International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18(15s) 2026 pp. 186-202
https://doi.org/10.70917/ijcisim-2026-4393
Article

Machine Learning and Explainable Artificial Intelligence for Early Heart Disease Prediction and Clinical Decision Support: A Systematic Review

Ronak Jain[1]*, Sachin Patel[2]

[1]SAGE University, Indore, Madhya Pradesh, India.
Email: ronak.jain.it@gmail.com
ORCID: 0009-0009-8216-5071
Scopus ID: 60616825200
[2]IET CSIT, SAGE University, Indore, Madhya Pradesh, India.
Email: sachin.patel@sageuniversity.in
ORCID: 0000-0003-0563-1752
Scopus ID: 57198335451
Corresponding Author: Ronak Jain

Abstract: Healthcare is a major concern, among which heart disease is considered as one of the most important diseases where community has a big concern, when it is matter of accounting the ratio of global mortality rate along with morbidity. As the people are more aware now, and there is also easy availability of data sets, there is possibility of the acceptance of machine learning (ML) methods that make improvements in computational intelligence and have enhanced complete diagnosis of cardiovascular disease prediction and solution. At the current time, the beginning of Explainable Artificial Intelligence (XAI) has solved a problem of limitation that conventional system has of black-box modelling by permitting transparency and interpretability in system that make clinical decision power strong. As healthcare applications demand both predictive accuracy and trustworthiness, the integration of ML and XAI has become an important area of research in intelligent cardiovascular care. This systematic review examines contemporary machine learning methods and explainable AI procedures engaged prediction of heart diseases and classification. By collecting databases from various scientific sources and also ensuing a well-structured examination and while keeping in mind following also all the protocol, deep study is performed using these datasets. Also, all the strategies that must be followed during pre-processing step also prepared along with the approaches that will be applied during feature selection. All the required classification algorithms will also be identified with the calculation for metrics and work on explainability methods will be performed in order to identify the gap in the previous work. Also, emphasis will be also put on use of traditional machine learning methods, deep learning approaches and ensemble learning methods, along with post-hoc explanation that are human-understandable for understanding complex AI models. Some examples of these explainability models are SHAP, Saliency Maps, LIME, Integrated Gradients, attention-based interpretability mechanisms, Grad-CAM etc. This study also reviews and work on strengths, restrictions, gaps and practical consequences of already researched applications that present in real-world for medical diagnosis for better healthcare environments. But this becomes now mandatory to identify main research challenges that researchers are facing due to heterogeneity nature of data, privacy and security issues, imbalance of class, model generalizability, adoption by clinical practitioners, interpretability-performance trade-offs etc. All knowledge can be only gained after studying old research papers and do findings, so this research paper talks about all new developing trends and discuss future probable and research directions for emerging system with more transparency, better reliability, that must focus to benefit patients for the prediction of cardiovascular systems for medical diagnosis. The main aim in this paper is to give platform to the researchers, clinical practitioners and healthcare professionals for knowing the current scenario of cardiovascular disease prediction using machine learning approaches and also by

Copyright: © 2026 by the authors

applying explainability AI heart disease prediction models that will ease the process and also make the upcoming system with better progress and more trustworthy, operative and unfailing clinical solutions that makes system decision support one.
Keywords: Healthcare Analytics; Clinical Decision Support Systems; Heart Disease Prediction; Explainable Artificial Intelligence Machine & Deep Learning; Cardiovascular Disease; SHAP; Grad-CAM; LIME;

1. Introduction

Heart plays a major role in our body, so Cardiovascular related issues are also measured as vital disease for the human being as essential public health which is to be worked on worldwide. Such research plays a big, that makes a starring role in terms of survival of the person, humanity, mortality and other related overheads in the medical field [1]. Heart disease is considered to be one of the most leading cause for death among all types of cardiovascular disorders, which has these days no age limit and found in every groups [1], [2]. Here breakdown is predominantly due to its compound problems that are neglected and then it progresses gradually in the body without indications. There are many factors which is responsible for heart related diseases that includes obesity, hypertension, chain smoking, diabetes, up and down in cholesterol level, may be genetic from generations, sedentary lifestyle etc. that all leads to the expansion of heart-related difficulties [1]. Many times it has been seen that cardiovascular situations do not show any sign in their early steps and continue to asymptomatic. This is also a very big risk for the human life and so early diagnosis play a vital role in saving patient life and also lessons the mortality rate [1], [2].

Researchers are working with digital transformation even in the healthcare mechanism also very fast. Due to this, it is possible to get records of patients through various electronics machines and laboratories experiments that are generating clinical data in big volumes through reports, electronic wearable devices and different monitoring systems applied [2], [3]. All the above data, which is easily available is now creating some great opportunities in the field of healthcare system for making a smart system as solutions to challenge in heart related disease predictions. This is best way to prevent from disease and take care of human health at early stages. The AI-ML concepts, plays a role here and emerging as a best tool for identification of many complex problems that can be observed by clinical data and predict the problem at early stages with good accuracy and outcome [1], [3].

All the conventional machine learning approaches can be used for the heart disease prediction commonly named as Support Vector Machines, Logistic Regression, Random Forests Decision Trees, and K-Nearest Neighbours classifiers [1], [3]. All the mentioned algorithms, shows a better result in analysing the data of health in a very structured way with a good performance and also help clinical practitioners to avoid high risk of patient's life. There is also ensemble based and deep learning methods which also proves themselves and got an attention for the detection of health issues by their way to judge complex models as well using complex nonlinear relationships and also remove hidden patterns from large volume of datasets [3], [13]. Many researchers working on this topic reported good predictive way with greater accuracy using such models; however, it sometimes observed that still the technique and number of availability of particular datasets can be improved [15], [21].

Study in future should focus on acceptance of machine learning models more practically for healthcare system which still remains a big challenge many a times despites of all work. Transparency must be maintained in the system while using predictive models and all advances study should be applied [2], [5]. Use of better algorithms results in achieving superior performance, and not all just remain as a black-box system which has less understanding by clinical professionals that how early diagnosis can be done by predicting the right outcomes [5], [6]. There must be clear cut diagnosis, plan of treatment in the clinical environments to guarantee answerability and faith of the patient [6], [7]. Less trusted models are not used by any clinical practitioners and prove to be fail in providing an understandable explanation which may not further used by al related stakeholders and professionals [2], [7]. To link between Explainable Artificial Intelligence which is combined with predictive frameworks shows an automated output with well-organized established medical logic [2], [8].

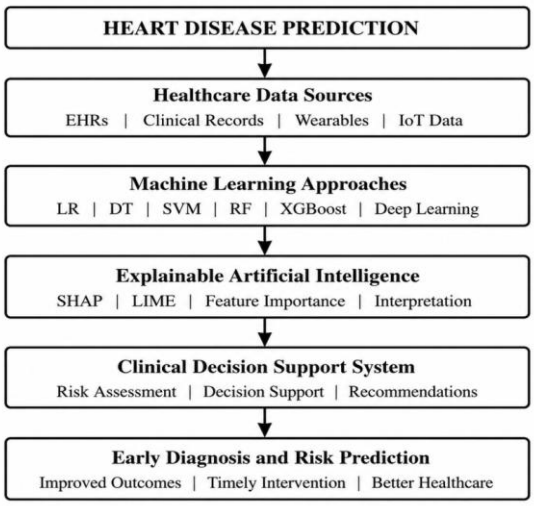

Figure 1. Framework of AI-driven heart disease prediction and clinical decision support.

Figure 1. Framework of AI-driven heart disease prediction and clinical decision support.

As the problem is challenging, so to avoid such challenges, introduction of XAI i.e. Explainable Artificial Intelligence has commanded over all present AI-ML Concepts that are really important research area to focus on for the use of different interpretability of machine learning models to make clear predictions with decisions. The intention of XAI approaches is to present reliable, trustworthy, transparent and most understandable result with delivering meaningful explanations for clinical showing model behaviour. This also predict outcomes by distinguishing the role of all distinguish features of the system and hence way to clear decision ways. XAI uses methods like SHapley Additive exPlanations, Local Interpretable Model-Agnostic Explanations, attention-based explanation mechanisms, Gradient-weighted Class Activation Mapping, feature importance analysis and other that is well-known for its efficacy to train the model for perfect transparency and help in supporting patients-centric decision-making systems. All clinicians believe in the work and style used in making such system and also understand the basic fundamentals applied for predictions, that how XAI like approaches benefits in building trust of them, confidence to use and making it answerable, for all AI-assisted healthcare systems [27], [28], [29].

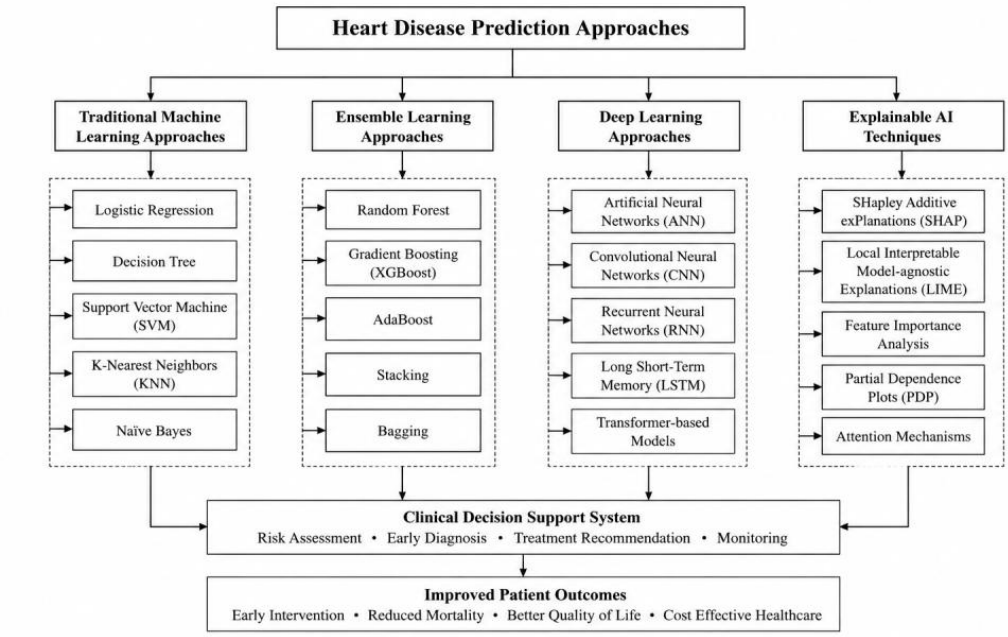

Figure 1. Taxonomy of machine learning and explainable AI approaches for heart disease prediction and clinical decision support.

Figure 2. Taxonomy of AI-based approaches for heart disease prediction.

By using the combination of technologies like machine learning, deep learning with the concepts of Explainable Artificial Intelligence, all the new potentials and opportunities are opening up that makes system, which is more acceptable, considerable for clinical decision support systems. Using Explainable AI approach for the prediction is benefiting not only by giving risk predictions and calculations accurately but also suggests and recommend all the powerful perceptions by working on factors that may resist and help in finding influences that indicate disease progression. In the medical systems, especially for cardiovascular- heart related diseases, such decision-making power is really very important, that helps for the survival and early diagnosis of patients and that make treatment easy to implement and also result with usefulness for future. Moreover, Explainability AI supplies ethical arrangement of technologies that provisions regular compliance of controls and also answerable for better use of healthcare tools. Though researchers have made noteworthy development in last few years, but still there are several challenges that is to be worked on. Many current studies depend on to work with less volume of data, but this can lead to undeveloped, incomplete trained model. There are few variances in pre-processing techniques of data, also changes in feature selection approaches, and some assessment procedures that also make work difficult in comparing of techniques and outcomes. Due to this, there are always issues that occur in terms of privacy of data, equality, biasness, and some clinical confirmation that not only affect outcomes but also hinder the better use of AI-driven healthcare solutions that can used world-wide. The nonexistence of standardised structures for the evaluation of explainability further highpoints the necessity for more wide-ranging approaches that must be also clinically relevant. This research paper is offerings an organized review of machine learning approaches with adding the concept of Explainable Artificial Intelligence methods for the better prediction of heart diseases along with the system that supports in clinical decision as well. The work suggested here critically observes all the existing approaches, used datasets, results, evaluation metrics, and explainability approaches that is useful for both understanding the strengths in the system and also to know limitations for the betterment. Here, work is done on the basis of the findings done from the review, also basic research gaps are recognized, and also survey of literature is performed to make a new outline of conceptual framework that is be designed which provision for transparency, reliability, and clinically significant for the prediction of heart disease. The outstanding section of this study is organised as follows. Section 2 is giving a complete review of existing literature that are associated to machine learning research with addition to explainable AI for cardiovascular healthcare. Section 3 designates the methodology that may help in a better performance. Section 4 delivers a reasonable analysis of all previous studies. Section 5 deliberates all recognized research gaps, suggestions for future works, and the work that be better from literature work done. Section 6 presents the conclusion of the presented work.

2. Literature Survey

Researchers find study on heart disease for the early prediction quite easier than before as this is one of the emerging topics of active interested domain in the field of healthcare system. This becomes relaxed with the help of easily available data that are recorded with the help of electronic health records systems applied to current patients. It uses advanced fundamentals of machine learning skills that helps in modelling various early predictive models that helps researchers in improving the accuracy for correct diagnosis and also supports for early intervention of the disease. Many researchers work signs that all current work follows the approaches of machine learning which found an effective way in finding the relationships i.e. between the clinical risk that are hidden and helps in analysing to clinical professionals for the diagnosis of disease and forecasting. **Rao et al. [9]** suggested a model that is XAI based where deep learning-based architecture is prepared for prediction of failure of heart and related incident. Authors have validated with the help of attention-based designs that has good potential for apprehending complex time-based relationships within the available clinical datasets. Also, **Ahsan and Siddique [3]** performed a review of literature in a systematic manner using machine learning centric methods for the diagnosis of heart disease and researched that the classifiers like decision trees, logistic regression, random forests, support vector machines etc. are the most frequently used techniques for the prediction of disease in cardiovascular research problems. New study started now focusing on use of engineering and optimised classification methods for the prediction. **Reddy et al. [10]** researched on machine learning methods which collaborated with attribute methods for evaluation. They compiled and compared the work done and found it to be effective as it gave the accuracy an improved one. Similarly, **Rani et al. [11]** established a better decision system that supports clinical in identifying high risks in patients by making an automated system with the help of machine learning approaches. **Gao et al. [12]** also made a system using ensemble learning methods which really outperform as it combines the various individual classifiers and also checked for the strengths of all those applied for prediction in healthcare. Here, the main theme is to improve the prediction of disease in the system with better reliability, so for the same researchers are working lot and tries to incorporate more than one method for better results using optimisation practices with the machine learning methods. **El-Shafiey et al. [13]** prepared a system based on hybrid technology using optimisation strategy (Particle Swarm Optimisation) and Genetic Algorithms for the prediction of heart diseases. They described that they found substantial developments using such technique in comparison with previous performances. **Gupta and Seth [1]** suggested a system and compared it with various machine learning and deep learning models and proved that deep learning methods are better in achieving higher predictive accuracy and they also clear that it is possible only with the help of greater size of data used and quality of datasets which made system effective. **Guleria et al. [8]** made a mark of the practicality of classification techniques when it is collective used with XAI for the risk prediction in health of cardiovascular diseases that achieved due to their interpretability. Another method, use of feature that is selected also plays a vital role in healthcare system for the prediction of heart diseases issues. **Dhumane et al. [14]** worked on cardiovascular disease prediction improvement methods by applying feature selection Relief and LASSO approach and recommended it due to as it witnessed better efficiency in the model and can predict for early diagnosis. Additionally, **Asif et al. [15]** suggested use of ensemble learning hybrid model that involved hyper-parameter optimisation approaches in it that also augmented accuracy of the system with better robustness for the prediction of heart disease in healthcare systems. It's obvious that only with the help of machine learning now it is not possible for researchers to work with but with the help of availability of large volume of datasets of patient's health has ignited researchers to use techniques of advanced machine learning that also uses new approaches like explainability artificial intelligence XAI fundamentals also. **El-Sofany et al. [16]** prepared a model for the prediction of diseases related to heart with the use of XAI and advance machine learning architectures, it improves the use of methods by the practitioners who better understand the related outcome of the disease by using multiple classification methods for healthier performance. Likewise, **Mienye and Jere [17]** prepared the model with the help of an optimization technique using ensemble learning approach that also uses XAL fundamentals to make a strong work that leads to better transparency in the model with the best reliability among all. This combination of XAI approaches with advance machine learning is playing a big role for the researchers for future work in last few years. **Shah et al. [18]** defined a hybrid model that inculcate approaches like ensemble learning along with use of explainability AI models for finding the risk occurrence in cardiovascular treatments and hence also proved with better developments in reliability for the disease prediction. **Kumar et al. [19]** suggested a complete literature of prediction techniques using machine learning centric systems for heart related health and also focused on how this research is reliable and trusted by many clinicals for the growth of the system while using in various applications related to healthcare. **Banerjee and Paçal [20]** suggested and notify that accuracy in the system that predicted is improved one in comparison with last researchers in few years but still there is gap to study and include interpretability and the adoption of system clinically which is still less acceptable. **Talukder et al. [2]** prepared a system named as XAI-HD for the prediction of heart disease recognition and showed that with the use of explainability AI like approaches system can give better insights to clinical as it represents better behaviour of model and also contributes best of features that also resolves the problem of class unbalancing very systematically and also suggests for missing

data. Similarly, **Ganie et al. [21]** prepared a combination of an explainable AI with ensemble learning approach that results in working for various cardiovascular diseases datasets and also improved the accuracy and interpretability at the same time by better XAI models. Researchers are also having a big concern for clinical that they must be getting better decision support systems which can make their work fast and also get a system that is having augmented transparency so that patients and doctors both satisfactory levels must be matched. **Bilal et al. [22]** suggested a model with XAI-based system that is well-intelligent and also able to generate required interpretable cardiovascular system for the prediction of risks on a large-scale dataset as well. Also, **Kailasanathan et al. [23]** made a framework that includes sensor based system which is called as a feature-sensitised interpretable model that emphasize on Internet of Medical Things sensors and shown that transparency based feature are highly required attribute for any mechanisms that really helps clinical for building a better confidence and also ready to use for future with more recommendations. All the current work is done with advanced explainability AI models for cardiovascular disease prediction. **Cenitta et al. [24]** prepared a framework based on ischemic heart disease prediction using XAI that added high squirrel search type attribute as feature selection. With this strategy and with the use of SHAP-based interpretation methods. **Kehkashan et al. [25]** also worked for the sustainable and interpretable growth in the healthcare field, that totally help practitioners for future systems to be prepared for prediction and also balances performance of prediction with higher transparency and justice. **Gul et al. [26]** also suggested a better review method that make a use of ensemble learning approaches as most prepared and an effective one for managing heterogeneous cardiovascular datasets. All the use of explainable artificial intelligence by the researchers gained better attention as it is theoretically much stronger and useful. **Tjoa and Guan [27]** suggested a system that emphasize on complete survey related to XAI & its use in medical diagnosis and also declared XAI as most appreciable and fundamentally a required one. It has gain reliability and trustworthiness in many healthcare systems based on AI. **Lundberg and Lee [28]** suggested SHAP, like technique that comes up as one of the most widely used approach of XAI that interpret with machine learning fundamentals. Equally, **Vilone and Longo [29]** suggested also a literature review on explainable artificial intelligence use methodologies systematically and emphasized about their importance and growth across healthcare applications. Many study have focused especially on explainable AI frameworks for the prediction of cardiovascular based diseases. **Moreno-Sánchez [30]** showed that how XAI methods will be used and augment the interpretability of chance of heart failure for the survival suggestion with the prediction in the model. **Ali et al. [31]** suggested a smart healthcare monitoring model that is implemented using an ensemble deep learning approach and also added feature fusion strategies for the prediction of cardiovascular disease. Currently, **Sourov et al. [32]** implemented an XAI-based system that uses machine learning framework and also improves the risk assessment of the patient, also **Sen and Bhattacharya [33]** prepared a model with explainable AI framework hybrid with Categorical Boosting with the Bees algorithm for the early heart disease detection. At the conclusion of review studied, it is suggested that all machine learning fundamentals is improving prediction of heart diseases with better competences, and also explainable artificial intelligence concepts when added includes more augmented transparency and trust in predictive systems. So, the development of clinically deployable frameworks always needs a balanced combination of predictive performance, trust, interpretability, reliability, robustness, and ethical based considerations. The reviewed studies disclose all frequent limitations that are related to explainable AI methods evaluation, using external validation, considering privacy preservation, and all clinical deployment for practitioners; these issues are solved in detail in coming Section 5.

3. Review Methodology

3.1 Review Design and Objectives

This study work brings about the systematic overview on recent trends in Machine Learning (ML) and Explainable Artificial Intelligence (XAI) in prediction of heart disease and a suitable for creating a detailed report to aid further decision-making. Although the expansion in the number of usages of smart health care systems offer numerous benefits in the areas of the predictive modelling; however, there exist a few challenges pertaining to interpretations judgement, and clinical transparency which require to be resolved for practical implement on medical solutions. Therefore, this present article is focused on comprehensively reviewing existing strategies to identify the newest trends that will allow for easy detection of the constraints and to expose potential opportunities in creating reliable cardiovascular condition prediction frameworks. This work attempts to deal with several of these above challenges by finding the solutions towards the following study concerns:

1. RQ1: What ML algorithms frequently utilized for heart failure projection?
2. RQ2: How the explanation from XAI models combined to cardiovascular disease prediction methodologies?

3. RQ3: What datasets, performance procedures, evaluation strategies utilised most regularly in research for cardiovascular diseases (CVD)?
4. RQ4: What are the foremost constraints and research limitations presented by prediction modelling systems which already existing? RQ5: What should be the promising future research for enhancing the interpretation, integrity as well as suitability for clinicians?

Research Question	Objective
RQ1	Identify commonly used ML algorithms for heart disease prediction
RQ2	Analyse the role of XAI techniques in healthcare applications
RQ3	Examine datasets, evaluation metrics, and validation methods
RQ4	Investigate limitations of existing studies
RQ5	Identify future research opportunities and challenges

Table 1 Research questions covered by the review

3.2 Literature Search Strategy

The study carried out systematic searching of relevant research published on heart disease prediction, machine learning, explainable artificial intelligence, and clinical decision support systems. The procedure involved exploring databases, fetching relevant articles via keyword search, screening and filtering articles based on their eligibility criteria. In general, we adhered to PRISMA guidelines for identification, screening, eligibility and inclusion of the research work. We discussed methodology of selection criteria, explainable mechanisms, research gap on various aspects based on selected publications. All publication data are gathered from the following six scientific repositories: Scopus, Web of Science, IEEE Xplore, ScienceDirect, SpringerLink, and PubMed. Those databases were included because all selected databases contain peer-reviewed publications that concern healthcare analytics, artificial intelligence, and medical informatics, and biomedical engineering research. We searched using an array of Boolean searches, combining the keywords: "heart disease prediction," "cardiovascular disease prediction," "machine learning," "deep learning," "explainable artificial intelligence," "XAI," "SHAP," "LIME", and "clinical decision support.

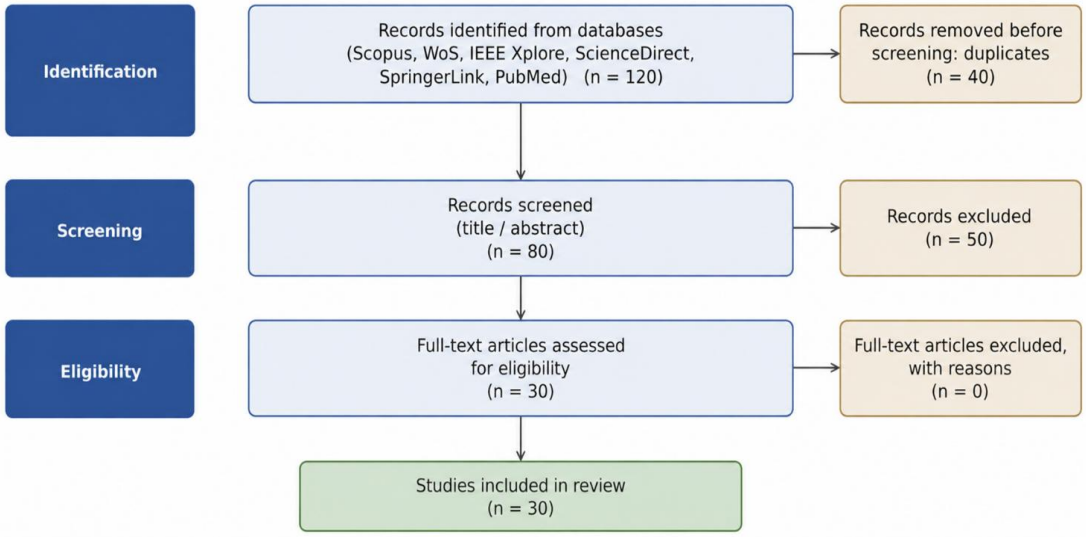

Figure 3. PRISMA 2020 flow diagram of study identification, screening, eligibility, and inclusion.

3.3 Study Selection Criteria

The following specific criteria for inclusion and exclusion were identified and used for uniformity and relatability purposes in our search: - Papers published between 2020 and 2026 inclusive. Only peer reviewed scientific journal papers or conference papers. The paper must aim to either predict heart disease or assess cardiovascular disease

risk. The methodology of the papers must involve machine learning (ML), deep learning (DL), or explainable artificial intelligence (XAI) methodologies or algorithms. The study must have conducted performance evaluations for their proposed models using metrics. Double publication entries. The article must be published in English language. Editorials, short communications and perspective papers should be excluded from the selected group. Any article whose scope of study is not the prediction of cardiovascular disease. Any publication that does not show any information about their methodologies and experiments.

3.4 Classification Framework

To assist with system analysis, the collected references were classified into five categories depending on their research methodologies and objectives.

Category	**Representative Techniques**
Traditional Machine Learning	Logistic Regression, Decision Tree, SVM, KNN
Ensemble Learning	Random Forest, XGBoost, AdaBoost
Deep Learning	ANN, CNN, RNN, Transformers
Explainable AI	SHAP, LIME, Feature Importance Analysis
Clinical Decision Support	Risk Assessment and Decision Support Systems

Table 2. A classification framework for reviewing literature.

3.5 Summary of Methodology

This review followed a planned methodology to select, assess, and categorize recent research studies about machine learning and explainable AI techniques for prediction of heart disease. All the collected relevant articles were searched using various scientific databases with the help of selected strategies and criteria, thereafter, divided based on their methodological aspect in order to allow comparison and finding research gap in the existing domain. Results of the selection and classification helped to build up comparative analysis and discussion as provided in following sections.

4. Comparative Analysis of Existing Studies

Machine learning, driven by the fast pace of its progress, has brought into use several effective frameworks for predicting heart disease and clinical decisions; however, the employed data sets vary considerably, and the feature engineering, model algorithms, explanation method, and valuation method vary considerably, the work carried out has been inconsistent. In this section, a comparative survey and the recent works are reviewed.

4.1 Comparative Analysis of Machine Learning Approaches

Underlying these modern approaches to the prediction of Cardiovascular Disease Machine learning models. Many basic classifiers: logistic regression, decision trees, support vector machines (SVM), k-nearest neighbours (KNN), are commonly used and are capable of modelling the structured health information, owing to their simplicity. These methods, the authors argued [3] will remain common use. The main reasons cited are, that this is so, since they give interpretable results and do not require high computing resources. Of the standard approaches, logistic regression is one of the widely adopted and is well known because it estimates the likelihood of the condition and explains the reasoning behind it but tends to give moderate results in modelling nonlinear relation, though decision trees offer greater readability. SVM are known to exhibit efficient high dimensional classification, though require appropriate parameter adjustment. More lately, ensemble methods are gaining considerable interest. Indeed Gao et al [12], Asif et al [15], and Ganie et al [21], showed that using ensemble methods, such as random forest and XGBoost (which were the most popular due to the robust performance, provision of Feature Importance measures and the ability to resist over fitting), provided high accurate prediction compared with a single predictor algorithm.

Algorithm	**Advantages**	**Limitations**
Logistic Regression	Simple, interpretable, fast training	Limited nonlinear learning capability

Decision Tree	Easy visualisation and interpretation	Overfitting risk
KNN	Simple implementation	Sensitive to noisy data
SVM	Effective in high-dimensional datasets	Computationally expensive
Random Forest	High accuracy and robustness	Reduced interpretability
XGBoost	Superior predictive performance	Complex parameter tuning
ANN	Learns complex patterns	Black-box behaviour

Figure 3. A Comparative Analysis of Different ML Algorithms

4.2 Comparative Analysis of Explainable AI Techniques

Machine learning algorithms achieved strong predictive accuracy but a significant impediment to their deployment in a clinical environment remains lack of interpretation of their decision processes. As a consequence, Explainable Artificial Intelligence (XAI) tools received a strong focus of attention in cardiovascular research and clinics. Shapley value based Explaining method (SHAP) presented by Lundberg and Lee [28] is one method that allows a global and local explanation of model behaviour by explaining the contribution of a given feature to an individual prediction thus can reveal dominant factors that contribute to risk factors for that specific patient. LIME a method that uses a local surrogate linear model to explain black-box predictions for an individual data instance. SHAP and LIME can both significantly explain models better and help doctors understand their machine learning outputs as per Talukder et al. [2], Cenitta et al. [24].

XAI Technique	**Characteristics**	**Advantages**	**Limitations**
SHAP	Game-theoretic explanation	Local and global interpretability	Computationally intensive
LIME	Local surrogate model	Model-agnostic	Limited global understanding
Feature Importance	Ranking-based explanation	Simple interpretation	Limited interaction analysis
Attention Mechanisms	Deep learning explanation	Useful for neural networks	Difficult to validate clinically
Counterfactual Explanations	What-if analysis	Human-friendly explanations	Computational complexity

Table 4. Comparability study of interpretability methods.

Comparability study of interpretability methods. Recent studies highlight that SHAP has become a common choice of interpretability methods, as it can systematically apply the given method to several machine learning models to yield full explanations.

4.3 Comparative Analysis of Datasets

The performance of the system for prediction of heart disease depends highly on the features of the dataset and sample diversity. Most of the articles utilize free access dataset for study, predominantly the Cleveland Heart Disease dataset, as a benchmark dataset. This helps the researchers to have easy comparison in terms of accuracy and computational cost, but the availability of few subjects or lack of variety makes them generalized only up to a certain extend.

Dataset	**Samples**	**Features**	**Application**
Cleveland	303	14	Heart disease classification
Statlog	270	13	Cardiovascular risk prediction

Hungarian	294	14	Disease diagnosis
Switzerland	123	14	Comparative studies
Framingham	4,000+	15+	Long-term cardiovascular risk

Table 5. Heart disease datasets

A variety of more recent works emphasize that the model performance and generalizability will be strengthened by obtaining the training data set of multi-institutional sources and higher volume. The work of Bilal et al. [22] is a recent, large dataset of 300,000 more cases, for example, reflecting that cardiovascular dataset sizes are increasing.

4.4 Comparative Analysis of Evaluation Metrics

The efficacy of models using machine learning approaches typically includes classification performance metrics to report performance and outcomes. As noted in this paper, accuracy is the most common metric used but can be misleading for use cases like imbalanced health care datasets where reporting on the outcomes such as precision, recall, F1-score, and ROC-AUC may provide a better, more accurate reflection of the classification performance. In the formulas listed, the terms TP, TN, FP, FN refer to the true positive, true negative, false positive and false negative terms, respectively.

Metric	**Definition**	**Formula**
Accuracy	Proportion of all correct predictions	(TP + TN) / (TP + TN + FP + FN)
Precision	Correct positives among predicted positives	TP / (TP + FP)
Recall (Sensitivity)	Correct positives among actual positives	TP / (TP + FN)
F1-Score	Harmonic mean of precision and recall	2 × (P × R) / (P + R)
ROC-AUC	Class-separation capability	Area under the ROC curve

Table 6. Performance metrics

4.5 Synthesis of Primary Studies

In order to facilitate direct comparison across the included primary research studies, a meta-synthesis for each paper's data set(s) used, main predictive algorithm, explainability technique, and validation procedure, is provided in landscape format in Tables 7, then a methodological quality and risk of bias analysis in Table 8. If the article presented only a single most performing algorithm in their data analysis, then this algorithm is stated here. If there was no method stated clearly that could uniquely verify accuracy, then cells have been labelled "verify from source" to be completed by article authors. It is not appropriate to directly compare accuracies across the dataset rows, as datasets, data pre-processing and, crucially, any data-balancing technique and validation method differ greatly.

Ref.	**Study (first author, year)**	**Best / principal model**	**Dataset(s)**	**XAI method**	**Validation**
[2]	Talukder et al., 2025	Hybrid ML + DL (XAI-HD)	CHD, FHD, SHD	SHAP	Multi-dataset; error reduced 20–25% vs. baselines
[3]	Ahsan & Siddique, 2021	Review (LR, SVM, DT, RF)	Multiple (review)	N/A (review)	Systematic review
[8]	El-Sofany et al., 2024	ML + XAI ensemble	Benchmark CVD	SHAP	Hold-out / CV (verify from source)

[9]	Rao et al., 2021	Transformer (attention)	EHR (longitudinal)	Attention	Temporal validation
[10]	Reddy et al., 2021	ML + attribute evaluators	Cleveland	Feature importance	Cross-validation
[11]	Rani et al., 2021	ML-based CDSS	Cleveland	Feature ranking	Hold-out
[12]	Gao et al., 2021	Ensemble	Cleveland / UCI	Feature importance	Cross-validation
[13]	El-Shafiey et al., 2022	RF + GA-PSO	Cleveland	N/A	Cross-validation
[15]	Asif et al., 2023	Tuned ensemble / MLP (CV acc. ≈ 87%)	Cleveland / UCI	N/A	Cross-validation + hyperparameter tuning
[16]	Mienye & Jere, 2024	Optimized ensemble + XAI	Benchmark CVD	SHAP / LIME	Cross-validation
[17]	Shah et al., 2025	Hybrid ensemble + XAI	Benchmark CVD	SHAP	Hold-out / CV (verify from source)
[20]	Ganie et al., 2025	Ensemble (multi-dataset)	Multiple CVD datasets	SHAP / LIME	Cross-dataset validation
[21]	Bilal et al., 2025	ML + XAI (large-scale)	Kaggle CVD (≈ 308,737 records)	SHAP	Hold-out on large dataset
[22]	Kailasanathan et al., 2025	Interpretable ML (IoMT)	IoMT sensor data	Feature-sensitised attribution	Hold-out
[23]	Cenitta et al., 2026	ML + enhanced squirrel-search FS	Ischemic HD dataset	SHAP	Cross-validation
[25]	Gul et al., 2026	Review (ensemble focus)	Multiple (review)	N/A (review)	Systematic review
[29]	Ali et al., 2020	Ensemble DL + feature fusion	Sensor + clinical	Feature fusion	Hold-out
[31]	Sen & Bhattacharya, 2025	CatBoost + Bees algorithm	Benchmark CVD	Hybrid XAI	Cross-validation
[32]	Fitriyani et al., 2020 (HDPM)	Ensemble (DBSCAN + SMOTE-ENN + XGBoost)	Statlog, Cleveland	N/A	Hold-out (CDSS prototype)

Table 7. Summary of assessed primary studies (datasets, models, explanation techniques, and verification).

The six evaluation points used to assess each included primary study were selected from published checklists on predictive-modelling and AI (identification of data source; transparency of preprocessing methods; rigour of model validation methods; reporting of several measures of performance including not just accuracy; assessment of explanatory properties of the model; external or clinical validity of model findings. Ratings of 'Y' represent 'yes, criteria met'; 'P' mean 'partially met', and 'N' represent 'no' or 'not reported'. Ratings for reviews are N/A where criteria were model level. They are this group of authors' assessment only and must be independently verified from a reading of the primary papers before they are sent to the journal.

Ref.	Study	Data source clear	Preprocessing transparent	Validation sound	Metrics beyond accuracy	Explainability evaluated	Ext./clin. valid.
[2]	Talukder et al., 2025	Y	Y	Y	Y	P	P
[3]	Ahsan & Siddique, 2021	N/A	N/A	N/A	N/A	N/A	N/A
[8]	El-Sofany et al., 2024	Y	P	P	Y	P	N
[9]	Rao et al., 2021	Y	P	Y	Y	P	P
[10]	Reddy et al., 2021	Y	Y	Y	P	N	N
[11]	Rani et al., 2021	Y	P	P	P	N	N
[12]	Gao et al., 2021	Y	P	Y	P	N	N
[13]	El-Shafiey et al., 2022	Y	P	Y	P	N	N
[15]	Asif et al., 2023	Y	Y	Y	Y	N	N
[16]	Mienye & Jere, 2024	Y	Y	Y	Y	Y	N
[17]	Shah et al., 2025	Y	P	Y	Y	Y	P
[20]	Ganie et al., 2025	Y	Y	Y	Y	Y	P
[21]	Bilal et al., 2025	Y	P	Y	Y	Y	P
[22]	Kailasanathan et al., 2025	Y	P	P	Y	Y	P
[23]	Cenitta et al., 2026	Y	Y	Y	Y	Y	N
[25]	Gul et al., 2026	N/A	N/A	N/A	N/A	N/A	N/A
[29]	Ali et al., 2020	Y	P	Y	Y	P	N
[31]	Sen & Bhattacharya, 2025	Y	Y	Y	Y	P	N
[32]	Fitriyani et al., 2020	Y	Y	Y	Y	N	P

Table 8. Methodological quality and risk-of-bias appraisal of included primary studies

Data sourcing and validation was generally poorly reported across the appraised studies but has been more thoroughly described recently. Most authors have reported on the external and clinical validation in their study, although all of the most recently appraised papers report either a single-data set or within-dataset cross-validation for external validation and the majority of most recently appraised papers (22 of the 33 from June to Dec 2021) used a single-site data set. Recently generated, but generally limited, was the explainability component, usually providing

either LIME or SHAP outputs without assessing the actual faithful and clinical value of such outputs. These observations align with the areas identified in Section 5 as lacking research development.

4.6 Quality and Risk-of-Bias Assessment

We assessed methodological rigour for each of the studies, using 6 criteria which have been inspired by recognised evaluation schemes for predictive modelling and for explainability research (i-vi) (in line with PROBAST, and some recent guidance for explainability research): (i) Description of the source dataset & population, (ii) Description of preprocessing, data preparation and handling of missingness / class imbalances, (iii) Description of validation method, (iv) Reporting of appropriate metrics other than accuracy, (v) Reporting the quality / faithfulness of the explanation, and (vi) Description of domain expertise involved in the study/validation of the outputs. In doing so we identify a consistent pattern in the studies within our review: issues around data specification, reporting metrics and method choice seem to be well covered and most XAI-focused research also clearly describes explainability methods and associated computations. Two key areas of concern appear frequently in the literature however: the quality/faithfulness of explanations generated seem to be less often assessed, with many studies reporting e.g. SHAP values without verifying these claims in terms of truthfulness [33], and secondly there is very limited use of clinical domain expert knowledge and clinical validation of model performance, as reported in [33]. High risks for bias stemming from dataset homogeneity are also noted as multiple studies used the same few, small, benchmark datasets (e.g. Cleveland, Statlog) leading to potential overestimates in performance and which fail to generalise across different populations, Table 8 lists this per study assessment using a coding schema where Y=satisfied, P=partial, N=not satisfied or not specified. In the assessment reported below (Table 8 below T8), we provide this data as we as interpreted directly from the papers themselves, and we advise that review authors would ideally crosscheck this with the source texts before final submission.

4.7 Summary of Comparative Findings

The comparative analysis reveals several important trends in current heart disease prediction research:

- Ensemble learning methods, particularly Random Forest and XGBoost, generally achieve superior predictive performance compared with traditional algorithms.
- Many machine & deep learning models prove that strong learning skills but still they grieve with the limitations of interpretability.
- SHAP and LIME ate the concepts from explainability AI techniques that are emerging as most significant approaches especially for cardiovascular like applications.
- Most readings of research also rely on datasets which are available with limitations due to demographic diversity that can lead to generative AI models.
- Accuracy is also one of the vital metrics that remains dominant, while all latest researches are increasingly putting highlight on values of precision, ROC-AUC, recall and F1-score.
- Though there are advancements in predictive performance, transparency, fairness, privacy, and clinical validation but these are to be worked on more for addressing them sufficiently.

All the above-mentioned findings highpoint for the need for explainable, trustworthy, and clinically deployable AI structures that make a balance predictive accuracy along with transparency and reliability.

5. Research Gaps, Future Directions, and Conceptual Framework

5.1 Research Gaps

The analysis and work of researchers review that AI-ML i.e. fundamentals of machine learning and deep learning methods have meaningfully amended predictive performance in the system, however there are several limitations that still needs to work on for the proper adopting it into real-world healthcare atmospheres. The most noteworthy challenge is the limitations of interpretability in the progressive prediction models: the solution to such problem uses models like ensemble and deep learning fundamentals that is attaining high accuracy but not able to deliver reasoning as required, which is creating the system difficult for medical clinicians to comprehend and authenticate recommendations for early diagnosis. There is also another limitation that there is the excessive presence of low number of datasets like Cleveland and Statlog due to which demographic diversity can affect the performance due to limited size that may also disturb generalizability through all the healthcare settings done for the patient populations to be served. Also, the present analysis shows great heterogeneity in feature selection; choice of metrics; validation procedures; hence the cross-study comparisons would result unavailing. Prediction accuracy alone has been

given preference over interpretability, fairness, robustness, clinical applicability; there is growing interest in data privacy and security that are hampered because heart and patients' personal information that should be properly stored, shared; the incorporation of privacy preserving learning and XAI in cardiovascular healthcare has been only little considered. Such issues have been pointed out in literature based on evaluations, the review of evaluation procedures in Cardiology indeed points out the relative few number of studies examining the quality and the actual reliability of explanations; similarly, only a limited number of works included doctors for the purpose of validating the concordance of AI explanatory models with existing medical logic, and no evaluation guidelines have yet been proposed specific for tabular data, nor for time series in Cardiology applications. Large models continue to be unable to comply with medicine requirements on transparency. Cumulatively, the main research gaps arising from the above is:

1. Opacity of predictive models: Difficulty to explain the logic behind a given prediction or prognosis.
2. Reliance on narrow and relatively homogenous databases: Risk of lack of generalizability and potential biases of training data.
3. Inadequate external clinical validation and clinical proof of real benefit.
4. No standardized explainability evaluation approaches: Lack of metrics or protocol to assess the level of explainability of a model and its impact on users.
5. Privacy-preserved AI and XAI not considered: Privacy considerations are often omitted in XAI studies, despite being relevant for clinical deployment.
6. Neglecting clinician-centered human factors and interfaces: Focus is often put on explainability alone, ignoring how clinicians want to interact with explainable models.
7. Lack of trust, clinical relevance, and implementable AI solutions.

5.2 Implications for Future Research

To address such issues, future work will have to place emphasis on clinician-centric performance metrics that characterize the trustworthiness and learnability of the explanations produced by AI systems [7], [33]. It would be very important to show through large-scale clinical studies how XAI based systems have the effect to enhance the clinician's decision process under natural environment [2], [21]. The computation cost of the explanation tool such as SHAP has to be taken into account, so that the implementation remains possible on devices under difficult computational constraints in real settings [24]. Standardization of measurement procedures on various datasets of cardiovascular cases should also lead to better results that hold for a wide range of cases and populations [18], [34].

5.3 Literature-Derived Conceptual Framework

Building on the review of evidence A Conceptual early Heart Disease Detection and Clinical Decision Support system derived from literature-based design was developed. This system is only conceptual, represents the culmination of a search for recurrent designs and as such, cannot be presented as having an experimentally validated system design. Machine-based predictions with some explainability elements were proposed in combination to facilitate trust of the physician in the system by rendering it reliable and transparent in six interrelated components.

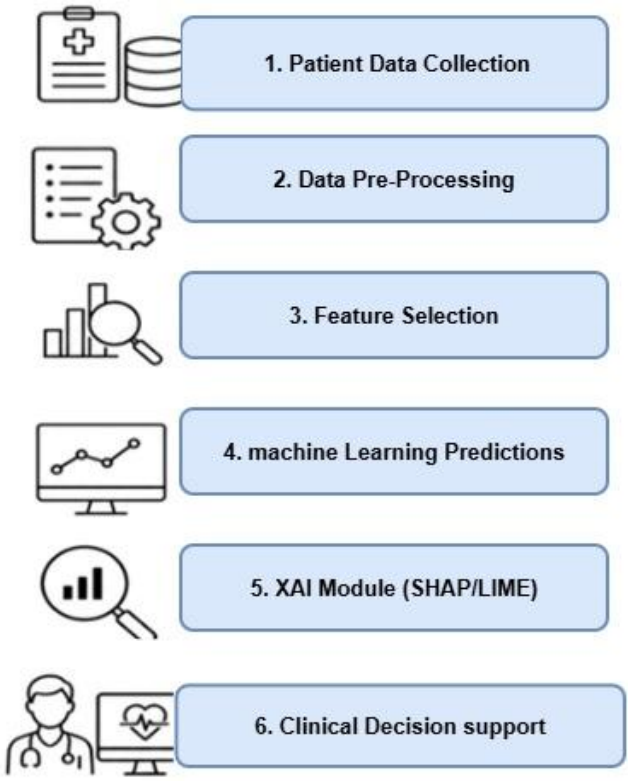

Figure 4. Concept Framework of the Literature for prediction of early heart disease and clinical support system.

Phase 1 - Collect Patient Health Data - First, information related to a patient's overall health record is procured, including general demographics, patient physical and vital signs, patient medical report data, cardiovascular risks (age, blood pressure, cholesterol, fasting blood sugar levels), heart rates, ECG (Electrocardiography), and features related to the onset and presence of chest pain.

Phase 2 - Data Preprocessing - The procured health data is cleaned and undergoes such steps as identifying and imputing values, dropping records, removing duplicated values, converting scales, and outlier treatment and transformation to fit the requirements for machine learning implementation.

Phase 3 - Feature selection - Through various methods for feature selection, more prominent and relevant predictive variables can be chosen; thereby reducing problem space dimension, simplifying computation, and enhancing classification by removing unrelated factors or less useful variables to assist ML models in better analysis.

Phase 4 - ML Predictions - Features derived using the selections described above are passed onto ML models, specifically, but not limited to; Logistic Regression, random forests, Support Vector Machine(SVM) and gradient based boosters like (XGBoost), ensemble machine learning, to make accurate classification of a patient as high-risk or low-risk against cardiovascular disease.

Phase 5 - Explainable AI Module - Prediction outcomes from models are then interfaced with tools for Explainable AI (XAI) technologies (like SHAP, LIME), which generate both global and local explanations of individual clinical features contributing to a prediction of a patient as low/high-risk. This functionality helps explain predictions to the doctor for decision making.

Phase 6 - Clinical Decision Support - The insights obtained from explainable AI are translated to concrete, understandable information; this would aid clinicians in providing clearer decision pathways to prescribe proper clinical decision strategies based on predicted risk level of a given patient and clinical interpretation is important [11], [22]. Sustainable, and explainable, models contribute to its usability in the real biomedical, clinical healthcare [25], [34]; Real time explainable risk forecasts improve early intervention and outcomes [35].

6. Conclusion

Heart disease is the most impactful of global health problems making a reliable early prediction system quite necessary. This systematic study has evaluated latest trends for ML, Explainable AI for prediction of heart disease &clinical support, highlighting the excellent performance in machine learning like ensemble techniques Random Forest and XGBoost for prediction of cardiovascular risk and in learning the correlations with complex relations in medical data for deep learning- based predictive techniques. However, the full potential of AI-enhanced healthcare technology, which makes predictive models explainable; has not been unlocked due to lack of transparency in models used in earlier days. This literature work here elucidates the enhanced integration of explainability methods such as SHAP and LIME to the machine learning- based predictive models and also sheds light on numerous challenges existing within system such as:- limited data samples availability, the absence of standardize valuation, minimal external validations, privacy problems, and also the lack of integration of explainability concept in real clinical environments to support medical decision-making. There is thus a pressing need for a literature-based conceptual structure that combines analysis of medical data and machine learning models, in a method supported by understandable explanation equipment for much better clinical judgment to enhance outcomes of healthcare delivery to heart patient, via providing better predictive and diagnostic aid with no compromises over clinical responsibility.. Future research should focus on developing robust, transparent, and clinically validated frameworks by incorporating large-scale healthcare datasets, federated learning, multimodal data analytics, and clinician-centred evaluation strategies, facilitating trustworthy and practical AI-assisted healthcare systems capable of supporting timely diagnosis, personalised treatment planning, and improved patient care.

References

1. P. Gupta and D. Seth, “Comparative analysis and feature importance of machine learning and deep learning for heart disease prediction,” Indonesian Journal of Electrical Engineering and Computer Science, vol. 29, no. 1, pp. 451–459, Jan. 2022, doi: 10.11591/ijeecs.v29.i1.pp451-459.
2. Md. A. Talukder, A. S. Talaat, M. Kazi, and A. Khraisat, “XAI-HD: an explainable artificial intelligence framework for heart disease detection,” Artificial Intelligence Review, vol. 58, no. 12, Oct. 2025, doi: 10.1007/s10462-025-11385-6.
3. M. M. Ahsan and Z. Siddique, “Machine learning-based heart disease diagnosis: a systematic literature review,” arXiv preprint, Dec. 2021, doi: 10.48550/arXiv.2112.06459.
4. S. Bharati, M. R. H. Mondal, and P. Podder, “A review on explainable artificial intelligence for healthcare: why, how, and when?,” IEEE Transactions on Artificial Intelligence, vol. 5, no. 4, pp. 1429–1442, Apr. 2023, doi: 10.1109/TAI.2023.3266418.
5. J. Andritsch and J. Dike, “XAI-driven explainability for cardiovascular diseases prediction,” Journal of Informatics and Web Engineering, vol. 5, no. 1, p. 167, Feb. 2026, doi: 10.33093/jiwe.2026.5.1.10.
6. P. Guleria, P. N. Srinivasu, S. Ahmed, N. Almusallam, and F. K. Alarfaj, “XAI framework for cardiovascular disease prediction using classification techniques,” Electronics, vol. 11, no. 24, p. 4086, Dec. 2022, doi: 10.3390/electronics11244086.
7. E. Tjoa and C. Guan, “A survey on explainable artificial intelligence (XAI): toward medical XAI,” IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 11, pp. 4793–4813, Nov. 2021, doi: 10.1109/TNNLS.2020.3027314.
8. H. F. El-Sofany, B. Bouallègue, and Y. M. A. El-Latif, “A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method,” Scientific Reports, vol. 14, no. 1, p. 23277, Oct. 2024, doi: 10.1038/s41598-024-74656-2.
9. S. Rao et al., “An explainable transformer-based deep learning model for the prediction of incident heart failure,” arXiv preprint, Jan. 2021, doi: 10.48550/arXiv.2101.11359.
10. K. V. V. Reddy, I. Elamvazuthi, A. A. Aziz, S. Paramasivam, H. N. Chua, and S. Pranavanand, “Heart disease risk prediction using machine learning classifiers with attribute evaluators,” Applied Sciences, vol. 11, no. 18, p. 8352, Sep. 2021, doi: 10.3390/app11188352.
11. P. Rani, R. Kumar, N. Ahmed, and A. Jain, “A decision support system for heart disease prediction based upon machine learning,” Journal of Reliable Intelligent Environments, vol. 7, no. 3, pp. 263–275, Jan. 2021, doi: 10.1007/s40860-021-00133-6.
12. X. Gao, A. A. Ali, H. S. Hassan, and E. M. Anwar, “Improving the accuracy for analyzing heart diseases prediction based on the ensemble method,” Complexity, vol. 2021, art. 6663455, Jan. 2021, doi: 10.1155/2021/6663455.
13. M. G. El-Shafiey, A. Hagag, E. A. El-Dahshan, and M. A. Ismail, “A hybrid GA and PSO optimized approach for heart-disease prediction based on random forest,” Multimedia Tools and Applications, vol. 81, no. 13, pp. 18155–18179, Mar. 2022, doi: 10.1007/s11042-022-12425-x.
14. A. Dhumane, P. Kaldate, A. Sawant, P. Kadam, and V. Chopade, “Efficient prediction of cardiovascular disease using machine learning algorithms with Relief and LASSO feature selection techniques,” in Lecture Notes in Networks and Systems, Springer, 2023, pp. 677–693, doi: 10.1007/978-981-99-3315-0_52.
15. D. Asif, M. Bibi, M. S. Arif, and A. Mukheimer, “Enhancing heart disease prediction through ensemble learning techniques with hyperparameter optimization,” Algorithms, vol. 16, no. 6, p. 308, Jun. 2023, doi: 10.3390/a16060308.

16. I. D. Mienye and N. Jere, “Optimized ensemble learning approach with explainable AI for improved heart disease prediction,” Information, vol. 15, no. 7, p. 394, Jul. 2024, doi: 10.3390/info15070394.
17. P. Shah, M. Shukla, N. H. Dholakia, and H. Gupta, “Predicting cardiovascular risk with hybrid ensemble learning and explainable AI,” Scientific Reports, vol. 15, no. 1, p. 17927, May 2025, doi: 10.1038/s41598-025-01650-7.
18. R. Kumar et al., “A comprehensive review of machine learning for heart disease prediction: challenges, trends, ethical considerations, and future directions,” Frontiers in Artificial Intelligence, vol. 8, p. 1583459, May 2025, doi: 10.3389/frai.2025.1583459.
19. T. Banerjee and İ. Paçal, “A systematic review of machine learning in heart disease prediction,” Turkish Journal of Biology, vol. 49, no. 5, pp. 600–634, Oct. 2025, doi: 10.55730/1300-0152.2766.
20. S. M. Ganie, P. K. D. Pramanik, and Z. Zhao, “Ensemble learning with explainable AI for improved heart disease prediction based on multiple datasets,” Scientific Reports, vol. 15, no. 1, p. 13912, Apr. 2025, doi: 10.1038/s41598-025-97547-6.
21. A. Bilal, A. Alzahrani, K. Almohammadi, M. Saleem, M. S. Farooq, and R. Sarwar, “Explainable AI-driven intelligent system for precision forecasting in cardiovascular disease,” Frontiers in Medicine, vol. 12, p. 1596335, Jul. 2025, doi: 10.3389/fmed.2025.1596335.
22. N. Kailasanathan, G. Ezhilarasan, S. Selvarajan, R. K. Dhanaraj, D. Pamucar, and N. S. Shankar, “Heart disease prediction with a feature-sensitized interpretable framework for the Internet of Medical Things sensors,” Frontiers in Digital Health, vol. 7, p. 1612915, Oct. 2025, doi: 10.3389/fdgth.2025.1612915.
23. D. Cenitta, N. Arul, R. V. Arjunan, K. Chadaga, and J. Andrew, “An explainable artificial intelligence framework for ischemic heart disease prediction using enhanced squirrel search feature selection,” Scientific Reports, vol. 16, no. 1, Apr. 2026, doi: 10.1038/s41598-026-46823-0.
24. T. Kehkashan et al., “Sustainable and interpretable heart disease prediction: a clinical decision support approach for biomedical healthcare applications,” Scientific Reports, Feb. 2026, doi: 10.1038/s41598-026-37840-0.
25. G. Gul et al., “Machine learning and ensemble methods for cardiovascular disease prediction: a systematic review of approaches, performance trends, and research challenges,” Computers, vol. 15, no. 1, p. 25, Jan. 2026, doi: 10.3390/computers15010025.
26. S. M. Lundberg and S.-I. Lee, “A unified approach to interpreting model predictions,” in Advances in Neural Information Processing Systems (NeurIPS), 2017, pp. 4765–4774, doi: 10.48550/arXiv.1705.07874.
27. G. Vilone and L. Longo, “Explainable artificial intelligence: a systematic review,” arXiv preprint, May 2020, doi: 10.48550/arXiv.2006.00093.
28. P. A. Moreno-Sánchez, “Improvement of a prediction model for heart failure survival through explainable artificial intelligence,” Frontiers in Cardiovascular Medicine, vol. 10, p. 1219586, Aug. 2023, doi: 10.3389/fcvm.2023.1219586.
29. F. Ali et al., “A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion,” Information Fusion, vol. 63, pp. 208–222, Jun. 2020, doi: 10.1016/j.inffus.2020.06.008.
30. M. E. A. Sourov et al., “An explainable AI-enhanced machine learning approach for cardiovascular disease detection and risk assessment,” arXiv preprint, 2025.
31. J. Sen and S. Bhattacharya, “An explainable hybrid framework for early detection of cardiovascular diseases using Categorical Boosting and the Bees algorithm,” Scientific Reports, vol. 15, no. 1, p. 45748, Dec. 2025, doi: 10.1038/s41598-025-28514-4.
32. N. L. Fitriyani, M. Syafrudin, G. Alfian, and J. Rhee, “HDPM: an effective heart disease prediction model for a clinical decision support system,” IEEE Access, vol. 8, pp. 133034–133050, Jan. 2020, doi: 10.1109/ACCESS.2020.3010511.
33. A. Salih et al., “A review of evaluation approaches for explainable AI with applications in cardiology,” Artificial Intelligence Review, vol. 57, no. 9, p. 240, Aug. 2024, doi: 10.1007/s10462-024-10852-w.
34. F. Di Martino and F. Delmastro, “Explainable AI for clinical and remote health applications: a survey on tabular and time series data,” Artificial Intelligence Review, vol. 56, no. 6, pp. 5261–5315, Oct. 2022, doi: 10.1007/s10462-022-10304-3.
35. S. Abhishek Sai, “Explainable AI (XAI) for a machine learning heart disease prediction model,” M.S. thesis, California State Univ., San Bernardino, CSUSB ScholarWorks, May 2025. [Online]. Available: https://scholarworks.lib.csusb.edu/etd/2241.