

A Review: Machine Learning and Data Mining Approaches for Cardiovascular Disease Diagnosis and Prediction

Gorapalli Srinivasa Rao^{1*}, G. Muneeswari²

^{1,2}School of Computer Science and Engineering, VIT-AP University, Amaravathi, Andhra Pradesh, India

Abstract

INTRODUCTION: Cardiovascular disease (CVD) is the most common cause of death worldwide, and its prevalence is rising in low-resource settings and among those with lower incomes.

OBJECTIVES: Machine learning (ML) algorithms are quickly evolving and being implemented in medical procedures for CVD diagnosis and treatment decisions. Every day, the healthcare business creates massive amounts of data. However, the majority of it is inadequately utilized. Efficient techniques for extracting knowledge from these datasets for clinical diagnosis or other uses are scarce.

METHODS: ML is being applied in the healthcare industry all over the world. In the health dataset, ML approaches useful in the prevention of locomotor disorders and heart disease.

RESULTS: The revelation of such vital information allows researchers to acquire significant insight into how to use the proper treatment and diagnosis for a specific patient. Researchers study enormous volumes of complex healthcare data using various ML approaches, which improves healthcare professionals in disease prediction.

CONCLUSION: The goal of this study is to summarize some of the current research on predicting heart diseases utilizing machine learning and data mining techniques, analyze the various mining algorithm combinations employed, and determine which techniques are useful and efficient. Future directions in prediction systems have also been considered.

Keywords: Heart Diseases, Machine Learning, Ensemble Models, Data Mining, Dataset, Classification

Received on 08 December 2023, accepted on 07 March 2024, published on 13 March 2024

Copyright © 2024 G. Srinivasa Rao *et al.*, licensed to EAI. This is an open access article distributed under the terms of the [CC BY-NC-SA 4.0](#), which permits copying, redistributing, remixing, transformation, and building upon the material in any medium so long as the original work is properly cited.

doi: 10.4108/eetpht.10.5411

1. Introduction

The heart is one of the most vital organs in human body. It is a muscular organ that pumps blood into the body and serves as the heart of the cardiovascular system. The cardiovascular system contains all blood vessels, including veins, capillaries, and arteries, which together constitute difficult blood circulation throughout the body [1-3]. Any restriction or aberration in normal blood circulation flow from the heart can lead to some significant heart disease complications. These are referred to as CVDs, and they are among the deadliest illness in the world. CVDs include heart disorders, vascular diseases of the brain, and blood vessel diseases [4, 5].

Although CVDs can be managed through lifestyle changes and other related measures, all indicators indicate that they are on the rise daily, as mentioned in several WHO studies. However, many WHO publications have shown a global increase in CVDs, which is highly concerning. CVDs kill more people than any other cause worldwide, killing an estimated 17.5 million individuals in 2012 [6, 7]. According to numerous WHO estimates, mortality from heart disease is on the rise, which is mostly linked to insufficient preventative actions despite rising risk factors.

Clinical information has revealed that some risk factors create a person's chances of building CVD. Some of these risk factors include a family heritage of cardiovascular disease, a bad LDL cholesterol level, a low good cholesterol level, a high-fat diet, hypertension, obesity, and a lack of active

*Corresponding author. Email: srinivasarao.22phd7042@vitap.ac.in

exercise. Smoking, diabetes, age, and gender are all risk factors [8]. Using these and other characteristics, physicians typically form diagnosis by examining a patient's present health status as well as previous diagnosis made on other patients in the same situation.

The increasing incidence of heart disease has become a popular problem. As an output, the healthcare industry must denote and enhance the way these illnesses are treated to reduce impact. Huge amounts of data are available in the healthcare industry, particularly data on CVD, which must be efficiently analyzed to make successful conclusions. According to information, statistics, hospital administration, and clinical records, medical data doubles every three years, making the health sector a multibillion-dollar domain. Medical data analysis and knowledge extraction rely heavily on machine learning and data mining methods [9, 10]. The rising mortality and morbidity rates from CVD have influenced authors to perform numerous studies to reduce the rates. Machine learning and data mining approaches have been commonly utilized in the development of the prediction of cardiac disease. Data mining applications are utilized to prevent clinical mistakes and enhance health policy, early illness identification, and preventable hospital fatalities. This review's critical contribution is,

- In this survey study, we review current research on heart disease prediction in different domains from 2021 to 2023.
 - The fundamental goal of problem-solving is to apply ML and DM approaches in heart disease recognition; IoT-based applications used in CVD are the primary goal of problem-solving.
 - We examined the performance of the heart disease recognition systems based on their accuracy.
- The following is the review's structure: Sections 2-6 outline the research methodologies utilized to choose the primary studies. Section 7 delves into and evaluates the CVD category review. Section 8 depicts CVD's research gap. Section 9 brings the paper to a close.

2. Machine Learning Based Heart Disease Prediction

Srinivas & Katarya [11] introduced HyOPTXg, an expert model that predicts heart illness using an optimized XGBoost classifier. They need good hyperparameter tweaking to create a better system with a classifier. As an outcome, they changed XGBoost's hyper-values and trained the model utilizing tuned parameters. OPTUNA (hyper-parameter optimization method) is the structure utilized for hyper-parameter tuning. This method has been evaluated on the Kaggle heart disease UCI, the Kaggle Heart Failure prediction dataset, and the Cleveland dataset.

Rani et al. [12] presented a hybrid method to aid in the early detection of CVD. To deal with missing values, the authors used multivariate imputation via chained equations. For the

selection of necessary attributes from the provided dataset, a hybrid attribute selection technique combining the GA and recursive attribute elimination was utilized. SMOTE (Synthetic Minority Oversampling Technique) and traditional scalar methods were also employed for data pre-processing. The researchers utilized LR, NB, RF, SVM, and Adaboost classifiers in the presented hybrid model. It was discovered that the RF classifier produced the best accurate results.

Nadeem et al. [13] presented an SVM-based structure for heart disease detection that is enabled by a fuzzy-based decision-level fusion. The preprocessing layer collected original data, which may contain missing values and noise. Different approaches, such as average, mode, and mean, are used in the pre-processing layer to forecast missing data and reduce noise using normalization. Furthermore, the processed data is received by the application layer and utilized to train the supervised ML approach known as SVM. Within the suggested structure, the same procedure is implemented in parallel.

In this research, Budholiya et al. [14] suggested a diagnostic system that predicts cardiac illness using an optimized XGBoost (Extreme Gradient Boosting) classifier. They utilized Bayesian optimization to optimize the hyper-parameters of XGBoost, which is a very effective approach for parameter optimization. Table 1 demonstrates the disadvantages and advantages of ML approaches in heart disease prediction.

Table 1. ADVANTAGES AND DISADVANTAGES OF ML APPROACHES IN HEART DISEASE PREDICTION.

References	Methods	Datasets	Advantages	Disadvantages	References
Srinivas & Katarya [11]	HyOPTXg	UCI open dataset			It predicts heart disease by training the model utilizing tuned hyperparameters to produce good outcomes than other systems.
Rani et al. [12]	SVM, NB, LR, RF, Adaboost	Cleveland Heart disease dataset			It can also assist the doctor in taking quick decisions.
Nadeem et al. [13]	SVM	Heart Disease Dataset kaggle			The fuzzy-based decision-level fusion has achieved above 96% accuracy.

Budholiya et al. [14]	Optimized XGBoost	Cleveland Heart Disease Dataset	The diagnostic procedure would increase the accuracy of decision-making during the heart disease diagnosis process. It predicts heart disease by training the model utilizing tuned hyperparameters to produce good outcomes than other systems.
Srinivas & Katarya [11]	HyOPTXg	UCI open dataset	

3. Heart Disease Prediction Based on Ensemble Models

Gao et al. [15] used supervised learning methods to predict cardiac disease in its early stages. The HD dataset is utilized to test and train systems. It has 1025 documents, 13 attributes, and 1 target column. The goal column has two classes: 1 for heart diseases and 0 for nonheart disorders. The attributes are scaled to be between [0, 1]. It should be noted that missing values are removed from the dataset. The extraction of the optimal features is critical since irrelevant features frequently influence the ML classifier's categorization effectiveness. To identify essential features from the dataset, LDA and PCA are used in this phase. To identify whether the people tested have heart disease or are healthy, ensemble methods and numerous algorithms such as SVM, NB, DT, KNN, and RF are employed.

Uddin & Halder [16] suggested a multilayer dynamic system (MLDS) based on an ensemble method that may increase its current knowledge in each layer. For attribute selection, the suggested approach employs the Extra Trees classifier (ETC), Gain Ratio Attribute Evaluator (GRAE), Correlation Attribute Evaluator (CAE), Lasso, and Information Gain Attribute Evaluator (IGAE). Finally, the ensemble technique for categorization in the structure is built using RF, NB, and GB classifiers. While the base classifiers stated above failed to identify accurately in any layer, the KNN technique was used to find the test information's neighbourhood data points.

Mahesh et al. [17] suggested employing ML to determine whether or not an individual has a cardiac illness. This work implements both forms of ensemble classifiers, namely heterogeneous and homogeneous classifiers. To deal with the class imbalance as well as noise, data mining preprocessing

utilized SMOTE was utilized. The planned work consists of two steps. In the first stage, SMOTE was employed to lessen the influence of data imbalance, and in the second stage, information is classified utilizing NB, DT, and their ensembles. Table 2 demonstrates the disadvantages and advantages of ensemble models in heart disease prediction.

Table 2. ADVANTAGES AND DISADVANTAGES OF ENSEMBLE MODELS IN HEART DISEASE PREDICTION.

References	Methods	Datasets	Advantages	Disadvantages	References
Gao et al. [15]	KNN, SVM, DT, NB, and RF		Heart Dataset Kaggle	Disease from	The bagging ensemble learning method using PCA and DT feature extraction approach performed well. The suggested framework can be utilized to forecast the risk of CVD and to provide useful clinical guidance.
Uddin & Halder [16]	RF, NB, and GB		Hungarian, Cleveland, and Cleveland-Hungary-Switzerland-Long Beach datasets	and	The performance of the model is high using the ensemble technique
Mahesh et al. [17]	NB, DT		UCI Heart Disease Dataset		

4. Data Mining-Based Heart Disease Prediction

Kavitha et al. [18] suggested an ML method for predicting cardiac disease. The suggested research makes use of the Cleveland heart disease dataset, as well as data mining methods including classification and regression. RF and DT machine learning algorithms are used. The ML model's innovative method was developed. Three ML algorithms are employed in the implementation using Hybrid methods (RF and DT). We employed hybrid methods of DT and RF to detect heart illness using the user's input parameters.

Yazdani et al. [19] used the heart illness dataset from the UCI ML Repository. During the pre-processing stage, missing

values were eliminated from the dataset's six instances. Using WARM, they explored every attribute as well as chosen important attributes. According to the findings, the important factors beat all other attributes with the greatest confidence score in forecasting HD.

Premsmith & Ketmaneechairat [20] suggested a methodology for detecting cardiac disease using data mining approaches. The Logistic Regression model and the Neural Network model are used by the data mining algorithm. This research dataset is based on heart disease data from UCI. In the United States, there are 303 Instances and 75 Attributes. The confusion matrix table is used to evaluate characteristics such as precision, accuracy, F-Measure, and recall. According to the results, the LR model outperforms the NN model.

Valarmathi & Sheela [21] presented a system for predicting heart disease. To enhance the performance of the XG Boost and RF classifier models, three HPO strategies were proposed: Randomized Search, Grid Search, and TPOT Classifier. The effectiveness of the two methods RF and XG Boost was compared to previous studies. The models' performance is examined using the publicly accessible datasets Z-Alizadeh Sani dataset and Cleveland Heart Disease Dataset (CHD).

Deepika & Balaji [22] suggested a model for accurately predicting CVD utilizing a attribute selection and categorization method. In this paper, an optimized unsupervised algorithm for attribute chosen and a MLP-EBMDA for CVD categorization are suggested. As input, the heart disease dataset is acquired, and pre-processing is performed. The optimized unsupervised method was utilized to choose attributes. The hybrid MLP-EBMDA approach was utilized to categorize HD based on the specified attributes. Table 3 demonstrates the disadvantages and advantages of data mining methods in heart disease forecasting.

Table 3. DISADVANTAGES AND ADVANTAGES OF DATA MINING METHODS IN HEART DISEASE FORECASTING.

Referen ces	Metho ds	Datas ets	Advantag es	Disadvant ages	Referenc es
Kavitha et al. [18]	DT and RF		Cleveland Heart Disease Dataset		This system was predicting heart disease by an automated machine learning diagnosis model with a hybrid model. And gave better prediction

Yazdani et al. [19]	WARM	UCI open dataset			outcomes The ML approaches utilized in this research's attribute selection phase are limited to the most common methods utilized in heart disease predictions.
Premsmith & Ketmaneechairat [20]	LR and NN	UCI Heart Disease Dataset			The model forecasts HD is better performance than neural networks
Valarmathi & Sheela [21]	RF and XG Boost	Cleveland Heart Disease Dataset			The model effectively predicts disease. MLP-EBMDA correctly detects patients heart problems and successfully monitors patients heart conditions over time.
Deepika & Balaji [22]	MLP-EBMDA	-			

5. IOT-Based Heart Disease Prediction

Nancy et al. [23] suggested a system gather data from IoT devices, and electronic clinical data relevant to patient history saved on the cloud is subjected to forecasting analytics. The suggested smart medical model for forecasting the risk of HD contains phases such as (1) data gathering, (2) pre-processing, and (3) illness prediction. At the Cloud layer, the information acquired from IoT sensors for HD risk forecasting is subjected to pre-processing steps such as filtering and cleaning. The resulting data is transmitted to the FIS for the first categorization. Then, the suggested Bi-LSTM algorithm was employed to accurately forecast a patient's risk of heart illness.

The IoT is an essential technology in a healthcare system. Malibari [24] suggested the EO-LWAMCNet framework to reliably forecast a patient's chronic health state in this study. A sensor implanted in the patient's body is capable of collecting all data and transmitting it to the cloud via a gateway. The EO-LWAMCNet framework begins the categorization process to forecast chronic disease based on the sensor data obtained. The model goes through testing and training. CKD and HD databases are used to forecast the disease. The processed information was used in the training stage for categorization. Following the completion of the training procedure, the CS sensor data was tested and classified as normal or abnormal. In the event of an aberrant result, the doctor receives an alert message to treat the patient.

In this paper, Yaqoob et al. [25] introduced a system with an M-ABC optimization method to eliminate privacy concerns and enhance the detection technique for heart disease identification. Using an updated federated learning method for user sites and the cloud, they create and suggest a privacy-aware system for predicting heart disease in healthcare. At the client end, the M-ABC optimizer was suggested for the optimal selection of features of heart illness information. A framework for a global cloud system is investigated using a federated matched averaging (FedMA)-based approach. Table 4 shows the disadvantages and advantages of IoT-based approaches in heart disease prediction.

Table 4. DISADVANTAGES AND ADVANTAGES OF IOT-BASED APPROACHES IN HEART DISEASE PREDICTION.

References	Methods	Datasets	Advantages	Disadvantages	References
Nancy et al. [23]	Bi-LSTM		Cleveland Hungarian Datasets	&	The IoT-based devices used for data acquisition and the proposed smart heart disease prediction system occur higher prediction accuracy

Malibari [24]	EO-LWAMCNet	CKD and HD Datasets			The model effectively predicts disease.
---------------	-------------	---------------------	--	--	---

6. Deep Learning-Based Heart Disease Prediction

Mehmood et al. [26] developed CardioHelp, a method that recognizes the existence of CVD in a patient by combining a Deep Learning (DL) technique known as convolutional neural networks (CNN). At its earliest stage, the suggested approach is concerned with temporal information modeling by employing CNN for HD forecasting. They produced the heart disease dataset and compared the outcomes using cutting-edge methodologies. In terms of performance assessment measures, experimental results reveal that the suggested strategy outperforms the existing approaches.

The real-time dataset and UCI heart disease dataset are utilized to compare DL approaches to traditional approaches. Dileep et al. [27] developed C-BiLSTM to increase the accuracy of existing approaches. The real-time and UCI heart disease datasets are utilized for performance findings, and both datasets are passed through the K-Means clustering method to remove duplicate information, and the HD predicted by the C-BiLSTM method. C-BiLSTM was compared to traditional classifier approaches such as RT, SVM, LR, KNN, Gated Recurrent Unit, and Ensemble. Table 5 shows the disadvantages and advantages of DL methods in HD prediction.

Table 5. DISADVANTAGES AND ADVANTAGES OF DL METHODS IN HD FORECASTING.

References	Methods	Datasets	Advantages	Disadvantages	References
Mehmood et al. [26]	Cardio Help			UCI Dataset	Accurately identify the widespread occurrences.
Dileep et al. [27]	C-BiLSTM			UCI Cleveland	Accurately identify the widespread occurrences.

7. Results and Discussion

This review focuses on the role of IoT, Data Mining, DL, and ML approaches in heart disease prediction. These algorithms were tested using public datasets like the UCI open dataset, the Heart Disease Dataset, the Cleveland heart disease dataset, and the Cardiovascular Disease Dataset, as well as local datasets. Table 6 shows the systems that were reviewed.

Table 6. DISADVANTAGES AND ADVANTAGES OF IOT-BASED APPROACHES IN HEART DISEASE PREDICTION.

Ref.	Methods	Dataset	Accuracy	Precision	Specificity	Recall	F1-Score
[11]	HyOPTXg	UCI open dataset	94.7%	87.03%	-	94%	90.38%
[12]	SVM, NB, LR, RF, and AdaBoost	Cleveland heart disease dataset	86.6%	88.46%	89.02%	84.14%	86.25%
[13]	SVM	Heart Disease Dataset 2019, Cardiovascular Disease Dataset 2019	96.23%	94.36%	95.64%	97.01%	-
[14]	Optimized XGBoost	Cleveland Heart Disease Dataset	91.8%	96.96%	-	85.71%	90.56%
[15]	KNN, SVM, DT, NB, and RF	Heart Disease Dataset	98.6%	97.8%	-	99%	98.5%
[16]	RF, NB, and GB	Hungarian, Cleveland, and Cleveland-Hungary-Switzerland-Long Beach datasets	98.88%	95.45%	91.11%	81.58%	-
[17]	NB, DT	UCI open dataset	95.47%	-	-	-	98%
[18]	RF and DT	Cleveland heart disease dataset	88.7%	-	-	-	-
[19]	WARM	UCI open dataset	98%	-	-	-	-
[20]	LR and NN	UCI Cleveland heart disease dataset	91.65%	95.45%	-	84%	89.36%
[21]	RF and XGBoost	Cleveland Heart Disease Dataset	97.52%	97%	97.70%	97.29%	97%
[22]	MLP-EBMDA	-	94.28%	96%	96%	96%	96%
[23]	Bi-LSTM	Cleveland and Hungarian dataset	98.86%	98.9%	98.89%	98.8%	98.86%
[24]	EO-LWAMCNet	HD dataset	94%	91.70%	-	91.82%	-
[25]	FedMA	Cleveland Heart Disease dataset	92.89%	94.2%	81.8%	96.6%	90.1%
[26]	CardioHelp	Heart Disease Dataset	97%	97.06%	-	96.35%	96.70%
[27]	C-BiLSTM	UCI open dataset	94.78%	-	94.18%	99.56%	-

According to the findings of the previous study, the suggested machine learning algorithms of Random Forest, Nave Bayes, and Gradient Boost produced higher detection accuracy than other studies. To choose the necessary features to predict heart disease, this article used GRAE, the CAE, Lasso, IGAE, ETC. The classifiers RF, NB, and GB were then used to determine whether the patient had heart disease or not. This technique's performance was examined using Hungarian, Cleveland, and Cleveland-Hungary-Switzerland-Long Beach datasets. Figure 1 shows the DL and ML techniques utilized in PA applications.

8. Limitations

The limitations and advantages of the reviewed HD prediction methods in the preceding literature have been summarized in Tables for a better understanding of the suggested methodology. All treatments used different methodologies to recognize HD in its early stages. However, all of these algorithms have low prediction accuracy and a long computation time for HD detection. The HD identification approaches prediction accuracy requires further refinement for efficient and precise detection at early stages for improved treatment and recovery. Thus, the key concerns with the existing methodologies are long computation times and low accuracy, which may be attributed to the usage of irrelevant attributes in the dataset. To overcome these issues, new approaches for identifying HD are required. Improving forecast accuracy is a significant challenge to the researchers.

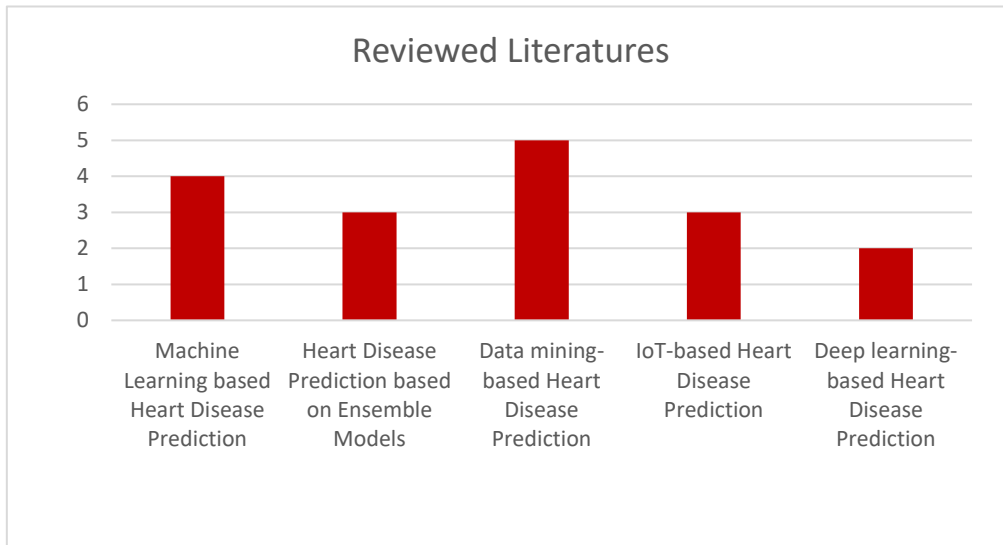


Fig. 1. DL and ML techniques utilized in PA applications

9. Conclusion

The use of several data mining and ML methods for predicting the occurrence of heart disease has been summarized. Determine every algorithm's prediction accuracy and apply the suggested approach to the required area. Improve algorithm accuracy by using more relevant attribute selection methods. If a patient is diagnosed with a specific type of cardiac disease, there are numerous treatments available. Data mining and ML may provide an enormous amount of information from an appropriate dataset. In conclusion, the literature survey revealed that only marginal success is reached in the development of predictive models for heart disease patients, indicating the need for combinational and more complex systems to improve the accuracy of identifying the early onset of heart disease. The more data that is supplied into the database, the more intelligent the system becomes.

An automated model that can assist in the choice of suitable treatment approaches for a heart disease patient may be developed in the future. A lot of studies have already been put into establishing systems that can detect whether an individual is going to develop cardiovascular disease or not. When an individual is identified with a specific form of heart disease, he or she has several treatment options. Machine Learning can be quite useful in selecting the type of treatment to be taken by retrieving information from such relevant databases. In addition, we are interested in classifying it as a multi-class problem to determine the disease's level. In order to more accurately and effectively anticipate heart problems, an ideal method needs to be developed.

References

- [1] U. Nagavelli, D. Samanta & P. Chakraborty, "Machine learning technology-based heart disease detection models," *Journal of Healthcare Engineering*, 2022.
- [2] A. Rath, D. Mishra, G. Panda & S. C. Satapathy, "Heart disease detection using deep learning methods from imbalanced ECG samples," *Biomedical Signal Processing and Control*, vol. 68, 102820, 2021.
- [3] S. Ketu & P. K. Mishra, "Empirical analysis of machine learning algorithms on imbalance electrocardiogram based arrhythmia dataset for heart disease detection," *Arabian Journal for Science and Engineering*, pp. 1-23, 2022.
- [4] R. Arnaout, L. Curran, Y. Zhao, J. C. Levine, E. Chinn & A. J. Moon-Grady, "An ensemble of neural networks provides expert-level prenatal detection of complex congenital heart disease," *Nature Medicine*, vol. 27(5), pp. 882-891, 2021.
- [5] F. HASSAN, A. U. Rahman, A. Javed, A. Alhazmi & M. Alhazmi, "CNN-CardioAssistant: Deep Convolutional Neural Network and Recursive Feature Elimination Method for Heart Disease Detection," 2022.
- [6] F. F. Firdaus, H. A. Nugroho & I. Soesanti, "Deep neural network with hyperparameter tuning for detection of heart disease," In 2021 IEEE Asia Pacific Conference on Wireless and Mobile (APWiMob) IEEE, pp. 59-65, 2021.
- [7] M. Alkhodari & L. Fraiwan, "Convolutional and recurrent neural networks for the detection of valvular heart diseases in phonocardiogram recordings," *Computer Methods and Programs in Biomedicine*, vol. 200, 105940, 2021.
- [8] T. S. Roy, J. K. Roy & N. Mandal, "Classifier identification using deep learning and machine learning algorithms for the detection of valvular heart diseases," *Biomedical Engineering Advances*, vol. 3, 100035, 2022.
- [9] A. Thakur, H. Kaur, N. Goel, P. Paul, P. Asopa, S. Goswami & M. K. Das, "A Hybrid Approach for Heart Disease Detection using K-Means and K-NN Algorithm," *American Journal of Electronics & Communication*, vol. 4(1), pp. 14-21, 2023.

- [10] I. D. Mienye & Y. Sun, "Improved heart disease prediction using particle swarm optimization based stacked sparse autoencoder," *Electronics*, vol. 10(19), 2347, 2021.
- [11] P. Srinivas & R. Katarya, "hyOPTXg: OPTUNA hyper-parameter optimization framework for predicting cardiovascular disease using XGBoost," *Biomedical Signal Processing and Control*, vol. 73, 103456, 2022.
- [12] P. Rani, R. Kumar, N. M. S. Ahmed & A. Jain, "A decision support system for heart disease prediction based upon machine learning," *Journal of Reliable Intelligent Environments*, vol. 7(3), pp. 263-275, 2021.
- [13] M. W. Nadeem, H. G. Goh, M. A. Khan, M. Hussain & M. F. Mushtaq, "Fusion-Based Machine Learning Architecture for Heart Disease Prediction. Computers," *Materials & Continua*, vol. 67(2), 2021.
- [14] K. Budholiya, S. K. Shrivastava & V. Sharma, "An optimized XGBoost based diagnostic system for effective prediction of heart disease," *Journal of King Saud University-Computer and Information Sciences*, vol. 34(7), pp. 4514-4523, 2022.
- [15] X. Y. Gao, A. Amin Ali, H. Shaban Hassan & E. M. Anwar, "Improving the accuracy for analyzing heart diseases prediction based on the ensemble method," *Complexity*, pp. 1-10, 2021.
- [16] M. N. Uddin & R. K. Halder, "An ensemble method based multilayer dynamic system to predict cardiovascular disease using machine learning approach," *Informatics in Medicine Unlocked*, vol. 24, 100584, 2021.
- [17] T. R. Mahesh, V. Dhilip Kumar, V. Vinoth Kumar, J. Asghar, O. Geman, G. Arulkumaran & N. Arun, "AdaBoost ensemble methods using K-fold cross validation for survivability with the early detection of heart disease," *Computational Intelligence and Neuroscience*, 2022.
- [18] M. Kavitha, G. Gnaneswar, R. Dinesh, Y. R. Sai, & R. S. Suraj, "Heart disease prediction using hybrid machine learning model," In 2021 6th international conference on inventive computation technologies (ICICT) IEEE, pp. 1329-1333, 2021.
- [19] A. Yazdani, K. D. Varathan, Y. K. Chiam, A. W. Malik & W. A. Wan Ahmad, "A novel approach for heart disease prediction using strength scores with significant predictors," *BMC medical informatics and decision making*, vol. 21(1), 194, 2021.
- [20] J. Premsmith & H. Ketmaneechairat, "A predictive model for heart disease detection using data mining techniques," *Journal of advances in Information Technology*, vol. 12(1), 2021.
- [21] R. Valarmathi & T. Sheela, "Heart disease prediction using hyper parameter optimization (HPO) tuning," *Biomedical Signal Processing and Control*, vol. 70, 103033, 2021.
- [22] D. Deepika & N. Balaji, "Effective heart disease prediction using novel MLP-EBMDA approach," *Biomedical Signal Processing and Control*, vol. 72, 103318, 2022.
- [23] A. A. Nancy, D. Ravindran, P. D. Raj Vincent, K. Srinivasan & D. Gutierrez Reina, "IoT-cloud-based smart healthcare monitoring system for heart disease prediction via deep learning," *Electronics*, vol. 11(15), 2292, 2022.
- [24] A. A. Malibari, "An efficient IoT-Artificial intelligence-based disease prediction using lightweight CNN in healthcare system," *Measurement: Sensors*, 26, 100695, 2023.
- [25] M. M. Yaqoob, M. Nazir, A. Yousafzai, M. A. Khan, A. A. Shaikh, A. D. Algarni & H. Elmannai, "Modified Artificial Bee Colony Based Feature Optimized Federated Learning for Heart Disease Diagnosis in Healthcare," *Applied Sciences*, vol. 12(23), 12080, 2022.
- [26] A. Mehmood, M. Iqbal, Z. Mehmood, A. Irtaza, M. Nawaz, T. Nazir & M. Masood, "Prediction of heart disease using deep convolutional neural networks," *Arabian Journal for Science and Engineering*, vol. 46(4), pp. 3409-3422, 2021.
- [27] P. Dileep, K. N. Rao, P. Bodapati, S. Gokuruboyina, R. Peddi, A. Grover & A. Sheetal, "An automatic heart disease prediction using cluster-based bi-directional LSTM (C-BiLSTM) algorithm," *Neural Computing and Applications*, vol. 35(10), pp. 7253-7266, 2023.