The Bioscan
AN INTERNATIONAL QUARTERLY JOURNAL OF LIFE SCIENCES
20(3): S.I (3), 268-276, 2025 www.thebioscan.com

A COMPREHENSIVE SURVEY ON HEART DISEASE PREDICTION USING MACHINE LEARNING AND DEEP LEARNING APPROACHES

[1*]S. Selvapriya, [2]Dr. M. Saranya

[1*]Research Scholar, P.K.R. ARTS COLLEGE FOR WOMEN, Gobichettipalayam, Erode (DT), TamiNadu, India.

[2]Research Supervisor, Associate Professor, P.K.R. ARTS COLLEGE FOR WOMEN, Gobichettipalayam, Erode (DT), TamiNadu, India.

E-mail: Id: [1*]selvapriya.2020@gmail.com, [2]drmsaranyamcapersona@gmail.com

DOI: 10.63001/tbs.2025.v20.i03.S.I(3).pp268-276

KEYWORDS
Machine Learning (ML), Deep Learning (DL), Long Short-term Memory (LSTM), Heart Diseases, Convolutional Neural Network (CNN), logistic regression, decision trees, support vector machines, random forests, neural networks.

Received on:

08-06-2025

Accepted on:

02-07-2025

Published on:

06-08-2025

ABSTRACT

Globally, the death rate is increased by one of the major conditions named heart disease (HD). This HD greatly impacts the global healthcare systems. For the purpose of enhancing outcomes of the patient and reducing medical challenges, the early detection (ED) and diagnosis of cardiovascular disease (CVD) is crucial. Then, the implementation of the artificial intelligence (AI), namely machine learning (ML) and deep learning (DL) techniques have revolutionized the predictive modelling of HD. A comprehensive insights regarding the recent developments in HD prediction with the application of the machine learning (ML) and deep learning (DL) algorithms, including logistic regression (LR), decision trees (DT), support vector machines (SVM), random forests (RF), neural networks (NN), convolutional NN (CNN), and long short-term memory (LSTM) models was offered in this study. Here, the commonly utilized datasets, feature selection (FS) strategies, data pre-processing approaches are all examined in this study. Then the study also analyses the assessment metrics that will helps in determining the accuracy (ACC) and dependability of the predictive models (PM). The benefits, drawbacks, and efficacy of every model is identified by the comparison of models. This survey also facilitates in resolving issues like data imbalance, model interpretability, privacy issues, and practical deployment limitations. Recommendations regarding future directions, like explainable AI, federated learning (FL), and the integration of multi-modal (MM) health data was also offered in this study, and it may help the experts in creating more clinically valuable and dependable prognostic tools. This comprehensive survey contributes the scholars and professionals in creating intelligent systems for HD diagnosis and risk assessment. So, this comprehensive survey is beneficial.

INTRODUCTION

According to the World Health Organisation (WHO), CVD, especially HD, is still the leading cause of mortality worldwide, taking the lives of around 17.9 million people year [1]. Heart failure (HF), arrhythmias, coronary artery disease, and other illnesses affecting the heart and blood vessels (BV) are included in this category. Even though conventional diagnostic techniques like electrocardiograms (ECG) [2], echocardiography [3], and angiography [4] have proven successful, they are frequently costly, time-consuming, and interpreters need to be trained. Then, after the disease progression in patient, the diagnosis will be given. This carelessness may reduce the possibilities in giving an effective treatment. So, in this case, the development of intelligent and automated solutions is demanded, and it is facilitated by the ED and precise prediction of HD. Hence, ED and precise prediction of the HD is crucial for giving a patient immediate and effective care [5].

The diagnosis and prediction of the cardiac-associated conditions are revolutionized by the application of the computational intelligence (CI), especially in the fields of ML [6] and DL [7]. When comparing these suggested models with conventional statistical methods, the recommended methods have a maximum potential in analyzing complex, non-linear (NL) relationships in clinical datasets, detecting hidden patterns, and generating predictive models (PM) than those conventional methods [8]. In a variety of structured and unstructured health data, the ML models like DT, RF, and SVM as well as DL architectures including Artificial NN (ANN), CNN, and Recurrent NN (RNN) [9] are used. Here, Medical imaging (Med Img), wearable device data, electronic health records (EHRs) [10] are some instances of structured and unstructured health data.

Here, pre-processing and selection of high-quality features from patient data are crucial. Because, these methods may help the experts in creating effective PM [11]. The following factors of the patient has to be considered, they are: age, gender, blood pressure (BP), cholesterol, smoking, family history, and ECG readings. Then, the publicly accessible datasets like the UCI HD dataset [12], and the Framingham Heart Study dataset [13] are utilized, and it may enhance algorithm benchmarking and performance comparison. These studies consistently report that hybrid models [14], combining multiple ML/DL approaches, tend to achieve higher accuracy (ACC), sensitivity (S), and robustness in detecting early signs of HD [15]. Despite its promising advancements, there are many risks that exist in this field. Then, the following issues like unbalanced datasets, overfitting, uninterpretable DL models limiting its widespread application.

This widespread application is also limited by the challenges in implementing real-time (RT) in clinical processes. Then, properly managing issues related to the confidentiality of the patient and transparency is crucial. For future research directions, a comprehensive review regarding the state-of-the-art (SOTA) ML and DL methods in HD prediction was offered by this survey. A more precise, interpretable, and scalable framework is needed, so this study combines the current data, and it may help the experts, scholars, and system developers for creating that precise model. This precise model thus contributes in ED and effectively controlling HD.

1. Literature Review

The Cleveland HD dataset was used by Ahmad and Polat [16] to develop an ML model that could predict HD with the best performance. The model was trained using the dataset's characteristics. The model's performance was significantly impacted by the choice of ML algorithm. The Jellyfish optimisation algorithm (JFOA) was used to reduce the Cleveland dataset to a lower-dimensional (LD) subspace in order to prevent overfitting caused by the curse of dimensionality, which is caused by the excessive amount of features in the dataset. The Jellyfish algorithm was adaptable in identifying the optimal features and had a high rate of convergence. Researchers investigated and compared the performance of the models that were produced by using several ML techniques to train the feature-selected dataset. With S, Specificity (SP), ACC, and Area Under Curve (AUC) of 98.56%, 98.37%, 98.47%, and 94.48%, the SVM classifier model that was trained on the dataset optimised using the Jellyfish algorithm had the best performance. By integrating the JFOA and SVM classifiers, an optimal efficiency for the HD prediction was attained, and it was demonstrated by the outcomes.

For diagnostic reasons, Ali et al. [17] sought to determine which ML classifiers had the highest ACC. Here, many supervised ML (SML) techniques are utilized for determining the ACC and efficiency of the method in predicting HD. Then, the feature significance scores (FIS) for each feature is computed by all applied methods with the exception of MLP and KNN. Then, ranking every feature based on its significance scores may help in determining the features that contributes most to obtain accurate HD predictions. 3-class classification on an HD dataset obtained from kaggle is used. In this execution, k-nearest neighbour (KNN), DT, and RF algorithms are used in this study. From the outcomes, it is clear that the RL method attains 100% of ACC, S, and SP. Thus, these values also indicate the potential of the SML model's in its practical applications, and its high ACC in predicting HD.

In order to create a precise ML system for predicting HD on early stage, El-Sofany et al. [18] sought to employ various FS methods. ANOVA, chi-square, and mutual information (MI) techniques were used for selecting features. SF-1, SF-2, and SF-3 were the names given to the three film groups that were produced. To identify the most accurate method and the best-matching feature subset, ten ML techniques were used. SVM, XGBoost (XGB), bagging, DT, and RF were among the classification techniques that were employed. A private dataset, a public dataset, and a variety of cross-validation (CV) techniques were used to assess the suggested HD prediction algorithm. To overcome data imbalance and determine which ML algorithm produced the best accurate HD predictions, the Synthetic Minority Oversampling Technique (SMOTE) was employed. This suggested approach helps the clinical specialists in the ED of HD in a rapid way. Based on the input symptoms, a mobile application was created, and it may help the clinicians in predicting HD in rapid manner via the support of the ML. Then, simulation was conducted, from the outcomes, it is clear that XGB executes better when it is used in the SF-2 feature subset and the combined datasets. This XGB also attains 97.57% ACC, 96.61% S, 90.48% SP, 92.68% F1 score (F1 measure), and 98% AUC, 95.00% P. To determine the way, the model decide its final predictions, an explainable AI technique based on SHAP methods was additionally developed.

To create a potential ML model to predict HD early on, Biswas et al. [19] use a variety of FS approaches to identify key features. Three different approaches: chi-square (CS), ANOVA, and MI were applied to FS.The chosen feature subsets were designated SF1, SF2, and SF3. LR (C1), SVM (C2), KNN (C3), RF (C4), Naive Bayes (NB) (C5), and DT (C6) were the six different ML models that were used to identify the most favorable model and the best-fit feature subset. In the end, the RF model yielded the most favorable outcomes for the SF3 feature subset, with a 94.95 area under the ROC curve (AUC-ROC), a 0.31 log loss, 94.51% ACC, 94.87% S, and 94.23% SP. Based on its performance and the selected variables, the recommended model have the potential to detect the ED of HD in a cheapest way with a limited period.

To enhance classification ACC, Bhatt et al. [20] suggested a technique that uses k-modes clustering with Huang initialisation. RF, DT classifier, multilayer perceptron (MLP), and XGB were among the models used.The outcomes were optimised by hypertuning the parameters of the used models using GridSearch (GS) CV. Seventy thousand real-world instances from Kaggle were used to test the suggested model. DT (86.37% with CV) and 86.53% without CV), XGBoost (86.87% with CV) and 87.02% without CV), RF (86.5% with CV) and 86.92% without CV), and MLP (87.28% with CV) and 86.94% without CV) were the accuracies attained after training the models using data split in an 80:20 ratio. The AUC values of DT of 0.94, XGBoost of 0.95, RF of 0.95, and MLP of 0.95 were also attained by the suggested models. The MLP model with CV achieved the maximum ACC of 87.28%, outperforming all other algorithms, according to the research's findings.

In order to identify imminent HD using ML approaches, Nashif et al. [21] suggested a tentative design for a cloud-based (CB) HD prediction system. Using the Java-based open-access (DM) data mining platform WEKA, an effective ML technique was used for precise detection. This method was developed through a comparative analysis of multiple algorithms. Two popular open-access databases were utilised to validate the suggested approach, and the HD detection performance was assessed using 10-fold CV. A 97.53% ACC, 97.50% S, and 94.94% SP were attained by the SVM algorithm. Additionally, an Arduino-based RT patient monitoring system was created to enable doctors or caretakers to continuously monitor HD patients. Body temperature, blood pressure (BP), humidity, and heartbeat (HB) were among the RT parameters that this device could detect. Then, once every ten seconds, it might send a central server the recorded data. In the event that they need to act quickly, doctors can then see the patient's RT sensor data using an application and initiate live video streaming. The system's ability to promptly notify the assigned doctor using GSM technology whenever any RT metric exceeded the preset threshold was another crucial component.

Based on a number of features, Saha et al. [22] suggested an ML method to identify individuals with CVD. To make it compatible with ML techniques, the HD dataset from the Centres for Disease Control and Prevention (CDC) underwent numerous pre-processing steps and attribute changes. The RF algorithm produced the greatest results across all performance parameters assessed out of the five SML algorithms that were put into practice. It obtained a AUC-ROC of 0.91 and a maximum ACC of 91%. Additionally, it achieved the lowest False Negative (FN) Rate (FNR) of 0.078 and the highest True Positive (TP) Rate (TPR) of 0.922, outperforming the other algorithms.

A novel self-attention (SA)-based transformer model was presented by Rahman et al. [23] to predict CVD risk by combining transformer networks with SA mechanisms (SAM). By gathering contextual information and creating representations, the SA layers were able to effectively model complex patterns in the data. By allocating attention weights (AW) to every element of the input sequence, the SAM made the data interpretable. In order to retrieve pertinent information, this required altering the attention processes, adding more layers, and adjusting the input layers (OL) and output layers (IL). This method may help the experts in determining the data features that impacted the framework's predictions. Then, the suggested method is assessed by using the Cleveland dataset, a benchmark dataset from the UCI ML repository. Thus simulation was conducted and it compares the suggested method with current SOTA methods. From the outcomes, it is clear that the suggested model attained greatest ACC of 96.51% in contrast to current SOTA methods. This value highlights the efficiency of the suggested model in HD prediction.

By combining the effectiveness of DT classification with a genetic algorithm (GA), Mandavkar et al. [24] suggested an ensemble model. The effectiveness of the suggested prediction system was increased by using the GA-identified features for ML classification.

This analysis utilizes Cleveland dataset. 5 class types, 14 attributes, and 303 instances are included in this Cleveland dataset that was accessible from UCI repository. 0 (Normal) to 4 (Stroke) were the class values. From the outcomes of the simulation, it is clear that the application of GA-based FS and classification delivers optimal outcomes for multi-class (MC) disease risk prediction. Thus, a web application is created, and it may help the patients to consult with clinicians and predict the status of patients.

A ML model for HD prediction based on pertinent features has been suggested by Sharma et al. [25]. A benchmark dataset of 14 different HD-related parameters was extracted from the UCI HD Prediction repository for this investigation. The model was created using ML methods like RF, SVM, NB, and DT. In order to forecast the likelihood of HD, the study additionally examined the relationships among a number of dataset features using standard ML approaches. The simulation was conducted by comparing RF with other ML methods. From the outcomes, it is clear that, the RF method attains maximum ACC in less time for prediction than the other Ml methods. For decision support system (DSS) in clinical settings, this model is considered effective, and it is beneficial for clinical experts.

To predict a patient's risk of getting HD, Pal and Parija [26] suggested the RF algorithm. The dataset, which included 303 samples with 14 features chosen as features, was retrieved from the Kaggle website. Python was used in the Jupyter Notebook environment to process the data. The RF ML technique was used to classify and analyse the dataset. ACC, S, and SP percentages were used to express the dataset's results. In predicting HD, the RF method attains an ACC of 86.9%, a S of 90.6%, and an SP of 82.7%. Based on the receiver operating characteristics (ROC), the diagnosis rate for HD prediction using RF was found to be 93.3%. The RF algorithm was used in the suggested system. Thus, it is considered to be the most effective technique for HD classification.

A unique ML method for HD prediction was put forth by Kavitha et al. [27].In addition to DM techniques like regression and classification, the Cleveland HD dataset was used in the investigation. RF and DT were among the ML techniques used. These methods were used to create a novel ML model. 3 ML algorithms: RF, DT, and a hybrid model (HM) that included DT and RF, were employed throughout implementation. From the outcomes of the simulation, the HM's ACC level for HD prediction was 88.7%. In order to predict HD, an interface was also created that used the hybrid model of DT and RF to gather user input parameters.

CardioHelp, a technique developed by Mehmood et al. [28], uses a CNN DL algorithm to estimate the likelihood that a patient would have CVD. The recommended method uses CNN to predict HF at the earliest stage, concentrating on temporal data modelling. Positive findings were obtained when the HD dataset was produced and compared using SOTA methods. In terms of performance evaluation metrics, simulation outcomes demonstrated that the suggested approach executed superior than the current approaches. The suggested approach has a 97% ACC.

For HD classification, an optimised unsupervised method for FS and a unique MLP for Enhanced Brownian Motion based on Dragonfly Algorithm (MLP-EBMDA) were suggested by Deepika and Balaji [29]. Preprocessing was done after obtaining the HD dataset as input. To select features, the best unsupervised approach was applied. Using the selected attributes, HD was categorised using the novel hybrid MLP-EBMDA approach. 94.28% was the outcome of comparing the suggested system's ACC to several other current systems. The study achieves a P of 96% in addition to the recall (R) and F1-score. When compared to different SOTA approaches, the recommended methodology's total performance study revealed better outcomes in terms of HD prediction.

To improve HD prediction performance, Gao et al. [30] suggested using ensemble learning (EL) techniques. Essential features were chosen from the dataset using two feature extraction (FE) techniques: Principal Component Analysis (PCA) and Linear Discriminant Analysis (LDA). Using the chosen features, a comparison between ML algorithms and EL techniques was carried out. ACC, R, P, F-measure, and ROC were among the evaluation metrics that were employed to evaluate the models. The bagging EL approach with a DT produced the best results, according to the outcomes.

Table 1. Comparison of Various HD Prediction Approaches

Author name & Year	Methods	Merits	Demerits
Ahmad and Polat [16]	SVM + Jellyfish Optimization	High accuracy (98.47%), avoids overfitting using feature reduction	Requires advanced optimization knowledge, limited generalizability
Ali et al., [17]	KNN, DT, RF	RF attained 100% ACC, S, and SP	Possible overfitting, lack of real-world validation
El-Sofany et al., [18]	XGBoost, SVM, Bagging, DT, RF + SF feature selection	Mobile app developed; XGBoost achieved 97.57% accuracy; SHAP for explainability	Performance depends on feature subset selection
Biswas et al., [19]	RF, DT, SVM, KNN, LR, NB + Chi2, ANOVA, MI	RF with SF3 gave 94.51% accuracy, greatest ability for medical applications	High complexity in choosing optimal subset and model
Bhatt et al., [20]	RF, DT, MLP, XGB + K-modes clustering	MLP with CV gave 87.28% accuracy; AUC up to 0.95	Slightly lower performance; limited clinical focus
Nashif et al., [21]	SVM in WEKA + Arduino Monitoring System	ACC of 97.53%; RT patient monitoring via GSM	High system complexity; hardware dependence
Saha et al., [22]	RF + CDC Dataset	RF gave 91% accuracy, TPR of 0.922	Limited to CDC data; moderate FNR
Rahman et al., [23]	Transformer + Self-Attention	Achieved 96.51% accuracy; interpretable model	Transformer models are resource-intensive
Mandavkar et al., [24]	GA + Decision Tree	Best for multi-class; integrated into web app	May need constant feature updates
Sharma et al., [25]	RF, SVM, NB, DT	RF showed better performance in less time	No deep learning comparison
Pal and Parija [26]	RF	Accuracy 86.9%, Sensitivity 90.6%, ROC 93.3%	Moderate specificity (82.7%)
Kavitha et al., [27]	RF, DT, Hybrid	Hybrid model gave 88.7% accuracy	Limited interpretability and scalability
Mehmood et al., [28]	CNN (CardioHelp)	Temporal modeling; 97% accuracy	High computational cost
Deepika and Balaji [29]	MLP-EBMDA + Unsupervised Feature Selection	Accuracy 94.28%; precision, recall, F1 all 96%	Novel method, needs more validation
Gao et al.,[30]	Bagging + LDA + PCA	Best performance using ensemble with DT	Requires complex preprocessing steps

2. Research Gap

Despite the extensive application of ML and DL techniques for HD prediction, several research gaps remain unaddressed. Most existing studies focus on achieving high ACC using specific datasets such as Cleveland or Kaggle, which limits the generalizability and clinical robustness of the models. Additionally, few studies explore RT, MM data integration from wearable devices or EHR, which is critical for practical deployment. FS methods vary widely, but there is a lack of standardized frameworks to evaluate their clinical relevance. While explainable AI approaches like SHAP have been introduced, their integration into clinical DSS remains limited. Moreover, hybrid models and ensemble methods show promise, but their complexity and interpretability present challenges for adoption in low-resource or point-of-care environments. Overall, there is a need for scalable, interpretable, and cross-population validated models that can support RT decision-making (DMak) in diverse healthcare settings.

3. Overview of HD Prediction

HF, arrhythmias, congenital heart abnormalities, coronary artery disease, and other disorders that impact the heart and BV are all included in HD, also known as CVD. According to the WHO, it continues to be the primary cause of illness and mortality globally, contributing to around 17.9 million deaths annually. In order to lower the death rate and enhance the quality of life for those who are impacted, ED and intervention are essential. Despite their clinical reliability, traditional diagnostic methods including ECG, echocardiography, angiography, and blood tests are frequently costly, time-consuming, and difficult to interpret [31]. For the purpose of resolving those issues, data-driven PM is often utilized nowadays. Then, with the support of the ML and DL methods, these data-driven PM may enhance in predicting HD.

The patients with a risk of developing or already having a CVD is examined by considering some clinical and demographics features like age, gender, BP, cholesterol, body mass index (BMI), blood sugar (BS), physical activity, smoking and alcohol use, family history, and ECG readings. By considering those features, one can effectively predict HD. There are 2 classes in the predictive task: binary classification (BC) problem (e.g., disease/no disease) or a MC classification problem (e.g., severity levels or types of HD). Here, the complicated and NL relationships are effectively identified by the integration of AI into this PM, and it may also facilitate high-dimensional (H-D) data analysis (DA). But, the conventional methods fail to detect them.

In HD prediction tasks, a robust efficiency was attained by the ML models such as DT, RF, SVM, LR, NB, and ensemble techniques like Gradient Boosting and XGB. Then, the DL models like ANN, CNN, and RNN are also used when managing time-series (TS) data or imaging inputs like ECG or echocardiogram scans. FS and dimensionality reduction (DR) techniques like PCA, CS, ANOVA, MI, and optimisation algorithms like GA and Particle Swarm Optimisation (PSO) are all utilized by many studies for the purpose of enhancing the efficiency and ACC of the PM.

The efficacy of the model is determined by the performance metrics like ACC, P, R, F1-score, S, SP, and the AUC-ROC. Some sophisticated models also integrate interpretability tools such as SHAP (SHapley Additive exPlanations) or LIME (Local Interpretable Model-Agnostic Explanations) are also integrated, and it may help the model in making DM, and improving transparency and confidence in AI-assisted clinical diagnostics.

Despite its promising advancements, there are many risks that exist in HD prediction. This widespread application is also limited by the challenges like absence of standardisation across datasets, class imbalance, noisy or missing data, and small sample sizes. The static datasets are mostly utilized by most of the current methods, so they lack RT adaptability or integration with wearable devices and EHR. Future systems must focus on addressing these gaps through FL, RT analytics, and explainable AI to develop comprehensive, interpretable, and deployable solutions for HD prediction in diverse clinical settings.

4. Dataset Used in HD Prediction

The quality and richness of the datasets used for training and evaluation have a significant impact on how well ML and DL models predict HD. Various public and private datasets have been employed in research, with the most commonly used ones sourced from institutions like the UCI ML Repository, Kaggle, and healthcare organizations. Clinical parameters that are known to correlate with cardiovascular risk, including as age, sex, cholesterol levels, resting BP, ECG findings, maximal heart rate, exercise-induced angina, fasting BS, and kind of chest pain, are commonly included in these datasets.

Some datasets are binary-class (presence or absence of HD), while others contain multi-class labels indicating severity or specific diagnoses. They differ in size, feature count, and completeness. Many researchers preprocess these datasets through techniques like normalization, imputation of missing values (MV), and FS to enhance model performance.

Below is a table summarizing commonly used datasets for HD prediction:

Dataset Name	Source	No. of Samples	No. of Features	Classes	Short Description
Cleveland HD	UCI Machine Learning Repository	303	14	5 (0–4; often binary)	Most widely used dataset for HD prediction. Contains numeric and categorical attributes.
Statlog Heart Dataset	UCI Repository	270	13	2	Derived from the Cleveland dataset. Used for binary classification (presence/absence).
Hungarian Heart Disease	UCI Repository	294	13	5	Similar to Cleveland dataset but less complete; often used after merging with other datasets.
Switzerland Dataset	UCI Repository	123	13	5	Smaller dataset with missing values; used in multi-center analysis.
Long Beach VA Dataset	UCI Repository	200	13	5	Part of the full heart disease dataset group from UCI; used for comparative analysis.
Framingham Heart Study	National Institutes of Health (NIH)	~4,000	15+	Binary (CVD/No CVD)	Longitudinal study with detailed patient history; useful for risk prediction models.
Heart Failure Dataset	UCI Repository	299	13	Binary (Yes/No)	Focused on heart failure rather than general CVD; includes ejection fraction and serum creatinine.
Kaggle Heart Disease	Kaggle	1,000+ (varies)	14+ (varies)	Binary or multi-class	Aggregated from multiple sources; widely used in competitions and student projects.
CDC Heart Disease Data	CDC BRFSS Survey Dataset	400,000+	18+	Binary	Large-scale survey data; includes behavioral and clinical features.

5. ML Methods for HD Prediction

Clinical parameters that are known to correlate with cardiovascular risk, including as age, sex, cholesterol levels, resting BP, ECG findings, maximal heart rate, exercise-induced angina, fasting BS, and kind of chest pain, are commonly included in these datasets. In HD prediction, ML methods are particularly valuable for risk stratification, identifying at-risk patients, and aiding clinicians in DM processes. These algorithms work by training on historical patient data, learning relationships between input features (e.g., BP, cholesterol levels, ECG outcomes) and the target variable (e.g., existence or nonexistence of HD), and applying the learned patterns to predict outcomes for new cases.

5.1. Supervised Learning (SL) Algorithms

The most popular application of SL algorithms is HD prediction. It consist of:

a. LR

One of the simplest but successful BC models is LR. The probability of a categorical dependent variable is computed by this LR. It is used to model the probability of illness present in HD prediction. For small datasets, it works well and is interpretable, even though it might miss intricate patterns.

b. DT

DT divides data according to feature thresholds using a framework resembling a flowchart. They are helpful for comprehending the DMak process and are quite interpretable. Overfitting occurs frequently with DT, especially when working with noisy data.

c. RF

RF is a DT ensemble that decreases overfitting and increases prediction ACC. It can handle big feature sets and does well in HD prediction tasks. Additionally, it offers feature importance rankings, which may be helpful in medical settings.

d. SVM

Finding the hyperplane (HP) that best divides classes is how SVMs classify data. When the quantity of features exceeds the quantity of samples, they function effectively in H-D spaces. H-D datasets with intricate associations can benefit from SVMs' ability to handle NL data via kernel functions.

e. KNN

A data point is classified by KNN according to the majority class of its KNN. It doesn't require any assumptions regarding the distribution of the data and is simple and non-parametric. KNN needs the value of k to be carefully chosen because it is sensitive to irrelevant features.

f. NB

The Bayes theorem is applied by NB classifiers, which presume feature independence. When features are conditionally independent, NB's performance in many HD prediction tasks is surprisingly good, despite its simplicity.

g. Gradient Boosting Procedures (e.g., XGB, LightGBM)

These models successively construct an ensemble of weak learners, typically DT. Every new model fixes errors in the ones that came before it. Because of its ACC, speed, and resilience in HD classification tasks, XGBoost in particular has gained popularity.

5.2. Unsupervised Learning in FS

In order to reduce dimensionality, visualise trends, or find clusters in patient data, preprocessing frequently employs unsupervised approaches such as PCA, k-means clustering (KMC), and t-SNE. They are not utilised directly for classification, but by removing superfluous or irrelevant features, they aid supervised models in performing better.

5.3. EL Methods

More predictive performance is achieved by combining many models using EL techniques than by using any single model. Methods including stacking, boosting, and bagging are commonly employed.

- **Bagging (Bootstrap Aggregating):** One well-known example is RF. By training several models on various data subsets, it lowers variance.
- **Boosting:** Models like AdaBoost and XGBoost sequentially train weak learners and focus on correcting previous errors, yielding high accuracy.
- **Stacking**: Combines multiple classifiers using a meta-classifier for improved performance. It is often used in competitive modeling environments like Kaggle.

6. DL Techniques for HD Prediction

A subset of ML is called DL. Due to its capacity to extract hierarchical representations from complex data, DL has become an essential tool in healthcare analytics [33]. DL models automatically extract pertinent patterns and features from raw input, in contrast to standard ML models that mostly rely on created features. These are especially appropriate for HD, time-series (TS), and MM data that are commonly used in cardiovascular diagnostics because of the DL.

In HD prediction, DL methods identifies better efficiency in detecting intricate patterns and associations within clinical records, physiological signals (e.g., ECG), Medical imaging (Med Img), and wearable device data. The following are the most widely used DL techniques in this domain:

6.1. ANN

ANN provide the basis of DL. They are made up of layers of interconnected nodes, or neurons. For predicting HD:

- ANNs are commonly applied to structured datasets (e.g., Cleveland dataset).
- They can model NL relationships between risk factors and outcomes.
- They achieve higher accuracy than shallow machine learning models, especially with sufficient data and tuning.

6.2. CNN

CNNs work very well with spatial data. CNNs are frequently employed in TS and image analysis:

- In HD prediction, CNNs are used to analyze ECG signals, CT angiograms, and cardiac MRI.
- CNNs can capture local temporal/spatial dependencies, creating it effective in identifying subtle patterns in ECG waveforms or imaging modalities.
- Some studies also apply CNNs to tabular data by converting features into 2D matrices.

6.3. RNN and LSTM

By keeping track of prior inputs, RNNs are made to process sequential data. Their enhanced version that can recognise long-term dependencies is called the LSTM.

- In HD, they are used for analyzing TS data like continuous ECG, BP trends, or wearable device outputs.
- LSTM-based models are ideal for predicting cardiac events from RT or longitudinal data.

6.4. Autoencoders (AE)

Autoencoders are unsupervised DL models used for dimensionality reduction (DR), anomaly detection (AD), and feature learning:

- In HD detection, they are used to compress patient data while retaining relevant clinical patterns.
- Variational AE (VAEs) can generate synthetic data to augment small HD datasets.

6.5. Deep Belief Networks (DBNs)

Multiple layers of Restricted Boltzmann Machines (RBMs) make up DBNs. For unsupervised pretraining, DBN is utilised:

- In cardiovascular studies, they have been applied to model latent structures in patient data.
- DBNs are beneficial when labeled data is limited but high-level (HL) abstraction is needed.

6.6. Hybrid Deep Learning Models

Many studies propose hybrid architectures by combining:

- **CNN-LSTM**: For ECG signal classification.
- **ANN-RF or CNN-SVM**: To combine the feature extraction strength of DL with the robustness of ML classifiers.
- **Transformer-based models**: For explainable predictions, using attention mechanisms to highlight important clinical features.

6.7. Explainable DL (XDL)

Clinical applications require interpretability:

- Methods like **SHAP**, **LIME**, and **attention mechanisms** are integrated into DL models.
- These approaches help clinicians understand which features influenced the prediction, improving trust and transparency.

Applications and Use Cases

- **ECG Analysis**: CNNs and LSTMs for arrhythmia detection and ischemia prediction.
- **EHR Mining**: ANNs and autoencoders for predicting heart failure or myocardial infarction.
- **Medical Imaging**: CNNs for segmentation and diagnosis from echocardiograms and MRI.
- **Wearable Data**: LSTM models for real-time risk assessment based on continuous monitoring.

DL methods are revolutionizing HD prediction through superior ACC and automated feature learning. Their application spans structured datasets, imaging, and sensor-based inputs. However, risks such as data scarcity, model interpretability, and clinical integration must be addressed. Future research should focus on building lightweight, explainable, and federated DL models for widespread clinical adoption.

7. FS Techniques in HD Prediction

In the medical field, where datasets frequently contain HD and possibly redundant information, FS is an essential pre-processing step in creating successful ML and DL models [34]. By choosing the most pertinent features, HD prediction can reduce computational costs, improve interpretability, decrease overfitting, and increase model ACC. Building effective models requires determining which of the factors that are commonly included in HD datasets such as age, gender, BP, cholesterol levels, ECG results, and more contribute significantly to the prediction task.

Filter methods, wrapper methods, and embedded methods are the three primary categories into which FS techniques are typically divided. For more sophisticated FS, hybrid strategies and optimization-based methods [35] have also gained popularity recently.

1. Filter Approaches

Regardless of any ML model, filter methods rate each feature using statistical measures and choose the highest-ranked ones. These techniques scale to HD data and are quick and easy to use.

- **CS Test (CST)**: Calculates the dependence between a feature and the target variable. Suitable for categorical data, it ranks features based on their CS statistics.
- **ANOVA (Analysis of Variance)**: Measures the variance between different groups of a feature with respect to the target class. Often used for continuous variables.
- **MI**: Calculates how much information a single feature adds to the target variable's prediction. Both classification and regression tasks benefit from its utilisation.
- **Pearson Correlation Coefficient**: It determines linear connections between the target and continuous features. Features that are highly connected might be eliminated if they are redundant.

2. Wrapper Methods

Wrapper approaches really train a model and measure its performance in order to evaluate subsets of features. They are computationally costly, but they take feature dependencies into account.

- **Recursive Feature Elimination (RFE)**: begins with each attribute and, using model coefficients, recursively eliminates the least significant ones until the ideal subset is discovered.
- **Forward FS**: adds features one at a time, starting with none and retaining those that enhance the model's functionality.
- **Backward Elimination**: At each stage, the least important feature is eliminated after starting with all of the features.
- **Exhaustive Search**: Evaluates all possible feature combinations (practical only for small feature sets).

3. Embedded Approaches

FS is incorporated into the model training procedure by embedded approaches [35]. These techniques frequently produce excellent results and need less computational power than wrapper techniques.

- **LASSO (Least Absolute Shrinkage and Selection Operator)**: A regression technique that adds a penalty for feature coefficients, shrinking less important ones to zero.
- **Ridge and ElasticNet Regression**: Similar to LASSO, but with different regularization strategies to handle multicollinearity.
- **Tree-Based Models (RF, XGBoost)**: Offers FIS during model training. Features contributing less to prediction accuracy can be removed.
- **Regularized LR**: Helps in identifying influential predictors through coefficient magnitude.

4. Optimization-Based and Metaheuristic Methods

To handle HD and complex datasets, researchers have turned to metaheuristic algorithms (MHA) for FS [36]. These approaches search for the optimal subset of features using biologically or physically inspired techniques.

- **GA**: Use natural selection principles to evolve the best feature subset across generations.
- **PSO**: Replicate social behavior of birds/fish to explore the feature space.
- **Ant Colony Optimization (ACO)**, **Grey Wolf Optimizer (GWO)**, **Dragonfly Algorithm (DA)**, and **JFOA**: Are also used for intelligent exploration and exploitation of feature space.

5. Hybrid Approaches

Hybrid FS combines multiple techniques (e.g., filter + wrapper or filter + optimization) to leverage their advantages and mitigate individual restrictions [37]. For example, initial filtering may reduce dimensionality, and then wrapper or optimization techniques can fine-tune the subset.

Example in Practice:

- A study might first apply MI to eliminate irrelevant features and then use GA or RFE for final selection.
- DL-based methods may use AE or attention layers for automatic feature learning.

In order to create reliable and understandable HD prediction models, FS is essential. The size of the dataset, the type of features, and the available computational resources all influence the technique selection. While filter methods offer speed and simplicity, wrapper and embedded methods yield better predictive ACC. Optimization algorithms, though computationally heavy, are excellent for fine-grained, HD problems. Effective FS not only expands efficiency of the frameworks but also aligns better with clinical relevance, enabling practitioners to focus on the most significant health signs.

8. Evaluation Metrics for HD Prediction

Particularly in medical applications like HD prediction, evaluation metrics are essential for evaluating how well ML and DL models work. In clinical settings, achieving high ACC is not sufficient; it is equally important to minimize FN (missed diagnoses) and FP (false alarms), as both can lead to serious health and economic consequences. To give a comprehensive outline of the framework's efficacy in many areas, a wide range of assessment metrics are employed.

1. ACC

ACC is a measure of the proportion of correctly predicted cases (both positive and negative) to all cases.

$$ACC = \frac{TP + TN}{TP + TN + FP + FN}$$

Here,

TP= True Positive (accurately predicted disease case)

TN= True Negative (accurately predicted non-diseases case)

FP= False Positive (non -disease cases inaccurately predicted as disease)

FN= False Negative (disease cases missed by the model)

2. P (Positive Predictive Value (PPV))

P is the proportion of TP predictions among all the positive predictions the model generates.

$$P = \frac{TP}{TP + FP}$$

3. R (S or TP Rate)

The percentage of TP that were accurately predicted is measured by R.

$$R = \frac{TP}{TP + FN}$$

4. SP (TN Rate)

The percentage of TN that the model accurately recognised is measured by SP.

$$SP = \frac{TN}{TN + FP}$$

5. F1 Score

The harmonic mean of P and R is the F1 Score. It gives the two metrics balance.

$$F1\ Score = 2 \times \frac{P\ \times R}{P\ + R}$$

6. AUC-ROC

Plotting the TP rate (R) against the FP rate (1-SP) at different threshold values is done by the ROC. This plot has been simplified into a single number by the AUC.

AUC∈ [0.1]. Here, 0.5 represents random guessing and perfect prediction are denoted by 1.

7. Confusion Matrix (CM)

A tabular illustration of actual vs predicted outcomes is called a CM.

	Forecast Positive	Forecast Negative
Actual Positive	TP	FN
Actual Negative	FP	TN

8. Log Loss LL (Logarithmic Loss)

The prediction uncertainty is measured by LL. High-confidence incorrect classifications are penalised more severely than low-confidence ones.

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

Here:

- For class 1, p_i is the predicted probability and y_i is the actual label.

Accurately evaluating HD PM's performance requires choosing the right evaluation metric. Metrics like P, R, F1 score, and AUC-ROC are more relevant in clinical settings or imbalanced datasets when the cost of misclassification is significant, even though ACC provides an overview of performance. A complete understanding of the model's advantages and disadvantages can be obtained by combining these metrics with the CM. It makes it easier to create diagnostic systems that are dependable and understandable.

9. Comparative Analysis of Models for HD Prediction

The use of different ML and DL algorithms has led to notable progress in the field of HD prediction. Regarding ACC, computational effectiveness, interpretability, and appropriateness for various data types (structured data, TS signals, or imaging data), each model offers unique advantages. To determine which approaches are best suited for particular use cases in clinical DSS, a comparative study of various models is essential.

This section presents a comparative overview of commonly used models based on multiple studies, highlighting their performance metrics, merits, and limitations.

1. Comparison Criteria

The frameworks are evaluated based on the following:

- **ACC**: The overall correct prediction rate.
- **S/R**: Ability to identify true HD cases.
- **Specificity**: Ability to correctly identify healthy cases.
- **Precision**: Accuracy of positive predictions.
- **F1 Score**: Balance between P and R.
- **AUC-ROC**: Frameworks ability to discriminate between classes.
- **Interpretability**: Ease of understanding model decisions.
- **Computational Cost**: Time and resources required.

Table 2. Comparative analysis of the existing model

Model	ACC	F1 Score	AUC-ROC	Merits	Demerits
LR	80-85%	0.78-0.83	0.85-0.89	Interpretable, fast, suitable for binary classification	Linear, may underfit complex relationships
DT	82-86%	0.80-0.85	0.87-0.91	Easy to visualize and interpret	Prone to overfitting unless pruned
RF	86-94%	0.88-0.92	0.90-0.95	High accuracy, handles feature interactions, reduces overfitting	Less interpretable, slower than simpler models
SVM	84-93%	0.86-0.91	0.89-0.94	Works well with high-dimensional data	Requires kernel tuning, less interpretable
KNN	80-88%	0.82-0.87	0.85-0.90	Simple, non-parametric	Sensitive to data scale and irrelevant features
NB	78-84%	0.76-0.82	0.84-0.88	Fast, good with small datasets	Assumes feature independence, may oversimplify
XGBoost	88-97%	0.90-0.94	0.94-0.98	Excellent performance, handles missing data, regularized boosting	High complexity, longer training time
ANN (Multilayer Perceptron)	85-92%	0.87-0.92	0.90-0.95	Learns complex non-linear relationships	Requires large data, less interpretable
CNN (for ECG/images)	90-97%	0.91-0.95	0.94-0.97	Ideal for spatial/time-series data	Needs large datasets, computationally expensive
LSTM (for sequential data)	89-96%	0.90-0.94	0.92-0.96	Effective for time-series ECG/wearable data	Longer training time, model complexity
Hybrid Models (e.g., CNN-LSTM, RF-GA)	91-98%	0.92-0.96	0.95-0.98	Combines strengths of individual models, often highest accuracy	High computational cost, requires careful integration
Transformer-Based Models	94-97%	0.94-0.96	0.96-0.98	High accuracy and interpretability via attention weights	Requires large computational resources, not yet widely deployed in clinics

3. Observations and Insights

- **RF and XGBoost** consistently overtake other ML frameworks by ACC and robustness, especially for structured tabular datasets like Cleveland or Kaggle heart disease data.
- **DL models** like CNNs and LSTMs are more effective for ECG signal and Med Img information, offering better performance on complex patterns but requiring larger datasets and higher computational resources.
- **Hybrid models**, particularly those combining DL with feature optimization (e.g., RF + Genetic Algorithm, CNN + LSTM), tend to provide the best results but at the cost of increased complexity.
- **Simple models** like LR and NB still hold value in primary care or low-resource environments due to their speed and interpretability.
- **Transformer models**, while relatively new in this domain, show exceptional potential by leveraging attention mechanisms, offering high accuracy and interpretability.

For HD prediction, there is no one-size-fits-all model. The particular application context, the kind of data accessible, the requirement for interpretability, and the computational resources at hand should all be taken into consideration when selecting a model. While ensemble and DL methods provide high predictive ACC, simpler models remain essential for rapid, interpretable decision support. Future work should focus on developing models that balance ACC, explain ability, and RT deployment feasibility, especially for integration into smart healthcare systems.

10. Challenges and Limitations

Even with the quick development of ML and DL methods for HD prediction, a number of obstacles still stand in the way of their broad use in actual medical environments:

1. **Data Quality and Availability**: Many HD datasets suffer from issues such as MV, noise, inconsistent formats, and class imbalance. Furthermore, privacy regulations and fragmented healthcare systems restrict access to sizable, varied, and excellent annotated datasets.
2. **Imbalanced Datasets**: Most clinical datasets are skewed, with significantly fewer positive (heart disease) cases compared to negative ones. This imbalance biases the model towards the majority class, reducing S and increasing FN.
3. **Overfitting and Generalization**: Models trained on small or homogeneous datasets often overfit and fail to generalize to different populations, geographic regions, or clinical settings.
4. **Interpretability**: Black boxes are used by many high-performing models, particularly DL models.This limits trust and acceptability by medical professionals.
5. **Integration with Clinical Workflow**: Few models are designed with interoperability standards, making it difficult to incorporate with EHR or hospital management systems.
6. **RT Prediction Constraints**: RT inference may be challenging on edge devices or in environments with limited resources because to the high processing requirements of DL models.
7. **Regulatory and Ethical Challenges**: Ensuring compliance with clinical procedures (e.g., HIPAA, GDPR) and maintaining patient data privacy while training and deploying models remain critical hurdles.

11. Research Gaps and Future Directions

Despite significant progress, the field still faces numerous unexplored areas and opportunities for innovation:

1. **Explainable AI (XAI)**: There is a need to integrate explainability mechanisms such as SHAP, LIME, and attention mechanisms into DL models to enhance clinician trust and regulatory transparency.
2. **Federated and Privacy-Preserving Learning**: FL frameworks should be investigated in future studies in order to train models across several hospitals without exchanging private patient information.
3. **MM Data Fusion**: Combining structured data (e.g., EHR), unstructured text (e.g., clinical notes), signal data (e.g., ECG), and image data (e.g., echocardiograms) can lead to more comprehensive and accurate models.
4. **Edge and Lightweight AI Models**: Developing lightweight, energy-efficient (EE) models suitable for deployment on mobile devices or wearable health monitors can support early detection in remote or low-resource settings.
5. **Benchmarking and Standardization**: There is a lack of standardized benchmarks and protocols for evaluating models across datasets and regions. Future work should focus on creating open, standardized benchmarks for fair comparison.
6. **Clinical Trials and Real-World Validation**: Most models are validated only on retrospective datasets. Prospective validation through medical test is crucial for evaluating real-world efficacy, safety, and usability.
7. **Adaptive and Continuous Learning**: Future systems should incorporate adaptive learning capabilities to evolve with new data and emerging patterns in heart disease diagnosis.

CONCLUSION

ML and DL have demonstrated significant ability in modernizing HD prediction, offering faster, more accurate, and cost-effective diagnostic support. Numerous models, including ensemble learners, NN, and hybrid architectures, have shown high predictive performance across benchmark datasets. However, challenges such as limited interpretability, imbalanced datasets, and integration difficulties with clinical systems continue to hinder real-world adoption. Addressing these issues through explainable AI, robust multimodal data integration, and privacy-preserving learning will be essential for future success. By focusing on generalizability, clinical interpretability, and RT applicability, next-generation AI systems can transform preventive cardiology and support early intervention, ultimately improving patient outcomes and saving lives.

REFERENCES

- Kundu, J. and Kundu, S., 2022. Cardiovascular disease (CVD) and its associated risk factors among older adults in India: Evidence from LASI Wave 1. *Clinical Epidemiology and Global Health*, *13*, p.100937.
- Wu, H., Patel, K.H.K., Li, X., Zhang, B., Galazis, C., Bajaj, N., Sau, A., Shi, X., Sun, L., Tao, Y. and Al-Qaysi, H., 2022. A fully-automated paper ECG digitisation algorithm using deep learning. *Scientific Reports*, *12*(1), p.20963.
- Ghorbani, A., Ouyang, D., Abid, A., He, B., Chen, J.H., Harrington, R.A., Liang, D.H., Ashley, E.A. and Zou, J.Y., 2020. Deep learning interpretation of echocardiograms. *NPJ digital medicine*, *3*(1), p.10.
- Gurav, A., Revaiah, P.C., Tsai, T.Y., Miyashita, K., Tobe, A., Oshima, A., Sevestre, E., Garg, S., Aben, J.P., Reiber, J.H. and Morel, M.A., 2024. Coronary angiography: a review of the state of the art and the evolution of angiography in cardio therapeutics. *Frontiers in Cardiovascular Medicine*, *11*, p.1468888.
- Bharti, R., Khamparia, A., Shabaz, M., Dhiman, G., Pande, S. and Singh, P., 2021. Prediction of heart disease using a combination of machine learning and deep learning. *Computational intelligence and neuroscience*, *2021*(1), p.8387680.
- Saboor, A., Usman, M., Ali, S., Samad, A., Abrar, M.F. and Ullah, N., 2022. A method for improving prediction of human heart disease using machine learning algorithms. *Mobile Information Systems*, *2022*(1), p.1410169.
- Barhoom, A.M., Almasri, A., Abu-Nasser, B.S. and Abu-Naser, S.S., 2022. Prediction of heart disease using a collection of machine and deep learning algorithms.
- Kumaragurubaran, T., Jaeyalakshmi, M. and Sheriff, M.A., 2024, January. A Comparative Analysis on the Prediction of Heart Failure using Machine Learning Algorithms. In *2024 5th International Conference on Mobile Computing and Sustainable Informatics (ICMCSI)* (pp. 206-211). IEEE.
- Saikumar, K., Rajesh, V. and Babu, B.S., 2022. Heart disease detection based on feature fusion technique with augmented classification using deep learning technology. *Traitement du Signal*, *39*(1), p.31.
- Bebortta, S., Tripathy, S.S., Basheer, S. and Chowdhary, C.L., 2023. Fedehr: A federated learning approach towards the prediction of heart diseases in iot-based electronic health records. *Diagnostics*, *13*(20), p.3166.
- Sami, O., Elsheikh, Y. and Almasalha, F., 2021. The role of data pre-processing techniques in improving machine learning accuracy for predicting coronary heart disease. *International Journal of Advanced Computer Science and Applications*, *12*(6).
- Srinivasan, S., Gunasekaran, S., Mathivanan, S.K., M. B, B.A.M., Jayagopal, P. and Dalu, G.T., 2023. An active learning machine technique based prediction of cardiovascular heart disease from UCI-repository database. *Scientific reports*, *13*(1), p.13588.
- Masih, N. and Ahuja, S., 2022. Application of data mining techniques for early detection of heart diseases using Framingham heart study dataset. *International journal of biomedical engineering and technology*, *38*(4), pp.334-344.

- Bhavekar, G.S. and Goswami, A.D., 2022. A hybrid model for heart disease prediction using recurrent neural network and long short term memory. *International Journal of Information Technology*, *14*(4), pp.1781-1789.
- Ahmed, M. and Husien, I., 2024. Heart disease prediction using hybrid machine learning: A brief review. *Journal of Robotics and Control (JRC)*, *5*(3), pp.884-892.

Literature review

- Ahmad, A.A. and Polat, H., 2023. Prediction of heart disease based on machine learning using jellyfish optimization algorithm. *Diagnostics*, *13*(14), p.2392.
- Ali, M.M., Paul, B.K., Ahmed, K., Bui, F.M., Quinn, J.M. and Moni, M.A., 2021. Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison. *Computers in Biology and Medicine*, *136*, p.104672.
- El-Sofany, H., Bouallegue, B. and El-Latif, Y.M.A., 2024. A proposed technique for predicting heart disease using machine learning algorithms and an explainable AI method. *Scientific Reports*, *14*(1), p.23277.
- Biswas, N., Ali, M.M., Rahaman, M.A., Islam, M., Mia, M.R., Azam, S., Ahmed, K., Bui, F.M., Al-Zahrani, F.A. and Moni, M.A., 2023. Machine Learning-Based Model to Predict Heart Disease in Early Stage Employing Different Feature Selection Techniques. *BioMed Research International*, *2023*(1), p.6864343.
- Bhatt, C.M., Patel, P., Ghetia, T. and Mazzeo, P.L., 2023. Effective heart disease prediction using machine learning techniques. *Algorithms*, *16*(2), p.88.
- Nashif, S., Raihan, M.R., Islam, M.R. and 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*(4), pp.854-873.
- Saha, S., Rahman, M.M., Suki, T.T., Alam, M.M., Alam, M.S. and Haque, M.A.S., 2024, April. Heart disease prediction using machine learning algorithms: Performance analysis. In *2024 3rd International Conference on Advancement in Electrical and Electronic Engineering (ICAEEE)* (pp. 1-6). IEEE.
- Rahman, A.U., Alsenani, Y., Zafar, A., Ullah, K., Rabie, K. and Shongwe, T., 2024. Enhancing heart disease prediction using a self-attention-based transformer model. *Scientific Reports*, *14*(1), p.514.
- Mandavkar, S., Shinde, P. and Pinjarkar, V., 2021, May. Heart diseases prediction system using classification and genetic algorithm. In *Proceedings of the 4th International Conference on Advances in Science & Technology (ICAST2021)*.
- Sharma, V., Yadav, S. and Gupta, M., 2020, December. Heart disease prediction using machine learning techniques. In *2020 2nd international conference on advances in computing, communication control and networking (ICACCCN)* (pp. 177-181). IEEE.
- Pal, M. and Parija, S., 2021, March. Prediction of heart diseases using random forest. In *Journal of Physics:* Conference *Series* (Vol. 1817, No. 1, p. 012009). IOP Publishing.
- Kavitha, M., Gnaneswar, G., Dinesh, R., Sai, Y.R. and Suraj, R.S., 2021, January. Heart disease prediction using hybrid machine learning model. In *2021 6th international conference on inventive computation technologies (ICICT)* (pp. 1329-1333). IEEE.
- Mehmood, A., Iqbal, M., Mehmood, Z., Irtaza, A., Nawaz, M., Nazir, T. and Masood, M., 2021. Prediction of heart disease using deep convolutional neural networks. *Arabian Journal for Science and Engineering*, *46*(4), pp.3409-3422.
- Deepika, D. and Balaji, N., 2022. Effective heart disease prediction using novel MLP-EBMDA approach. *Biomedical Signal Processing and Control*, *72*, p.103318.
- Gao, X.Y., Amin Ali, A., Shaban Hassan, H. and Anwar, E.M., 2021. Improving the accuracy for analyzing heart diseases prediction based on the ensemble method. *Complexity*, *2021*(1), p.6663455.
- Ahsan, M.M. and Siddique, Z., 2022. Machine learning-based heart disease diagnosis: A systematic literature review. *Artificial Intelligence in Medicine*, *128*, p.102289.
- Jindal, H., Agrawal, S., Khera, R., Jain, R. and Nagrath, P., 2021. Heart disease prediction using machine learning algorithms. In *IOP conference series: materials science and engineering* (Vol. 1022, No. 1, p. 012072). IOP Publishing.
- Badawy, M., Ramadan, N. and Hefny, H.A., 2023. Healthcare predictive analytics using machine learning and deep learning techniques: a survey. *Journal of Electrical Systems and Information Technology*, *10*(1), p.40.
  Dissanayake, K. and Md Johar, M.G., 2021. Comparative study on heart disease prediction using feature selection techniques on classification algorithms. *Applied Computational Intelligence and Soft Computing*, *2021*(1), p.5581806.
- Manigandan, A., & Saranya, M. Secretary Bird Optimization with Differential Evolution (SBODE) and Trust Energy Aware Clustering Routing (TREACR) Protocol for Wireless Sensor Network (WSN). *International Journal of Computer Networks and Applications (IJCNA), 12(2), pp.291-306.*
- Mhamdi, L., Dammak, O., Cottin, F. and Dhaou, I.B., 2022. Artificial intelligence for cardiac diseases diagnosis and prediction using ECG images on embedded systems. *Biomedicines*, *10*(8), p.2013.
- Ay, Ş., Ekinci, E. and Garip, Z., 2023. A comparative analysis of meta-heuristic optimization algorithms for feature selection on ML-based classification of heart-related diseases. *The Journal of Supercomputing*, *79*(11), pp.11797-11826.
- Pavithra, V. and Jayalakshmi, V., 2023. Hybrid feature selection technique for prediction of cardiovascular diseases. *Materials today: proceedings*, *81*, pp.336-340.