Ensemble Learning for Heart Disease Prediction: a Review

Rasha Alquhali

September 30, 2024

Date of publication xxxx 00, 0000, date of current version xxxx 00, 0000.

Digital Object Identifier 10.1109/ACCESS.2024.Doi Number

Ensemble Learning for Heart Disease Prediction: A Review

RASHA ALQUHALI

Faculty of Engineering and Information Technology,Amran University,Amran,Yemen

Corresponding (alquhalirasha@gmail.com).

ABSTRACT Humanity has been affected by various diseases throughout history, which have killed many lives. One of the deadliest diseases that humanity has seen in the modern age and is still acknowledged today is heart disease. Heart disease is on the rise as a result of the spread of unhealthy behaviors including smoking, overeating, and inactivity. This paper examined the machine learning (ML), deep learning (DL), and ensemble learning methods (ELMs) utilized in heart disease prediction research, as well as how they are being implemented. Searches were carried out on the Google Scholar online datasets. Sixty-five studies were included, with ML methods making up most of the studies with 28 (43%), and ELMs were the next single largest group with 24 (37%). DL methods were the smallest single group with 13 (20%). The Cleveland dataset was used in most studies. The result shows that over the last 5 years, there has been a growing desire of leveraging ML, and DL techniques to help further the understanding of heart disease prediction, whether it be by expanding the knowledge of the physiological changes or by improving the accuracies of models to help improve the treatments and disease management.

INDEX TERMS Review**,** machine learning, deep learning, ensemble learning, heart disease prediction.

I. INTRODUCTION

Heart disease has been the main cause of death worldwide over the last decade. The world health organization (WHO) estimated that around 23.6 million people die largely from cardiovascular diseases each year, with coronary artery disease and brain stroke accounting for 82 percent of these deaths. Such reasons include elevated blood pressure, high cholesterol, diabetes, obesity, smoking, and a history of heart disease in the family[1].

Modern technology, including robotics, computers, and mobile phones, as well as the field of health care, nearly everywhere uses machine learning (ML) (i.e., disease diagnosis, safety). ML is becoming more and more popular in a wide range of sectors, including healthcare and disease diagnosis[2]. It is a method that aids the system in picking up knowledge from earlier data samples. In many fields, ML is essential. It also demonstrates how it affects the prediction of heart disease[3].Deep Learning (DL) is a component of artificial intelligence (AI), which is also a subset of ML. Also, it is an increasingly common ML method [4]. Numerous more study fields can benefit from the use of DL. It is used to predict heart disease as well[3].

The ensemble is a technique that is used to improve the classifier's accuracy. It is combining weak classifiers with strong learners to boost the effectiveness of the weak classifiers. So, merging different classifiers is getting improved performance over each classifier working alone[5].

This study highlights ML and DL methods that are used in heart disease prediction (HDP). It started by outlining several ML and DL methods that are used in common recent studies

This study is to provide insights to recent and future researchers and practitioners regarding ML and DL-based heart disease prediction (MLDL-B-HDP) that will aid and enable them to choose the most appropriate and superior machine learning/deep learning methods. Additionally, it aims to identify potential studies related to the MLDL-B-HDP. In general, the scope of this study is to provide the proper explanation for the following questions:

Which ML-DLHDP datasets are the most widely used?

Which ML and DL approaches are presently used in health care to classify heart diseases?

How is the model's performance evaluated? Is that sufficient?

This study summarized different ML and DL methods utilized in HDP models. The remainder of the paper is structured as follows. In Section 2, the background and overview of ML and DL are discussed, whereas Section 3 is showed the method used to select the studies. Section 4 is presented the results obtained. Finally, Section 5 is concluded the article with a general conclusion.

II. BASICS AND BACKGROUND

ML is a branch of AI that uses numerical calculations and statistical computations to perform analysis. It was coined in 1959 by Arthur Samuel. It requires methods that help the data be processed and generate the final results. It is based on creating software programmers that gain knowledge from data and increase precision without being programmed, over time. ML methods can work with large datasets and make decisions and predictions.

DL is a part of the broad area of AI as well as part of ML, in which suitable methods are augmented by layers of neurons of brain function and structure called artificial neural networks (ANNs). DL replicates the functions of the brain when analyzing and processing data to make decisions. It performs a deep analytical procedure to assiduously learn a dataset using hierarchical layers of ANN. DL Processing data uses a non-linear method to connect and associate all inputs to produce the optimal output. A neural network's first layer gathers input data, analyses it, and transmits the results to the second layer as output. Before making decisions and providing outcomes, the next layer of neurons in a deep neural network processes earlier data [6].

To improve accuracy, the ensemble methods combine multiple classifiers into a single model. There are three types of ensembles learning methods. The first one is bagging that aggregates similar classifiers by the voting method. The second is "boosting," which is similar to "bagging," but new models are influenced by the results of previous models. The third is stacking. It is an advanced ensemble method that is aggregated different classifiers to build the model.

Machine Learning Algorithms	• Decision Tree (DT) • Support Vector Machine (SVM) • K-Nearest Neighbor (KNN) • Naïve Bayes (NB) • Logistic Regression (LR) • Artificial neural network (ANN)
Deep Learning Algorithms	• Multi-layer Perceptron (MLP) • Convolutional Neural Network (CNN) • Recurrent neural network (RNN) • Generative Adversarial Networks (GANs)

Figure 1. Most ML and DL algorithms used in HDP.

A. *MACHINE LEARNING ALGORITHMS*

This section presents the most ML methods that are used in HDP.

1) DECISION TREE

Decision Tree (DT) is a supervised method for machine learning. It is employed to find continuous solutions to classification and regression issues. dividing data based on specific criteria The distribution of the data resembles a tree. Decisions are made in the leaves, which are separated into nodes. This process iterates across the features, and the leaf nodes deliver the final result. The classification tree's decision variable is categorical, but the regression tree's decision variable is continuous (the result is yes or no) [7].

2) SUPPORT VECTOR MACHINE

Both classification and regression issues can be solved with Support Vector Machine (SVM). It classifies data into two classes over a hyperplane. Keep comparable data of one type on one side and comparable data of a different type on the other side of the hyperplane. In order to reduce misclassification, it aims to maximize the separation between each class's two closest data points and the hyperplane. The hyperplane should define what the decision's borders are. In order to divide a group of objects from various classes, you need a decision plan [8].

SVM can be utilized for classification. A hyperplane can divide disease classes in the case of the HDP such that one side of the edge has heart disease while the other does not. Linear and nonlinear SVM are additional divisions of SVM. A nonlinear SVM is used when the data cannot be linearly separated using a line, as opposed to a linear SVM, which can do so. A nonlinear SVM kernel function is used because the data may be complex and cannot be separated using a linear SVM [9].

3) K-NEAREST NEIGHBOR

K-Nearest Neighbor (KNN) is a method for supervised learning that may be applied to boz th regression and classification issues. The k nearest data points in the training set are found if the KNN is missing a target value for a given data point, and the estimated average value of collected data points is calculated. The mean of the k labels is returned by regression, whereas the mean of the k labels is either assigned or returned by classification. When prior knowledge of the data is unavailable, KNN is the default classification method employed. The closest data points can be determined using the Manhattan distance and Euclidean distance, two distance measures. Even with noisy and large amounts of data, it can produce better outcomes and forecasts [9].

4) NAÏVE BAYES

Naïve Bayes (NB) is a probabilistic method that uses the Bayes theorem in application and makes strong (naive) assumptions about the independence of feature pairs. Simple Bayesian models are particularly useful in medicine for diagnosing heart disease patients because they are easy to build without complex iterative parameter estimation. Although being simple, naive Bayesian classifiers are widely used because they often perform surprisingly well and outperform more complex classification methods. The posterior probability can be calculated according to the Bayes theorem: P(X|Y) from P(X), P(Y), and P(Y|X). The Naive Bayes assumes that the influence of the value of a predictor variable (X) on a particular class (Y) is independent of the values of other predictor variables. This assumption is called class independence.

$$P(X/Y) = \frac{P(Y/X) \times P(X)}{P(Y)} \quad (1)$$

- P(X|Y) is the posterior likelihood of the (target) class given the predictor (attribute).
- P(X) is the prior likelihood of class.
- P(Y|X) is the probability which is the probability of the predictor given class.
- P(Y) is the prior likelihood of the predictor

Where X and Y are two events. This method works well with categorical data, but poorly if the training dataset has numeric data [10].

5) LOGISTIC REGRESSION

Both classification and regression problems are resolved using logistic regression (LR). To predict the result, input values may be linearly combined with a logistic or sigmoid function and coefficient values. Given the value of (0 or 1) of the input variable, it provides a binomial result, indicating the probability that the event will occur. There are different types of logistic regression results, like binomial, ordinal (classifications with ordering), and polynomial (classifications without ordering). This model is simple to use and can make accurate predictions. To predict the values of continuous variables, linear regression is used [7].

6) Artificial Neural Networks (ANNs)

Artificial neural network (ANN) is a field of machine learning in neural networks. ANNs are similar to that of Human brain function. The cell is a simulation of a human neuron, it is Similar to how a cell processes information and responds. ANN learns from data, categorizes it, and anticipates an output. It is a nonlinear statistical architecture for discovering complex problem solutions. It contains three layers: an input layer a hidden layer, and an output layer with many nodes that resemble neurons in the human brain. The nodes of the ANN act as inputs for the input layer as neurons converse with one another. Data from the outside world is transferred to the concealed layer through the input layer. Here, the hidden layer analyses the data and makes some computations to search for patterns. Pass the classified data to the output layer after processing. Input functions are converted into output functions using activation functions. There are various varieties, including logistic, tanh, sigmoid, linear, and more. These days, ANNs are widely employed in different industries, including health, image identification, speech recognition, and face recognition [9].

B. DEEP LEARNING ALGORITHMS

This section presents a review of the most commonly used DL algorithms in HDP.

1) MULTI-LAYER PERCEPTRON

Multi-layer perceptron (MLP) is a form of supervised learning approach and an ANN. This is also referred to as deep learning's fundamental architecture or deep neural network (DNN).

A basic MLP is made up of just three layers: an input layer that accepts input data, and an output layer that decides what to do with the input signal. Between these two there may be one or more hidden layers that serve as the network's processing units. MLPs' output is calculated using a different of activation functions, including Tanh, Sigmoid, and Softmax, rectified linear unit (ReLU), and numerous optimization techniques, including limited memory BFGS (L-BFGS), adaptive moment estimation (Adam), and stochastic gradient descent (SGD) are used throughout the training phase. MLP needs tuning many hyperparameters, such as hidden layers, neurons, and several iterations, so complex models can be solved computationally intensively[11].

2) CONVOLUTIONAL NEURAL NETWORK

Convolutional Neural Network (CNN) is a well-known supervised DL architecture. It learns directly from inputs without needing to extract features. A CNN with numerous convolutional and pooling layers is demonstrated in

Figure 2. As a result, CNNs enhance the architecture of conventional ANNs such as controlled MLP networks. Each CNN layer considers optimal parameters to produce meaningful output while reducing the complexity of the model. CNNs also use a dropout layer that can address the overfitting problem that can occur in traditional networks. The ability to automatically detect key features from inputs without requiring human interaction makes them more effective than traditional networks. In the visual geometry group (AlexNet, Xception, ResNet, etc.), depending on their learning capacity, different CNN variations can be used in various application fields[11].

3) RECURRENT NEURAL NETWORK

A recurrent neural network (RNN) is one more common neural network that processes sequential or time-series data. It provides the current stage with the output of a previous stage as input. Similarly, to CNNs and feedforward, recurrent networks learn from their training inputs. However, they have a different memory. For that, the information from previous inputs can be used to influence current inputs and outputs. In contrast to a normal DNN, which assumes that the input and output are independent of one another, in RNN, the output is dependent on the prior element in the sequence. Standard recurrent networks, on the other hand, contain vanishing gradients, which makes it challenging to train extended data sequences. A Feed-Forward Neural Network can be transformed into RNN. A single layer of RNNs is created by compressing the nodes from the neural network's input, hidden, and output layers. A, B, and C are the parameters of the RNN network.

Below, are some common variants of recurrent networks that are minimally problematic and work well in many domains of real-world application:

- The LSTM is a well-liked RNN architecture that employs specialized units to address the vanishing gradient issue. In LSTM devices, memory cells have a long-term data storage capacity. Three gates control how information enters and exit the cell. For example, the "forget gate" determines what information is preserved from the cell in the previous state and removes information that is no longer needed. The "input gate" determines what information is put into the cell state. The "output gate" determines and controls the output. LSTM networks are considered one of the most successful RNNs for solving the problem of training recurrent networks.
- Another popular version of recurrent networks called GRU uses gating techniques to regulate and manage the information flow between neural network cells. GRUs are similar to LSTMs but have reset and update gates but no output gates and fewer parameters. GRU and LSTM vary primarily in that GRU only has two gates (the reset and update gates), while LSTM has three gates (input, forget, and output gates). The GRU's structure enables it to record dependencies on lengthy data sequences in an adaptive manner without losing information from previous segments of the sequence.
- Recurrent networks have the fundamental characteristic of having at least one feedback link that permits looping activations. As a result, the network can carry out temporal processing and sequence learning tasks including sequence duplication or detection, temporal association or prediction, etc. Recurrent networks have certain common applications in speech recognition, machine translation, natural language processing, prediction difficulties, and text summarization[11].

4) GENERATIVE ADVERSARIAL NETWORK

In order to generate new believable patterns on demand, generative modeling uses a form of neural network architecture called generative adversarial networks (GANs). By automatically detecting regularities and patterns in the incoming data, the model can be utilized to generate new instances from the original dataset.

GAN contains two neural networks. A discriminator D forecasts the possibility that successive samples will be produced from the real data instead of the created data that was produced using the generator. A generator G generates new data with attributes comparable to the original data. As a result, both generators and discriminators in GAN modeling are trained to compete with one another. The deployment of GAN networks is designed for unsupervised learning tasks.

By producing more realistic data, the generator may attempt to deceive and perplex the discriminator. Healthcare, data augmentation, picture analysis, video generation, audio generation, traffic control, pandemics, cyber security, etc. are just a few of the fast-expanding application areas for GAN networks. In general, GANs have become a significant autonomous data augmentation field and a solution to issues that call for generative approaches [11].

III. METHOD

A. SEARCH STRATEGY AND SELECTION PROCESS

In this study, four key search terms (heart disease, machine learning, ensemble learning) across reputable 5-9 journals like IEEE, ACM, Springer, ScienceDirect, and Emerald. To focus on more recent advancements in the field, it was only applicable to articles from 2018 to 2023. There were 100 results in total for original articles with keywords mentioned above, after reviewing articles, case reports and meta-data analysis articles were filtered out. Looking into the content of each publication, it was identified that some of the results were not related to heart diseases and/or machine learning. In the end, 65 publications are relevant to this study. These publications were elected to subcategorize under the results section concerning the disease. It aims to diagnose to provide a more comprehensive comparison and a more coherent reading experience for readers. **Figure** shows a flowchart detailing the process of how relevant studies are obtained, classified, and explored. **Table 1** outlines the criteria used to define the search term and where, within the manuscript, each term focuses.

Table 1. Criteria used to build the literature search.

	Criteria	Term Location
A	"Machine Learning" OR "Deep Learning" OR "Ensemble Learning"	Anywhere within the manuscript
B	Heart OR Cardiac Disease	Anywhere within the manuscript

1) SEARCH RESULT

The search process is detailed in ***Figure*** based on the search criteria, 97 total studies were found on the scientific search engines. The unique studies were subsequently extracted, which left a total of 97 studies.

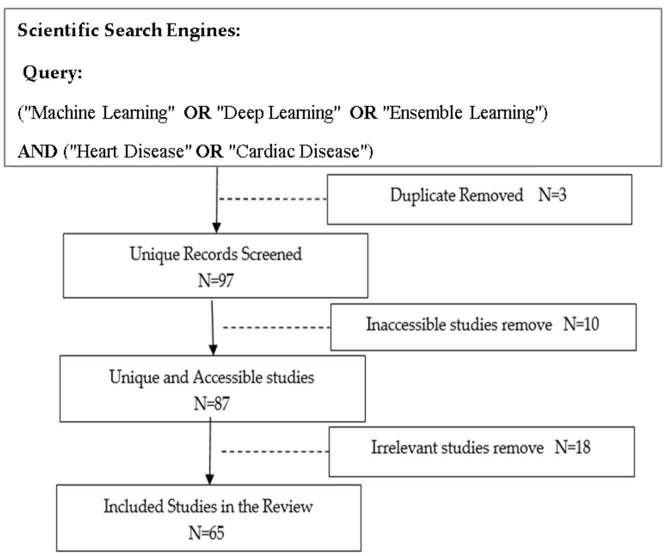

Figure 2. Flow chart of studies selection.

Among the 97 studies that remained, 10 studies were excluded due to issues with accessing the full manuscript, leaving 87 studies to be included for full-text readings and to form the dataset for this study. However, during the full-text readings, a further 18 studies were excluded. After all the exclusions had been applied, this left a final total of 65 studies that were considered for this study [3], [4], [12]- [13].

IV. RESULTS

Of the 65 studies, several different approaches were taken. The studies were clustered into three subgroups of methods: ML, DL, and ELMs. Each study was then assigned to one of these three groups using the criteria outlined in Table 1. Machine learning methods made up most of the studies with 28 (43%) [3] [12]-[14] being assigned to this group. ELMs were the next single largest group with 24 (37%) studies [15] - [13]. Deep learning methods were the smallest single group with 13 (20%) studies [4], [16] - [17]

A. HEART DISEASE WITH MACHNE LEARNING METHODS

The studies within this group are focused on using SVM, KNN, DT, RF, NB, LR, ANN, and J48 methods. Most of the 28 studies used more than the methods then compared between them as [12], [18]- [19]- [20], [3], [21]- [22], [14]. Two studies that are used one method [23], [24].

Year	Ref	Dataset	Methods	Best Accuracy/Results	Future Work/Limitation
2018	[12]	Cleveland	DT, J48, LMT, RF, NB, KNN, & SVM	NB is the best classifier	Use combinational models.
2018	[23]	Cleveland & Statlog	ANN	Best of PCA (94.7%, 97.7%)	
2018	[18]	Cleveland & Statlog	RF, DT, & NB	RF with perfect results	Apply genetic method.
2018	[25]	Framingham	LR, RF, KNN, SVM & DT	LR (88.86%)	
2018	[26]	Cleveland & Statlog	Cloud 4-Tier Arch (ANN, SVM, RF, NB, & DT)	ANN (86%)	
2019	[19]	Combined (Statlog, Switzerland, Hungarian, V.A. Medical)	NN, NN, SVM, NB & RF	90–95 %	
201	[27	Cleveland	Hybrid RF with a Linear Model (HRFLM)	88.7%	

9	]				
2019	[20]	Cleveland	KNN, SVM, NB, RF, MLP, ANN optimized by PSO & ACO	KNN (99.65%), RF (99.6%)	
2019	[28]	Cleveland	KNN with SBS feature selection	90%	
2019	[29]	Cleveland & IOT Sensors	Cloud and IoT model using a set of classifiers J48, LR, MLP & SVM	J48 classifiers is the best	
2019	[3]	Kaggle	LR, KNN, AdaBoost, DT, NB, RF, SVM, Extra Tree Classifier (ETC) & Gradient Boosting	The best: SVM, RF, ETC	
2020	[21]	Cleveland	LR, KNN, RF, DT & SVM with grid search for tunning hyperparameter	KNN with grid search (91.80%)	Feature selection methods with different techniques.
2020	[30]	Cleveland	RF, NB, SVM, DT, Hoeffding Trees & LMT	RF(95.08%)	Add more attributes and analyze with proposed models.
٢٠٢٠	[31]	Cleveland	SVM, NN, DT, and LR	KNN is the best	
2020	[32]	Cleveland	KNN, NB, DT, and RF	KNN (90.78%)	Incorporating other data mining technique

					s.
2020	[33]	Cleveland	LR, NB, SVM, KNN, DT, RF, & XGBoost	RF (86.89%)	
2020	[34]	Cleveland	SVM, RF, NB, DT with Weka	RF (99%)	
2021	[35]	Framingham	(RF, LR, SVM) using Linear Kernel Function; SVM (Radial Basis Kernel Function, NB)	RF (84.81**%**)	
2021	[36]	Cleveland	(hybrid GA and RFE) & (NB, SVM, LR, RF, AdaBoost)	RF (86.60%)	Use ACO & PSO as feature selection methods.
2021	[37]	Cleveland	KNN, LR, RF	KNN (88.52%)	
2021	[38]	Svetlana Ulianova 2019	KNN, RF, DT, and SVM	NB is the best	
2021	[39]	Kaggle	KNN, RF, and DT	RF (100%)	
2021	[40]	Cleveland	ANN, DT, NB, RF, LR, SVM & XG Boost	RF (95.08%)	
2021	[41]	40 thousand ECGs	XGBoost for training, Optuna for tuning parameters	F1 Scores (0.93 – 0.99)	
2021	[42]	Data on people's tests	NN, SVM, & KNN	NN (93%)	Feature selection methods, increase

					the dataset size.	Dataport dataset	- using medical IoT devices and sensors for collection the clinical parameters
2022	[22]	Cleveland	KNN, DT, LR, NB, & SVM	LR (92.30%)	Furthe analysis methods.		
2022	[24]	Cleveland & Statlog	SVM for classification and χ2 statistical optimum for feature selection	(89.47, 89.7%)			
2022	[14]	Cleveland, Hungarian, Switzerland, & Long Beach	RF, DT, AB, & KNN	KNN (100%, 97.82%)			
2023	[43]	Cleveland and 95% for the IEEE	soft voting classifier combining all ML method	93.44%	- limited amount of patient data		

summarizes the studies that used machine learning methods.

Table 2. Summary of ML methods used in heart disease prediction (HDP).

Year	Ref	Dataset	Methods	Best Accuracy/Results	Future Work/Limitation
2018	[12]	Cleveland	DT, J48, LMT, RF, NB, KNN, & SVM	NB is the best classifier	Use combinational models.
2018	[23]	Cleveland & Statlog	ANN	Best of PCA (94.7%, 97.7%)	
2018	[18]	Cleveland & Statlog	RF, DT, & NB	RF with perfect results	Apply genetic method.
2018	[25]	Framingham	LR, RF, KNN, SVM & DT	LR (88.86%)	
2018	[26]	Cleveland & Statlog	Cloud 4-Tier Arch (ANN, SVM, RF, NB, & DT)	ANN (86%)	
2019	[19]	Combined (Statlog, Switzerland, Hungarian, V.A. Medical)	NN, NN, SVM, NB & RF	90–95 %	
2019	[27]	Cleveland	Hybrid RF with a Linear Model (HRFLM)	88.7%	
2019	[20]	Cleveland	KNN, SVM, NB, RF, MLP, ANN optimized by PSO & ACO	KNN (99.65%), RF (99.6%)	
2019	[28]	Cleveland	KNN with SBS feature selection	90%	
2019	[29]	Cleveland & IOT	Cloud and IoT model using a set of	J48 classifiers is the best	

		Sensors	classifiers J48, LR, MLP & SVM		
2019	[3]	Kaggle	LR, KNN, AdaBoost, DT, NB, RF, SVM, Extra Tree Classifier (ETC) & Gradient Boosting	The best: SVM, RF, ETC	
2020	[21]	Cleveland	LR, KNN, RF, DT & SVM with grid search for tunning hyperparameter	KNN with grid search (91.80%)	Feature selection methods with different techniques.
2020	[30]	Cleveland	RF, NB, SVM, DT, Hoeffding Trees & LMT	RF(95.08%)	Add more attributes and analyze with proposed models.
٢٠٢٠	[31]	Cleveland	SVM, NN, DT, and LR	KNN is the best	
2020	[32]	Cleveland	KNN, NB, DT, and RF	KNN (90.78%)	Incorporating other data mining techniques.
2020	[33]	Cleveland	LR, NB, SVM, KNN, DT, RF, & XGBoost	RF (86.89%)	
2020	[34]	Cleveland	SVM, RF, NB, DT with Weka	RF (99%)	
2021	[35]	Framingham	(RF, LR, SVM) using Linear Kernel Function; SVM (Radial Basis Kernel Function, NB)	RF (84.81**%**)	
2021	[36]	Cleveland	(hybrid GA and RFE) & (NB, SVM, LR, RF, AdaBoost)	RF (86.60%)	Use ACO & PSO as feature selection methods.
2021	[37]	Cleveland	KNN, LR, RF	KNN (88.52%)	
2021	[38]	Svetlana Ulianova 2019	KNN, RF, DT, and SVM	NB is the best	
2021	[39]	Kaggle	KNN, RF, and DT	RF (100%)	
2021	[40]	Cleveland	ANN, DT, NB, RF, LR, SVM & XG Boost	RF (95.08%)	
2021	[41]	40 thousand ECGs	XGBoost for training, Optuna for tuning parameters	F1 Scores (0.93 – 0.99)	
2021	[42]	Data on people's tests	NN, SVM, & KNN	NN (93%)	Feature selection methods, increase the dataset size.
2022	[22]	Cleveland	KNN, DT, LR, NB, & SVM	LR (92.30%)	Furthe analysis methods.
2022	[24]	Cleveland & Statlog	SVM for classification and $\chi 2$ statistical optimum for feature selection	(89.47, 89.7%)	
2022	[14]	Cleveland, Hungarian, Switzerland, & Long Beach	RF, DT, AB, & KNN	KNN (100%, 97.82%)	
2023	[43]	Cleveland and 95% for the IEEE Dataport dataset	soft voting classifier combining all ML method	93.44%	- limited amount of patient data - using medical IoT devices and sensors for collection the clinical parameters

methods. Five of these studies used DNN as [4],[16]-[44].

B. HEART DISEASE WITH DEEP LEARNING METHODS

The studies within this group are focused on using DNN, CNN, GRU, LSTM, RNN, and BiLSTM

Table **3.** Summary of DL methods used in heart disease prediction (HDP). summarizes studies that used deep learning methods.

	Ref	Dataset	Methods	Best Accuracy/Results	Future Work/Limitation
2018	[16]	Cleveland	Fve layer DNN architecture	99% accuracy	
2018	[45]	Cleveland	DNN	93.51% accuracy	Use LSTM, RNN, and CNN.
2019	[46]	Cleveland	χ2 statistical model & DNN	93.33% accuracy	Use GA with ANN & DNN.
2019	[44]	Multiple datasets	DNN	87.64% accuracy	Use other DL networks.
2019	[47]	Cleveland & Physionet	CNN	97%	
2019	[48]	Cleveland	MLPNN with Back-propagation	94% accuracy	
2020	[49]	Not mention	DLMNN	92% accuracy	
2020	[50]	Cleveland	CNN - GRU	94% accuracy	
2020	[4]	Cleveland	DNN using Talos optimization	90.78% accuracy	
2021	[51]	MIMIC-II	RNN, LSTM, GRU, & BI-LSTM	GRU with 3 layers is the best	
2021	[52]	Kaggle	Enhanced RNN	91% accuracy	
2021	[53]	Cleveland	LASSO & CNN	97%	
2022	[17]	Cleveland & Hungarian	Fuzzy inference system with Bi-LSTM	98.85% accuracy	
2023	[54]	Cleveland, Hungarian, Long Beach and Switzerland	CNN	83%	Test CNN model on structured and unstructured data to improve it
2023	[55]	: Cleveland, Hungarian, Switzerland, Long Beach, stalog	CNN with SAE	90.08%	

Table 3. Summary of DL methods used in heart disease prediction (HDP).

	Ref	Dataset	Methods	Best Accuracy/Results	Future Work/Limitation
2018	[16]	Cleveland	Fve layer DNN architecture	99% accuracy	
2018	[45]	Cleveland	DNN	93.51% accuracy	Use LSTM, RNN, and CNN.
2019	[46]	Cleveland	χ2 statistical model & DNN	93.33% accuracy	Use GA with ANN & DNN.
2019	[44]	Multiple datasets	DNN	87.64% accuracy	Use other DL networks.
2019	[47]	Cleveland & Physionet	CNN	97%	
2019	[48]	Cleveland	MLPNN with Back-propagation	94% accuracy	
2020	[49]	Not mention	DLMNN	92% accuracy	
2020	[50]	Cleveland	CNN - GRU	94% accuracy	
2020	[4]	Cleveland	DNN using Talos optimization	90.78% accuracy	
2021	[51]	MIMIC-II	RNN, LSTM, GRU, & BI-LSTM	GRU with 3 layers is the best	
2021	[52]	Kaggle	Enhanced RNN	91% accuracy	
2021	[53]	Cleveland	LASSO & CNN	97%	
2022	[17]	Cleveland & Hungarian	Fuzzy inference system with Bi-LSTM	98.85% accuracy	
2023	[54]	Cleveland, Hungarian, Long Beach and Switzerland	CNN	83%	Test CNN model on structured and unstructured data to improve it
2023	[55]	: Cleveland, Hungarian, Switzerland, Long Beach, stalog	CNN with SAE	90.08%	

C. HEART DISEASE WITH ENSEMBLE LEARNING METHODS

The studies

within this group are focused on using Majority Voting, Bagging, Boosting, Classification and Regression Tree (CART), Weighted Aging Classifier Ensemble (WAE), AdaBoost, and Stacking methods Also some of these studies were combined and integrated between two or more ML and DL methods. Three studies used more than one ensemble method and compared them [56], [57], [58].

	Ref	Dataset	Methods	Best Accuracy/Results	Future Work/Limitation
2018	[16]	Cleveland	Fve layer DNN architecture	99% accuracy	
2018	[45]	Cleveland	DNN	93.51% accuracy	Use LSTM, RNN, and CNN.
2019	[46]	Cleveland	χ2 statistical model & DNN	93.33% accuracy	Use GA with ANN & DNN.
2019	[44]	Multiple datasets	DNN	87.64% accuracy	Use other DL networks.
2019	[47]	Cleveland & Physionet	CNN	97%	
2019	[48]	Cleveland	MLPNN with Back-propagation	94% accuracy	
2020	[49]	Not mention	DLMNN	92% accuracy	
2020	[50]	Cleveland	CNN - GRU	94% accuracy	
2020	[4]	Cleveland	DNN using Talos optimization	90.78% accuracy	
2021	[51]	MIMIC-II	RNN, LSTM, GRU, & BI-LSTM	GRU with 3 layers is the best	
2021	[52]	Kaggle	Enhanced RNN	91% accuracy	
2021	[53]	Cleveland	LASSO & CNN	97%	
2022	[17]	Cleveland & Hungarian	Fuzzy inference system with Bi-LSTM	98.85% accuracy	
2023	[54]	Cleveland, Hungarian, Long Beach and Switzerland	CNN	83%	Test CNN model on structured and unstructured data to improve it
2023	[55]	: Cleveland, Hungarian, Switzerland, Long Beach, stalog	CNN with SAE	90.08%	

Table **3.** Summary of DL methods used in heart disease prediction (HDP).

summarizes studies that used ELMs.

Table 4. Summary of ensemble learning methods (ELMs) used in heart disease prediction (HDP).

Year	**Ref**	**Dataset**	**Ensemble Learning Methods (ELMs)**	**Best Accuracy/Results**	**Future Work/Limitation**
2018	[15]	Cleveland	KNN, NB, DT, Majority Voting	90% accuracy	
2018	[59]	Cleveland & Hungarian	RF trees, SVM, NB, NN, LR	ELM is a superior approach	Use another dataset.
2018	[60]	Cleveland	NB, LR, NN	91.26% accuracy	Use other data mining algorithms with greater medical data.

2018	[61]	SPECT	Hybrid ELMs	96% accuracy	Extending Hybrid ELMs for other diseases.
2019	[62]	Kaggle	Boosting-based ELMs (AdaBoost, GBM, XGBoost, LGBM, CatBoost)	Tuning parameters improved the algorithms	
2019	[63]	Medical database	Ensemble DT with GA	85.37% accuracy	Use hybrid generic intelligent systems.
2019	[64]	Cleveland	Voting ELMs (LR, RF, KNN & SGD)	90% accuracy	
2019	[5]	Cleveland	ELMs for NB, Bayes Net, C4.5, Multilayer Perceptron, and PART	Increase of 7% accuracy	
2019	[65]	Statlog	Voting ELMs (LR, NB & MLP)	88.88%	
2019	[56]	Kaggle	Bagging, Boosting, RSM, RUS Boos	Bagging (99.3%)	
2020	[50]	Sensor data & electronic medical records (EMRs)	Using ensemble DL and feature fusion	98.5%	
2020	[66]	Cleveland & Framingham	CART & WAE	Cleveland (93%), Framingham (91%)	
2020	[67]	Cleveland	ELMs (RF, KNN, SVM, LSTM, GRU)	85.71%	
٢٠٢٠	[68]	Medica Norte Hospital in Mexico	CNN-MLP And LSTM, GRU, BiLSTM, BiGRU	91%-96%	Use other NN such as GAN or RNN.
2020	[69]	Cleveland	Hybrid gradient boosting DT with LR	91.8%;	
2020	[57]	Cleveland	AdaBoost, Bagging & Stacking	AdaBoost is the best	Apply genetic method for AdaBoost parameters fine-tuning.
2020	[70]	Cleveland	ANN, KNN, SVM & majority voting	Majority Voting (61.16% multiclass class.),	

				(87.37% binary class.)	
2021	[71]	Tunisian biotechnology center & Cleveland	ELMs (SVM, KNN DT C4.5, Bagging and Adaptive boosting)	ELMs improved performance	
2021	[72]	Kaggle	LR, CART, LDA, KNN, SVM, GB, and ELMs	86.32%	
2021	[73]	Statlog, SPECTF	a hybrid ELM with GA-LDA	93.65%	Use (PSO, ACO, Firefly).
2021	[74]	Cleveland	Combination of ML and DL	94.2%	Increase dataset size with other techniques.
2022	[75]	Kaggle	Ensemble Stacked ML (XGB, KNN, DT) and DL (DNN, KDNN)	88.70%	
2022	[58]	Kaggle	ML with (Majority Voting, Stacking, Bagging)	98.38%	
2022	[13]	Cleveland & others	CNN-LSTM and CNN-GRU	98.41%	
2023	[76]	Cleveland and a large public dataset	CNN-LSTM model	97.75% with Cleveland 98.86% with arge dataset	- lack of comprehensive testing on real-world datasets - lack of deep ensemble learning methods - In the future, generalize the system

V. DISCUSSION

According to the papers covered in this study, there is a clear intent to use ML and DL in the field of heart disease prediction research. This is demonstrated by 65 papers that either use ML, DL, or ensemble approaches to build high-accuracy models, evaluate how ML or DL are being used, or compare them [29] - [38], [40], [22], [14], [51], [59], [56], [57], [58]. Most studies applied ML and DL methods to classify and detect heart disease at early stages.

Most of the studies used UCI ML Repository especially the Cleveland dataset with 303 records and 14 features. The results obtained depend on the dataset [24], [66], [13]. As an average of accuracies, the studies that applied DL methods present higher accuracy than studies that applied ML methods. Figure 2 shows the average accuracies of ML and DL for the studies.

Figure 2. The average accuracy of the studies.

Also ensemble models provide better results and improve the performance than individual models [59], [64], [71].

.

Acknowledgments:

Conflicts of Interest:

References

1. Ramesh, T., et al., *Predictive analysis of heart diseases with machine learning approaches.* Malaysian Journal of Computer Science, 2022: p. 132-148.
2. Ahsan, M.M., S.A. Luna, and Z. Siddique. *Machine-learning-based disease diagnosis: A comprehensive review*. in *Healthcare*. 2022. MDPI.
3. Lakshmanarao, A., Y. Swathi, and P.S.S. Sundareswar, *Machine learning techniques for heart disease prediction.* Forest, 2019. **95**(99): p. 97.
4. Sharma, S. and M. Parmar, *Heart diseases prediction using deep learning neural network model.* Interna-tional Journal of Innovative Technology and Exploring Engineering (IJITEE), 2020. **9**(3): p. 124-137.
5. Latha, C.B.C. and S.C. Jeeva, *Improving the accuracy of prediction of heart disease risk based on ensemble classification techniques.* Informatics in Medicine Unlocked, 2019. **16**: p. 100203.
6. Krishnan, S., P. Magalingam, and R. Ibrahim, *Hybrid deep learning model using recurrent neural network and gated recurrent unit for heart disease prediction.* International Journal of Electrical & Computer Engineering (2088-8708), 2021. **11**(6).
7. Akgül, M., Ö.E. Sönmez, and T. Özcan. *Diagnosis of heart disease using an intelligent method: A hybrid ANN–GA approach*. in *International conference on intelligent and fuzzy systems*. 2019. Springer.
8. Animesh Hazra, et al., *Heart disease diagnosis and prediction using machine learning and data mining techniques: a review.* Advances in Computational Sciences and Technology, 2017. **10**(7): p. 2137-2159.
9. Katarya, R. and S.K. Meena, *Machine learning techniques for heart disease prediction: a comparative study and analysis.* Health and Technology, 2021. **11**(1): p. 87-97.
10. Vembandasamy, K., R. Sasipriya, and E. Deepa, *Heart diseases detection using Naive Bayes algorithm.* International Journal of Innovative Science, Engineering & Technology, 2015. **2**(9): p. 441-444.
11. Sarker, I.H., *Deep learning: a comprehensive overview on techniques, taxonomy, applications and research directions.* SN Computer Science, 2021. **2**(6): p. 1-20.
12. Kumar, M.N., K. Koushik, and K. Deepak, *Prediction of heart diseases using data mining and machine learning algorithms and tools.* International Journal of Scientific Research in Computer Science, Engineering and Information Technology, 2018. **3**(3): p. 887-898.
13. Almulihi, A., et al., *Ensemble Learning Based on Hybrid Deep Learning Model for Heart Disease Early Prediction.* Diagnostics, 2022. **12**(12): p. 3215.
14. Absar, N., et al. *The efficacy of machine-learning-supported smart system for heart disease prediction*. in *Healthcare*. 2022. MDPI.
15. Kurian, R.A. and K. Lakshmi, *An ensemble classifier for the prediction of heart disease.* International Journal of Scientific Research in Computer Science, 2018. **3**(6): p. 25-31.
16. Tomov, N.-S. and S. Tomov, *On deep neural networks for detecting heart disease.* arXiv preprint arXiv:1808.07168, 2018.
17. Nancy, A.A., et al., *Iot-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning.* Electronics, 2022. **11**(15): p. 2292.
18. David, H. and S.A. Belcy, *HEART DISEASE PREDICTION USING DATA MINING TECHNIQUES.* ICTACT Journal on Soft Computing, 2018. **9**(1).
19. Reddy, N.S.C., et al., *Classification and feature selection approaches by machine learning techniques: Heart disease prediction.* International Journal of Innovative Computing, 2019. **9**(1).
20. Khourdifi, Y. and M. Bahaj, *Heart disease prediction and classification using machine learning algorithms optimized by particle swarm optimization and ant colony optimization.* International Journal of Intelligent Engineering and Systems, 2019. **12**(1): p. 242-252.
21. Hashi, E.K. and M.S.U. Zaman, *Developing a hyperparameter tuning based machine learning approach of heart disease prediction.* Journal of Applied Science & Process Engineering, 2020. **7**(2): p. 631-647.
22. Gupta, C., et al. *Cardiac Disease Prediction using Supervised Machine Learning Techniques*. in *Journal of Physics: Conference Series*. 2022. IOP Publishing.
23. Awan, S.M., M.U. Riaz, and A.G. Khan, *Prediction of heart disease using artificial neural network.* 2018.

24. Sarra, R.R., et al., *Enhanced heart disease prediction based on machine learning and χ2 statistical optimal feature selection model*. Designs, 2022. **6**(5): p. 87.
25. Chauhan, Y.J., *Cardiovascular disease prediction using classification algorithms of machine learning.* International Journal of Science and Research (IJSR), ISSN, 2018: p. 2319-7064.
26. Ahmed, M.R., et al. *A cloud based four-tier architecture for early detection of heart disease with machine learning algorithms*. in *2018 IEEE 4th International Conference on Computer and Communications (ICCC)*. 2018. IEEE.
27. Mohan, S., C. Thirumalai, and G. Srivastava, *Effective heart disease prediction using hybrid machine learning techniques.* IEEE access, 2019. **7**: p. 81542-81554.
28. Haq, A.U., et al. *Heart disease prediction system using model of machine learning and sequential backward selection algorithm for features selection*. in *2019 IEEE 5th International Conference for Convergence in Technology (I2CT)*. 2019. IEEE.
29. Ganesan, M. and N. Sivakumar. *IoT based heart disease prediction and diagnosis model for healthcare using machine learning models*. in *2019 IEEE International Conference on System, Computation, Automation and Networking (ICSCAN)*. 2019. IEEE.
30. Motarwar, P., et al. *Cognitive approach for heart disease prediction using machine learning*. in *2020 International Conference on Emerging Trends in Information Technology and Engineering (ic-ETITE)*. 2020. IEEE.
31. Singh, A. and R. Kumar. *Heart disease prediction using machine learning algorithms*. in *2020 international conference on electrical and electronics engineering (ICE3)*. 2020. IEEE.
32. Shah, D., S. Patel, and S.K. Bharti, *Heart disease prediction using machine learning techniques.* SN Computer Science, 2020. **1**(6): p. 1-6.
33. Anbuselvan, P., *Heart disease prediction using machine learning techniques.* Int. J. Eng. Res. Technol, 2020. **9**: p. 515-518.
34. Sharma, V., S. Yadav, and M. Gupta. *Heart disease prediction using machine learning techniques*. in *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)*. 2020. IEEE.
35. Rubini, P., et al., *A cardiovascular disease prediction using machine learning algorithms.* Annals of the Romanian Society for Cell Biology, 2021: p. 904-912.
36. Rani, P., et al., *A decision support system for heart disease prediction based upon machine learning.* Journal of Reliable Intelligent Environments, 2021. **7**(3): p. 263-275.
37. Jindal, H., et al. *Heart disease prediction using machine learning algorithms*. in *IOP conference series: materials science and engineering*. 2021. IOP Publishing.
38. Yadav, A., L. Gediya, and A. Kazi, *Heart disease prediction using machine learning.* International Research Journal of Engineering and Technology (IRJET, 2021. **8**(09).
39. Ali, M.M., et al., *Heart disease prediction using supervised machine learning algorithms: Performance analysis and comparison.* Computers in Biology and Medicine, 2021. **136**: p. 104672.
40. Williams, R., et al. *Heart disease prediction using machine learning techniques*. in *2021 International Conference on Data Analytics for Business and Industry (ICDABI)*. 2021. IEEE.
41. Bertsimas, D., L. Mingardi, and B. Stellato, *Machine learning for real-time heart disease prediction.* IEEE Journal of Biomedical and Health Informatics, 2021. **25**(9): p. 3627-3637.
42. Salhi, D.E., A. Tari, and M.-T. Kechadi. *Using machine learning for heart disease prediction*. in *Advances in Computing Systems and Applications: Proceedings of the 4th Conference on Computing Systems and Applications*. 2021. Springer.
43. Chandrasekhar, N. and S. Peddakrishna, *Enhancing heart disease prediction accuracy through machine learning techniques and optimization.* Processes, 2023. **11**(4): p. 1210.
44. Ashraf, M., M. Rizvi, and H. Sharma, *Improved heart disease prediction using deep neural network.* Asian Journal of Computer Science and Technology, 2019. **8**(2): p. 49-54.
45. Miao, K.H. and J.H. Miao, *Coronary heart disease diagnosis using deep neural networks.* international journal of advanced computer science and applications, 2018. **9**(10).
46. Ali, L., et al., *An automated diagnostic system for heart disease prediction based on ${\chi^{2}}$ statistical model and optimally configured deep neural network.* Ieee Access, 2019. **7**: p. 34938-34945.
47. Mukherjee, S. and A. Sharma, *Intelligent heart disease prediction using neural network.* Int. J. Recent Technol. Eng., 2019. **7**(5): p. 402-405.
48. Subhadra, K. and B. Vikas, *Neural network based intelligent system for predicting heart disease.* International Journal of Innovative Technology and Exploring Engineering, 2019. **8**(5): p. 484-487.
49. Sarmah, S.S., *An efficient IoT-based patient monitoring and heart disease prediction system using deep learning modified neural network.* Ieee access, 2020. **8**: p. 135784-135797.

50. Ali, F., et al., *A smart healthcare monitoring system for heart disease prediction based on ensemble deep learning and feature fusion.* Information Fusion, 2020. **63**: p. 208-222.
51. Alharbi, A., et al., *Real-time system prediction for heart rate using deep learning and stream processing platforms.* Complexity, 2021. **2021**: p. 1-9.
52. Bavani, B., et al., *Classification of Arrhythmia Disease using Enhanced RNN Model.* Design Engineering, 2021. **1**(7): p. 4062-4072.
53. Mehmood, A., et al., *Prediction of heart disease using deep convolutional neural networks.* Arabian Journal for Science and Engineering, 2021. **46**(4): p. 3409-3422.
54. Almazroi, A.A., et al., *A clinical decision support system for heart disease prediction using deep learning.* IEEE Access, 2023. **11**: p. 61646-61659.
55. García-Ordás, M.T., et al., *Heart disease risk prediction using deep learning techniques with feature augmentation.* Multimedia Tools and Applications, 2023. **82**(20): p. 31759-31773.
56. Kamley, S., *Performance of hybrid ensemble classification techniques for prevalence of heart disease prediction.* Int. J. Innovative Technol. Exploring Eng, 2019. **8**(10): p. 1875â.
57. David, H.B.F., *Impact of ensemble learning algorithms towards accurate heart disease prediction.* ICTACT J. Soft Comput, 2020. **10**: p. 2084-2089.
58. Gupta, P. and D. Seth, *Improving the Prediction of Heart Disease Using Ensemble Learning and Feature Selection.* International Journal of Advances in Soft Computing & Its Applications, 2022. **14**(2).
59. Jan, M., et al., *Ensemble approach for developing a smart heart disease prediction system using classification algorithms.* Research Reports in Clinical Cardiology, 2018: p. 33-45.
60. Rahman, M.J.-U., et al. *Ensemble of multiple models for robust intelligent heart disease prediction system*. in *2018 4th international conference on electrical engineering and information & communication technology (ICEEiCT)*. 2018. IEEE.
61. Nikookar, E. and E. Naderi, *Hybrid ensemble framework for heart disease detection and prediction.* International Journal of Advanced Computer Science and Applications, 2018. **9**(5).
62. Habib, A.-Z.S.B., T. Tasnim, and M.M. Billah. *A study on coronary disease prediction using boosting-based ensemble machine learning approaches*. in *2019 2nd International Conference on Innovation in Engineering and Technology (ICIET)*. 2019. IEEE.
63. Chandra Shekar, K., P. Chandra, and K. Venugopala Rao. *An ensemble classifier characterized by genetic algorithm with decision tree for the prophecy of heart disease*. in *Innovations in Computer Science and Engineering: Proceedings of the Sixth ICICSE 2018*. 2019. Springer.
64. Atallah, R. and A. Al-Mousa. *Heart disease detection using machine learning majority voting ensemble method*. in *2019 2nd international conference on new trends in computing sciences (ictcs)*. 2019. IEEE.
65. Raza, K., *Improving the prediction accuracy of heart disease with ensemble learning and majority voting rule*, in *U-Healthcare Monitoring Systems*. 2019, Elsevier. p. 179-196.
66. Mienye, I.D., Y. Sun, and Z. Wang, *An improved ensemble learning approach for the prediction of heart disease risk.* Informatics in Medicine Unlocked, 2020. **20**: p. 100402.
67. Javid, I., A.K.Z. Alsaedi, and R. Ghazali, *Enhanced accuracy of heart disease prediction using machine learning and recurrent neural networks ensemble majority voting method.* International Journal of Advanced Computer Science and Applications, 2020. **11**(3).
68. Baccouche, A., et al., *Ensemble deep learning models for heart disease classification: A case study from Mexico.* Information, 2020. **11**(4): p. 207.
69. Yuan, K., et al. *Heart disease prediction algorithm based on ensemble learning*. in *2020 7th International Conference on Dependable Systems and Their Applications (DSA)*. 2020. IEEE.
70. Mehanović, D., Z. Mašetić, and D. Kečo. *Prediction of heart diseases using majority voting ensemble method*. in *CMBEBIH 2019: Proceedings of the International Conference on Medical and Biological Engineering, 16–18 May 2019, Banja Luka, Bosnia and Herzegovina*. 2020. Springer.
71. Sammout, R., et al. *A proposal of clinical decision support system using ensemble learning for coronary artery disease diagnosis*. in *Wireless Mobile Communication and Healthcare: 9th EAI International Conference, MobiHealth 2020, Virtual Event, November 19, 2020, Proceedings 9*. 2021. Springer.
72. Puvar, P., et al., *Heart Disease Detection using Ensemble Learning Approach.* Int. Res. J. Eng. Technol.(IRJET), 2021. **8**: p. 2395-0072.
73. Jothi Prakash, V. and N. Karthikeyan, *Enhanced evolutionary feature selection and ensemble method for cardiovascular disease prediction.* Interdisciplinary Sciences: Computational Life Sciences, 2021. **13**(3): p. 389-412.
74. Bharti, R., et al., *Prediction of heart disease using a combination of machine learning and deep learning.* Computational intelligence and neuroscience, 2021. **2021**.

75. Alqahtani, A., et al., *Cardiovascular Disease Detection using Ensemble Learning.* Computational Intelligence and Neuroscience, 2022. **2022**.

76. Al Reshan, M.S., et al., *A robust heart disease prediction system using hybrid deep neural networks.* IEEE Access, 2023.