

Heart Disease Prediction with Pairwise and Three Model Stacking – An Analysis

Gnana Prakasi Oliver Sirya Pushpam*, Diana Jeba Jingle Issac,
Kanmani Palanisamy, Pari Agarwal, Rashi Dubey and Aryaman Kant

Department of Computer Science and Engineering, Christ University, Bengaluru, India.

*Corresponding Author E-mail: gnana.prakasi@christuniversity.in

<https://dx.doi.org/10.13005/bpj/3507>

(Received: 15 April 2025; accepted: 02 April 2026)

Coronary heart disease is one of the most common causes of death in the world. Early diagnosis helps us to provide better treatment, which results in the reduction of the mortality rate. Though many machine learning algorithms are in research to predict heart disease, the accuracy in the prediction is still low. To overcome these drawbacks, we apply bagging and boosting techniques to these Machine Learning algorithms. In addition, we hybridize these models by pairwise stacking and three-layer stacking to improve the performance of predication and classification. Results show that the stacking of Machine Learning algorithms with Random Forest gives better results in terms of prediction and classification.

Keywords: Bagging, Boosting, Coronary Heart Disease, Decision Tree, Machine Learning, Stacking.

Coronary heart disease is an important public health issue since it is one of the top causes of mortality in all ages worldwide. Early detection of heart disease is required for an early treatment process and to reduce the complexity of treatment. These heart diseases are of different forms and the severity also changes with respect to time. In order to start the treatment at right time, it is necessary to classify the type of heart disease and its severity. Most of the analysis and prediction are mainly based on the description of the basic parameters from the clinical record. This restricted emphasis makes it more difficult to offer comprehensive insights into the wide range of heart-related conditions and prevents patients from receiving individualized treatment plans.

The commonly used machine learning algorithms in heart disease predication are Support Vector Machines (SVM), Decision Trees, and Logistic Regression. But it is difficult for these models to accurately classify the different types of heart diseases and the severity of this heart disease. So, there is a need to optimize these algorithms by combining multiple base models using bagging, boosting, and stacking techniques. The advanced hybrid models help to refine the individual models and classify in a better way by reducing their variance, enhancing the weak models, learning from overfitting and bias etc.

This study explores reducing the gap in existing heart disease prediction models by implementing ensemble methods and analyzing

the various combinations of these algorithms specifically pairwise stacking and three-model stacking, to predict and classify heart disease with higher accuracy than traditional single models.

The current scenario of various machine learning algorithms is explained in this section. An XGBoost classifier was demonstrated using telemonitored data for the cardiac decompensation events (CDEs) in patients with chronic heart failure.¹ New visions and attributes for cyber physical system-based homecare robotic systems (CPS-HRS) were described to improve home care by patient monitoring and customized care by means of smart automation and integration of advanced technologies.² Stackelberg pricing scheme and a medical coin were introduced to secure, private, and efficient management of electronic medical records (EMRs).³ Feature selection strategy was introduced using this three-feature group. These identified features were analyzed to improve the performance.⁴ A data driven machine health monitoring method was described to address the challenges in extracting heterogeneous features.⁵ A more robust and flexible security measures for IoT systems through machine learning algorithms were discussed to secure the patient information.⁶ A localization machine learning solution for ischemic heart disease detection using magneto cardiography (MCG) was proposed to improve the effectiveness and accuracy of cardiac diagnosis which helps in early treatment.⁷ A machine learning model for early detection of high-risk or coronary artery disease (CAD) from patient information was introduced using Pearson correlation feature selection method to prevent cardiovascular morbidity and mortality.⁸ Machine learning methods for the diagnosis of heart disease and synthesizing evidence for early risk analysis was examined using fine-tuning predictive model.⁹ e-healthcare,¹⁰ a machine learning classification to detect heart disease from clinical and telemonitoring systems that enable patients to self-check their heart condition. ThunderGBM,¹¹ a GPU-based implementation to train Gradient Boosting Decision Trees (GB- DTs) was discussed, which speed up 10x compared to top CPU and GPU libraries without sacrificing model quality. The usage of machine learning (ML) in the identification and assessment of multiple features to predict heart disease was analyzed with seven

ML classifiers: Logistic Regression, Random Forest, Decision Tree, Naive Bayes, k-nearest Neighbours, Neural Networks, and Support Vector Machine (SVM).¹² A supervised machine learning model was introduced for timely intervention and to reduce cardiovascular morbidity and mortality in coronary heart disease (CHD) using the biological, clinical, and environmental factors of the patient information.¹³ Predict cardiovascular disease (CVD),¹⁴ an early stage using classifiers, multi-layer perceptron (MLP) and K- nearest neighbour (K-NN) with the significant health predictors such as hypertension, hyperlipidemia, and coronary heart disease. The complexity of applying machine learning in heart disease prediction with large-scale healthcare data such as medical history, physiological measurements, and lifestyle variables was analyzed by Naive Bayes with Genetic Algorithm, Decision Trees, and Artificial Neural Networks and evaluates their performance in enhancing diagnostic accuracy.¹⁵ Both Deep learning (DL) methods and enhanced deep learning (ETDL) techniques were analyzed in early detection, timely intervention, and personalized treatment of cardiovascular disease.¹⁶ WT-CNN,¹⁷ a hybrid machine learning using the combination of wavelet transform (WT) and convolutional neural networks (CNN) for early prediction of cardiovascular disease. AI models were evaluated to predict cardiovascular disease (CVD) and address limitations of regular models in replicating intricate interactions between risk factors.¹⁸ Meta-analysis was introduced to uncover the challenges associated with imbalanced data in heart disease predictions.¹⁹ Suggestions are provided to handle imbalanced data. A hybrid quantum machine learning (QML),²⁰ method with ensemble-based machine learning models for the early diagnosis of coronary heart disease (CHD). Voting and stacking were applied to various non-ensemble machine learning methods such as KNN, random forest, XGB and GBM, and the performance was analyzed.²¹ A comparative analysis based on supervised machine learning methods are decision tree (DT), random forest (RF), Support vector Machine (SVM), Principal Component Analysis (PCA) was experimented for the prediction of heart diseases.²² Explored three feature selection strategies using chi-square, ANOVA, and mutual information methods to

produce an accurate ML algorithm for early heart disease prediction.²³ Categorical Boosting (CatBoost) and Light Gradient Boosting Machine (LGBM) showed better performance for the heart disease prediction with the key features Age, and Fasting Blood Sugar, correlation between Maximum Heart Rate, Gender, Chest Pain Type, Exercise Angina, peak slope of exercise ST segment.²⁴ implemented various predictive analytic methods such as Decision tree, Random Forest, Logistic regression and KN neighbor to predict the risk of heart disease.²⁵

Worked on the hazard of heart assault utilizing distinctive learning calculations to select the proper one using Calculated Relapse, Irregular Timberland, K Neighbors.²⁶ The performance of Logistic regression was improved compared with supervised classifiers and further ensemble techniques showed accuracy improvement compared with single classifier.^{27,28} Proposed k-modes clustering with Huang starting that improve classification accuracy resulted in optimized result when used in various machine learning algorithms.²⁹ Helped to improve the prediction of atherosclerotic cardiovascular disease (ASCVD) from the electronic health record (EHR) database using the machine learning algorithm.³⁰

From the survey, it is clear that using single machine learning algorithms for practical application in real-world settings is not enough for clinical decisions. So, this paper addresses these issues by reviewing the existing research to identify the performance of the individual existing models, develop a Pairwise stacking and Three-layer stacking models and evaluating the performance of the hybrid models in the healthcare sector to predict and classify heart disease.

MATERIAL AND METHODS

The optimization process began with the exploration of machine learning algorithms in neutral and synthesized models. In this paper, Logistic Regression is selected for a neutral model which served as a foundational baseline for the analysis. In addition, various algorithms such as Decision Tree, Random Forest, Support Vector Machines (with RBF and Sigmoid kernels), XGBoost, Gaussian Naïve Bayes, CatBoost, AdaBoost, Gradient Boosting, Ridge Classifier,

and Multi-layer Perceptron (MLP) were selected for this analysis from the synthesized models.

Once the individual models are identified, each model is evaluated individually for its performance and then ensembling techniques such as bagging, boosting and stacking are implemented to improve prediction accuracy.

This structured methodology helps us in analysis of both individual and hybrid models, leading to the identification of the most accurate model combination in the predication of heart diseases. The workflow—from preprocessing to model evaluation—is depicted in Figure 1.

Data Collection & Preprocessing

A publicly available heart disease dataset from Kaggle was used in this study.³¹ The dataset includes a comprehensive set of patient attributes relevant to cardiovascular health. The parameters included for the analysis are (i) Demographic Information: Age, Gender, (ii) Medical History: Blood Pressure, Cholesterol Level, Fasting Blood Sugar, (iii) Electrocardiogram (ECG) Data: Resting ECG Results, ST Depression, (iv) Physical Attributes: Chest Pain Type, Maximum Heart Rate Achieved.

To ensure data quality and consistency, the data is pre-processed by handling Missing Values, Outliers. In addition, the numerical features are normalized to optimize the model performance.

Model Training & Evaluation

In this research, the selected neutral and synthesized models were trained individually and then ensemble learning strategies were applied to improve the performance of these algorithms.

Bagging

In this research, bagging is implemented in Random Forest, where multiple decision trees were trained independently, and their outputs were averaged for the decision making.

Boosting

Technique in which weak models are trained sequentially to improve performance in term of classification. In this paper, the boosting technique is implemented in AdaBoost, Gradient Boosting, XGBoost, and CatBoost, to improve classification accuracy.

Stacking

As an alternative, rather than depending completely on one model, grouping of multiple models may increase accuracy. To analyze, we

examine stacking in three levels as follows.

- (i) Single Model Comparison: Evaluated as individual classifiers from the selected neutral and synthesized model to determine the baseline performance of each model.
- (ii) Pairwise Stacking: In this, two models were combined and evaluated to leverage their individual strengths and to improve the overall performance.
- (iii) Three Layer Stacking Model: In this, three models were combined with a meta-classifier to further enhance prediction accuracy in the health care sector.

Combining these models using stacking helps in identification of one of the most accurate hybrid models. In this, the combination of Random Forest and Logistic Regression and Decision Tree give better than other combinations. The following section will elaborate the same.

Stacking-Based Hybrid Models

Individual Models

In this research, the twelve machine learning algorithms were selected, trained and tested independently. Logistic Regression (LC), Decision Tree (DT), Random Forest (RF), SVM with RBF kernel, SVM with Sigmond Kernel,

XGBoost, Gaussian navies Bayes (GNB), CatBoost (CB), AdaBoost (AB), Gradient Boosting (GB), Ridge Classifier (RC), Multi-layer Perceptron (MLP). From the analysis, the results show that Random Forest performed the best, making it the strongest individual model.

Pairwise Stacking

In this, we paired models together and tested their combined performance based on the individual strength. The various combination used in this research are (i) Logistic Regression (LR) and Random Classifier (RC) (ii) CatBoost (CB) and Adaboost(AB), (iii) SVM with RBF and SVM with Sigmoid (iv) Random Forest (RF) and Gradient Boosting (GB), (v) Logistic Regression (LR) and Decision Tree (DT), (vi) Logistic Regression (LR) and Random Forest (RF), (vii) Random Forest (RF) and Gaussian navies Bayes (GNB), (viii) Random Forest (RF) and Multi-layer Perceptron (MLP), (ix) Random Forest (RF) and Decision Tree (DT), (x) SVM with RBF kernel and CatBoost (CB), (xi) AdaBoost (AB) and Gradient Boosting (GB). Results show that Random Forest and Logistic Regression provided the best efficiency and accuracy and their interpretability.

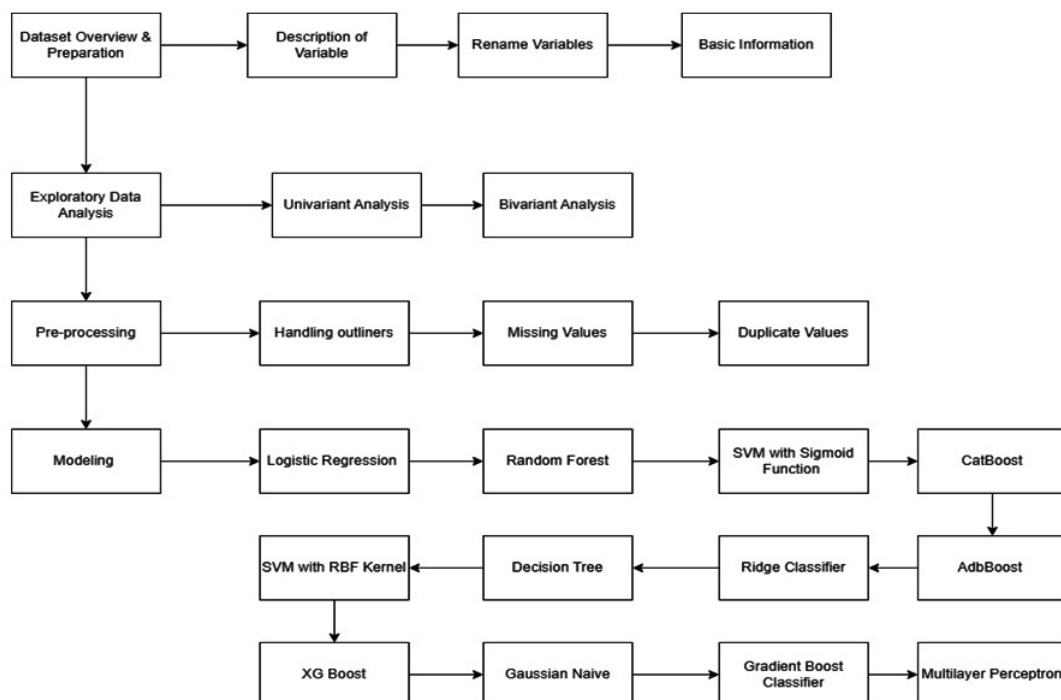
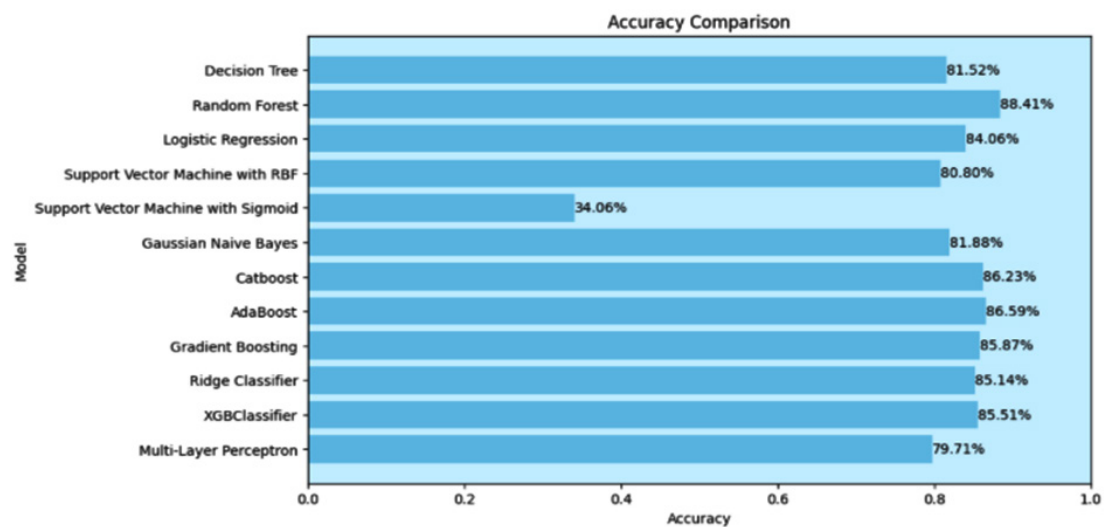


Fig. 1. Process Flow

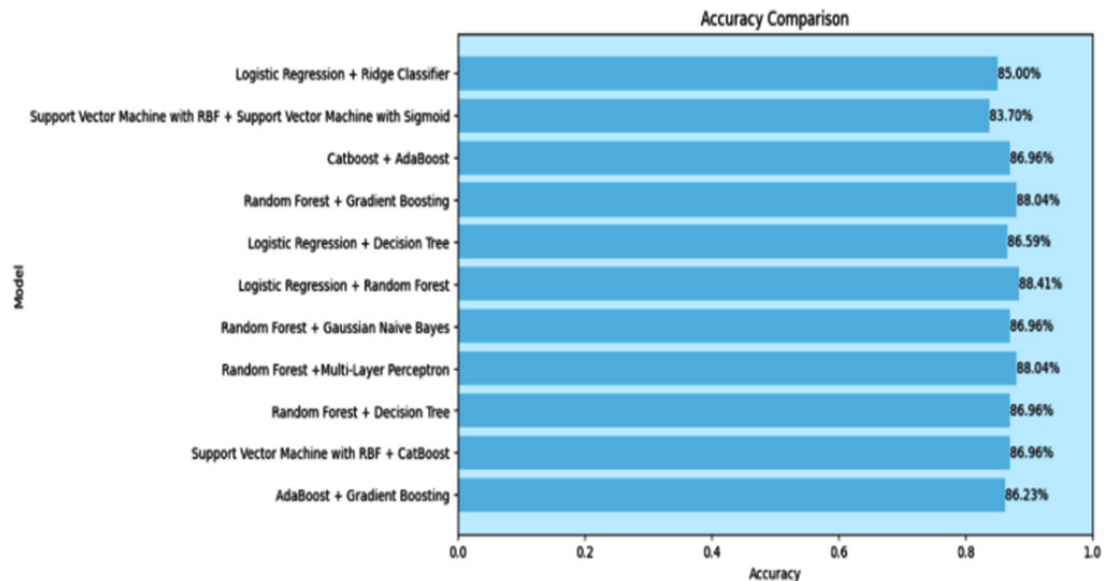
Three Layer Stacking Model

Last, the analysis includes combining these models as a stack of three with a meta-classifier as Random Forest to further enhance prediction accuracy. The combinations we analyzed are (i) Logistic Regression (LR) with Random Classifier (RC) and Gradient Boosting (GB) (ii) CatBoost (CB) with Adaboost (AB) and Random Forest (RF), (iii) Random Forest (RF) with Gradient Boosting (GB) and Logistic Regression (LR), (iv) Gaussian Naives Bayes (GNB) with Logistic Regression (LC) and Random Forest (RF), (v) Gaussian

Naive Bayes (GNB) with Decision Tree (DT) and Random Forest (RF), (vi) Multi-layer Perceptron (MLP) with Random Forest (RF) and Logistic Regression (LC), (vii) Multi-layer Perceptron (MLP) with Random Forest (RF) and Decision Tree (DT), (viii) Multi-layer Perceptron (MLP) with Logistic Regression (LR) and Decision Tree (DT) (ix) AdaBoost (AB) with Gradient Boosting (GB) and CatBoost (CB), (x) Decision Tree (DT) with Logistic Regression (LR) and Random Forest (RF). The most effective models from the pairwise stage was further stacked to form the



Graph 1. Individual algorithm comparison



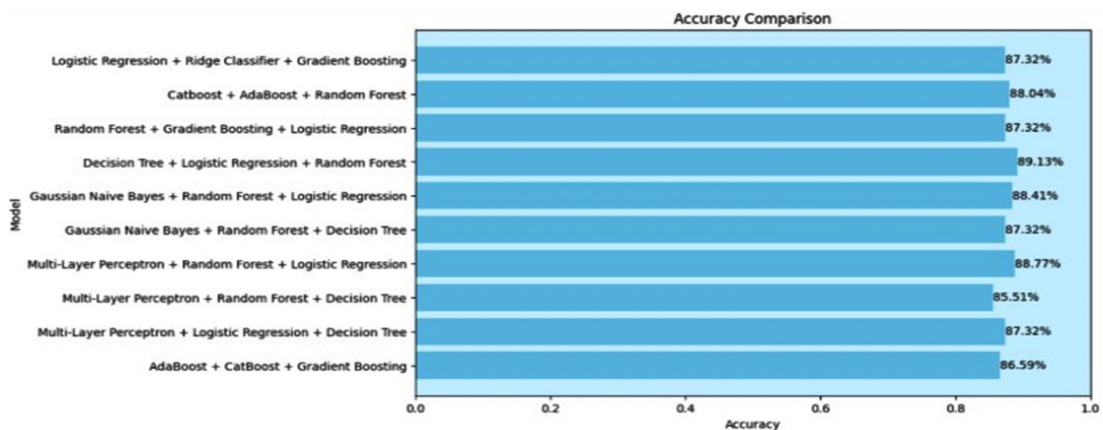
Graph 2. Pairwise Hybrid Model Comparison

Table 1. Performance Metrics for Individual Models

	Precision	Recall	F1-Score
Logistic Regression (LC)	0.879	0.789	0.832
Decision Tree (DT)	0.837	0.782	0.808
Random Forest (RF)	0.949	0.811	0.875
SVM with RBF Kernel	0.793	0.834	0.813
SVM with Sigmoid	0.322	0.289	0.305
Gaussian Naive Bayes (GNB)	0.843	0.782	0.812
CatBoost(CB)	0.923	0.789	0.851
AdaBoost (AB)	0.868	0.812	0.838
XGBoost (XB)	0.882	0.818	0.849
Ridge Classifier(RC)	0.907	0.782	0.840
Gradient Boosting (GB)	0.896	0.811	0.851
Multi-layer Perceptron (MLP)	0.753	0.884	0.813

Table 2. Confusion Matrix for Individual Models

	Predicted Normal	Predicted Heart Disease	Actual Normal	Actual Heart Disease
LR	123	15	29	109
DT	117	21	30	108
RF	132	6	26	112
SVM RBF	108	30	23	115
SVM Sigmoid	54	84	98	40
XGB	123	15	25	113
CB	129	9	29	109
GNB	118	20	30	108
AB	121	17	26	112
GB	125	13	26	112
RC	127	11	30	108

**Graph 3.** Three-layer Stacking Model Comparison

three-model stage. Results show that Random Forest and Logistic Regression and Decision Tree gave the highest accuracy in terms of the three-layer stacking model.

RESULTS

The accuracy of the algorithms were computed, analyzed and the results are shown for individual algorithms, pair-wise stacking and three-layer stacking model. The outcomes of our heart

Table 3. Performance Metrics for Pairwise Hybrid Models

Hybrid Model	Precision	Recall	F1-Score
LR + RC	0.905	0.803	0.851
SVM with RBF + SVM with Sigmoid	0.884	0.775	0.826
CB + AB	0.925	0.804	0.860
RF + GB	0.933	0.818	0.872
LR + DT	0.910	0.811	0.858
LR + RF	0.934	0.826	0.876
RF + GNB	0.925	0.804	0.860
RF + MLP	0.894	0.862	0.878
RF + DT	0.918	0.811	0.861
SVM with RBF + CB	0.925	0.804	0.860
AB + GB	0.903	0.811	0.854

Table 4. Confusion Matrix for Pairwise Hybrid Models

	Predicted Normal	Predicted Heart Disease	Actual Normal	Actual Heart Disease
LR+RC	127	18	26	129
SVM RBF + Sigmoid	124	14	31	107
CB+AB	129	9	27	111
RF+GB	130	8	25	113
LR+DT	127	11	26	112
LR+RF	130	8	24	114
RF+GNB	129	9	27	111
RF+MLP	124	14	19	119
RF+DT	128	10	26	112
SVM RBF+CB	129	9	27	111
AB+GB	126	12	26	112

Table 5. Performance Metrics for Three-layer stacking Hybrid Models

Hybrid Model	Precision	Recall	F1-Score
LR + RC + GB	0.918	0.818	0.865
CB + AB + RF	0.933	0.818	0.872
RF + GB + LR	0.932	0.804	0.863
DT + LR + RF	0.935	0.940	0.885
GNB + RF + LR	0.934	0.826	0.876
GNB + RF + DT	0.940	0.797	0.862
MLP + RF + LR	0.921	0.847	0.883
MLP + RF + DT	0.871	0.833	0.851
MLP + LR + DT	0.887	0.855	0.870
AB + CB + GB	0.917	0.804	0.857

Table 6. Confusion Matrix for three-layer stacking Hybrid Models

	Predicted Normal	Predicted Heart Disease	Actual Normal	Actual Heart Disease
LR + RC + GB	128	10	25	113
CB + AB + RF	130	8	25	113
RF + GB + LR	130	8	27	111
RF + DT + LR	130	8	22	116
GNB + RF + LR	130	8	24	114
GNB + DT + RF	131	7	28	110
RF + MLP + LR	128	10	21	117
RF + LP + DT	121	17	23	115
LR + MLP + DT	123	15	20	118
AB + GB + CB	128	10	27	111

disease prediction model demonstrate how stacking techniques and ensemble learning greatly improve prediction accuracy to predict heart disease. The findings contribute to the ongoing research in heart disease prediction showcasing the potential of hybrid machine learning approaches in medical diagnostics.

Individual Algorithms

Individual machine learning models were first evaluated for its accuracy, precision score and confusion matrix. Results show that Random Forest is the best among them with the highest accuracy of 88.41% while SVM with Sigmoid Kernel performs the worst of 34.06% as shown in graph1.

Pair Wise Stacking

As a study of ensemble learning technique, we analyse the pairwise stacking by grouping two algorithms in a hybrid model. The graph2 compares the accuracy of pair-wise hybrid model for heart disease prediction. From the graph, it has been identified that Logistic Regression and Random Forest achieve the highest accuracy 88.41% while SVM with RBF Kernel and SVM with Sigmoid achieves 83.70%. Though the performance of SVM with Sigmoid is low when compared with other algorithms, the performance is increased.

Three-layer Stacking model

Further refinement using three-layer model stacking was implemented and analyzed. The three-model stacking of Random Forest, Logistic Regression and Decision Tree achieves the highest accuracy of 89.1%, shown graph 3, indicating that combining diverse models enhances predictive capability.

DISCUSSION

This study aims to improve the accuracy and reliability of heart disease prediction using machine learning. However, these individual models often struggle to achieve the requisite high accuracy for the complex task of accurately classifying both the type and severity of heart disease, leading to high variance, overfitting, and suboptimal clinical utility. For this, the study demonstrates the value of ensemble learning and stacking techniques in improving the accuracy and reliability of heart disease prediction models and help the doctors in early diagnosis of this disease and to start treatment in early. Though individual algorithms like Random Forest or Logistic Regression perform well, combining multiple models through ensemble stacking was explored to enhance performance further in terms of prediction accuracy.

Individual Algorithms

Table 1 and table 2 shows the performance metrics and confusion matrix of individual model. Tables show that Random Forest has 0.949 as the precision score SVM with Sigmoid Kernel has 0.322 as the precision score that reflects the true positive and false positive in table 2.

Pair Wise Stacking

Table 3 and table 4 show the performance metrics and confusion matrix of the pair-wise stacking model. Tables show that Logistic Regression with Random Forest shows 0.934 as the precision score, while SVM with RBF Kernel and SVM with Sigmoid achieves 0.884 as the

precision score that reflects the true positive and false positive in table 4.

Three-layer Stacking model

Table 5 and table 6 shows the performance metrics and confusion matrix of the three-layer stacking model. Tables show that Random Forest and Logistic Regression and Decision Tree shows 0.935 as the precision score that reflects the true positive and false positive in table 6.

The key insights of the above results are

- Stacked models consistently outperformed individual algorithms.
- Pairwise and three-model stacking significantly enhanced precision and recall metrics, indicating better handling of both false positives and false negatives.
- Model diversity in stacking by mixing tree-based model with linear models contributed to improved generalization and robustness.

CONCLUSION

In this study, accurate prediction of heart disease was achieved by analyzing various machine learning algorithms. To enhance predictive performance, two hybrid stacking methods, pairwise and three-model ensemble stacking, were implemented. A comparative analysis of these hybrid stacking methods shows that the Random Forest, Logistic Regression, and Decision Tree as the most effective combination, achieving a maximum prediction accuracy of 89.1%.

These techniques significantly improved classification reliability and can be integrated into real-world clinical decision support systems, enabling early detection and management of cardiovascular disease.

Future work will focus on expanding the dataset to include diverse populations for improved model generalization. Deep learning techniques like CNNs and RNNs will be explored to capture complex patterns. The integration of lifestyle, genetic, and real-time monitoring data aims to enhance personalized heart disease prediction.

ACKNOWLEDGEMENT

We the authors acknowledge Christ University, Bangalore, for providing this opportunity.

Funding Sources

The author(s) received no financial support for the research, authorship, and/or publication of this article.

Conflict of Interest

The author(s) do not have any conflict of interest.

Data Availability Statement

The manuscript incorporates the readily available Kaggle data set: <https://www.kaggle.com/code/chanchal24/heart-disease-eda-prediction-7-models/input?select=heart.csv>.

Ethics Statement

This research did not involve human participants, animal subjects, or any material that requires ethical approval.

Informed Consent Statement

This study did not involve human participants, and therefore, informed consent was not required.

Clinical Trial Registration

This research does not involve any clinical trials

Permission to reproduce material from other Sources

Not Applicable.

Author contributions

Gnana Prakasi Oliver Sirya Pushpam: Conceptualization, Supervision, Project Administration, Writing – Original Draft; Diana Jeba Jingle Issac: Interpretation of Results and Final Draft; Kanmani Palanisamy: Interpretation of Results and Final Draft; Pari Agarwal : Conceptualization and Implementation; Rashi Dubey: Conceptualization and Implementation; Aryaman Kant : Conceptualization and Implementation

REFERENCES

1. Kerexeta J, Larburu N, Escolar V, et.al. Prediction and Analysis of Heart Failure Decompensation Events Based on Telemonitored Data and Artificial Intelligence Methods. *J Cardiovasc Dev Dis.* 2023, 10(2):48.
2. Yang G, Pang Z, Jamal Deen M, Dong M, Zhang Y Lovell N. Homecare Robotic Systems for Healthcare 4.0: Visions and Enabling Technologies. *IEEE J Biomed Health Inform.* 2020 ;24(9):2535-2549
3. Xu J, Xue K, Li S, Tian H, Hong J Hong P. Healthchain: A Blockchain-Based Privacy

- Preserving Scheme for Large-Scale Health Data. *IEEE Internet of Things Journal*. 2019; 6(5); 8770-8781.
4. El-Sofany H, Bouallegue B, El-Latif YMA. A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method. *Scientific Reports*. 2024; 14 (23277);
5. Cheng C, Ding J, Zhang Y. A Koopman operator approach for machinery health monitoring and prediction with noisy and low-dimensional industrial time series. *Neurocomputing*. 2020; 406; 204-214.
6. Sun P, Li J, Bhuiyan M Z, Wang L, Li B. Modeling and clustering attacker activities in IoT through machine learning techniques. *Information Sciences*. 2019; 479; 456-471.
7. Tao R, Zhang Y, Zhang R, et al. Magnetocardiography-based ischemic heart disease detection and localization using machine learning methods. *IEEE Trans Biomed Eng*. 2019;66(6):1658-1667.
8. Sayadi M, Varadarajan V, Sadoughi F, Chopannejad S, Langarizadeh M. A machine learning model for detection of coronary artery disease using noninvasive clinical parameters. *Life (Basel)*. 2022;12(11):1933.
9. Hussain A, Kumar P, Chaurasia P, Sabri M. A review on heart disease prediction using machine learning techniques. *SSRN Electron J*. 2019;9:208-224.
10. Li JP, Haq AU, Din SU, Khan J, Khan A, Saboor A. Heart disease identification method using machine learning classification in e-healthcare. *IEEE Access*. 2020;8:107562-107582.
11. Wen Z, Liu H, Shi J, Li Q, He B, Chen J. ThunderGBM: fast GBDTs and random forests on GPUs. *J Mach Learn Res*. 2020;21(108):1-5.
12. Zhao L, Ni L, Hu S, Chen Y, Zhou P, Xiao F. InPrivate digging: enabling tree-based distributed data mining with differential privacy. In: *IEEE INFOCOM 2018 - IEEE Conference on Computer Communications*; 2018:2087-2095.
13. Krishnani D, Kumari A, Dewangan A, Singh A, Naik NS. Prediction of coronary heart disease using supervised machine learning algorithms. In: *TENCON 2019 - 2019 IEEE Region 10 Conference*; 2019:367-372.
14. Pal M, Parija S, Panda G, Dhama K, Mohapatra R. Prediction of cardiovascular disease using machine learning classifiers. *Open Med (Wars)*. 2022;22.
15. Gour S, Panwar P, Dwivedi D, Mali C. A machine learning approach for heart attack prediction. In *Intelligent Sustainable Systems*. 2022; 741-747.
16. Wu Y. Deep learning for cardiovascular disease prediction: recent advances, challenges and future directions. *Theor Nat Sci*. 2024;62:24-32.
17. Mohammad F, Al-Ahmadi S. WT-CNN: a hybrid machine learning model for heart disease prediction. *Mathematics*. 2023;11(22):4681.
18. Cai Y, Cai YQ, Tang LY, et al. Artificial intelligence in the risk prediction models of cardiovascular disease and development of an independent validation screening tool: a systematic review. *BMC Med*. 2024;22:56.
19. Ahsan MM, Siddique Z. Machine learning-based heart disease diagnosis: a systematic literature review. *Artif Intell Med*. 2022;128:102289.
20. Banday M, Zafar S, Agarwal P, Alam MA, Abubeker KM. A hybrid quantum machine learning model for early diagnosis of coronary heart disease. *IEEE J Biomed Health Inform*. 2024;28(1):88-99.
21. Moumin ZM, Ecemip YN, Karhan M. Heart disease detection using ensemble and non-ensemble machine learning methods. *Eur Phys J Spec Top*. 2024.
22. Singh A, Mahapatra H, Biswal AK, Mahapatra M, Singh D, Samantaray M. Heart disease detection using machine learning models. *Procedia Comput Sci*. 2024;235:937-947. doi:10.1016/j.procs.2024.04.089
23. El-Sofany H, Bouallegue B, Abdel-Latif YM. A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method. *Sci Rep*. 2024;14:23277. doi:10.1038/s41598-024-74656-2
24. Hossain MR. Optimizing heart disease prediction: A comparative study of machine learning techniques. *Asian J Cardiol Res*. 2024;7(1):348-359. doi:10.9734/ajcr/2024/v7i1238
25. Telsang S, Pokale T, Posugade A, Patwal P, Shirode P, Popade O. Heart disease prediction using machine learning. In: *Proceedings of the 2024 IEEE North Karnataka Subsection Flagship International Conference (NKCon)*; 2024 Aug; Bagalkote, India. IEEE; 2024. p. 1-6. doi:10.1109/NKCon62728.2024.10774628
26. Kanimozhi S, Sivanandan N. Machine learning based heart disease prediction system. *E3S Web Conf*. 2024;491:03013. doi:10.1051/e3sconf/202449103013
27. Gupta, C., Saha, A., Reddy, N. V. S., & Acharya, U. D. (2022). Cardiac disease prediction using supervised machine learning techniques. *Journal of Physics: Conference Series*, 2161(1), Article 012013. <https://doi.org/10.1088/1742-6596/2161/1/012013>
28. Dritsas, E., Alexiou, S., & Moustakas, K. (2022). Cardiovascular disease risk prediction with supervised machine learning techniques. In

- Proceedings of the 8th International Conference on Information and Communication Technologies for Ageing Well and e-Health (ICT4AWE) – Volume 1: ICT4AWE* (pp. 315–321). SciTePress. <https://doi.org/10.5220/0011088300003188>
29. Bhatt, C. M., Patel, P., Ghetia, T., & Mazzeo, P. L. (2023). Effective heart disease prediction using machine learning techniques. *Algorithms*, *16*(2), 88. <https://doi.org/10.3390/a16020088>
30. Ward, A., Sarraju, A., Chung, S., Li, J., Harrington, R., Heidenreich, P., Palaniappan, L., Scheinker, D., & Rodriguez, F. (2020). Machine learning and atherosclerotic cardiovascular disease risk prediction in a multi-ethnic population. *NPJ Digital Medicine*, *3*, 125. <https://doi.org/10.1038/s41746-020-00331-1>
31. <https://www.kaggle.com/code/chanchal24/heart-disease-eda-prediction-7-models/input?select=heart.csv>