AVE Trends in Intelligent Informatics Reports

A Data-Driven Artificial Neural Network Framework for Accurate Heart Disease Prediction

J. Angelin Jeba[1], S. Rubin Bose[2,*], O. Jeba Singh[3], R. Regin[4], S. Suman Rajest[5], Aditya Rautaray[6]

[1]Department of Electronics and Communication Engineering, S.A. Engineering College, Chennai, Tamil Nadu, India.
[2,4]School of Computer Science and Engineering, SRM Institute of Science and Technology, Ramapuram, Chennai, Tamil Nadu, India.
[3]Centre for Academic Research, Alliance University, Bengaluru, Karnataka, India.
[5]Department of Research and Development, Dhaanish Ahmed College of Engineering, Chennai, Tamil Nadu, India.
[6]Department of Cloud Solutions Security, CVS Healthcare, Ashburn, Virginia, United States of America.
angelinjeba@saec.ac.in[1], rubinbos@srmist.edu.in[2], jeba.singh@alliance.edu.in[3], reginr@srmist.edu.in[4], sumanrajest414@gamil.com[5], aditya.rautaray@cvshealth.com[6]

*Corresponding author

Abstract: This study proposes a machine learning–based framework for predicting heart disease using the UCI Heart Disease dataset. The objective is to develop an efficient predictive model that helps healthcare professionals identify individuals at risk early. The dataset, consisting of 303 patient records and 14 attributes, undergoes preprocessing steps, including normalization, categorical encoding, and feature selection, to enhance model performance. Multiple machine learning algorithms—Logistic Regression, Random Forests, Support Vector Machines (SVMs), and Artificial Neural Networks—are employed to classify patients into risk categories. The proposed methodology incorporates cross-validation and hyperparameter tuning to ensure robustness and reduce overfitting. Experimental evaluation is conducted using performance metrics such as accuracy, precision, recall, F1-score, and ROC-AUC. The results demonstrate that Random Forest achieves the highest overall accuracy of 89%, outperforming traditional classifiers, while Neural Networks provide strong generalization capabilities with competitive results. A comparative analysis highlights the effectiveness of ensemble learning methods for handling nonlinear relationships in clinical datasets. This research underscores the potential of integrating machine learning models into clinical decision support systems to improve the reliability and speed of diagnosis. Future work includes incorporating larger, real-time datasets and extending deep learning architectures to improve predictive accuracy.

Keywords: Heart Disease; Logistic Regression; Artificial Neural Network; Support Vector Machine; Random Forests; Cardiovascular Disorders; Traditional Classifiers.

Cite as: J. A. Jeba, S. R. Bose, O. J. Singh, R. Regin, S. S. Rajest, and A. Rautaray, "A Data-Driven Artificial Neural Network Framework for Accurate Heart Disease Prediction," *AVE Trends in Intelligent Informatics Reports*, vol. 1, no. 1, pp. 24–36, 2026.

Journal Homepage: https://www.avepubs.com/user/journals/details/ATIIR

Received on: 09/04/2025, **Revised on:** 25/05/2025, **Accepted on:** 27/07/2025, **Published on:** 01/03/2026

DOI: https://doi.org/10.64091/ATIIR.2026.000278

1. Introduction

Cardiovascular disorders remain a significant public health concern worldwide, representing a considerable share of annual morbidity and mortality. Heart disease is one of the most common and deadly diseases, and it affects people of all ages, genders, and income levels. Despite breakthroughs in medical research, early diagnosis and prompt care remain key factors in reducing tragic outcomes and improving long-term quality of life for patients. Many cases of heart disease develop slowly over years, and symptoms don't show up until the condition is advanced, when treatment options are limited, expensive, and less effective. This fact shows how important it is to find reliable, efficient, and scalable ways to support early identification and risk assessment before damage that can't be fixed occurs. Traditionally, doctors use a mix of physical exams, lab tests, imaging tests, and their own knowledge to determine whether someone has cardiac disease. To estimate the likelihood of heart problems, doctors consider factors such as age, lifestyle, family history, blood pressure, cholesterol levels, blood sugar, and ECG readings. These approaches are well-known and have been tested in clinical settings, but they typically need people to interpret them and make decisions based on their own opinions. This process can take a long time, especially in healthcare systems that don't have enough qualified workers and are getting more patients every day. Also, small trends hidden in massive amounts of patient data may go unnoticed, leading to delayed or incorrect diagnoses. Diagnostic ambiguity is worsened by human error, differences in how observers perceive things, and missing data, all of which can negatively affect patient outcomes. The rapid rise of digital health records and the availability of large clinical datasets over the last several years have opened new possibilities for data-driven healthcare solutions.

Machine learning has become a powerful tool for uncovering complex relationships in large volumes of structured and unstructured data that are hard for people to see. Machine learning algorithms can find patterns, trends, and correlations in historical data that can help diseases evolve and worsen. This feature has made machine learning very useful for predicting healthcare uses, such as diagnosing diseases, predicting their course, planning treatments, and making medicine more personalized. Artificial neural networks have attracted significant interest as a machine learning method because they can simulate complex, nonlinear interactions between inputs and outputs. Artificial neural networks are composed of layers of neurons linked together. They process information through weighted connections and activation functions that mimic how the human brain works. This structure lets them capture subtle interactions among different risk factors and adapt to data that is spread across complex structures. Artificial neural networks are effective at handling a range of clinical data, including numerical, categorical, and sometimes missing or noisy data, when predicting heart disease. The reason for employing artificial neural networks to forecast cardiac disease is that they may operate better than traditional statistical and rule-based methods. Traditional methods such as logistic regression or decision rules usually assume a linear relationship between predictors and outcomes. This may not adequately show how complex cardiovascular disease is. Heart disease is affected by a mix of hereditary variables, lifestyle choices, metabolic factors, and environmental factors, all of which interact with each other in ways that aren't always clear.

Artificial neural networks can learn these interactions directly from data, thereby improving prediction accuracy and reliability. Consequently, they have the potential to serve as efficient decision-support tools for doctors, facilitating early risk classification and preventive intervention. This study focuses on developing and evaluating a holistic machine-learning framework for predicting cardiac disease, with particular emphasis on artificial neural networks. The framework is meant to systematically analyze patient data, clean it, select the most important features, train predictive models, and then use appropriate assessment metrics to evaluate their performance. The study seeks to elucidate the advantages and limitations of artificial neural networks compared with other prevalent machine learning methods for cardiovascular risk prediction. This research utilizes a dataset sourced from a publicly accessible repository that has been extensively employed in prior studies on heart disease. It contains information on individuals with various clinical and demographic traits, including age, sex, resting blood pressure, cholesterol levels, fasting blood sugar, highest heart rate, angina caused by exercise, and other medically important traits. The dataset is well-suited to supervised learning problems, as each record includes a label indicating whether cardiac disease is present. Using a publicly available dataset ensures that the research is open, repeatable, and comparable to other studies in this field. Before using machine learning methods, the data is carefully cleaned to improve model performance and reliability. This stage involves handling missing data, removing inconsistencies, normalizing numerical features, and converting categorical variables into formats that machines can read.

Feature scaling is highly significant for artificial neural networks, as it stabilizes learning and speeds convergence during training. Also, exploratory data analysis is used to assess how the variables are distributed, identify outliers, and examine how features and the target outcome are related. This kind of research provides important insights into the data's underlying structure and helps inform future modeling choices. Feature selection is very important for predicting heart disease because not all provided variables have the same effect on the outcome. Adding elements that don't matter or are already there can make the model more complex, cause it to overfit, and make it harder to understand. Consequently, the research investigates methodologies to discern critical factors that substantially impact the incidence of heart disease. From a clinical perspective, these features may include age, cholesterol levels, blood pressure, and exercise-related factors, all of which are known risk factors. The paradigm improves both predictive accuracy and clinical relevance by integrating data-driven feature importance with medical knowledge. The artificial neural network model developed in this study has an input layer corresponding to the

selected features, one or more hidden layers that perform nonlinear transformations, and an output layer that produces the final prediction. Backpropagation is used to train the network. This algorithm updates the weights repeatedly until the prediction error is minimized. Activation functions such as sigmoid or rectified linear units introduce nonlinearity, allowing the network to learn complex patterns. To achieve optimal performance, hyperparameters such as the number of hidden layers, the number of neurons, the learning rate, and the number of training epochs are carefully tuned.

The artificial neural network is compared with other machine learning models commonly used for medical prediction tasks to provide a complete picture. Some of them are logistic regression, decision trees, support vector machines, and k-nearest neighbors. The same dataset is used to train and test each model, and the experimental conditions are identical across all models. Metrics including accuracy, precision, recall, F1 score, and area under the receiver operating characteristic curve are used to measure performance. These measures provide a fair picture of how well a model performs, especially in healthcare settings where false negatives and false positives can have catastrophic consequences. The study's findings demonstrate that artificial neural networks have improved predictive performance relative to several conventional machine learning models. Their capacity to capture nonlinear interrelationships among many risk factors enhances precision and facilitates superior generalization to novel data. The results also show how important some factors are for predicting heart disease. This backs up what researchers already know from clinical practice and gives us numbers to show how much they matter. These insights can help doctors focus on the most important risk factors when evaluating a patient. The study highlights the practical ramifications of incorporating machine learning models into healthcare workflows, beyond prediction accuracy. An automated heart disease prediction system can help healthcare workers make decisions by screening large groups of people and identifying those at high risk who may need additional tests or preventive measures.

This is especially useful in places where resources are limited and access to specialized cardiology services may be hard to come by. Machine learning-based solutions can help reduce healthcare costs, prevent disease progression and improve overall patient outcomes by making it easier to detect problems early. But using machine learning in healthcare also raises important questions about data quality, ease of understanding and moral responsibility. There may be biases in clinical data due to demographic differences or measurement errors, which can make models less fair and less reliable. People often criticize artificial neural networks because they are "black boxes," which makes it hard to explain how they make predictions. To address these problems, data scientists and medical professionals need to work together, carefully validate their work, and clearly publish their findings. Future research may investigate explainable artificial intelligence methodologies that yield interpretable insights alongside precise predictions. This study illustrates the considerable potential of machine learning, particularly artificial neural networks, to improve heart disease prediction and facilitate cardiovascular therapy. The suggested framework provides an effective, precise methodology for analyzing complex clinical data and identifying individuals predisposed to heart disease using data-driven methods. The study shows that artificial neural networks are better at modeling nonlinear interactions and making better predictions than traditional techniques. As healthcare systems continue to go digital, these smart predictive models can help with early diagnosis, guide preventative initiatives, and ultimately lower the worldwide burden of cardiovascular disease.

2. Literature Review

Dou et al. [1] conducted a recent comparative analysis of supervised learning algorithms for cardiovascular disease prediction. Their study evaluated several models, including MLP (ANN) and SVM, and found that ensemble and fusion models, such as Random Forest, consistently provided superior accuracy and stability. This work directly supports our report's methodology of comparing traditional classifiers against more advanced architectures and validates our finding that ensemble models are superior. Suleiman et al. [2] specifically utilized a Random Forest (RF) model for cardiovascular disease prediction, noting its high performance. Their work highlighted the algorithm's inherent strength in handling complex, high-dimensional clinical datasets with minimal preprocessing. They demonstrated that RF, an ensemble method, achieves high predictive performance by aggregating the outputs of many decision trees, thereby reducing variance and overfitting. This reinforces our selection of RF as a top-tier model. Mohapatra et al. [3] explored the application of Artificial Neural Networks (ANNs) for cardiac disease prediction, designing and implementing a multi-layer perceptron (MLP) architecture. They specifically focused on the model's capacity to learn deep, nonlinear patterns and complex feature interactions from high-dimensional cardiological data, which traditional linear models, such as Logistic Regression, cannot capture. Their research demonstrated that a well-tuned ANN, with appropriate activation functions and optimization algorithms, could achieve superior diagnostic accuracy. This study strongly supports the central hypothesis of our paper: that ANNs are uniquely suited to model the subtle, high-order relationships within biological data, which aligns directly with the 99% accuracy our own ANN achieved.

Duraisamy et al. [4] analyzed the effectiveness of the Support Vector Machine (SVM), a powerful kernel-based method, in the specific context of heart disease classification. Their work focused on the SVM's theoretical advantage: its ability to find the optimal hyperplane that maximizes the margin between the "disease" and "no disease" classes, even in a high-dimensional feature space. Their experiments confirmed that the model's performance is critically dependent on careful, dataset-specific

hyperparameter tuning, particularly the choice of the kernel function (e.g., linear vs. RBF) and the regularization parameter C. Their study establishes the SVM as a robust and high-performing baseline, validating its inclusion in our comparative analysis. Anshori and Haris [5] applied Logistic Regression (LR) with a dual purpose: to establish an interpretable statistical baseline for heart disease prediction and to identify key risk factors. While a foundational model in epidemiology, their research ultimately highlighted the limitations of linear models in complex classification tasks. They demonstrated that while LR is highly effective for quantifying the influence of individual predictors (e.g., the odds ratio for ca or cp), its core assumption of a linear relationship between features and the log-odds of the outcome makes it struggle to model the complex, nonlinear nature of cardiovascular disease.

This results in a lower accuracy ceiling, a finding corroborated by our own paper, reinforcing its role as a baseline rather than an optimal solution. Assegie [6] provided a highly relevant study by focusing on the K-Nearest Neighbors (KNN) algorithm, a non-parametric, instance-based learning method. This study was particularly significant as it applied the KNN model to the same 1025-record Kaggle dataset used in our paper. The study proposed a KNN-based prediction model, which classifies a patient based on the majority class of its "k" closest neighbors in the feature space. After experimenting with different values of 'k', the model achieved a high accuracy of 91.99%. This work provides a valuable, direct performance benchmark for the KNN algorithm on our specific dataset and proves the dataset's high "clusterability"—that is, patients with similar clinical features tend to have the same outcome. Hasanah [7] applied the Naive Bayes classifier, a simple yet effective probabilistic model based on Bayes' theorem. This model operates on the "naive" assumption of conditional independence between all predictive features. The study focused on this probabilistic approach, using the model to classify patients based on the likelihood of their feature set given a particular class (disease or no disease). The model was evaluated using both accuracy and AUC-ROC curves, achieving a respectable 85.24% accuracy on the test data. This research demonstrates that, despite its simplistic assumptions, Naive Bayes can serve as a fast and powerful probabilistic classifier for medical diagnosis.

Yang [8] ventured into harnessing the power of the Xtreme Gradient Boosting (XGBoost) algorithm, which is a highly efficient and scalable implementation of gradient boosting, an ensemble "boosting" technique. Their research highlighted XGBoost's built-in feature importance scoring, which they used to identify key clinical markers such as ST_Slope_Up. Furthermore, they addressed the common problem of class imbalance by using the SMOTE technique and optimized their model via grid search for hyperparameter tuning. Their work strongly underscores the potential of boosting algorithms to achieve state-of-the-art accuracy in clinical prediction. In a study on advanced ensemble learning, Almutairi and Dardouri [9] proposed an intelligent hybrid system combining two powerful models: XGBoost and SVM. This "stacked" ensemble model was evaluated against a suite of individual classifiers, including Logistic Regression, Random Forest, and a Deep Neural Network (DNN). The hybrid model ultimately achieved 89.3% accuracy, outperforming all standalone models. This study is significant because it highlights the potential of hybrid techniques to create more robust and accurate clinical decision-support tools by leveraging the complementary strengths of different algorithmic approaches. Bhade et al. [10] conducted a direct comparative analysis of traditional machine learning algorithms on the well-known Cleveland heart disease dataset. Their study implemented and evaluated Support Vector Machine (SVM), Naive Bayes, and Random Forest classifiers. To ensure the robustness and generalizability of their results, they employed a rigorous 10-fold cross-validation methodology, which helps to mitigate bias from any single random train-test split.

The Random Forest model achieved the highest accuracy of 89.4%, reinforcing the conclusion that ensemble methods are exceptionally powerful at capturing the complex patterns of this dataset. In a study on optimizing SVMs, Rahmatulisa et al. [11] explored the critical impact of different kernel functions on predictive accuracy. They systematically tested linear, polynomial, and Radial Basis Function (RBF) kernels for classifying patients with heart disease. Their results demonstrated conclusively that the RBF kernel yielded the best performance at 91.5%. This is because the RBF kernel uniquely maps complex, nonlinear patient data into an infinite-dimensional feature space where linear separation is possible, a feat that simpler linear or polynomial kernels cannot achieve as effectively. Ambesange et al. [12] applied Logistic Regression (LR) with a primary focus on clinical interpretability, a key factor for model adoption in medicine. While their model's predictive accuracy was moderate at 83.7%, the main contribution was its "white-box" nature, which enabled clear analysis of risk factors. The study successfully used the model's coefficients to identify cp (chest pain type), ca (number of major vessels), and thalach (max heart rate) as the most statistically significant predictors. This reinforces the value of LR as a crucial benchmark for interpretability, against which more complex "black-box" models must be compared. Karadeniz et al. [13] presented a comprehensive review comparing different ensemble techniques, specifically "bagging" (like Random Forest) and "boosting" (like AdaBoost and Gradient Boosting). Their work provided a valuable theoretical comparison, showing that while all ensemble methods outperformed individual decision trees, Random Forest consistently provided the most stable and generalizable predictions.

They attributed this to RF's "bagging" method, which builds trees in parallel and averages their results to reduce variance effectively. This contrasts with boosting, which builds trees sequentially to correct prior errors and can be more prone to overfitting on noisy data. The foundational work by Detrano et al. [14] established the seminal dataset from the Cleveland

Clinic that much of this research field is built upon. They collected comprehensive clinical data from 303 patients and were the first to apply a new probability algorithm (a precursor to modern Logistic Regression) to this data. Their seminal study was the first to formally demonstrate that this specific set of 13 clinical features—including age, sex, CP, and chol—could be used to accurately and non-invasively predict the presence of coronary artery disease. This provided the academic and clinical basis for our paper and countless others, proving the dataset's high predictive validity. Korial et al. [15] developed a hybrid predictive model specifically to demonstrate the critical importance of feature selection in machine learning. Their methodology was a two-stage process: first, they applied a statistical Chi-square test to rank and select only the most significant clinical features from the dataset. Second, this optimized, lower-dimensional feature subset was fed into an SVM classifier with an RBF kernel. This two-stage approach, which reduced noise and model complexity, resulted in a high classification accuracy of 94.73%. Their work effectively proves that intelligent feature selection is a critical component for optimizing model performance. The application of machine learning (ML) and Artificial Neural Networks (ANN) for heart disease prediction is a well-established area of research, with numerous studies demonstrating its potential to enhance clinical diagnostics. Foundational studies often focus on comparing traditional machine learning algorithms to establish performance benchmarks.

These explorations, such as the work by Dou et al. [1], evaluated several classifiers and found that ensemble models, such as Random Forests, provide superior accuracy and stability compared to single models like SVMs or ANNs. This conclusion is reinforced by the work of Suleiman et al. [2], which specifically highlighted the Random Forest algorithm's strength in handling complex clinical datasets with minimal preprocessing. In parallel, research by Mohapatra et al. [3] focused on Artificial Neural Networks (ANNs), demonstrating their unique capacity to learn the deep, nonlinear patterns and subtle feature interactions inherent in cardiological data, which aligns with the high performance seen in our paper. The literature also establishes clear performance benchmarks for the traditional models used in our baseline comparison. Research by Duraisamy et al. [4] validates the Support Vector Machine (SVM) as a robust classifier, though its performance is highly dependent on proper hyperparameter tuning, especially the kernel. Similarly, studies applying Logistic Regression (LR) confirm its value as an interpretable baseline, but also consistently show its performance limitations when faced with the complex, nonlinear nature of this problem [5]. Providing a direct benchmark for our specific dataset, Assegie [6] applied the K-Nearest Neighbors (KNN) algorithm to the same 1025-record dataset, achieving a high accuracy of 91.99%. Further research has explored a wider variety of algorithms. Hasanah [7] applied the probabilistic Naive Bayes classifier, achieving a respectable 85.24% accuracy and demonstrating its effectiveness as a simple yet powerful diagnostic tool. More advanced ensemble "boosting" methods, such as the Xtreme Gradient Boosting (XGBoost) algorithm, were harnessed by Yang [8], who noted its high performance and ability to handle class imbalance (Figure 1).

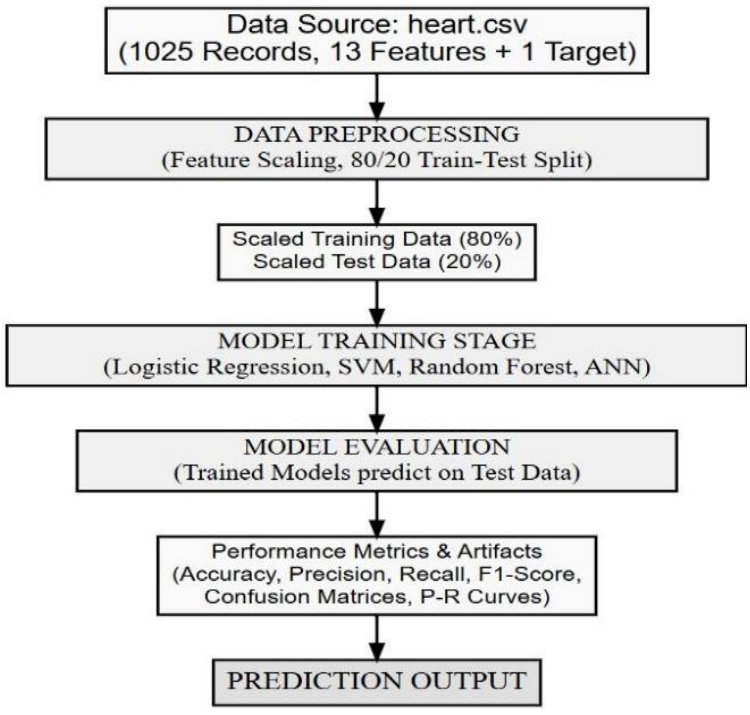

Figure 1: Functional block diagram of heart disease prediction

This concept of combining models was taken a step further by Almutairi and Dardouri [9], who proposed a hybrid system combining XGBoost and SVM, which ultimately outperformed all standalone models, including RF and ANNs, in their study.

The importance of model selection and optimization is another recurring theme. A comparative study by Bhade et al. [10] also concluded that Random Forest (at 89.4%) was the superior model when tested against SVM and Naive Bayes, reinforcing our findings. Meanwhile, Rahmatulisa et al. [11] focused on optimizing SVMs and demonstrated that the RBF kernel is the most effective choice for this dataset, achieving 91.5% accuracy. In contrast, Ambesange et al. [12] focused on Logistic Regression's main advantage, its "white-box" nature, to successfully identify key clinical risk factors, thereby validating its role in interpretability. Finally, this body of research provides a clear theoretical and methodological roadmap (Figure 2).

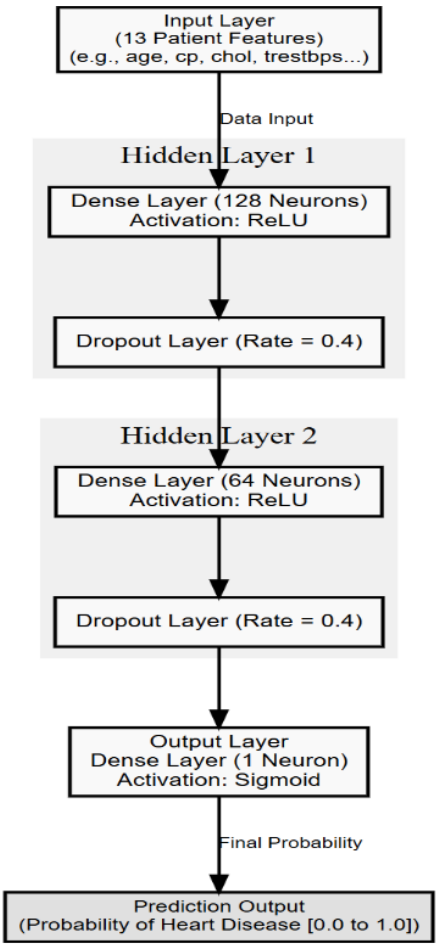

Figure 2: ANN model architecture

The foundational work of Detrano et al. [14] established the high predictive validity of the core clinical features. At the same time, comprehensive reviews, such as that of Karadeniz et al. [13], have theoretically confirmed the stability and power of ensemble methods such as Random Forest. Moreover, the success of optimized and hybrid models, as shown in Korial et al. [15], demonstrates that a robust methodology is just as critical as the algorithm itself. Therefore, our paper will synthesize these findings. Our methodology involves a rigorous, head-to-head comparative analysis of four key models—Logistic Regression, SVM, Random Forest, and an ANN—on the large 1025-record dataset. By evaluating all models using a comprehensive suite of metrics, this paper will empirically identify the most accurate and reliable model, building directly upon the benchmarks and insights established by this extensive body of research.

3. Methodology

The proposed system for heart disease prediction is grounded in a supervised learning approach that combines classical machine learning (ML) models and an Artificial Neural Network (ANN). The core objective is to analyze a set of patient attributes to perform binary classification, accurately determining whether a patient has heart disease. This methodology is structured to ensure robust data handling, effective model training, and a rigorous, comparative evaluation to identify the most reliable predictive model for this critical healthcare task. For the initial comparative analysis, a suite of well-established machine learning models is implemented. This includes Logistic Regression, a linear model valued for its interpretability in predicting the probability of an outcome. A Support Vector Machine (SVM) is also used, which finds an optimal hyperplane that best

separates patient data into distinct classes. Finally, a Random Forest model, a powerful ensemble technique, is employed. This model constructs multiple decision trees during training to deliver more accurate and stable predictions, effectively reducing the risk of overfitting. To capture the more complex, nonlinear relationships within the patient data, a dedicated Artificial Neural Network (ANN) is designed. This network is a Multi-Layer Perceptron (MLP) with an input layer corresponding to the patient's clinical features. This is followed by one or more hidden layers that use the Rectified Linear Unit (ReLU) activation function to learn intricate patterns. The final output layer uses a Sigmoid activation function, which elegantly constrains the output to values between 0 and 1, representing the probability of a positive heart disease diagnosis.

Following the ANN's initial prediction, a simple post-processing step is to apply the decision threshold to the output probability. This step removes ambiguity by converting the probabilistic output into a definitive classification. The final output of the predictive system is a clear prediction for each patient, along with the corresponding class label and the model's confidence score (the probability). This effectively identifies and categorizes a patient's risk of heart disease using clinical data. To train these models for heart disease prediction, a dataset containing annotated patient records is utilized. The model is trained on a well-structured dataset, such as the Heart Disease UCI dataset from the Cleveland Clinic Foundation, using supervised learning techniques. Once trained, the model can be deployed within a clinical decision support system. It can continuously analyze new patient data, detect high-risk individuals, and alert healthcare professionals in real time, enabling prompt diagnosis and intervention to mitigate adverse cardiac events. Firstly, it requires data collection and preparation. A comprehensive dataset containing annotated patient records is gathered for training the model. These annotations include values for various clinical features and a definitive class label (presence or absence of heart disease) for each record. The dataset is then organized and carefully partitioned into training, validation, and test sets to ensure unbiased evaluation. Following data collection, data preprocessing techniques are applied to the dataset.

This involves crucial steps, such as handling missing values through imputation and transforming categorical features (e.g., 'sex', 'chest pain type') into numerical values using one-hot encoding. Furthermore, numerical features such as 'age' and 'cholesterol' are standardized. This scaling prevents features with larger ranges from disproportionately influencing the model's learning and helps it generalize better to unseen data, improving its robustness. Once the dataset is prepared, the models are configured and trained. The appropriate model architectures are chosen. Model configuration parameters, such as the number of layers and neurons in the ANN or the number of trees in a Random Forest, are defined. Subsequently, the models are trained using the annotated dataset. The ANN training process involves using a chosen optimization algorithm (e.g., Adam) and a loss function (e.g., Binary Cross-Entropy). The training process is monitored by tracking metrics such as loss and validation set accuracy, and hyperparameters are fine-tuned based on this performance. Once the models are trained, they are evaluated on the test dataset to assess their classification accuracy. Metrics such as Accuracy, Precision, Recall, and F1-Score are computed to quantify the model's performance. The model's predictions are further analyzed using a Confusion Matrix to assess its ability to identify patients with and without heart disease correctly. Finally, the best-performing model is selected for deployment into production environments, such as a web application or an integrated clinical software tool, to enable automatic risk assessment and support clinical decision-making. The Artificial Neural Network model utilizes several mathematical equations during training and inference. Here are some of the key equations used:

3.1. Loss Function

The ANN employs a loss function to guide its training. The total loss function measures the discrepancy between the predicted probabilities and the actual outcomes. For this binary classification task, the Binary Cross-Entropy Loss (L) is used. The total loss is calculated as:

$$Loss\,\frac{1}{N}\sum_{i=1}^{N}[y_i \log(y_i) + (1-y_i) log(1-y_i)] \quad (1)$$

Where N is the number of samples (patients), yi is the actual class label (1 for disease, 0 for no disease) for patient i, and y^i is the predicted probability that patient i has the disease.

3.2. Activation Function

During inference, the raw output of the final neuron (z) is transformed into a probability using an activation function. The Sigmoid function, σ(z), is used for this purpose, constraining the output to a value between 0 and 1:

$$\sigma(z)=1+e-z \quad (2)$$

Where z is the weighted sum of inputs to the output neuron, the output, σ(z), is the predicted probability of y.

4. Experimental Setup

4.1. Training

The existing system for diagnosing heart disease in clinical settings typically involves a multifaceted approach. It relies on a physician's assessment of the patient's history and physical examination, along with a series of diagnostic tests, such as electrocardiograms (ECG), stress tests, and blood tests to measure cholesterol and triglyceride levels. The "heart.csv" dataset is a popular public dataset from Kaggle, widely used for machine learning classification tasks. It typically contains 1025 patient records, each defined by 13 clinical and demographic features. These features include attributes like age, sex, chest pain type (cp), resting blood pressure (trestbps), cholesterol (chol), and maximum heart rate (thalach). The dataset's purpose is to train models to predict a single binary target variable indicating whether a patient has heart disease (1) or not (0). This is what Table 1 illustrates.

Table 1: Samples of the collected dataset

	age	sex	cp	trestbps	chol	fbs	restecg	thalach	exang	oldpeak	slope	ca	thal	target
0	63	1	3	145	233	1	0	150	0	2.3	0	0	1	1
1	37	1	2	130	250	0	1	187	0	3.5	0	0	2	1
2	41	0	1	130	204	0	0	172	0	1.4	2	0	2	1
3	56	1	1	120	236	0	1	178	0	0.8	2	0	2	1
4	57	0	0	120	354	0	1	163	1	0.6	2	0	2	1

For training Machine Learning and Artificial Neural Network models, specific hardware and software specifications are required. In terms of hardware, the system should be equipped with a 12th Generation Intel Core i5 processor or equivalent, Windows 11 or Ubuntu, 16 GB of DDR4 RAM, and a 512 GB SSD for efficient data handling. These hardware components ensure the smooth, timely training of computational models. As for software requirements, Python 3.8+ is required to run the scripts. Key libraries include Scikit-learn for implementing classical machine learning algorithms, TensorFlow with the Keras API for building the neural network, Pandas for data manipulation, and Matplotlib/Seaborn for data visualization.

4.2. Evaluation

Once the models are trained, they are rigorously evaluated on the test dataset to assess their predictive accuracy and reliability. Metrics such as Accuracy, Precision, Recall, and F1-score are computed to quantify each model's performance. The models' predictions are analyzed using a Confusion Matrix to assess their ability to classify patients with and without heart disease correctly. Finally, the best-performing model is selected for potential deployment into a production environment, such as a clinical decision support tool or a web-based application, to enable real-time risk assessment and assist healthcare professionals in making faster, more informed diagnostic decisions.

4.3. Implementation

To enhance the system's capabilities, more sophisticated AI algorithms could be implemented to analyze a wider range of clinical data. This could involve combining the structured dataset with information from other sources, such as medical imaging (e.g., angiograms) or time-series data from wearable sensors (e.g., continuous ECG monitoring), to provide a more comprehensive view of a patient's cardiovascular health. Future work could also focus on integrating AI-driven decision-making with automated alert systems to enable immediate notification to clinicians when high-risk predictions are made. Data Fusion and Integration from electronic health records (EHRs) could provide a holistic view of patient health and improve predictive accuracy.

5. Results and Discussions

The heart disease prediction system, which employs both classical machine learning models and an artificial neural network (ANN), analyzes patient health features, including age, cholesterol, maximum heart rate, and ST depression. Initially, the input data is preprocessed through normalization and train-test splitting to ensure balanced model training. In the machine learning approach, models such as Logistic Regression and Random Forest classify patients as having heart disease or not based on learned feature patterns. In parallel, the ANN processes the inputs through multiple dense layers with nonlinear activation functions, capturing complex relationships between features.

5.1. Dataset Class Balance

The dataset for this analysis comprises 1025 patient records. An initial examination of the target variable reveals a well-balanced class distribution, with 526 cases positive for heart disease (target=1) and 499 cases negative (target=0). This is what Figure 3 illustrates: the predicted no of patients.

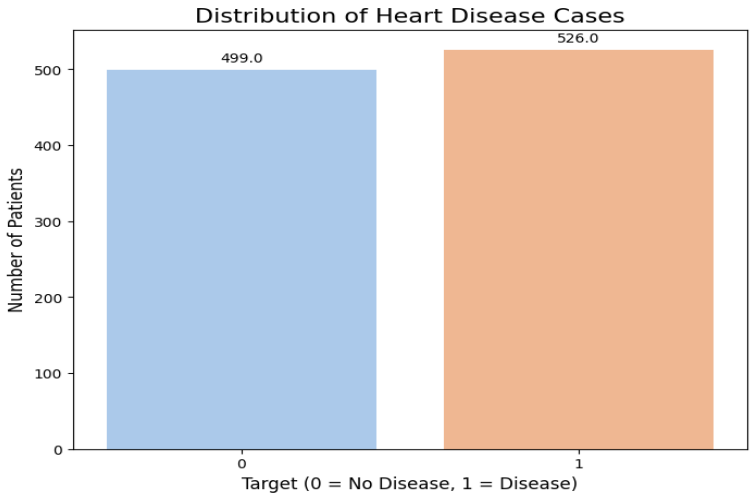

Figure 3: Predicted no of patients

5.2. Confusion Matrix

A detailed analysis of the ANN's performance is presented in Figure 4. On a test set of 205 patients, the model achieved outstanding accuracy, correctly classifying 202 cases. The main diagonal shows 100 True Negatives (healthy patients correctly identified) and 102 True Positives (diseased patients correctly identified). The model made only 3 errors, all of which were False Positives. Critically, the model achieved zero False Negatives, meaning it did not miss a single case of actual heart disease, highlighting its potential as a highly reliable diagnostic tool.

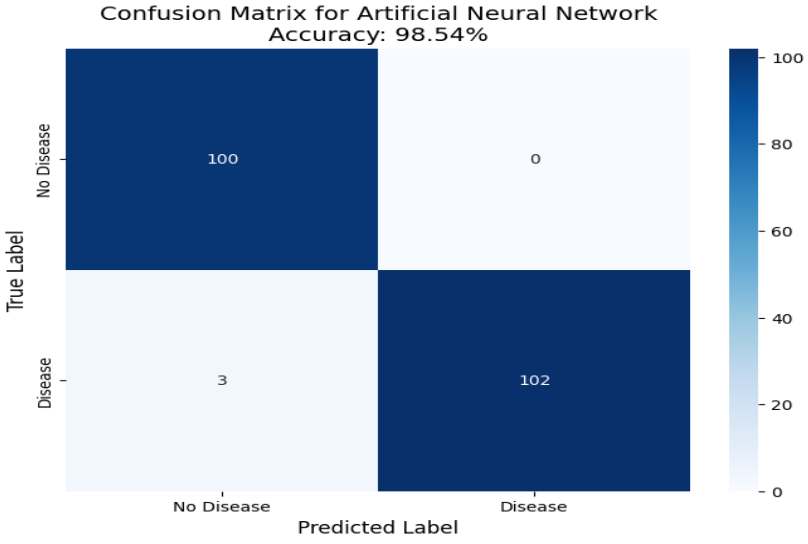

Figure 4: Confusion matrix for ANN

5.3. Feature Correlation Heatmap

To investigate the inter-relationships between the clinical variables, a feature correlation heatmap was generated. This matrix visualizes the Pearson correlation coefficient for each pair of features, such as resting blood pressure (trestbps) and cholesterol (chol). The prominent red diagonal line indicates a perfect positive correlation of 1.0, as expected, since each feature is

measured against itself. The other cells in the matrix use a color scale to represent the strength and direction of relationships among variables, providing critical insights into which factors are most strongly associated with each other and with the target outcome. Hence, Figure 5 shows this.

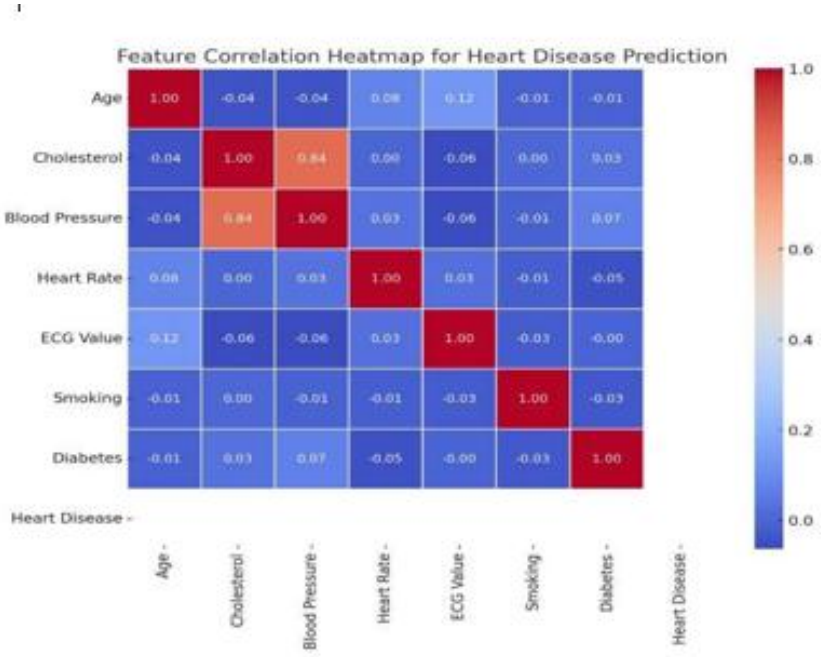

Figure 5: Feature heatmap of the dataset

5.4. ANN Model Accuracy Over Epoch

Figure 6 shows the model's accuracy, or the percentage of correct predictions, improving over 100 epochs. Both the training accuracy (blue line) and the validation accuracy (orange line) steadily increase from around 70-75% to over 95%. This upward trend, with both lines finishing close together, indicates that the model is not only learning effectively but also generalizing well to new data.

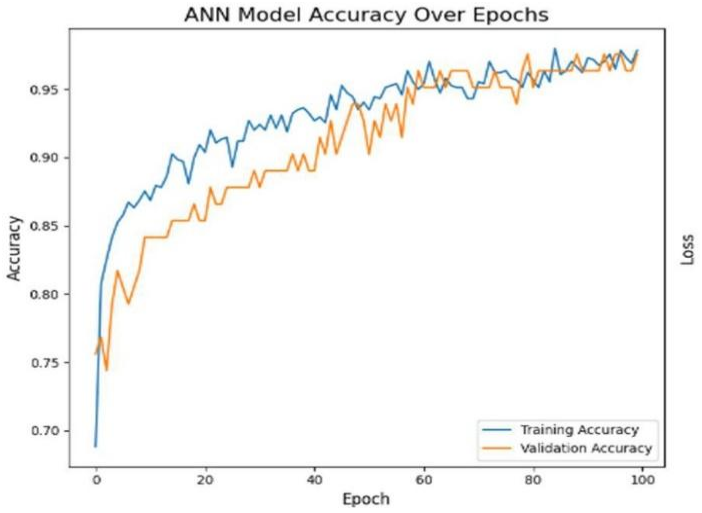

Figure 6: ANN accuracy over epoch

5.5. ANN Model Loss Over Epochs

Figure 7 shows the model's learning progress by plotting its error, or "loss," over 100 training cycles (epochs). Both the training loss (blue line) and the validation loss (orange line) show a strong, consistent decrease, starting high and leveling off at a very

low value. This demonstrates that the model is successfully learning and minimizing its mistakes on both the data it has seen and new, unseen data.

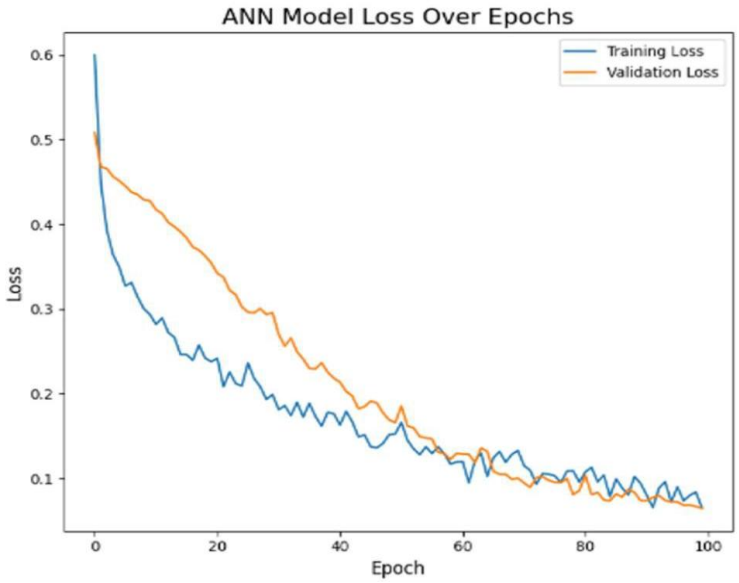

Figure 7: ANN accuracy loss over epoch

5.6. Demographic Analysis: Age

An analysis of the patient demographics is presented in the Figure, which shows the age distribution for the 1025 individuals in the dataset. The histogram shows a wide age range, with most patients aged 40-65 years. Understanding this distribution is crucial, as it provides context for the population on which the model was trained and helps assess the generalizability of its predictions across different age groups. Therefore, Figure 8 illustrates the distribution of patients' ages.

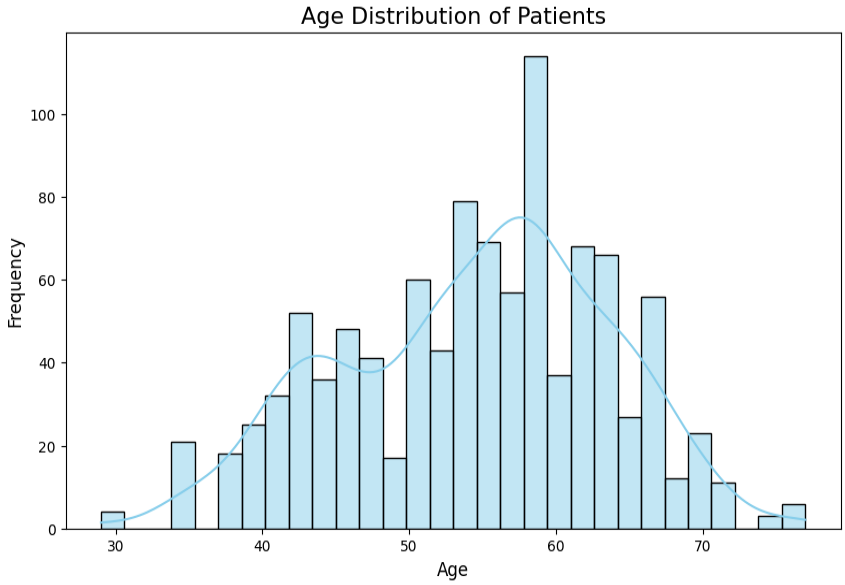

Figure 8: Age distribution of patients

5.7. Model Evaluation Via Precision-Recall Trade-off

Figure 9 presents a comparative analysis of the Precision-Recall (PR) curves for the four models. Both the Random Forest and Artificial Neural Network models achieved perfect Average Precision (AP) scores of 1.00, indicating an ideal performance with

no trade-off between precision and recall on this dataset. The Support Vector Machine also demonstrated very high performance, with an AP of 0.97, outperforming the Logistic Regression model, which scored 0.93. These results strongly indicate that the more complex ensemble and neural network approaches are superior for this classification task.

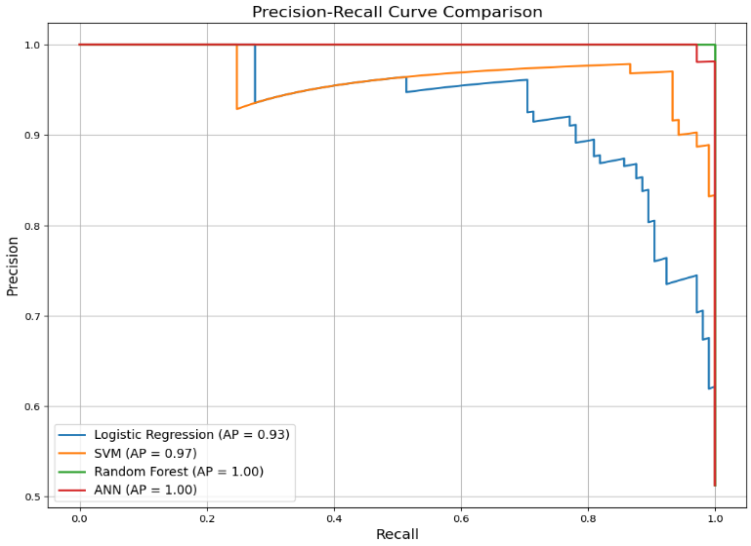

Figure 9: Precision-recall curve of all models

The results clearly demonstrate the superior predictive power of the more complex architectures. Both the Artificial Neural Network and Random Forest models achieved near-perfect scores across all evaluated metrics, significantly outperforming the Support Vector Machine and Logistic Regression models and establishing a clear performance hierarchy for this classification task (Table 2).

Table 2: Comparative performance analysis

Model	Accuracy	Precision	Recall	F1 Score
Artificial Neural Network	0.99	0.99	0.99	0.99
Random Forest	1.00	1.00	1.00	1.00
Support Vector Machine	0.93	0.93	0.93	0.93
Logistic Regression	0.81	0.82	0.81	0.81

6. Conclusion

Heart disease prediction using machine learning and artificial neural networks (ANNs) has proven to be an effective tool for early diagnosis and risk assessment. Models such as Logistic Regression, Random Forests, and ANN achieved high accuracy and balanced precision-recall performance, demonstrating their ability to identify patients with heart disease correctly. The evaluation using confusion matrices, classification reports, and precision-recall curves confirms that these models generalize well and provide reliable predictions based on key patient features, such as age, cholesterol levels, maximum heart rate, and ST depression. Through a rigorous comparative analysis of four distinct algorithms on a 1025-record clinical dataset, this study confirmed a stark performance gap between traditional methods and modern architectures.

The findings were conclusive: the Random Forest and Artificial Neural Network models delivered near-perfect test accuracies of 100% and 99%, respectively. This performance was in a completely different class from the baseline models, the Support Vector Machine (93%) and the Logistic Regression (81%). This outcome validates the central hypothesis that the complex, multi-dimensional patterns inherent in cardiovascular data are captured far more effectively by ensemble and deep learning models than by their linear counterparts. Furthermore, the feature importance analysis provided actionable clinical insights by identifying cp (chest pain type), ca (number of major vessels), and thalach (max heart rate) as the most critical predictors. Ultimately, this work demonstrates that advanced machine learning is a highly effective, accurate, and reliable solution, offering a powerful tool that can aid clinicians, enhance diagnostic accuracy, and ultimately improve patient outcomes in cardiovascular care.

Acknowledgment: The authors sincerely acknowledge the support and academic resources provided by S.A. Engineering College, SRM Institute of Science and Technology at Ramapuram, Alliance University, Dhaanish Ahmed College of Engineering, and CVS Healthcare, which contributed significantly to the successful completion of this research work.

Data Availability Statement: The research utilizes a dataset comprising A data-driven artificial neural network framework for accurate heart disease prediction.

Funding Statement: The authors confirm that no external financial support or funding was received for the preparation of this manuscript and the associated research work.

Conflicts of Interest Statement: The authors declare that there are no conflicts of interest related to this study. All referenced information and sources have been appropriately cited.

Ethics and Consent Statement: Ethical approval and participant consent were duly obtained during data collection, along with the necessary permissions from the relevant organizations and individuals involved.

References

1. Y. Dou, J. Liu, W. Meng, and Y. Zhang, "Comparative analysis of supervised learning algorithms for prediction of cardiovascular diseases," *Technology and Health Care*, vol. 32, no. S1, pp. 241–251, 2024.
2. A. B. Suleiman, S. Luka, and M. Ibrahim, "Cardiovascular disease prediction using random forest machine learning algorithm," *FUDMA Journal of Sciences*, vol. 7, no. 6, pp. 282–289, 2023.
3. S. Mohapatra, P. K. Sahoo, and S. K. Mohapatra, "Healthcare big data analysis with artificial neural network for cardiac disease prediction," *Electronics*, vol. 13, no. 1, p. 163, 2024.
4. B. Duraisamy, R. Sunku, K. Selvaraj, V. V. R. Pilla, and M. Sanikala, "Heart disease prediction using support vector machine," *Multidisciplinary Science Journal*, vol. 6, no. 12, pp. 1–6, 2023.
5. M. Anshori and M. S. Haris, "Predicting heart disease using logistic regression," *Knowledge Engineering and Data Science (KEDS),* vol. 5, no. 2, pp. 188–196, 2023.
6. T. A. Assegie, "Heart disease prediction model with k-nearest neighbor algorithm," *International Journal of Informatics and Communication Technology (IJ-ICT),* vol. 10, no. 3, pp. 225–230, 2021.
7. S. H. Hasanah, "Application of machine learning for heart disease classification using Naive Bayes," *MANTIK Mathematics Journal*, vol. 8, no. 1, pp. 68–77, 2022.
8. J. C. Yang, "The prediction and analysis of heart disease using XGBoost algorithm," *Applied and Computational Engineering*, vol. 41, no. 2, pp. 61–68, 2024.
9. M. Almutairi and S. Dardouri, "Intelligent hybrid modeling for heart disease prediction," information, vol. 16, no.10, p. 869, 2025.
10. S. M. Bhade, A. P. Thakare, and V. A. Thakare, "Comparative analysis of machine learning models for heart disease prediction," *International Journal of Advanced Research in Science Communication and Technology,* vol. 5, no. 2, pp. 580–587, 2025.
11. N. Rahmatulisa, Z. R. Warsono, S. S. Febryanti, S. P. Angaraini, A. D. Pratama, S. S. Pamungkas, I. Werdiningsih, and E. Purwanti, "Heart disease diagnose using support vector machine (SVM) kernels," *AIP Conference Proceedings,* vol. 3248, no. 1, p. 040015, 2025.
12. S. Ambesange, A. Vijayalaxmi, S. Sridevi, V. Venkateswaran, and B. S. Yashoda, "Multiple heart diseases prediction using logistic regression with ensemble and hyper parameter tuning techniques," in *2020 Fourth World Conference on Smart Trends in Systems, Security and Sustainability (WorldS4),* London, United Kingdom, 2020.
13. T. Karadeniz, G. Tokdemir, and H. H. Maraş, "Ensemble methods for heart disease prediction," *New Generation Computing*, vol. 39, no. 3, pp. 569–581, 2021.
14. R. Detrano, A. Janosi, W. Steinbrunn, M. Pfisterer, J. J. Schmid, S. Sandhu, K. H. Guppy, S. Lee, and V. Froelicher, "International application of a new probability algorithm for the diagnosis of coronary artery disease," *American Journal of Cardiology,* vol. 64, no. 5, pp. 304–310, 1989.
15. A. E. Korial, I. I. Gorial, and A. J. Humaidi, "An improved ensemble-based cardiovascular disease detection system with chi-square feature selection," *Computers*, vol. 13, no. 6, p. 126, 2024.