IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, May 2025
Impact Factor: 7.67

A Review of Heart Attack Prediction Systems Using Machine Learning Techniques

Supriya S. Patil and Prof. S. R. Shegar
Department of Computer Engineering
Samarth College of Engineering & Management, Belhe, India
profsupriyapatil@gmail.com and hodcompcoe@gmail.com

Abstract: *Cardiovascular illnesses (CVDs), specifically heart attacks, are still the most common cause of death globally. It is vital to identify at-risk individuals at an early stage in order to facilitate early intervention and minimize fatal cases. Conventional methods are very useful but tend to fail in their ability to foretell heart attacks at an early stage because they are based on linear models and very few interaction features. With the high-rate increase in health care information and computing capabilities, machine learning (ML) is a new-age technology for ensuring the improved precision and efficacy of heart attack forecasting systems. This review delves into the existing state of ML methods used for heart attack prediction. It investigates a broad spectrum of algorithms such as logistic regression, decision trees, support vector machines, random forests, artificial neural networks, and ensemble methods. The paper discusses widely used datasets, performance metrics, and the importance of feature selection and preprocessing. It also emphasizes the strengths and weaknesses of different models and presents a comparative study of recent research in the area. Challenges like data imbalance, explainability of the model, generalizability, and clinical workflow integration are also discussed. The review summarizes by naming upcoming trends and avenues for future research, such as the application of deep learning, wearable devices, explainable AI, and federated learning for personalized and secure healthcare solutions. Through the synthesis of current literature, this paper seeks to present researchers and practitioners with a full understanding of the strengths and limitations of ML-based heart attack prediction systems, eventually helping to drive predictive cardiology forward.*

Keywords: Heart Attack Prediction, Cardiovascular Disease, Machine Learning, Artificial Intelligence, Healthcare Analytics, Predictive Modeling, Feature Selection, Deep Learning

I. INTRODUCTION

Cardiovascular disorders (CVDs), consisting of heart attack, stroke, and other cardiac and vascular disorders, are the major cause of mortality in the world. As per the World Health Organization (WHO), CVDs cause around 17.9 million deaths annually, accounting for 31% of total global deaths. Of these, it has been estimated that 85% result from heart attacks and strokes. The high risk of death and long-term effect on quality of life make prevention and early detection of cardiovascular events an international health priority.

Heart attacks, or myocardial infarctions (MIs), result when blood supply to the heart muscle is interrupted, typically as a result of a deposit of fat, cholesterol, and other material in the coronary arteries. This interruption can cause damage to or destruction of heart tissue, resulting in potentially life-threatening complications or death. Detection of heart attack risk at an early stage is important to allow preventive intervention, including lifestyle modification, medication, or surgery.

The rise in the prevalence of heart disease is fueled by numerous factors such as an older population, physically inactive lifestyles, poor diet, smoking, diabetes, hypertension, and stress. Symptoms in many instances may only emerge after major damage has taken place, pointing towards the utility of predictive models to detect at-risk individuals before the occurrence of a clinical event.

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429 Volume 5, Issue 1, May 2025 Impact Factor: 7.67

A. Traditional Methods of Heart Attack Detection

The prediction and diagnosis of heart attacks have been based on clinical evaluation and diagnostic procedures like electrocardiograms (ECG), echocardiography, stress tests, blood tests for cardiac biomarkers (such as troponin), and coronary angiography. Although these methods are useful for diagnosing established heart disease, they are not as appropriate for early-stage prediction, particularly in asymptomatic patients.

Risk assessment algorithms, including the Framingham Risk Score and Systematic Coronary Risk Evaluation (SCORE), have been designed to predict the probability of cardiovascular events within a certain timeframe (usually 10 years). These algorithms take into account variables such as age, gender, blood pressure, cholesterol, smoking status, and diabetes. Nevertheless, they can be imprecise as a result of oversimplifying assumptions and the inability to represent complex, non-linear associations among risk factors. Moreover, conventional models might not generalize across populations because of ethnic, genetic, and lifestyle variations.

B. The Rise of Machine Learning in Healthcare

Machine learning (ML), a specialized form of artificial intelligence (AI), has been a robust tool in healthcare analytics in recent years, with promising applications for disease prediction, diagnosis, treatment planning, and personalized medicine. ML is characterized by algorithms that learn from data and refine their performance over time without explicit programming. These models can identify faint correlations and interactions between variables that may go undetected by conventional statistical methods.

Machine learning methods have already shown substantial success in applications including medical imaging, genomics, and electronic health record (EHR) analysis. Their capacity to handle large amounts of heterogeneous data—clinical measurements, imaging findings, genetic data, and lifestyle information—makes them an excellent fit for multifactorial, complex diseases like heart disease. When used for the prediction of heart attack, ML models can scrutinize patient data to detect concealed risk factors and make more accurate predictions than traditional scoring systems. They can also enable real-time decision-making and assist clinical workflows by generating automatic alerts or recommendations.

C. Motivation for Machine Learning-Based Heart Attack Prediction

The motivation for developing machine learning-based heart attack prediction systems stems from the limitations of existing methods and the urgent need for proactive interventions. ML models can potentially:

- Improve accuracy and precision in identifying at-risk individuals.
- Handle high-dimensional data from multiple sources, including wearable devices and EHRs.
- Adapt to new data over time, improving predictions with additional training.
- Provide personalized risk assessments by accounting for individual differences.
- Enable earlier intervention, reducing the likelihood of adverse cardiovascular events.

Integrating ML into preventive cardiology can therefore redefine the paradigm from reactive treatment to proactive health management. As more healthcare data becomes available and computational capabilities increase, the integration of ML into clinical practice becomes a reality.

D. Overview of Machine Learning Techniques

Several machine learning methods have been investigated for heart attack prediction, from classical classifiers such as logistic regression and decision trees to more complex models such as artificial neural networks (ANNs) and ensemble techniques. The methods differ in terms of their complexity, explainability, and performance.

- Logistic Regression (LR) is a robust yet straightforward linear model that tends to be a baseline in classification tasks. It has probabilistic outputs and is easy to interpret but can't capture intricate data relationships.
- Decision Trees (DT) represent decisions in the form of a tree structure, thus being easy to comprehend and understand. They tend to overfit, though, which can limit generalization.

- Random Forests (RF) and Gradient Boosting Machines (GBM) are ensemble techniques where several decision trees are used collectively to enhance prediction accuracy and minimize variance. They are sturdy and popularly applied in the medical field.
- Support Vector Machines (SVM) decide on the optimal hyperplane to distinguish classes in high-dimensional space. SVM is effective in non-linear classification problems but can be computationally costly.
- K-Nearest Neighbors (KNN) is a non-parametric algorithm that assigns a sample to the most common class of its neighbors. It is simple but may be computationally costly for large data.

All these algorithms have been employed in a range of studies with differing levels of success based on factors such as data quality, feature selection, and preprocessing techniques.

E. Importance of Data and Feature Engineering

The performance of ML models strongly relies on data quality and size used for model training. Large publicly available datasets, like the Cleveland Heart Disease dataset from the UCI Machine Learning Repository, have been the benchmark for developing and testing algorithms. These data sets often feature age, sex, type of chest pain, blood pressure, cholesterol, and ECG values.

Feature engineering—choosing, converting, and building useful input variables—is another key step in building good models. Irrelevant or redundant features may add noise and reduce performance, whereas well-designed features can increase predictive capability. Dimensionality reduction methods such as Principal Component Analysis (PCA) and feature selection methods like Recursive Feature Elimination (RFE) are typically used.

F. Challenges in Developing ML-Based Heart Attack Prediction Systems

Despite the promise of ML in heart attack prediction, several challenges remain:

- Data Imbalance: In many datasets, the number of heart attack cases is much smaller than non-cases, leading to biased models. Techniques such as SMOTE (Synthetic Minority Over-sampling Technique) are used to address this issue.
- Generalization: Models trained on specific datasets may not generalize well to other populations due to demographic and clinical variability.
- Interpretability: Complex models, especially deep learning algorithms, are often difficult to interpret, which hinders clinical adoption. Explainable AI (XAI) methods are being developed to address this issue.
- Integration with Clinical Workflows: For ML systems to be useful in practice, they must be integrated with existing clinical decision support systems and be user-friendly for healthcare providers.
- Privacy and Security: The use of sensitive patient data raises ethical and legal concerns, necessitating robust data governance and privacy-preserving methods like federated learning.

G. Objectives of This Review

Given the rapid growth of research in this field, this review aims to provide a comprehensive overview of machine learning approaches for heart attack prediction. Specifically, it will:

- Summarize the current state of the art in ML-based heart attack prediction systems.
- Compare and evaluate different algorithms, datasets, and performance metrics.
- Discuss the challenges, limitations, and opportunities for future research.
- Highlight potential directions for clinical implementation and real-world applications.

By integrating current literature, the paper aims to inform researchers, clinicians, and policymakers regarding the strengths and limitations of ML for cardiovascular risk prediction. Ultimately, the paper's aim is to improve the formulation of accurate, interpretable, and scalable prediction systems capable of saving lives with early intervention and tailored care.

II. LITERATURE SURVEY

Mohammad Alshraideh et al. (2024) [1] suggested an intelligent heart attack prediction system based on a blend of machine learning classifiers i.e., SVM, Random Forest, Naive Bayes, Decision Tree, and KNN with Particle Swarm Optimization (PSO) for feature selection. The study employed a real-world dataset consisting of 486 patient records of Jordan University Hospital, which had 58 medical and demographic features. PSO simplified these characteristics to 19 strongly applicable attributes, enhancing model effectiveness. The research proved that SVM, in combination with PSO, exhibited the best classification accuracy of 94.3%, surpassing other algorithms. Not only did it minimize processing time and computation, but it also enhanced diagnostic performance, providing excellent assistance with the early diagnosis and personalized therapy of heart disease. The conclusions highlight the potential of hybrid AI approaches in enhancing clinical decision-making, especially in settings with a high cardiovascular death rate such as in Jordan.

Thomas et al. (2024) [2] examined the application of nonlinear machine learning to identify adaptive divergence in tropical fish Melanotaenia splendida splendida across environmental gradients within the Australian Wet Tropics. Employing convolutional neural networks (CNNs), the research identified genetic loci related to environmental variation more effectively than through conventional linear approaches. By combining genomic, spatial, and ecological information, the study uncovered strong genotype–environment correlations, especially with temperature, precipitation, and stream flow. Morphological traits, such as body shape, were also genetically and environmentally organized, validating the parallel phenotypic and genetic adaptation hypothesis. The research illustrates the improved ability of nonlinear models such as CNNs to identify intricate adaptive patterns in natural populations. These results are of relevance both from the perspective of ecological adaptation and conservation of biodiversity, and they highlight the importance of state-of-the-art machine learning software in evolutionary biology.

Barrière et al. (2023) [3] studied the problem of topology discovery in wireless sensor networks (WSNs) and proposed distributed algorithms enabling nodes to locally infer the network topology on their own. The authors tackled theoretical models, primarily Unit Disk Graphs, to develop and analyze the time and message complexity of the algorithms in synchronous and asynchronous settings. They are also bounds on algorithm performance and message complexity vs. convergence time trade-off study. Their findings emphasize the local communication constraint as a factor that affects discovery efficiency and scalability. These distributed algorithms provide a foundation for enabling routing, monitoring, and self-organization in WSNs, particularly in systems with no centralized coordination. The study makes significant theoretical contributions to network science and enhances effective design of communication protocols in low-power systems.

Sharma et al. (2022) [4] explored the use of machine learning (ML) methods to forecast the probability of heart attacks to enable early diagnosis and preventive treatment. The authors employed the UCI Heart Disease dataset, which includes important factors like age, sex, chest pain type, cholesterol, and more, to train and test different ML models. Their approach involved preprocessing techniques such as data normalization, feature selection, and comparison of models via classifiers such as K-Nearest Neighbors (KNN), Decision Tree, Logistic Regression, Random Forest, and Support Vector Machine (SVM). A few of the major contributions were comparative performance evaluation where accuracy was highest by KNN (around 91%), followed by Decision Tree and Random Forest models in close second. The study highlights the ability of ML models to support clinical decision-making by identifying individuals at risk through non-invasive analysis of data. The findings illuminate the growth of intelligent healthcare systems, which offers a cost-effective and scalable solution for the assessment of cardiac risk.

Md. G. Harinadha Babu et al. (2023) [5] proposed a heart disease prediction model by employing different machine learning classifiers with special focus on the Random Forest algorithm. The research activity conducted at KKR & KSR Institute of Technology and Sciences tested classification methods such as Naive Bayes, Logistic Regression, Decision Tree, and SVM and ensemble methods such as AdaBoost and XGBoost. The research emphasized the application of Random Forest in the management of large volumes of health data, predictive accuracy improvement, and reduction of clinical time and complexity. The system utilized feature selection techniques and demonstrated that ensemble classifiers were potentially more accurate than individual models. The model effectively offers early detection of heart disease, which is critical in preventive care. It is particularly useful in making diagnosis cost-effective and affordable for patients and in alleviating doctors' workload.

Ekta Maini et al. (2021) [6] proposed a machine learning-based system for predicting heart disease among the Indian population using clinical data from a South Indian tertiary hospital. Five ML algorithms were used in the study work—k-Nearest Neighbors, Naïve Bayes, Logistic Regression, AdaBoost, and Random Forest—the maximum accuracy being offered by Random Forest with 93.8% accuracy. The model was tested on 1670 anonymized data with 70:30 train-to-test and was highly sensitive (92.8%) and specific (94.6%). The model was also made available as a cloud application, which can be easily accessed, especially in rural areas. The project demonstrates the effectiveness and efficiency of ML in addressing India-specific healthcare challenges, delivering timely diagnosis and potentially reducing cardiovascular mortality.

K. Babu et al. (2025) [7] suggested a cloud-disease prediction system for cardiovascular disease that uses machine learning methods to facilitate improved early diagnosis. It follows a four-step cloud-based approach for managing patient health data and classification models for detecting the presence of disease. Cloud computing is being used in research because it is scalable and allows monitoring of real-time data, enhancing the accuracy of diagnosis with performance-measuring metrics like precision, recall, and specificity. The device, as proposed, is able to distinguish between healthy and vulnerable individuals and can be made low-cost and long-lasting. It has the capability to transform the healthcare monitoring market through the provision of scalable and precise diagnostic solutions.

Karna Vishnu Vardhana Reddy et al. (2021) [8] developed a heart disease prediction model using ten machine learning classifiers and three attribute evaluators on the Cleveland dataset. They identified that the Sequential Minimal Optimization (SMO) algorithm generated the optimal prediction accuracy of 86.47% with the chi-squared attribute set. In addition, bagging with logistic regression generated an ROC area of 0.91. The study aims at the necessity of hyperparameter tuning and feature selection to improve prediction performance, specifically that results show that hybrid models using optimal parameters yield improved accuracy and generalizability in heart disease diagnosis.

Mohammed Amine Bouqentar et al. (2024) [9] have suggested a stable machine learning model to predict heart disease at an early stage based on the Cleveland and Statlog datasets. Several ML algorithms such as Decision Tree, Random Forest, SVM, Logistic Regression, AdaBoost, and KNN were implemented. The system achieved accuracy of 92% on the Cleveland dataset and 91.18% on the Statlog dataset through hyperparameter tuning and feature manipulation. The model performance was verified using 10-fold cross-validation, and results point towards the future possibilities of ML methods in reducing the diagnostic inaccuracies as well as early treatments in cardiovascular disease.

Shadman Nashif et al. (2018) [10] proposed a two-component heart disease detection framework that integrates machine learning classifiers and a real-time cardiovascular health monitoring system. When model evaluation was performed using WEKA, SVM provided the highest accuracy of 97.53% with sensitivity and specificity of 97.50% and 94.94%, respectively. In addition to this, the study used an IoT-based patient monitoring system with Arduino and physiological sensors that refresh vital signs every 10 seconds. Doctors receive notifications through GSM when thresholds are crossed. The hybrid AI-IoT model demonstrates high promise for real-time, continuous, and precise heart disease prediction.

Amrit Singh et al. (2024) [11] introduced a heart disease diagnosis system using Decision Tree, Random Forest, Support Vector Machine, and PCA methods. The research aimed to solve the overfitting issue and improve generalizability on training and testing data. Their diagnostic system obtained competitive performance while ensuring interpretability and low computational complexity. Through comparative accuracy analysis, the paper demonstrated the merits of supervised learning in cardiovascular healthcare, inviting future use in real-world diagnostic applications.

Aashish Gnanavelu et al. (2025) [12] developed a heart disease prediction system based on machine learning models on the Kaggle Heart Disease dataset. The research used Decision Tree, K-Nearest Neighbors, Naive Bayes, XGBoost, and Random Forest algorithms. Out of these, XGBoost gave the highest accuracy of 93%. Age, sex, BMI, and lifestyle were the most contributing features. An interactive dashboard was also developed to allow user-friendly predictions. The article highlights the applied function of ML in facilitating clinical diagnosis and early interventions via explainable, accessible tools.

III. SYSTEMATIC REVIEW

The mounting load of cardiovascular diseases, especially heart attacks, has compelled extensive research into predictive methods and early detection. With the advancement in computational methods and increasing availability of clinical data, machine learning (ML) has become a strong tool in the creation of prediction models for heart disease. A critical review of literature is a thorough understanding of methodologies, datasets, and algorithms used in heart attack prediction systems. This section is intended to critically analyze recent studies, emphasizing the advantages and limitations of different ML methods in terms of accuracy, feature selection, and real-world applicability. By integrating findings from several sources, the review establishes emerging trends, shared challenges, and research gaps, hence providing insights that can inform future development of stable, interpretable, and clinically feasible heart disease prediction models

Table 1: Systematic Review

Sr. No.	Author(s)	Title of the Project	Methodology	Technical Aspect	Overview
1	Mohammad Alshraideh et al. (2024)	Intelligent Heart Attack Prediction System	ML classifiers (SVM, RF, NB, DT, KNN) + PSO for feature selection	Used real-world hospital dataset with 58 features; PSO reduced to 19; SVM with PSO achieved 94.3% accuracy	Hybrid AI method improved diagnostic precision and reduced computational complexity
2	Thomas et al. (2024)	Adaptive Divergence in Tropical Fish	Convolutional Neural Networks (CNNs)	Genomic + environmental data integration; identified genotype–environment associations	Showcased power of nonlinear ML in evolutionary biology (non-medical application)
3	Barrière et al. (2023)	Topology Discovery in WSNs	Distributed algorithms on Unit Disk Graphs	Message complexity analysis; synchronous/asynchronous models	Focus on network theory and protocol design in WSNs (non-healthcare study)
4	Sharma et al. (2022)	Heart Attack Prediction Using ML	KNN, DT, LR, RF, SVM on UCI dataset	Preprocessing + feature selection; KNN achieved ~91% accuracy	Demonstrated effectiveness of standard classifiers in heart disease prediction
5	Md. G. Harinadha Babu et al. (2023)	Heart Disease Prediction System	RF, NB, LR, DT, SVM, AdaBoost, XGBoost	Emphasis on ensemble models and feature selection	RF showed strong performance, supporting early detection and ease of clinical use
6	Ekta Maini et al. (2021)	India-Specific Heart Disease Prediction	RF, KNN, NB, LR, AdaBoost	Used 1670 patient records; RF had 93.8% accuracy, high sensitivity/specificity	Tailored ML system deployed as a cloud app for accessibility in rural India
7	K. Babu et al. (2025)	Cloud-Based Cardiovascular Prediction	Cloud computing + ML classifiers	4-step cloud framework with real-time analytics	Designed for low-cost, scalable, real-time heart disease monitoring
8	Karna Vishnu Vardhana Reddy et al.	Predictive Modeling Using SMO	10 classifiers + 3 attribute evaluators (Chi-squared, Gain Ratio)	SMO with chi-squared achieved 86.47% accuracy; Bagging+LR ROC = 0.91	Highlighted importance of hyperparameter tuning and hybrid models

IJARSCT
International Journal of Advanced Research in Science, Communication and Technology
International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal
ISSN: 2581-9429
Volume 5, Issue 1, May 2025
Impact Factor: 7.67

	(2021)				
9	Mohammed Amine Bouqentar et al. (2024)	Robust Early Heart Disease Prediction	DT, RF, SVM, LR, AdaBoost, KNN	Accuracy of 92% (Cleveland), 91.18% (Statlog); 10-fold CV used	Reinforced value of feature engineering and model tuning
10	Shadman Nashif et al. (2018)	AI-IoT Heart Disease Detection	ML models + IoT monitoring (Arduino, sensors)	SVM accuracy 97.53%; GSM alerts; 10s data update	Real-time health system combining ML and IoT for continuous monitoring
11	Amrit Singh et al. (2024)	PCA-Based Heart Disease Detection	DT, RF, SVM + PCA	Focused on reducing overfitting and improving generalizability	Balanced model interpretability with accuracy and speed
12	Aashish Gnanavelu et al. (2025)	ML-Driven Heart Disease Dashboard	DT, KNN, NB, XGBoost, RF	Kaggle dataset; XGBoost accuracy = 93%; interactive dashboard built	Built explainable and user-accessible system for clinical decision support

IV. RESEARCH GAP

Despite the significant progress in developing machine learning (ML)-based heart attack prediction systems, several critical gaps remain unaddressed in the current literature. First, most research uses public datasets (e.g., Cleveland, UCI, Statlog), which, while convenient for benchmarking purposes, tend to be limited by a lack of diversity in demographic characteristics, geographic variability, and clinical presentation. This restricts the ability to generalize trained models to populations and healthcare environments.

Second, although different ML algorithms were tested individually or against each other, there is less investigation of ensemble or hybrid models specifically for certain clinical situations. While some work (e.g., Alshraideh et al., 2024; Harinadha Babu et al., 2023) utilized optimization algorithms such as Particle Swarm Optimization or ensembles, the level of investigation of model interpretability and integration with real-time clinics is superficial.

Another important gap is the low utilization of real-time or streaming health data. There were very few studies (e.g., Nashif et al., 2018) that integrated IoT-based health monitoring systems, but this has not been broadly followed or tested through clinical trials. This limits the use of such systems for round-the-clock patient monitoring and intervention on time.

In addition, explainability and transparency—important for clinical adoption—are frequently neglected. Numerous top-performing models such as SVMs, Random Forests, and deep networks are black boxes, not producing interpretable outputs that clinicians can rely upon and take action on with confidence.

Finally, ethical, privacy, and data security issues—especially in cloud-based or federated ML models—are not sufficiently explored, although they become increasingly significant in actual deployment.

V. DISCUSSION

The systematic analysis of current research in heart attack prediction through machine learning presents promising developments along with its serious limitations that need to be further explored. The literature covered illustrates an unremitting tide towards the accuracy of the model, with classifiers like Support Vector Machine (SVM), Random Forest (RF), K-Nearest Neighbors (KNN), and ensemble methods (e.g., AdaBoost, XGBoost) showing great predictive capability. Exactly, models involving optimization methods—like Particle Swarm Optimization (PSO) or feature selection processes—were found to surpass (e.g., Alshraideh et al., 2024; Bouqentar et al., 2024).

However, while technical accuracy is an important metric, clinical application demands more balanced assessment of performance. Sensitivity, specificity, and interpretability are just as important, especially in life-critical applications like cardiovascular disease diagnosis. Significantly, few studies (e.g., Maini et al., 2021) presented detailed evaluation metrics, which detracts from confidence in their generalizability to other populations and health systems.

Another significant finding is the dominance of structured, static datasets—primarily the UCI and Cleveland datasets. Although these resources are valuable for academic exploration, they lack diversity in terms of patient demographics and health conditions, which restricts their applicability in real-world, heterogeneous environments. Furthermore, the absence of longitudinal or real-time patient data in most models limits their ability to adapt to temporal variations in cardiovascular risk factors.

Explainability emerged as a crucial challenge. High-performing black-box models, such as SVM and deep neural networks, often lack transparency, making it difficult for clinicians to trust and act on their outputs. Despite this, few studies have embraced explainable AI (XAI) techniques to bridge this gap. This lack of interpretability hampers adoption in clinical workflows, where explainability is not optional but a regulatory and ethical requirement.

In terms of system integration, only a minority of studies, such as Nashif et al. (2018), explored real-time monitoring through IoT and sensor-based platforms. This points to a broader oversight: the integration of ML models with hospital information systems (HIS), electronic health records (EHRs), and mobile health applications. Bridging this gap could significantly improve patient monitoring, early detection, and personalized treatment plans.

Additionally, readiness for deployment is a challenge. Although cloud-based and web-based prediction tools have been suggested (e.g., Babu et al., 2025; Maini et al., 2021), thorough validation within clinical settings is uncommon. Most of the systems are still at the prototyping level with no empirical validation on actual hospital data or in combination with healthcare professionals. This absence of clinical translation highlights the importance of inter-disciplinary collaboration among clinicians, data scientists, and healthcare policymakers.

Ethical concerns—such as data privacy, bias, and informed consent—are also minimized. With systems moving towards cloud computing and mobile health apps, patient data protection and compliance with regulatory structures like HIPAA or GDPR are a concern. Federated learning and differential privacy could be solutions but have not been used extensively in the field of cardiovascular ML systems.

Finally, the absence of standard evaluation procedures and benchmarking protocols hinders uniform comparison across studies. Varying train-test splits, validation methods (e.g., k-fold vs. hold-out), and performance measures hinder equitable assessment and reproducibility. This diversity emphasizes the necessity for a shared evaluation pipeline to facilitate reproducibility and regulatory approval.

Finally, although ML-based heart attack prediction systems hold vast potential for enhancing cardiovascular risk stratification, clinical adoption on a large scale has to resolve key concerns with respect to data diversity, model explainability, real-time adaptability, system integration, ethical alignment, and benchmarking. Future studies need to overcome accuracy-focused methods and adopt balanced, ethically sound, and clinically tested models in order to unlock the full power of artificial intelligence in cardiovascular medicine.

VI. CONCLUSION

Heart disease, especially heart attack, is among the top causes of death globally, requiring timely diagnosis and effective preventive measures. This review has discussed the current developments in machine learning (ML) models for predicting heart attack, and discussed a wide range of algorithms such as Support Vector Machines, Random Forests, K-Nearest Neighbors, Naive Bayes, Decision Trees, Logistic Regression, and ensemble methods such as AdaBoost and XGBoost. Such models have shown excellent accuracy in predicting heart disease, particularly when combined with optimization algorithms like Particle Swarm Optimization and using the application of sophisticated feature selection techniques.

Nevertheless, in spite of considerable advances, important challenges still persist. Most research is carried out on narrow and frequently old datasets that do not capture the diversity of actual populations. Key clinical characteristics are frequently omitted, and most models value accuracy above more clinically appropriate measures like sensitivity, specificity, and interpretability. The absence of explainability in top-performing models poses a significant impediment to clinical adoption since doctors need transparency and accountability for AI-based decisions. In addition, it is only a limited number of models that have been validated in actual live clinical environments or even connected to hospital

IJARSCT

International Journal of Advanced Research in Science, Communication and Technology

International Open-Access, Double-Blind, Peer-Reviewed, Refereed, Multidisciplinary Online Journal

ISSN: 2581-9429 Volume 5, Issue 1, May 2025 Impact Factor: 7.67

systems, calling into question their real-world applicability. Additionally, critical factors like ethics compliance, data privacy, and real-time decision support systems are in an early stage in this research area. A definitive need for interdisciplinary methodologies that merge ML know-how with clinical understanding to build solid, interpretable, and scalable prediction systems exists. In conclusion, although machine learning has tremendous promise in transforming the prediction of heart attacks and enhancing patient outcomes, future work will need to overcome existing constraints by addressing diverse datasets, model interpretability, real-world testing, and ethical considerations. Only then are such smart systems able to transition from theoretical potential to real-world effect in global healthcare.

REFERENCES

[1]. Alshraideh, M., Alshraideh, N., Alshraideh, A., Alkayed, Y., Al Trabsheh, Y., & Alshraideh, B. (2024). Enhancing heart attack prediction with machine learning: A study at Jordan University Hospital. *Applied Computational Intelligence and Soft Computing, 2024*, Article ID 5080332. https://doi.org/10.1155/2024/5080332

[2]. Thomas, L., Coleman, M. A., Cooke, I. R., et al. (2024). Nonlinear machine learning uncovers signals of adaptive divergence associated with ecological gradients in a rainforest fish. *Scientific Reports, 14*(1), 74656. https://doi.org/10.1038/s41598-024-74656-2

[3]. Barrière, L., Fraigniaud, P., & Pignolet, Y. A. (2023). Topology discovery in wireless sensor networks: Distributed algorithms and complexity. *Algorithms, 16*(2), 88. https://doi.org/10.3390/algorithms16020088

[4]. Sharma, S., Verma, O. P., Tripathi, R., & Rautaray, S. S. (2022). A machine learning approach for heart attack prediction: Comparative study and analysis. *Procedia Computer Science*, 213, 146–155. https://doi.org/10.1016/j.procs.2022.10.132

[5]. Babu, G. H., Jayasree, G., Ashika, C., Ahalya, V., & Niroopa, K. A. (2023). Heart Disease Prediction System Using Random Forest Technique. *International Journal for Research in Applied Science & Engineering Technology*, 11(1).

[6]. Maini, E., Venkateswarlu, B., Maini, B., & Marwaha, D. (2021). Machine learning-based heart disease prediction system for Indian population: An exploratory study done in South India. *Medical Journal Armed Forces India*, 77(2), 302–311. https://doi.org/10.1016/j.mjafi.2020.10.013

[7]. Babu, K., Chandar, A. G., & Kannadhasan, S. (2025). Prediction and diagnosis of cardiovascular disease using cloud and machine learning design. *Journal of Cloud Computing*, 14(3). https://doi.org/10.1186/s13677-024-00720-x

[8]. Reddy, K. V. V., Elamvazuthi, I., Aziz, A. A., Paramasivam, S., Chua, H. N., & Pranavanand, S. (2021). Heart Disease Risk Prediction Using Machine Learning Classifiers with Attribute Evaluators. *Applied Sciences*, 11(18), 8352. https://doi.org/10.3390/app11188352

[9]. Bouqentar, M. A., Terrada, O., Hamida, S., Saleh, S., Lamrani, D., Cherradi, B., & Raihani, A. (2024). Early heart disease prediction using feature engineering and machine learning algorithms. *Heliyon*, 10, e38731. https://doi.org/10.1016/j.heliyon.2024.e38731

[10]. Nashif, S., Raihan, M. R., Islam, M. R., & Imam, M. H. (2018). Heart Disease Detection by Using Machine Learning Algorithms and a Real-Time Cardiovascular Health Monitoring System. *World Journal of Engineering and Technology*, 6, 854–873. https://doi.org/10.4236/wjet.2018.64057

[11]. Singh, A., Mahapatra, H., Biswal, A. K., Mahapatra, M., Singh, D., & Samantaray, M. (2024). Heart Disease Detection Using Machine Learning Models. Procedia Computer Science, 235, 937–947. https://doi.org/10.1016/j.procs.2024.04.089

[12]. Gnanavelu, A., Venkataramu, C., & Chintakunta, R. (2025). Cardiovascular Disease Prediction Using Machine Learning Metrics. *Journal of Young Pharmacists*, 17(1), 226–233. https://doi.org/10.5530/jyp.20251231