International Journal of Artificial Intelligence
p-ISSN: 2407-7275, e-ISSN: 2686-3251

Original Research Paper

Cardiovascular Disease Prediction Using Machine Learning

Atharva Mangeshkumar Agrawal[1]

[1] *Computer Science, University of Florida. United States.*

Article History
Received:
22.05.2023

Revised:
28.06.2023

Accepted:
04.07.2023

***Corresponding Author:**
Atharva Mangeshkumar Agrawal
Email:
atharva9111@gmail.com

Abstract: Heart disease-related deaths have become a big issue in today's world, with one person dying from the disease every minute. It considers both male and female groups, and the ratio varies by location. It is also used for the 25-69 age group. This isn't to say that people of all ages will be affected by heart disease. This condition could start in the early stages of life, and predicting the source and sickness is currently a huge challenge. Heart disease is one of the world's most fatal problems, one that cannot be seen with the naked eye and manifests itself as soon as it reaches its limits. As a result, precise diagnosis at the right moment is necessary. Every day, the health-care business generates massive amounts of patient and illness-related data. Researchers and practitioners, on the other hand, do not make appropriate use of this data. Despite its lack of knowledge, the healthcare business now has a wealth of data. In data mining and machine learning, there are a variety of approaches and tools for extracting usable information from databases and using that information to make more accurate diagnoses and decisions. So, in order to detect such disorders in time for adequate treatment, a reliable, precise, and feasible approach is required. In the realm of medicine, machine learning algorithms and approaches have been used to process enormous data sets. Researchers employ a variety of data mining and machine learning approaches to analyse large data sets and aid in the accurate prediction of cardiac illnesses. This research compares and contrasts the Nave Bayes, Help Vector Machine, Random Forest, and supervised learning models to find the most successful algorithm. When compared to other algorithms, Random Forest has 95.08 per cent more precision.

Keywords: Cardiovascular Diseases, K-Nearest Neighbour, Naive Bayes, Random Forest, Support Vector Machines.

International Journal of Artificial Intelligence, vol. 10, no. 2, pp. 59-67, December 2023. **DOI: 10.36079/lamintang.ijai-01002.542**

1. Introduction

A defect could result in an untimely death. Heart disease is currently the country's leading cause of mortality. The World Health Organization (WHO) estimates that 12 million people die each year from heart disease around the world. In 2008, 17.3 million people died of heart disease. Heart disease is responsible for more than 80% of all fatalities worldwide. Heart disease is expected to kill 23.6 million people by 2030, according to the WHO. This is one of the reasons why researchers are concentrating their efforts on building an intelligent system capable of accurately diagnosing heart disease and preventing misdiagnosis [1] - [5]. Many people, in reality, have no idea how to deal with heart disease. More patient deaths may be avoided if cardiovascular disease could be predicted earlier, and a more reliable and effective treatment plan could be devised. Because the heart is such an important organ of our bodies, life depends on it functioning properly. When the heart isn't working properly, it can cause problems in other regions of the body, such as the liver and kidneys. Cardiac illness is a condition that causes heart surgery to fail. Heart and blood vessel disease, often known as heart disease, encompasses a wide range of issues, many of which are linked to a condition known as atherosclerosis. When a compound called plaque builds up in the walls of the arteries, it is known as atherosclerosis [6] - [9].

Blood flow can be cut off if blood clots form. A stroke or a heart attack can occur as a result of this. Heart disease is made more deadly by a number of circumstances. High blood cholesterol, triglyceride levels, high blood pressure, diabetes and pre-diabetes, overweight and obesity, smoking, lack of physical activity, unhealthy diet, and tension are all linked to high blood cholesterol, triglyceride levels, high blood pressure, diabetes and pre-diabetes, overweight and obesity, smoking, lack of physical activity, unhealthy diet, and tension. Heart disease is currently the biggest cause of death worldwide, according to a World Health Organization report [10] [11], which states that 12 million people die each year as a result of heart disease. Cardiovascular diseases account for half of all deaths in the United States and other industrialised countries. In several underdeveloped nations, it is also the main cause of death. It is generally accepted as the leading cause of adult death. Heart disease kills one person every 34 seconds in the United States. Patients are handled correctly and care is successful when they receive exemplary health services. Lower-quality medical judgments can have devastating and unacceptably negative effects [12] - [17]. The health-care industry should also make an effort to reduce the number of tests required to detect the disease. All of this is possible with the installation of an appropriate decision support system. In today's world, most healthcare institutions employ hospital information systems to manage patient data [18].

Both tools are intended to aid in the accounting, inventory management, and statistical computations processes. Regrettably, the information isn't used to make decisions. This enormous amount of data can be utilised to address questions like "can you predict the risk of heart disease in patients?" Medical diagnosis is seen as a critical but difficult process that must be completed precisely and efficiently. In our daily lives, we are subjected to a normal and busy schedule, which causes tension and anxiety. Furthermore, the number of obese and cigarette-addicted persons is rapidly increasing. This adds to diseases including heart disease and cancer, among others. The problem with these illnesses is that they are difficult to predict. Everyone's heart rate and blood pressure are different. Nonetheless, the pulse rate must be between 60 and 100 beats per minute, and the blood pressure must be between 120/80 and 140/90, according to scientific evidence. One of the top causes of death worldwide is heart disease [19] [20]. Heart disease is becoming more common in both men and women of all ages, although other factors such as gender, diabetes, and BMI can also contribute to this ailment. We attempted to predict and evaluate heart disease using characteristics such as age, gender, blood pressure, heart rate, diabetes, and so on in this work. Heart disease is difficult to anticipate since it is caused by a combination of causes. The following are some of the most common heart attack symptoms [21] - [29]:

1. Tightness in the chest.
2. Breathing problems.
3. Nausea, indigestion, heartburn, or stomach pain are all possible symptoms.
4. Sweating and Fatigue are two words that come to mind while thinking about sweating.
5. A squeezing sensation in the upper back the pain is spreading to the arm.

International Journal of Artificial Intelligence, vol. 10, no. 2, pp. 59-67, December 2023. **DOI: 10.36079/lamintang.ijai-01002.542**

The following are the several types of cardiac disease: The word "cardio" implies "heart." As a result, cardiovascular disease groups include all heart disorders. The following are some of the different types of heart disease:

1. Coronary heart disease is a type of heart disease that affects both men and women.
2. Angina pectoris is a type of heart attack.
3. Heart failure due to congestive heart failure.
4. Cardiomyopathy is a disease that affects the heart.
5. Congenital heart disease is a term used to describe a condition in which a person's heart

Coronary heart disease, also known as coronary artery disease, is a constriction of the coronary arteries. The Coronary Arteries supply oxygen and blood to the heart. This causes a large number of individuals to become ill or die. It is one of the most common types of cardiovascular illness. High blood glucose levels in diabetics can harm the blood vessels and nerves that control the heart and blood vessels. If a person has diabetes for a long time, he or she is more likely to develop heart disease in the future. Other factors that lead to heart disease include diabetes. They smoke, which raises their chances of having heart disease.

Excessive blood pressure makes it more difficult for the heart to pump blood, putting strain on the heart and causing damage to the blood vessels, while high cholesterol levels can contribute to heart disease and obesity. Heart disease can also be caused by a family history of cardiac disease. This article, however, does not include this history for predicting heart disease. Age, sex, stress, and a bad nutrition are the other risk factors. As a person gets older, their chances of developing a cardiac issue increase. Men are more likely than women to get heart disease. Women, too, have the same probability after menopause. Living a hectic life can potentially damage the arteries and increase the risk of coronary heart disease.

The heart is a vital organ in all living creatures, and it is responsible for pumping blood via the circulatory system's blood vessels to the rest of the body's organs. Heart disorders are a group of illnesses that affect the heart and are the greatest cause of death in the globe. Heart disease and stroke are the country's two most dangerous and costly health problems today. The aortic, mitral, pulmonary, and tricuspid valves are four valves in your heart that open and close to regulate blood flow into your heart. Valves can be harmed by a variety of diseases that include constriction, leakage (regurgitation or insufficiency), or faulty closing [1]. While heart disease is typically referred to as a "man's disease," heart disease kills roughly the same number of women and men in the United States each year. In the first year after a heart attack, one out of every four women will die, compared to one out of every five males.

Diabetes, obesity, poor diet, physical inactivity, and heavy alcohol use are all conditions and lifestyle behaviours that increase a person's chance of getting heart disease [7] [13]. One of the major organs responsible for the functioning of blood in our human body and for all parts of the body is the heart; heart attack is one of the most common diseases in India; if the heart stops working, the entire blood circulation system in our body stops working, which can lead to death and serious health problems. The following are the most common types of heart illness found around the world, with Cardio Vascular Disease referring to a group of disorders that damage the heart or blood vessels. Angina and myocardial infarction are symptoms of coronary artery disease (CAD), often known as heart attack. The other type of heart disease [4] is coronary heart disease, which is caused by plaque build-up. The coronary arteries expand as a result, restricting blood flow to the heart.

2. Literature Review

Disease diagnosis is the most important responsibility in the healthcare industry. Many lives can be spared if a disease is detected early. Machine learning classification approaches have the potential to considerably help the medical industry by allowing for more accurate and timely disease diagnosis. As a result, both doctors and patients will save time. Heart disease is the world's leading cause of death today, making it one of the most challenging diseases to diagnose. In this study, we look at various machine learning algorithms for detecting heart illness and compare the outcomes using various performance criteria, such as accuracy, precision, and recall [20] - [27]. We used python and panda’s operations to do heart disease classification using data collected from the UCI repository after reviewing the results from previous approaches. It gives a simple visual depiction of the dataset, working environment, and predictive analytics construction. The machine learning process begins

International Journal of Artificial Intelligence, vol. 10, no. 2, pp. 59-67, December 2023. **DOI: 10.36079/lamintang.ijai-01002.542**

with data pre- processing, followed by feature selection based on data cleaning, classification, and evaluation of modelling performance. To improve the accuracy of the outcome, the random forest technique is applied. The proposed research investigates the four classification algorithms listed above and performs performance analysis to predict heart disease. The goal of this study is to accurately predict whether or not the patient has heart disease. The input values from the patient's health report are entered by the health professional. The information is incorporated into a model that forecasts the likelihood of developing heart disease.

3. Methodology

Figure 1 shows the research methodology.

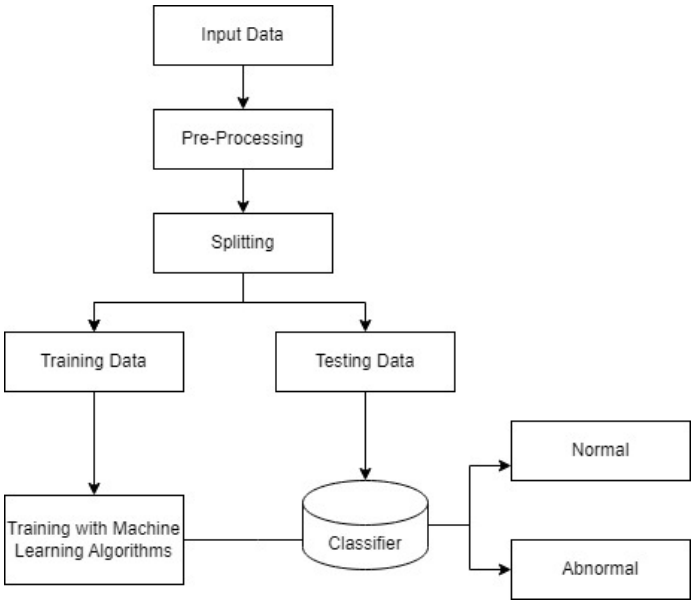

Figure 1. The Research Methodology

4. Finding and Discussion

4.1. Finding

Figure 2 shows the proposed method diagram.

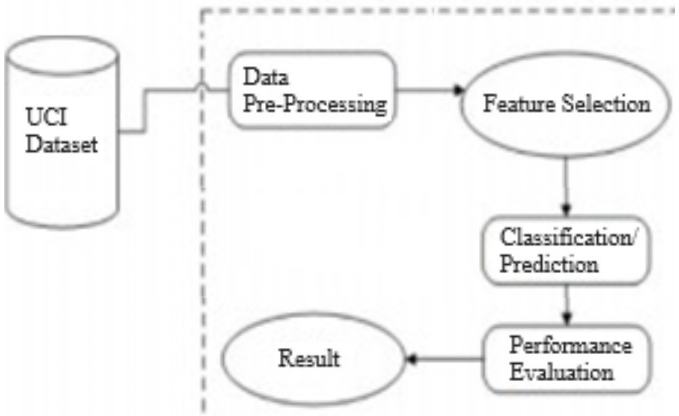

Figure 2. Diagram of the Proposed Method

International Journal of Artificial Intelligence, vol. 10, no. 2, pp. 59-67, December 2023. **DOI: 10.36079/lamintang.ijai-01002.542**

4.1.1. Advantages

There are some advantages of the proposed method:

1. Increased accuracy for a more accurate diagnosis of cardiac disease
2. Random forest technique and feature selection are used to handle the largest (enormous) amount of data.
3. Doctors' time complexity should be reduced.
4. Patients will save money.

4.1.2. Approaches

There are some approaches of the proposed method:

1. Data Pre-processing
   Data about heart disease is pre-processed using a variety of records. There are a total of 303 patient records in the dataset, with 6 of them having some missing values. The remaining 297 patient records are used in pre-processing after those 6 records were excluded from the dataset.
2. Feature Selection and Detection
   Two qualities related to age and sex are utilised to identify the patient's personal information among the 13 attributes in the data set. The remaining qualities are significant because they provide critical clinical information. Clinical data are essential for determining the degree of cardiac disease and diagnosing it.

4.1.3. Classification Modelling

The clustering of datasets is done using Decision Tree (DT) features as variables and criteria. The classifiers are then used to estimate the performance of each clustered dataset. Based on their low rate of error, the best performing models are identified from the given results.

1. Decision Trees
   The trees for training samples of data D are built using entropy inputs. A top-down recursive divide and conquer (DAC) strategy is used to build these trees. On D, tree pruning is done to get rid of the samples that aren't relevant.

   $$\text{Entropy} = -x_{mj} = 1\ p_{ij} \log 2\ p_{ij} \quad (1)$$

   Decision Tree-Based Partitioning Algorithm

   Require:
   D dataset – features with a goal class for what features do Execute the Decision Tree algorithm for each sample. Identify the feature space $f_1, f_2, \ldots, f_x$ of the dataset UCI. Calculate the total number of leaf nodes $l_1, l_2, l_3, \ldots, l_n$, taking into account the limitations. Based on the leaf nodes constraints, divide the dataset D into $d_1, d_2, d_3, \ldots, d_n$. Partition datasets d_1, d_2, and d_3 as output.
2. Random Forests
   To achieve the optimum result, this ensemble classifier constructs many decision trees and combines them. It primarily uses bootstrap aggregation or bagging for tree learning. For a given set of data, $x = x_1, x_2, x_3, \ldots, x_n$, $y = x_1, x_2, x_3, \ldots, x_n$, which repeats the bagging from $b = 1$ to B.
   KNN, Logistic Regression, and Random Forest Classifier are the classifiers utilised in the proposed model to categorise the pre-processed data once they have been pre-processed. Finally, we put the suggested model to the test, evaluating it for accuracy and performance using a variety of performance indicators. Using several classifiers, an effective Heart Disease Prediction System (EHDPS) has been built in this model. For prediction, this model incorporates 13 medical characteristics including chest pain, fasting sugar, blood pressure, cholesterol, age, and sex.

4.2. Discussion

The goal of this study is to compare the performance of various classification algorithms in order to determine which one is the most effective at predicting whether or not a patient will acquire heart disease. Nave Bayes, Support Vector Machine, and Random Forest approaches were used to analyse the dataset. The dataset has been divided into training and test data, and the models have been trained, with the accuracy measured using Python.

International Journal of Artificial Intelligence, vol. 10, no. 2, pp. 59-67, December 2023. **DOI: 10.36079/lamintang.ijai-01002.542**

The performance of the algorithms is compared in the table below, which includes their accuracy ratings, recall, accuracy, and F1 score. Each algorithm's outputs and accuracy levels are examined, and the results are displayed. The method with the best accuracy yields correct results in this model. This section displays the outcomes of using Random Forest, Decision Tee, Naive Bayes, and Logistic Regression. Accuracy score, Precision (P), Recall (R), and F-measure are the metrics used to assess the algorithm's performance.

The precision metric (given in equation (2)) offers a correct measure of positive analysis. The measure of correct actual positives is defined by recall [as specified in equation (3)]. The F-measure [given in equation (4)] evaluates precision.

$$\text{Precision} = (TP) / (TP + FP) \quad (2)$$

$$\text{Recall} = (TP) / (TP + FN) \quad (3)$$

$$\text{F Measure} = (2 * \text{Precision} * \text{Recall}) / (\text{Precision} + \text{Recall}) \quad (4)$$

Table 1. Naïve Bayes, SVM, Random Forest Comparison Table

Algorithm	F1 Score	Recall	Precision	Accuracy
Naïve Bayes	0.87	0.91	0.84	85.25%
Support Vector Machine	0.85	0.88	0.81	81.97%
Random Forest	0.95	0.91	1	95.08%

The algorithms we utilised are more accurate, save a lot of money (i.e., they are cost- effective), and are faster than the algorithms used by earlier studies. Furthermore, the maximum accuracy attained by KNN and Logistic Regression is 88.5 per cent, which is greater or almost equal to prior study accuracies. As a result of the enhanced medical information, we used from the dataset, we may conclude that our accuracy has improved. In the prediction of patients diagnosed with heart disease, our experiment also shows that Logistic Regression and KNN outperform Random Forest Classifier.

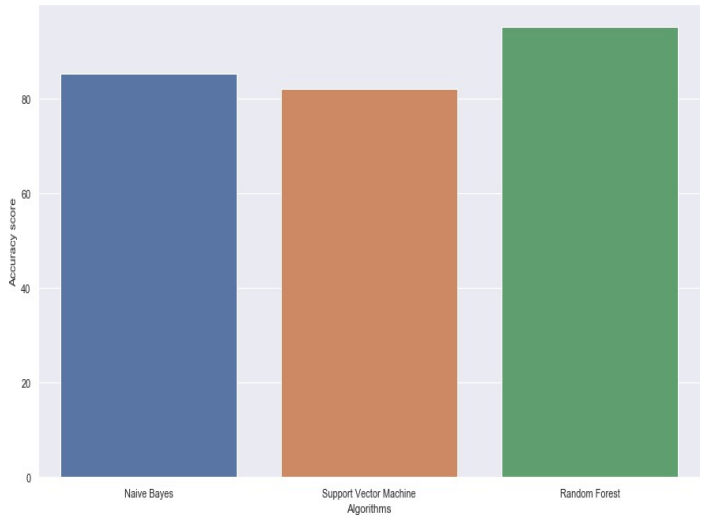

Figure 3. Comparison Graph

5. Conclusion

On the basis of the above analysis, it is clear that most machine learning algorithms perform well in the prediction and diagnosis of cardio vascular and heart diseases; however, some algorithms may

perform poorly in efficiency and accuracy tests. For example, the Random Forest algorithm works well on over-fitting data, whereas the support vector machine and the Naive Bayes algorithm work well on under-fitting data. We will apply neural networks to improve accuracy and broaden the usage of these methods to high-dimensional data in the future. We suggested a strategy for predicting heart disease using machine learning techniques in this research, and the results revealed a high accuracy threshold for providing a better estimation result. We solve the problem of rate prediction without equipment by providing a new proposed Random Forest classification and provide a method to estimate heart rate and condition. We find the information from the above input by ML Techniques. Sample results of heart rate are to be taken at different stages of the same subjects. First, we introduced a dataset-based support vector classifier.

With the rising number of deaths due to heart disease, it is becoming increasingly important to build a system that can effectively and accurately forecast heart disease. The goal of the research was to discover the most effective machine learning system for detecting cardiac problems. Using the UCI machine learning repository dataset, this study analyses the accuracy scores of Decision Tree, Logistic Regression, Random Forest, and Naive Bayes algorithms for predicting heart disease. According to the findings of this study, the Random Forest algorithm is the most efficient algorithm for predicting heart disease, with an accuracy score of 90.16 per cent. We've compiled a list of different types of machine learning algorithms for heart disease prediction. We developed a number of machine learning algorithms and used feature analysis to determine which one was the best. In different conditions, each algorithm has produced varied results.

Three ML classification modelling techniques were used to create a cardiovascular disease detection model. This study predicts people who will develop cardiovascular disease by extracting patient medical history that leads to a deadly heart illness from a dataset that includes patients' medical history, such as chest pain, sugar level, blood pressure, and so on. This Heart Disease Detection System aids a patient based on clinical information from a previous heart disease diagnosis. Logistic regression, Random Forest Classifier, and KNN [22] are the methods used to create the provided model. Our model has an accuracy of 87.5 per cent.

References

[1] B. Kaur, W. Singh, "Review on Heart Disease Prediction System using Data Mining Techniques," *International Journal on Recent and Innovation Trends in Computing and Communication*, vol. 2, no. 10, pp. 3003-3008, Oct 2014.

[2] P. Anooj, "Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules and decision tree rules," *Open Computer Science*, vol. 1, no. 4, Jan 2011, doi: 10.2478/s13537-011-0032-y.

[3] K. Srinivas and G. Raghavendra Rao, "Survey on Prediction of Heart Morbidity Using Data Mining Techniques," *International Journal of Data Mining & Knowledge Management Process*, vol. 1, no. 3, pp. 14–34, May 2011, doi: 10.5121/ijdkp.2011.1302.

[4] P. K. Sahoo and P. Jeripothula, "Heart Failure Prediction Using Machine Learning Techniques," *SSRN Electronic Journal*, Dec 2020, doi: 10.2139/ssrn.3759562.

[5] D. Gianola and C.-C. Schön, "Cross-Validation Without Doing Cross-Validation in Genome-Enabled Prediction," *G3: Genes|Genomes|Genetics*, vol. 6, no. 10, pp. 3107–3128, August 2016, doi: 10.1534/g3.116.033381.

[6] M. R, "Prediction of Diabetes Disease Using Classification Data Mining Techniques," *International Journal of Engineering and Technology*, vol. 9, no. 5, pp. 3610– 3614, Oct 2017, doi: 10.21817/ijet/2017/v9i5/170905319.

[7] C. Furlanello, "Disease networks and predictive methods for clinical data analytics," *Toxicology Letters*, vol. 258, pp. S33, September 2016, doi: 10.1016/j.toxlet.2016.06.1226.

[8] K. K. Y. A. Badholia and A. Sharma, "Heart Disease Prediction Using Machine Learning Techniques," *Information Technology in Industry*, vol. 9, no. 1, pp. 207–214, February 2021, doi: 10.17762/itii.v9i1.120.

[9] M. M. Bhajibhakare, "Heart Disease Prediction using Machine Learning," *International Journal for Research in Applied Science and Engineering Technology*, vol. 7, no. 12, pp. 455–460, Dec 2019, doi: 10.22214/ijraset.2019.12075.

[10] K. S. Rani, M. S. Manoj, and G. S. Mani, "A Heart Disease Prediction Model using Logistic Regression," *International Journal of Trend in Scientific Research and Development*, vol. 2, no. 3, pp. 1463–1466, April 2018, doi: 10.31142/ijtsrd11401.

[11] S. P. S. and H. Dr.M., "Heart Disease Prediction Using Integer-Coded Genetic Algorithm (ICGA) Based Particle Clonal Neural Network (ICGA-PCNN)," *Bonfring International Journal of Industrial Engineering and Management Science*, vol. 8, no. 2, pp. 15–19, April 2018, doi: 10.9756/bijiems.8394.

[12] S. K. Devi, S. Krishnapriya, and D. Kalita, "Prediction of Heart Disease using Data Mining Techniques," *Indian Journal of Science and Technology*, vol. 9, no. 39, Oct 2016, doi: 10.17485/ijst/2016/v9i39/102078.

[13] K. S. Rani, M. S. Manoj, and G. S. Mani, "A Heart Disease Prediction Model using Logistic Regression," *International Journal of Trend in Scientific Research and Development*, vol. 2, no. 3, pp. 1463–1466, April 2018, doi: 10.31142/ijtsrd11401.

[14] R. M. Shaikh, "Cardiovascular Diseases Prediction Using Machine Learning Algorithms," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 6, pp. 1083–1088, April 2021, doi: 10.17762/turcomat. v12i6.2426.

[15] S. K. Devi, S. Krishnapriya, and D. Kalita, "Prediction of Heart Disease using Data Mining Techniques," *Indian Journal of Science and Technology*, vol. 9, no. 39, Oct 2016, doi: 10.17485/ijst/2016/v9i39/102078.

[16] V. Krishnaiah, G. Narsimha, and N. Subhash, "Heart Disease Prediction System using Data Mining Techniques and Intelligent Fuzzy Approach: A Review," *International Journal of Computer Applications*, vol. 136, no. 2, pp. 43–51, February 2016, doi: 10.5120/ijca2016908409.

[17] A. Makwana and J. Patel, "Decision Support System for Heart Disease Prediction using Data Mining Techniques," *International Journal of Computer Applications*, vol. 117, no. 22, pp. 1–5, May 2015, doi: 10.5120/20683-3496.

[18] A. Pandita, "Prediction of Heart Disease using Machine Learning Algorithms," *International Journal for Research in Applied Science and Engineering Technology*, vol. 9, no. 6, pp. 2422–2429, June 2021, doi: 10.22214/ijraset.2021.3412.

[19] N. Masih and S. Ahuja, "Prediction of Heart Diseases Using Data Mining Techniques," *International Journal of Big Data and Analytics in Healthcare*, vol. 3, no. 2, pp. 1–9, July 2018, doi: 10.4018/ijbdah.2018070101.

[20] P. Kaur and K. Singh, "Detection of Heart Diseases using Machine Learning and Data Mining," *International Journal of Computer Applications*, vol. 178, no. 31, pp. 34–40, July 2019, doi: 10.5120/ijca2019919183.

[21] P. Santhi, "A Survey on Heart Attack Prediction Using Machine Learning," *Turkish Journal of Computer and Mathematics Education (TURCOMAT)*, vol. 12, no. 2, April 2021, doi: 10.17762/turcomat. v12i2.1955.

[22] V. Poornima and D. Gladis, "Analysis and Prediction of Heart Disease Aid of Various Data Mining Techniques: A Survey," *International Journal of Business Intelligence and Data Mining*, vol. 1, no. 1, pp. 1, 2018, doi: 10.1504/ijbidm.2018.10014620.

[23] P. Anooj, "Clinical decision support system: risk level prediction of heart disease using weighted fuzzy rules and decision tree rules," *Open Computer Science*, vol. 1, no. 4, January 2011, doi: 10.2478/s13537-011-0032-y.

[24] M. S. Amin, Y. K. Chiam, and K. D. Varathan, "Identification of significant features and data mining techniques in predicting heart disease," *Telematics and Informatics*, vol. 36, pp. 82–93, March 2019, doi: 10.1016/j.tele.2018.11.007.

[25] S. Shah, S. Batool, I. Khan, M. Ashraf, S. Abbas, and S. Hussain, "Feature extraction through parallel Probabilistic Principal Component Analysis for heart disease diagnosis," *Physica A: Statistical Mechanics and its Application*s, vol. 482, pp. 796–807, Sept 2017, doi: 10.1016/j.physa.2017.04.113.

[26] M. Butwall and S. Kumar, "A Data Mining Approach for the Diagnosis of Diabetes Mellitus using Random Forest Classifier," *International Journal of Computer Applications*, vol. 120, no. 8, pp. 36–39, June 2015, doi: 10.5120/21249-4065.

[27] M. T., D. Mukherji, N. Padalia, and A. Naidu, "A Heart Disease Prediction Model using SVM-Decision Trees-Logistic Regression (SDL)," *International Journal of Computer Applications*, vol. 68, no. 16, pp. 11–15, April 2013, doi: 10.5120/11662-7250.

[28] K. S. Rani, M. S. Chaitanya, and G. S. Kiran, "A Heart Disease Prediction Model using Logistic Regression by Cleveland DataBase," *International Journal of Trend in Scientific Research and Development*, vol. 2, no. 3, pp. 1467–1470, April 2018, doi: 10.31142/ijtsrd11402.

[29] M. Rathi, D. Garg, M. Goel, and V. Singhal, "Prediction of Health Fitness and Heart Status Using Data Mining Techniques," *SSRN Electronic Journal*, pp. 26-27, April 2018, doi: 10.2139/ssrn.3170518.