Aditi Kumari, 2026, 14:6
ISSN (Online): 2348-4098
ISSN (Print): 2395-4752

International Journal of Science, Engineering and Technology
An Open Access Journal

Early Cardiovascular Disease Prediction Using Machine Learning

Aditi Kumari[1], Anamika Kumari[2], Suraj Gangawane[3], Prof. Vaishali Suryawanshi[4]
Jspm Rajarshi Shahu College Of Engineering, Tathawade

Abstract — Cardiovascular disease remains a major global cause of mortality, although timely detection may prevent many adverse outcomes. Machine learning (ML) can analyse medical history, demographic details, lifestyle factors and electrocardiogram (ECG) signals to identify individuals at risk before their condition becomes severe. However, prior studies use varied datasets, input variables and evaluation measures, making it hard to identify consistently reliable approaches. This review examines 25 studies published between 2019 and 2024 in four categories: traditional ML, ensemble learning, deep learning and explainable AI. Datasets, features and reported results are compared. Ensemble and boosting approaches performed strongly, with reported accuracy of roughly 88–99% on public test datasets. Deep learning has also analysed millions of ECG recordings, while explainability is increasingly important for clinical trust. Recurring limitations include limited external validation, inconsistent datasets and metrics, class imbalance and hard-to-interpret models. The review concludes with recommendations for more robust and trustworthy prediction systems.

Keywords — Heart Disease Prediction, Machine Learning, Deep Learning, Ensemble Learning, Explainable AI, ECG, Systematic Review.

I. INTRODUCTION

Cardiovascular disease is one of the leading causes of death worldwide. Timely diagnosis and treatment could help prevent many of these deaths. Machine learning (ML) offers a useful approach by analysing routinely collected information—including age, blood pressure, cholesterol, ECG signals and lifestyle factors—through techniques such as Logistic Regression, Random Forest, deep learning and explainable AI. As healthcare datasets grow in size and complexity, these models may identify relationships that conventional statistical approaches fail to capture.Despite these advances, many studies assess models on a few public datasets under different experimental conditions. Although accuracy above 90–95% is reported on familiar datasets, it is uncertain whether such performance would hold among larger, more diverse populations or in everyday hospital settings. This review therefore consolidates ML-based early heart disease prediction research to identify promising, dependable approaches.

This paper makes three contributions: (i) it brings together 25 studies and sorts them into four categories; (ii) it compares datasets, features and reported performance through tables and figures; and (iii) it outlines limitations and next steps toward an improved model.

1. Research Objectives

- Examine ML techniques applied to early heart disease prediction.
- Identify the datasets and input features most used in existing research.
- Compare the reported performance of ML approaches.
- Identify research limitations, including low interpretability, reliance on a single data type and class imbalance.
- Recommend ways to build a more reliable prediction model.

2. Traditional And Ensemble Learning

Karthick et al. found that Random Forest (RF) outperformed SVM, Logistic Regression and Naïve Bayes, with accuracy of about 88.5% [8]. Mohan et al. also found RF best, with about 88.7% accuracy against Decision Tree, Logistic Regression and Naïve Bayes [21]. Gupta et al. reported RF above 90%, surpassing SVM and KNN [22], while Ansari and Nazeer recorded results

© 2026 Aditi Kumari, This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

near 99% for KNN and RF [10]. Gradient boosting [9], stacked approaches [17] and RF-based frameworks [6], [23] similarly suggest that combining models beats relying on one algorithm.

3. Feature Selection And Hybrid Models

Islam et al. integrated PCA, a Genetic Algorithm and K-Means, reaching about 94.06% accuracy [3]. Theerthagiri and Vidya combined Recursive Feature Elimination with Gradient Boosting, reporting 89.7% accuracy and a higher AUC (a measure of how well a model separates affected from unaffected patients) [5]. Peng et al. identified age, blood pressure, cholesterol and BMI as informative predictors [7]. Selecting relevant features can improve performance.

4. Deep Learning

A one-dimensional CNN reported about 96% test accuracy, outperforming several traditional techniques [4]. Feature augmentation raised a deep model's precision to around 90% [18]. A multimodal approach combining ECG and heart-sound signals beat either source alone [13]. A large interpretable deep-learning model trained on over 2.3 million ECGs gave strong diagnoses and flagged mortality-linked patterns [15].

5. Explainable Ai

Explainable AI makes a model's reasoning more accessible. An RF model with SHAP highlighted the most influential ECG features [11], while another framework identified key clinical variables without a substantial accuracy drop [24]. Interpretable ECG techniques may build clinicians' confidence [12], and multimodal systems can offer visual explanations [13], [14].

II. LITERATURE REVIEW

Research on ML-based heart disease prediction from 2019 to 2024 falls into five themes.

1. Systematic Reviews

Previous reviews reported promising performance on ECG and clinical datasets, but frequently observed class imbalance and weaker generalisation to unfamiliar data [1], [2]. A review covering 33 survival-analysis studies found that Random Survival Forests and deep learning offered better long-term risk prediction than conventional statistical techniques [16]. A meta-analysis of electronic health record (EHR) models reported pooled AUC values reaching 0.865 [19]. Other reviews emphasised limited external validation, inconsistent reporting, the need for interpretable predictions and the value of common evaluation standards [20], [25].

2. Main Ml Methods In Simple Words

Logistic Regression estimates a boundary between patients with and without heart disease. A Decision Tree predicts through a series of questions, such as whether cholesterol is above a threshold. Random Forest combines many trees and takes their vote. Gradient Boosting builds trees in turn, fixing earlier errors.SVM finds a boundary that maximises the margin between patient groups. KNN predicts for a new patient by comparing them with similar existing cases, whereas Naïve Bayes bases predictions on probabilities derived from the observed features.Deep-learning models learn subtle patterns in signals or images, which suits ECG analysis. SHAP and LIME show how each input influences a prediction; visual methods can also highlight the ECG or image regions that drew attention.

III. RESEARCH METHODOLOGY

This study follows a structured review of previously published research. It relies only on secondary sources, such as academic papers and databases; no patient data was collected.



Fig. 1. Methodology of the systematic review

1. Search And Screening

Google Scholar, IEEE Xplore and ScienceDirect were searched using terms including "heart disease prediction," "machine learning," "ECG classification" and "cardiovascular disease detection." Publications were screened for topic relevance, publication date (primarily 2021–2024, together with one 2019 comparative study using standard datasets) and the use of clinical or ECG data. The final selection comprised 25 studies with verifiable findings, including peer-reviewed journal articles, established conference papers and preprints.

2. Data Extraction And Analysis

A review matrix in MS Excel recorded each paper's objective, results, conclusion, technology, research gap and future work. Studies were grouped into traditional ML, ensemble learning, deep learning and explainable AI, then compared by methodology, data sources, features and reported accuracy.

3. Scope And Limitations

Performance figures are those reported by the original authors; models were not rerun. Differences in data prep preparation and measures differ between studies, the figures and metrics mean the figures give only a rough overview and are not directly comparable. No formal meta-analysis was done.

4. Measures Used To Judge Models

Most studies assess models with standard metrics. Accuracy is the share of all cases classified correctly. Precision is the fraction of predicted-ill patients who truly have the condition, while recall is how many genuinely affected patients are found. AUC summarises how well affected and unaffected patients are separated. For screening, recall matters most, since missing a diseased patient is worse than a false alarm, so accuracy alone is not enough.

IV. RESULTS AND DISCUSSION

Of the 25 studies, 18 are original research and 7 are reviews. As Fig. 2 shows, publications rose over time, with 10 in 2024. Table I summarises all studies.

TABLE I. SUMMARY OF THE REVIEWED STUDIES

#	Study	Year	Method / Technology	Key Result	Main Limitation
1	Ahsan & Siddique	2022	Review: RF, DL	High accuracy on ECG and clinical data	Unbalanced data; poor on new data
2	Osamika et al.	2021	Review: RF, DT, neural networks (EHR)	Consistently high performance	Not tested on outside data
3	Islam et al.	2021	PCA + hybrid GA + K-Means	94.06% accuracy	Few datasets
4	Hussain et al.	2021	1D CNN	≈96% test accuracy	Needs large training data
5	Theerthagiri & Vidya	2021	RFE + gradient boosting	89.7% accuracy; better AUC	Single dataset
6	Leo et al.	2023	RF, SVM, LR, feature selection	High accuracy; fewer misclassifications	Depends on the data used
#	Study	Year	Method / Technology	Key Result	Main Limitation
7	Peng et al.	2023	Feature selection + ensemble learning	Age, BP, cholesterol, BMI key predictors	Data from one group only
8	Karthick et al.	2022	RF vs SVM	RF best, ≈88.5%	Mainly Cleveland dataset
9	Jawalkar et al.	2023	Stochastic gradient boosting	Beat several other models	Needs more hospital data

© 2026 Aditi Kumari, This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly credited.

7	Peng et al.	2023	Feature selection + ensemble learning	Age, BP, cholesterol, BMI key predictors	Data from one group only
10	Ansari & Nazeer	2023	KNN, RF, SVM, DT, NB, LR	≈99% (KNN, RF)	Standard test data only
11	Lakshmanan et al.	2024	RF on ECG (EWT, DWT)	High accuracy; clear key features	Standard ECG data only
12	Ayano et al.	2023	Review: interpretable ML for ECG	More trust from doctors	No standard to judge explanations
13	Oliveira et al.	2024	Multimodal 1D/2D CNN + Grad-CAM	Joining signals beats one signal	Small dataset
14	Goettling et al.	2024	(interpretable CNN)	Strong, interpretable ECG analysis	Needs wider testing
15	Lu et al.	2024	Deep learning on 2.3 M ECGs	Excellent diagnosis; death-risk patterns	Needs strong computers; hard to use
16	Teshale et al.	2024	Review of 33 survival-AI studies (RSF, DL)	Better long-term risk prediction	Not tested on outside data
17	Tiwari et al.	2023	Stacking (Extra Trees, RF)	92.34% accuracy	Standard test data only
18	García-Ordás et al.	2024	Deep learning + feature augmentation	≈90% precision	Needs more varied data
19	Liu et al.	2024	Meta-analysis of ML on EHR	Combined AUC up to 0.865	Studies differ; some results unpublished
20	Wang et al.	2024	Review + validation screening tool	Strong predictive performance	Not tested on outside data
21	Mohan et al.	2019	RF, DT, LR, NB	RF best, ≈88.7%	Small dataset
22	Gupta et al.	2022	RF, SVM, KNN	> 90% accuracy	Not enough variety
23	Hasan et al.	2023	RF, DT, gradient boosting	RF best; high precision/recall	Standard test data only
24	Alotaibi et al.	2024	RF,KNN	Key clinical features identified	Few datasets
25	Krittanawong et al.	2024	Review: AI across ECG, imaging, EHR	Better diagnosis; explainability is key	No common rules

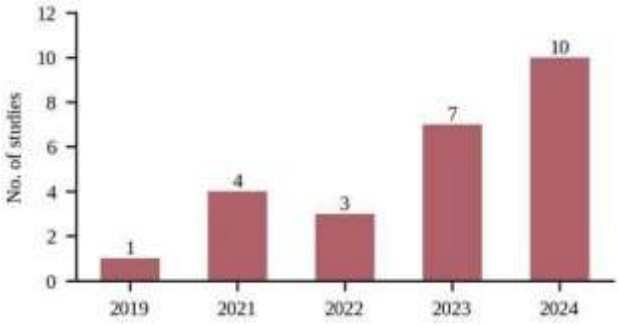


Fig. 2. Number of reviewed studies by publication year.

1. Performance Comparison

Fig. 3 shows accuracy reported by selected studies, from about 88.5% to 99%. Ensemble methods such as Random Forest, boosting and stacking often ranked among the strongest on structured clinical datasets, and a 1D CNN reached about 96%. The top result, near 99%, came from a standard test dataset rather than hospital data [10]. As datasets differ, scores are not a direct ranking.

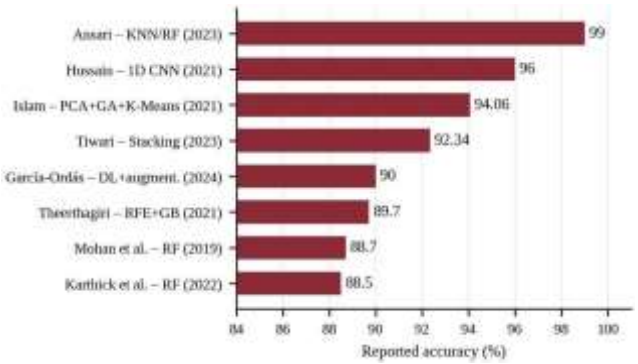


Fig. 3. Accuracy reported by selected studies (not directly comparable).

2. Datasets And Features

Three data types are used most: (i) public tabular datasets such as Cleveland [8], [10], [17]; (ii) electronic health records(EHR) [2], [19]; and(iii) physiological signals, mainly ECG [11], [12], [14], [15] and heart sounds [13]. Common variables include age, blood pressure, cholesterol and BMI [7], plus ECG-derived measures [11]. Feature selection such as PCA, RFE and genetic algorithms improved results [3], [5], [9].

3. Technique-Wise Synthesis

Traditional ML (LR, SVM, KNN, NB, DT) is simple

and light, but ensembles often do better, especially on structured clinical data. Deep learning performs strongly on large datasets and signals but needs substantial data and computing power [4], [15]. Explainable AI adds the transparency clinicians need, often with little accuracy loss [11], [12], [24].

TABLE II. STRENGTHS AND LIMITS OF EACH METHOD GROUP

Method group	Strengths	Limits
Traditional ML	Simple, fast and easy to follow	Usually less accurate than ensembles
Ensemble learning	Most dependable on table-style clinical data	Harder to explain than a single model
Deep learning	Strong on ECG and very large data	Needs a lot of data and computing power
Explainable AI	Shows why a decision was made, which builds trust	Few standard ways to judge explanations

1. Practical Uses

The studies suggest several clinical uses. Models built on health records may serve as an initial screening aid, flagging higher-risk individuals [2], [19]. ECG-based systems can help clinicians process many recordings and flag patterns linked to cardiac conditions or mortality risk [11], [15]. Explainable tools can show why a person was classed as high risk so clinicians can verify it [12], [14]. Such systems should assist, not replace, clinicians, and external validation is needed before deployment.

2. Key Findings in Brief

- On structured clinical datasets, ensembles such as RF and boosting look particularly reliable.
- Deep learning does well on ECG and large datasets but needs much training data and computing power.
- Explainability can raise clinician confidence with little reported loss in performance.
- Accuracy near 99% mostly comes from public test data and needs cautious reading.
- Validation on independent, diverse populations remains a major need.

Enhancements Suggested by the Reviewed Studies

A recurring recommendation is to evaluate models on larger clinical datasets from several hospitals [1], [3], [5], [7], [10], [13], [14], [21], [23]. Several studies propose combining explainable AI with high-performing methods [1], [2], [5], [6], [9], [17], [21]. Other directions include real-time hospital data and wearable devices [2], [7], [9], [14], [15], [22], and multimodal approaches combining ECG, heart sounds, images and clinical records [11], [13], [18], [25].

Others suggest shared benchmarks and evaluation measures, plus prospective studies that follow patients over time before models enter clinical practice [10], [12], [16], [19], [20].

V. RESEARCH GAPS IDENTIFIED

- Limited external validation: many models are tested on only one established dataset [1], [2], [5], [20].
- Inconsistent evaluation: differences in datasets, features, preprocessing and metrics hinder direct comparison [16], [19], [25].
- Limited interpretability: high-performing ensemble and deep-learning models can operate as black boxes, and no widely accepted standard exists for assessing their explanations [12].
- Small or imbalanced samples: data from a narrow population may not represent other groups [1], [7], [18].
- Deployment barriers: high computing demands and little long-term testing on real patients [15], [25].

1. Challenges In Real Hospital Use

Strong test-dataset performance alone does not show that a model suits hospital use. Clinical records may have missing or inconsistent entries, and patient populations vary by region, age and lifestyle. ECG devices may differ from those used for training data. Large deep-learning systems need computing resources some clinics lack [15], [25]. Because patient data is confidential, sharing between hospitals needs care; federated learning is one proposed alternative [25]. Clinicians must also be able to verify outputs before relying on them [12], [14].

2. Future Directions

To improve reliability, future models should: (1) be evaluated on multi-centre hospital datasets and longitudinal studies; (2) include explainable AI; (3) combine sources such as ECG, heart sounds, electronic records and wearable data, possibly through federated learning [25]; (4) address class imbalance; and (5) adopt common benchmarks and reporting practices.

Plan for the Next Phase

Building on this review, the next project stage will develop and evaluate a prediction model through these steps:

- Select and clean public heart disease datasets, including Cleveland and a larger dataset.
- Address class imbalance; pick key features.
- Train and evaluate Random Forest, SVM, KNN, Logistic Regression and a basic neural network.
- Assess each with accuracy, precision, recall, AUC.
- Apply SHAP or LIME to the best model.
- Test the final model on a separate dataset to check performance on new patients.

VI. CONCLUSION

Machine learning shows considerable potential for early heart disease prediction, with many reviewed studies reporting accuracy above 85–95%. Ensemble and boosting methods appear dependable for structured clinical data, deep learning is promising for ECG and large-scale analysis, and explainability matters for clinical acceptance. However, evidence is spread across different datasets, methods and evaluation practices, and few studies test models on diverse populations, so real-world reliability remains uncertain. This review consolidates findings and outlines steps toward a robust, trustworthy model.

REFERENCES

1. M. M. Ahsan and Z. Siddique, “Machine learning-based heart disease diagnosis: A systematic literature review,” Artif. Intell. Med., vol. 128, 2022, doi: 10.1016/j.artmed.2022.102289.
2. D. Osamika, B. S. Adelusi, M. C. Kelvin-Agwu, et al., “Machine learning models for early detection of cardiovascular diseases: A systematic review,” Iconic Res. Eng. J., 2021.
3. M. T. Islam, S. R. Rafa, and M. G. Kibria, “Early prediction of heart disease using PCA and hybrid genetic algorithm with K-Means,” 2021.
4. S. Hussain, S. K. Nanda, S. Barigidad, et al., “Novel deep learning architecture for heart disease

prediction using convolutional neural network," 2021.
5. P. Theerthagiri and J. Vidya, "Cardiovascular disease prediction using recursive feature elimination and gradient boosting classification techniques," 2021.
6. M. Leo, P. Carcagnì, C. Distante, et al., "Effective heart disease prediction using machine learning techniques," Algorithms (MDPI), 2023.
7. M. Peng, F. Hou, Z. Cheng, et al., "Prediction of cardiovascular disease risk based on major contributing features," Sci. Rep., 2023.
8. K. Karthick, S. K. Aruna, R. Samikannu, et al., "Implementation of a heart disease risk prediction model using machine learning," Comput. Math. Methods Med., 2022.
9. A. P. Jawalkar, P. Swetcha, N. Manasvi, P. Sreekala, et al., "Early prediction of heart disease with data analysis using supervised learning with stochastic gradient boosting," J. Eng. Appl. Sci., 2023.
10. S. Ansari, J. Nazeer, et al., "Performance evaluation of machine learning techniques for heart disease prediction," Comput. Math. Methods Med., 2023.
11. K. Lakshmanan et al., "Explainable AI-driven machine learning for heart disease detection using ECG signal," Appl. Soft Comput., 2024.
12. Y. M. Ayano, F. Schwenker, et al., "Interpretable machine learning techniques in ECG-based heart disease classification: A systematic review," Diagnostics, 2023, doi: 10.3390/diagnostics13010111.
13. B. Oliveira, A. Lobo, C. I. Costa, R. Fontes-Carvalho, M. Coimbra, and F. Renna, "Explainable multimodal deep learning for heart sounds and electrocardiogram classification," in Proc. IEEE EMBC, 2024.
14. M. Goettling, A. Hammer, H. Malberg, M. Schmidt, et al., "xECGArch: A trustworthy deep learning architecture for interpretable ECG analysis considering short-term and long-term features," Sci. Rep., 2024.
15. L. Lu, A. L. P. Ribeiro, et al., "Decoding 2.3 million ECGs: Interpretable deep learning for advancing cardiovascular diagnosis and mortality risk stratification," Eur. Heart J. Digit. Health, 2024, doi: 10.1093/ehjdh/ztae014.
16. A. B. Teshale, H. L. Htun, A. J. Owen, R. Freak-Poli, et al., "A systematic review of artificial intelligence models for time-to-event outcome applied in cardiovascular disease risk prediction," J. Med. Syst., 2024.
17. A. Tiwari, A. Chugh, and A. Sharma, "Ensemble framework for cardiovascular disease prediction," arXiv, 2023.
18. M. T. García-Ordás, M. Bayón-Gutiérrez, C. Benavides, J. A. Benítez-Andrades, et al., "Heart disease risk prediction using deep learning techniques with feature augmentation," arXiv, 2024.
19. T. Liu, A. Krentz, L. Lu, and V. Curcin, "Machine learning based prediction models for cardiovascular disease risk using electronic health records data: Systematic review and meta-analysis," Eur. Heart J. Digit. Health, 2024.
20. Y. Wang, J. Zhang, 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.
21. S. Mohan, C. Thirumalai, and G. Srivastava, "Heart disease prediction using machine learning techniques: Performance evaluation and comparison," IEEE Access, 2019.
22. A. K. Gupta, R. Kumar, and S. Arora, "Heart disease prediction using machine learning algorithms," Int. J. Intell. Syst. Appl., 2022.
23. T. Hasan, M. Islam, and A. Rahman, "Machine learning-based prediction of cardiovascular disease using clinical data," Healthcare (MDPI), 2023.
24. M. Alotaibi, F. Alharbi, and H. Alshammari, "An explainable artificial intelligence framework for cardiovascular disease prediction," Front. Artif. Intell., 2024.
25. Krittanawong et al., "Artificial intelligence for cardiovascular disease detection: A comprehensive review," Eur. Heart J. Digit. Health, 2024.