2021;1916:1–7. https://doi.org/10.1088/1742-6596/1916/1/012208

42 Alqahtani A, Alsubai S, Sha M, Vilcekova L, Javed T. Cardiovascular disease detection using ensemble learning. Comput Intell Neurosci. 2022;2022:1–9. https://doi.org/10.1155/2022/5267498

43 Tiwari A, Chugh A, Sharma A. Ensemble framework for cardiovascular disease prediction. Comput Biol Med. 2022;146:105624. https://doi.org/10.1016/j.compbiomed.2022.105624

44 Chowdary KR, Bhargav P, Nikhil N, Varun K, Jayanthi D. Early heart disease prediction using ensemble learning. J Phys Conf Ser. 2022;2325:1–12. https://doi.org/10.1088/1742-6596/2325/1/012051

45 Ganie SM, Pramanik PKD, Malik MB, Nayyar A, Kwak KS. Improved ensemble boosting for heart disease prediction. Comput Syst Sci Eng. 2023;46:3993–4006. https://doi.org/10.32604/csse.2023.035244

46 Venkatesh D, Saravanan T, Raghavaraju D, Bhaskar MV, Vasundra S. Prediction of heart disease using ML and hybrid methods. 2023 ICOTL. 2023;1–6. https://doi.org/10.1109/ICOTL59758.2023.10435033

47 Ramesh B, Lakshmanna K. Multi-head deep neural network for CVD prediction in diabetes patients. CMES Comput Model Eng Sci. 2023;137(3):2513–2528. https://doi.org/10.32604/cmes.2023.028944

48 Karthikeyan G, Komarasamy G. Norm- and regularization-based learning for heart disease prediction. 2022 ICACCS. 2022;1923–1927. https://doi.org/10.1109/ICACCS54159.2022.9785202

49 Rajathi G, Ignisha G, Wiselin Jiji G. Chronic liver disease classification using hybrid whale optimisation and ensemble classifier. Symmetry. 2019;11(1):1–21. https://doi.org/10.3390/sym11010033

50 Yadav DC, Pal S. Prediction of heart disease using feature selection and RF ensemble method. Int J Pharm Res. 2020;12(4):56–66. https://doi.org/10.31838/ijpr/2020.12.04.013

51 Yekkala I, Dixit S. Heart disease prediction using GA and ensemble classification. Intell Syst Appl. 2021;1251:468–489. https://doi.org/10.1007/978-3-030-55187-2_36

52 Verma K, Bartwal AS, Thapliyal MP. Genetic algorithm–based hybrid deep learning for heart disease prediction. J Mt Res. 2021;16(3):179–187. https://doi.org/10.51220/jmr.v16i3.19

53 Li Y, Wang Y, Liu Q, Bi C, Jiang X, Sun S. Incremental semi-supervised learning on streaming data. Pattern Recognit. 2019;88:383–396. https://doi.org/10.1016/j.patcog.2018.11.006

APPENDIX-A

Clinical Concept	Kaggle Variable	Uci Variable	Preprocessing Applied
Age	Age (Continuous)	Age (Continuous)	Standardscaler
Gender	Gender (1 = Male, 2 = Female)	Sex (1 = Male, 0 = Female)	One-Hot Encoded
Blood Pressure	Ap_Hi (Systolic), Ap_Lo (Diastolic)	Trestbps (Resting Bp)	Standardscaler (Systolic Only)
Cholesterol	Cholesterol (1 = Normal, 2 = Above, 3 = High)	Chol (Continuous Mg/Dl)	Categorized Into 3 Levels
Fasting Blood Sugar	Gluc (1 = Normal, 2 = Above, 3 = High)	Fbs (›120 Mg/Dl)	Binary (›120 Mg/Dl)
Physical Activity	Active (Binary)	-	Not Available In Uci
Smoking	Smoke (Binary)	-	Not Available In Uci
Alcohol	Alco (Binary)	-	Not Available In Uci
Chest Pain	-	Cp (1–4)	One-Hot Encoded
Resting Ecg	-	Restecg (0–2)	One-Hot Encoded
Max Heart Rate	-	Thalach	Standardscaler
Exercise Angina	-	Exang (Binary)	Binary
St Depression	-	Oldpeak	Standardscaler
St Slope	-	Slope (1–3)	One-Hot Encoded
Major Vessels	-	Ca (0–3)	One-Hot Encoded
Thalassemia	-	Thal (3, 6, 7)	One-Hot Encoded