STM JOURNALS

Title of the STM Journals

http://stmjournals.com/journals-index.html

ISSN: 2278-2257 (Online)
ISSN: 2348-9553 (Print)
Volume 10, Issue 1, 2021
DOI (Journal): 10.37591/ABBREV

Type of Article ABBREVIATION

Trends in Mechanical Engineering and Technology
ISSN: 2231-1793
Article type: Review
Volume:14 Issue: 2
Received Date: 9 May 2024
Accepted Date: 2 July 2024
Published Date: 17 July 2024

Heart Disease AI-Based Prediction: A Comparative Analysis

Shivam Saini[1*], Ujjwal Thakur[1], Shashwat Verma[1], Anand Sehgal[2]

[1]Student, Department of computer science and engineering, The Northcap University, Gurgaon, Haryana, India

[2]Assistance Professor, Department of computer science and engineering, The Northcap University, Gurgaon, Haryana, India

E-mail: shivam.saini.02707@gmail.com

Abstract

The present investigation looks at how well various machine learning algorithms predict cardiac disease. Since heart disease is one of the major causes of death worldwide, early detection and precise diagnosis are essential for managing and treating the condition. Our goal is to enhance diagnostic processes and improve patient outcomes by leveraging machine learning techniques.

Six widely-used machine learning algorithms are evaluated in this research paper. These algorithms were selected due to their established effectiveness in classification tasks. Numerous patient characteristics, including age, gender, blood pressure, cholesterol, and other pertinent medical data, are included in the dataset. To evaluate each algorithm's performance, we separated the dataset into subsets for testing and training.

Our study's primary goal is to evaluate each algorithm's predictive power for heart disease. Various performance indicators are employed for both the training and testing datasets, such as accuracy, precision, recall, and F-measure. These metrics give a comprehensive view of each model's predictive capabilities and how well they generalize.

This research has significant implications for the use of machine learning in healthcare. By identifying the strengths and weaknesses of different algorithms in predicting heart disease, we provide valuable insights that can help develop more reliable and accurate prediction models. This study not only broadens our understanding of machine learning applications in healthcare but also sets the stage for future research aimed at improving patient care through advanced predictive analytics.

Our study underscores the potential of machine learning to revolutionize diagnostic processes and improve outcomes for patients at risk of heart disease. By using these advanced algorithms, we can move towards more effective and personalized healthcare solutions.

Keywords: Machine Learning, Heart Disease, Neural Network, Kaggle, Evaluation.

INTRODUCTION

Cardiac disease remains a significant cause of mortality globally, imposing a substantial burden on both the healthcare system and society. Timely and accurate heart disease prediction can lead to early intervention and preventative measures, potentially saving lives and reducing medical costs.

In this research study focused on Assessing heart disease prediction through a comparative analysis of various machine learning (ML) algorithms. is conducted. These algorithms commonly comprise Logistic Regression, Support Vector Machine (SVM), k-Nearest Neighbors (kNN), Naive Bayes (NB), Decision Tree (DT), and Multilayer Perceptron (MLP). The objective of this comparison is to offer insights into the effectiveness, advantages, and constraints of each algorithm for this specific task.

Each selected algorithm offers unique features and mechanisms for pattern recognition and classification. Logistic Regression utilizes statistical methods for binary classification, while SVM optimizes hyperplanes to segregate data points effectively. kNN relies on similarity principles, Naive Bayes on probabilistic models, Decision Tree on hierarchical structures, and Neural Networks on mimicking the brain's architecture. Understanding each of their performances can provide valuable information to healthcare professionals and researchers, aiding them in selecting the most appropriate approaches for various scenarios.

Moreover, this paper delves into the evaluation metrics commonly employed to analyze prediction models, this encompasses metrics such as accuracy, recall, precision, and the F-measure, offering a thorough insight into the model's performance, allowing for comparisons across various algorithms and facilitating informed decision-making regarding their deployment in clinical practice.

In summary, this research aims to enhance the existing knowledge on AI-based heart disease prediction by undertaking a thorough comparison examination of several ML algorithms. The findings of this research hold significant potential to offer valuable insights for healthcare professionals, researchers, and policymakers in developing effective strategies for early identification and treatment of cardiac disease. By leveraging the outcomes of this study, it becomes possible to enhance patient outcomes and public health. For information on the prediction system's architecture, Figure 1 .

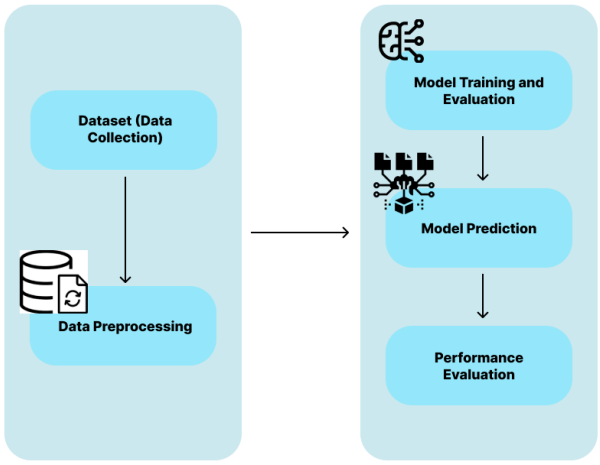

Figure 1. Architecture of the Prediction System.

LITERATURE REVIEW

Cardiac disease is a major global health concern, with its incidence steadily rising around the world. Worldwide, heart disease carries a substantial economic burden. Healthcare expenditures and productivity losses due to heart disease amount to billions of dollars annually [1][2].

Artificial intelligence (AI) and machine learning (ML) have become much more popular recently. Techniques in healthcare has shown promising results across various domains, including illness prognosis and diagnosis. A study shows that using AI and ML algorithms in predicting cardiac disease has received significant interest because of their ability to enhance accuracy and efficiency in assessing risk [3].

Many research efforts have delved into the utilization of machine learning (ML) algorithms for predicting cardiac disease. For instance, [4] explored the utilization of ML algorithms within the realm

of cardiovascular medicine. The purpose was to analyze genotypes and phenotypes, enhance the quality of patient treatment, decrease readmission and death rates, and improve cost-effectiveness. Similarly, a thorough examination was carried out by [5] to evaluate the potential advantages of AI in the fields of cardiology and cardiac imaging, with the aim of enhancing patient care[6].

Metrics for evaluation are essential for determining how well machine learning algorithms forecast cardiac disease.. Machine learning greatly improves the precision of cardiovascular disease prediction, therefore finding more patients who might benefit from preventative medication, while also preventing unneeded treatment for others [7].

In summary, the literature study emphasizes the significance of choosing the right algorithms, developing effective features, and using appropriate assessment metrics to make judgments on the accuracy of predictions. The objective of this research is to make a significant contribution to ongoing efforts aimed at developing efficient strategies for early detection. This will be achieved through a comparative analysis of multiple machines learning algorithms, evaluating their performance using standardized metrics[8].

METHODOLOGY / EXPERIMENTAL

This section is a comprehensive explanation of the methods used in this study. The process includes choosing algorithms, selecting datasets, applying preprocessing methods, defining parameters, and employing evaluation metrics to gauge the effectiveness of each method(FIG.2).

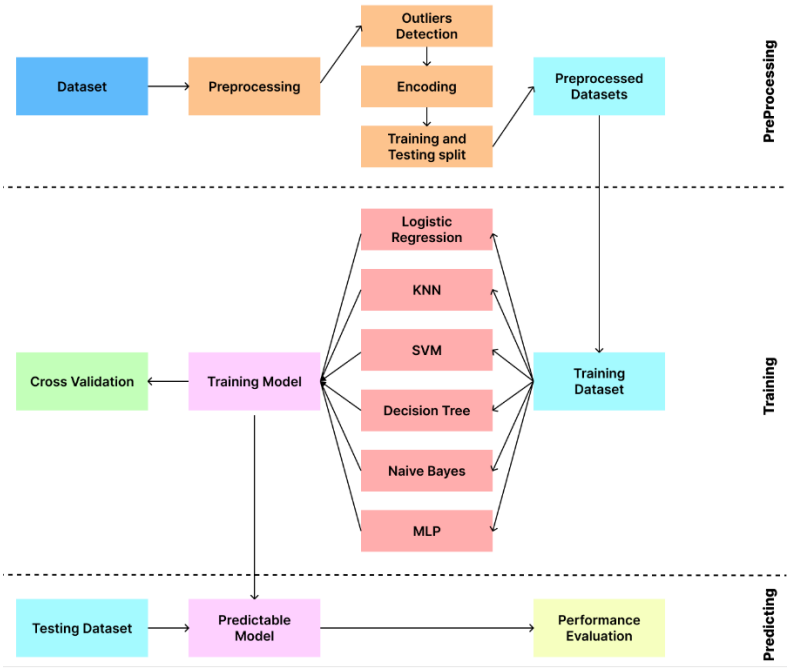

Figure 2. A Detailed Procedure's Steps.

Selected Algorithms

Accurate prediction of heart disease relies heavily on the utilization of machine learning algorithms. These algorithms offer extensive coverage of predictive methodologies and adeptly capture a broad spectrum of data points. In the following section, concise explanations of each selected algorithm will be provided[9].

Logistic Regression

One statistical technique that is commonly used for binary classification tasks is logistic regression. It models the likelihood that an event will occur based on input features by fitting a logistic curve to the available data[10].

Support Vector Machine (SVM)

The provided description appears to be referring to Support Vector Machine (SVM), rather than Logistic Regression.

k-Nearest Neighbors (kNN)

It functions by identifying the K nearest data points to a given query point and then utilizes the majority class (for classification) or the average values (for regression) of those neighbors to generate recommendations. The selection of K impacts the model's adaptability since lower values provide more intricate models while bigger values produce smoother decision limits[11].

Naive Bayes

The provided description accurately characterizes the Naive Bayes (NB) algorithm. The algorithm computes the likelihood of each category based on a given set of input characteristics and determines the category with the highest likelihood as the prediction. Although it is simple and may not adhere to certain assumptions, it frequently achieves good results in real-world scenarios, particularly in jobs involving text categorization.

Decision Tree

It is a technique that constructs a hierarchical structure in the form of a tree to make judgments by evaluating the values of input attributes. The process involves dividing the feature space into subsets repeatedly, using a specified criterion such as information gain or Gini impurity to optimize each split. Decision trees are valued for their intuitive interpretation, as they provide clear decision rules, and its capability to capture nonlinear relationships within data.

Multilayer Perceptron (MLP)

Multilayer Perceptron (MLP) is a form of artificial neural network (ANN) characterized by multiple layers of interconnected neurons. Neurons within the network are typically connected to neurons in adjacent layers, allowing for complex interconnectedness. Allowing for complex patterns to be learned and represented. They are excellent at collecting complex data patterns and are widely used in healthcare and other fields.

Dataset Description

The dataset utilized in this study was obtained from Kaggle, a well-known platform for hosting data science competitions and datasets. This dataset contains a wide range of clinical characteristics pertinent to heart disease prediction. The attributes included in the dataset are listed in Table 1:

Table 1. Feature Description.

Algorithm	Description
Age	Patient's age in years.
Sex	The gender is classified as either Male (M) or Female (F).
ChestPainType	The patient's chest pain can be classified into four categories: Typical Angina (TA), Atypical Angina (ATA), Non-Anginal pain (NAP), or Asymptomatic (ASY).
RestingBP	Resting blood pressure is expressed in millimeters of mercury (mm Hg).
Cholesterol	The serum cholesterol level is measured in milligrams per deciliter (mg/dl).
FastingBS	The fasting blood sugar level is represented by a binary variable, where a value of 1 denotes a fasting blood sugar level above 120 mg/dl, and a value of 0 represents a fasting blood sugar level less or equal to 120 mg/dl.
RestingECG	The resting ECG results can be classified as normal, showing ST-T wave abnormalities (ST), or left ventricular hypertrophy (LVH).
MaxHR	The maximum heart rate reached during testing is expressed as a numerical number ranging from 60 to 202.
ExerciseAngina	Indication of angina caused by exercise, presented as Yes or No.
Oldpeak	Exercise-induced ST depression, measured as a numerical value compared to rest. .
ST_Slope	The incline of the peak workout ST phase can be categorised as either up-sloping (upward), flat, or down-sloping (downward).
HeartDisease	It indicates the presence (1) or absence (0) of heart disease.

This comprehensive collection of attributes offers useful data for studies on predicting heart disease.

Pre-Processing

This section discusses how we applied several techniques to prepare the datasets for analysis. In particular, we tackled outliers utilizing the Interquartile Range (IQR) method. This involved identifying and removing data points that fell outside the range defined by the IQR, helping to ensure the robustness and reliability of our analysis, Additionally, we encoded categorical columns, as certain algorithms exclusively operate on numerical data[12].

Parameter Tuning

We iteratively adjusted the input parameters of the comparison algorithms until we saw a decline in model performance, to determine the most optimal values. The parameter combinations for each method are specified in Table 2:

Table 2. Parameter Configuration of Algorithms

Algorithm	Parameter Configuration	Optimal Value
Logistic Regression	C: The regularization strength, which controls the inverse of regularization strength. (Default Value: 1.0) penalty: It refers to the kind of regularization employed in the model. 'l1' denotes L1 regularization, whereas 'l2' signifies L2 regularization. (Default Value: 'l2') solver: The optimization method employed in the logistic regression model. (Default Value: 'lbfgs')	C: 0.8 penalty: l1 solver: liblinear
SVM	C: It is the regularization parameter. It's the trade-off between a smooth decision boundary and accurately identifying the training points. (Default Value: 1.0)	C: 0.8
KNN	n_neighbors: It indicates the amount of neighboring entities taken into account for categorization. (Default: 5)	n_neighbors: 3
Naive Bayes	No configurable parameters	
Decision Tree	splitter: The method used to decide the split at each node. 'random' selects the best random split, while 'best' chooses the most optimal split. (Default: 'best') max_depth: The maximum depth of the tree. (Default: None) min_samples_split: The minimum number of samples required to split an internal node. (Default: 2) min_samples_leaf: The minimum number of samples required for a node to be considered a leaf. (Default: 1)	splitter: random max_depth: 6 min_samples_split: 5 min_samples_leaf: 3
Neural Network (MLP)	hidden_layer_sizes: The tuple represents the amount of neurons in every hidden layer. alpha: It's the L2 penalty regularization parameter. A regularization parameter that helps mitigate overfitting by penalizing large weights. (Default: 0.0001) max_iter: The maximum number of iterations allowed for optimization. (Default: 200)	hidden_layer_sizes: (100, 50) alpha: 0.8 max_iter: 1000

Evaluation Metrics

This section outlines the evaluation of prediction model accuracy, primarily utilizing the confusion matrix method. The confusion matrix summarizes the counts of right predictions (True Positives (TP) and True Negatives (TN)) as well as incorrect predictions (False Positives (FP) and False Negatives (FN)) when true labels are known.

Accuracy measures the proportion of correctly classified instances out of the total instances. Accuracy can be defined as follows (1):

$$\text{Accuracy} = \frac{TP+TN}{TP+FP+TN+FN} \quad (1)$$

A model's recall quantifies its capacity to accurately recognize every positive case. The definition of recall is as follows: (2):

$$\text{Recall} = \frac{TP}{TP+FN} \qquad (2)$$

The precision metric quantifies the percentage of actual positive occurrences among all the instances that the model predicts as positive. One definition of precision is (3):

$$\text{Precision} = \frac{TP}{TP+FP} \qquad (3)$$

It might be difficult to compare models with different accuracy and recall values. To solve this, the F-measure creates a unified assessment where a higher number denotes a better result by combining accuracy and recall into a single statistic. F-measure can be defined as (4):

$$\text{F-measure} = \frac{2 * Precision * Recall}{Precision + Recall} \qquad (4)$$

The above evaluation metrics evaluate which model is better on average.

EXPERIMENTAL RESULTS

The six machine learning (ML) algorithms that we compared and contrasted were Decision Tree (DT), Naive Bayes (NB), k-Nearest Neighbors (kNN), Support Vector Machine (SVM), and Logistic Regression (LR) (MLP). In order to assess it, the dataset was split into training and testing subsets. Next, using a k-fold cross-validation procedure with k set to 10, the algorithms were trained on the training subset. The trained models were subsequently employed to make predictions on both the training and testing datasets, followed by the computation of evaluation metrics. Figures 3 and 4 show the generated confusion matrices for the training and testing datasets..

True Class

Predicted Class	Logistic Regression	Positive	Negative
	Positive	280	53
	Negative	50	341

True Class

Predicted Class	Naive Bayes	Positive	Negative
	Positive	279	54
	Negative	53	338

True Class

Predicted Class	SVM	Positive	Negative
	Positive	288	45
	Negative	38	353

True Class

Predicted Class	Decision Tree	Positive	Negative
	Positive	286	47
	Negative	44	347

True Class

Predicted Class	kNN	Positive	Negative
	Positive	290	43
	Negative	35	356

True Class

Predicted Class	MLP	Positive	Negative
	Positive	319	14
	Negative	16	375

Figure 3. Confusion Matrix for the Training Set.

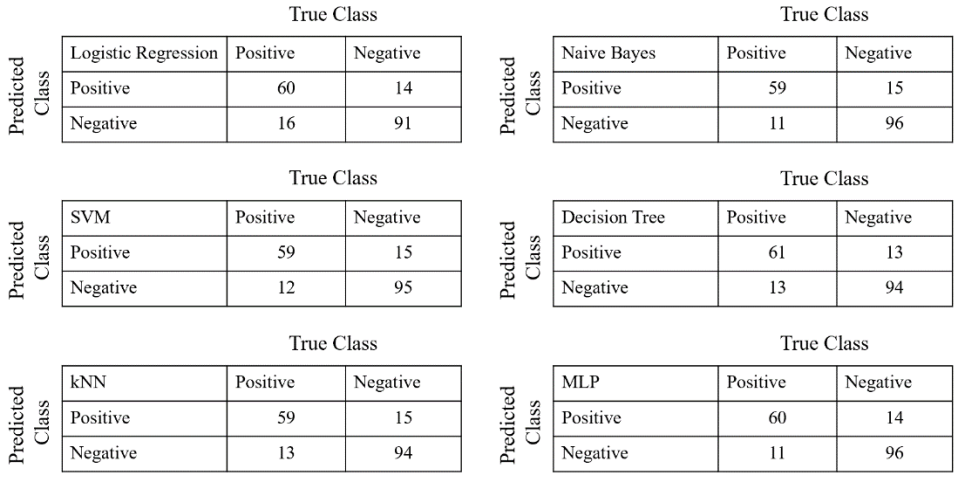

Figure 4. Confusion Matrix for the Testing Set.

Table 3 and Table 4 present the results of heart disease prediction models, highlighting the most outstanding performance values in bold. In comparison to the techniques for Logistic Regression (LR), k-Nearest Neighbors (KNN), Support Vector Machine (SVM), Naive Bayes (NB), and Decision Tree (DT) in both datasets, the Multi-layer Perceptron (MLP) method performed better. In the training dataset, the model achieved an accuracy of 0.934, a precision of 0.936, a recall of 0.945, and an F-measure of 0.941. On the testing dataset, the corresponding values were 0.861 for accuracy, 0.872 for precision, 0.897 for recall, and 0.884 for the F-measure.

Table 3. Results of Compared Algorithms (Training Set).

Classifier	Accuracy	Precision	Recall	Optimal Value
Logistic Regression	0.857	0.865	0.872	0.868
KNN	0.892	0.892	0.910	0.901
SVM	0.885	0.886	0.902	0.894
Naïve Bayes	0.852	0.862	0.864	0.863
Decision Tree	0.887	0.880	0.887	0.884
MLP	***0.958***	***0.964***	***0.959***	***0.961***

Table 4. Results of Compared Algorithms (Testing Set).

Classifier	Accuracy	Precision	Recall	Optimal Value
Logistic Regression	0.834	0.866	0.850	0.858
KNN	0.845	0.862	0.878	0.870
SVM	0.850	0.863	0.887	0.875
Naïve Bayes	0.856	0.864	0.897	0.880
Decision Tree	0.856	0.870	0.878	0.878
MLP	***0.861***	***0.872***	***0.897***	***0.884***

CONCLUSION

In this research, we carried out a comprehensive comparative analysis of six algorithms for heart disease prediction using the Kaggle dataset. Our findings revealed notable variations in the predictive capabilities of the algorithms. Logistic regression demonstrated robust performance in certain metrics, while support vector machines (SVM) exhibited superior performance in others. K-nearest neighbors (KNN) showcased competitive accuracy, albeit with some limitations in scalability. Naive Bayes showcases simplicity and effectiveness, although its dependency on assumptions may not consistently match real-world scenarios. Decision trees offered interpretability but faced challenges in handling complex data relationships. Lastly, neural networks exhibited remarkable flexibility and potential for capturing intricate patterns, albeit with higher computational requirements.

Multilayer Perceptron (MLP) emerged as the most promising algorithm, boasting superior accuracy (0.861), precision (0.872), recall (0.897), and F1-score (0.884) compared to others. This highlights the potential of deep learning techniques in enhancing predictive capabilities and emphasizes the importance of exploring advanced methodologies for more accurate heart disease prediction models.

Our study underscores the significance of leveraging innovative algorithms like MLP to enhance heart disease prediction accuracy, thereby facilitating more effective patient care and management strategies in clinical settings.

FUTURE WORK

In future research, expanding the scope of the study to encompass larger and more diverse datasets would be valuable. This could entail collaborating with multiple healthcare institutions to access comprehensive patient records, enabling a more robust assessment of the MLP algorithm's performance across various demographic profiles and risk factors.

Moreover, there's an opportunity to explore the efficacy of alternative machine learning algorithms beyond MLP. This exploration could yield deeper insights into the accuracy of heart disease prediction. Additionally, the integration of additional relevant features or biomarkers into the predictive model may offer avenues for further improving its performance.

Conducting longitudinal studies to evaluate the models' long-term predictive capabilities could provide valuable insights into their reliability over time. Furthermore, assessing the feasibility of deploying the developed model in real-time clinical settings and examining its impact on patient outcomes and healthcare delivery are critical steps for practical implementation.

Overall, expanding the study's scope to include larger and more diverse datasets, exploring alternative models, and investigating various factors such as longitudinal trends and real-world deployment scenarios are key avenues for advancing heart disease prediction research and improving clinical outcomes.

REFERENCES

1. World Health Organization. (2020). Cardiovascular diseases (CVDs). Retrieved from https://www.who.int/en/news-room/fact-sheets/detail/cardiovascular-diseases-(cvds).
2. S. S. Virani, A. Alonso, H. J. Aparicio, E. J. Benjamin, M. S. Bittencourt, C. W. Callaway, A. P. Carson, A. M. Chamberlain, S. Cheng, F. N. Delling, M. S. V. Elkind, K.R. Evenson, J.F. Ferguson, D. K. Gupta, S. S. Khan, B. M. Kissela, K. L. Knutson, C. D. Lee, & C. W. Tsao (2020). Heart disease and stroke statistics—2020 update: A report from the American Heart Association. Circulation, 141(9), e139–e596.
3. Z.I. Attia, S. Kapa, F. Lopez-Jimenez, P.M. McKie, D.J. Ladewig, G. Satam, P.A. Pellikka, M. Enriquez-Sarano, P.A. Noseworthy, & T.M. Munger, (2019). Screening for cardiac contractile dysfunction using an artificial intelligence–enabled electrocardiogram. Nature Medicine, 25(1), 70–74.
4. C. Krittanawong, H. Zhang, Z. Wang, M. Aydar, & T. Kitai, T. (2018). Artificial intelligence in precision cardiovascular medicine. Journal of the American College of Cardiology, 71(23), 2668–2679.
5. D. Dey, P. J. Slomka, P. Leeson, D. Comaniciu, S. Shrestha, P.P. Sengupta, T.H. Marwick, A. Banerjee, J. K. Min, & Y. Chandrashekhar (2019). Artificial intelligence in cardiovascular imaging: JACC state-of-the-art review. Journal of the American College of Cardiology, 73(11), 1317–1335.
6. M. Motwani, D. Dey, D. S. Berman, G. Germano, S. Achenbach, M. H. Al-Mallah, D. Andreini, M. J. Budoff, F. Cademartiri, T. Q. Callister, H. J. Chang, K. Chinnaiyan, B. J. Chow, R. C. Cury, A. Delago, M. Gomez, H. Gransar, M. Hadamitzky, J. Hausleiter, & P. J. Slomka (2020). Machine learning for prediction of all-cause mortality in patients with suspected coronary artery disease: A 5-year multicentre prospective registry analysis.

European Heart Journal, 41(25), 4428–4437.
7. S. F. Weng, J. Reps, J. Kai, J. M. Garibaldi, & N. Qureshi (2017). Can machine learning improve cardiovascular risk prediction using routine clinical data? PLOS ONE, 12(4), e0174944.
8. T. Amarbayasgalan, V.-H. Pham, N. Theera-Umpon, Y. Piao, & K. H. Ryu (2021). An Efficient Prediction Method for Coronary Heart Disease Risk Based on Two Deep Neural Networks Trained on Well-Ordered Training Datasets. IEEE, 135210 - 135223.
9. J. Raval, J.P. Verma, S.N.M. Islam, R. Jain, N. Thakur (2022). AI Based Prediction for Heart Disease: A Comparative Analysis and an Improved Machine Learning Approach. IEEE.
10. A. Gavhane, G Kokkula, I Pandya and K Devadkar, "Prediction of heart disease using machine learning in: 2018 second international conference on electronics communication and aerospace technology (ICECA)", IEEE, pp. 1275-1278, 2018.
11. https://towardsdatascience.com/heart-disease-prediction-73468d630cfc
12. https://www.analyticsvidhya.com/blog/2022/02/heart-disease-prediction-using-machine-learning-2/

