International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 640-647 www.ijaem.net ISSN: 2395-5252

Machine Learning-Driven Prediction of Heart Strokes

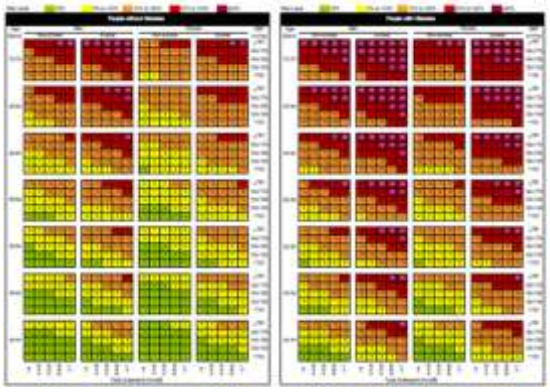
Selma Vreto Vatrić[1], Emir Beba[1], Dželila Mehanović[1]
[1] International Burch University, Faculty of Engineering, Natural and Medical Sciences, Sarajevo, Bosnia and Herzegovina
Corresponding Author: Selma Vreto Vatrić

Date of Submission: 20-05-2025 Date of Acceptance: 30-05-2025

ABSTRACT: Globally, heart attacks are a serious health concern. Timely prediction can significantly improve patient outcomes and healthcare resource allocation. This study aims to harness machine learning techniques to develop efficient predictive models for early detection of heart strokes.
Research is based on a dataset created by combining different (five) datasets. The dataset encompasses patient demographics, clinical measurements, and historical medical records. Various machine learning algorithms were employed to analyze the data, including KNearest Neighbor, Logistic Regression, Support Vector Machine, Decision Tree Classifier, and Random Forest Classifier.Cross-validation and relevant performance metrics, such as accuracy and F1-score, are employed to assess model performance.
The author's goal was to develop a strong prediction model that would reliably identify those who are more likely to suffer a heart attack. The Random Forest Classifier (RFC) is the best-performing algorithm among the models tested, with an accuracy of 86.96%, precision of 85%, recall of 91.89%, and an F1 score of 88.31%. Logistic Regression is the next top performer, with an accuracy of 84.06% and an F1 score of 85.43%. However, based on the presented metrics, the Random Forest Classifier remains the most successful option.
This tool has the potential to enable early interventions and preventive measures, thereby reducing the burden of heart strokes on healthcare systems and improving patient care. Limitations encompass data quality and availability, potential bias in healthcare records, and privacy concerns related to patient data.
KEYWORDS: heart stroke prediction, machine learning, Decision Tree Classifier (DTC), Random Forest Classifier (RFC), K-Nearest Neighbor (KNN), Logistic Regression (LR), Support Vector Machine (SVM), data, dataset

I. INTRODUCTION

The context for this study lies in the growing importance of applying machine learning to improve healthcare. Cardiovascular conditions, especially heart attacks, remain a primary source of morbidity and mortality around the world. According to the World Health Organization, more than 17.9 million individuals died from cardiovascular conditions in 2019, with heart attacks and strokes accounting for 85% of these deaths. Fortunately, it is possible to prevent 80% of early heart attacks and strokes. Early detection of heart strokes is crucial for timely intervention and improved patient outcomes [17].



A comparative analysis of cardiovascular disease risk based on laboratory data in Central Europe, distinguishing between individuals with and without diabetes.

In this research, machine learning is employed, specifically several distinct machine learning algorithms, to evaluate performance while predicting heart attacks. By using methods such as Support Vector Machine (SVM), Logistic Regression (LR), K-Nearest Neighbors (KNN), Decision Tree Classifier (DTC), and Random Forest Classifier (RFC), the author aims to create a

prediction model that can identify people who are more likely to have heart attacks, therefore lessening the impact of this illness and meeting the urgent need for prompt intervention and better patient care.

The research aims to clarify key elements of machine learning model optimization for heart attack prediction. Authors initially focus on the critical function of precise feature selection, indicating that finding the most valuable variables considerably improves model accuracy. Moving on, the emphasis changes to the importance of data cleaning and preprocessing, which are critical steps in ensuring high-quality training data and improving precision for heart attack prediction models. Moreover, research investigates the impact of machine learning algorithm selection on accuracy, with hypothesis positing that the decision is influenced by the dataset's attributes and the desired balance of interpretability and prediction accuracy. Finally, research investigates how machine learning deployment affects patient outcomes, predicting asignificant impact on reducing heart stroke incidence and enhancing overall cardiovascular health through prompt interventions. The following sections will thoroughly review the existing literature, offering a comprehensive analysis of relevant research. We will also elaborate on the research questions and hypotheses that will guide our study. Furthermore, we will consider themethodology that will be employed in our research. Moving forward, present and review the data and delve into the insightful findings that have emerged from this study. Finally, the significant contributions of our research to the field of healthcare, and the benefits of this work for both science and health service.

II. LITERATURE REVIEW

To contextualize the present study and highlight its relevance, this section begins with a thorough review of the existing literature in the area of machine learning techniques to develop efficient predictive models for the detection of heart strokes.

In [1] authors aimed to predict heart attacks accurately using machine learning techniques. The research highlights the importance of data - the lic Health Heart Attack dataset - preprocessing and early heart attack detection through Machine Learning algorithms. The researchers compared results from five machine learning approaches within Weka - MLP, RBF (Radial Basis Function Network), SVM, and RF. Weka was utilized to configure a Multi-Layer Perceptron model with specified parameters. To train these models, a 10-fold cross-validation technique was implemented, where the dataset was partitioned into ten equally sized subgroups. One subgroup served as the test set while the remaining nine acted as the training set in each iteration.

The results obtained from this study:
MLP: Accuracy 91.6%, Precision 91.9%
RBF: Accuracy 88.7%, Precision 90.1%,
SVM: Accuracy 85.4%, Precision 84.3%
KNN: Accuracy 95.2%, Precision 96.4%
RF: Accuracy 90.4%, Precision 92.1%

Research [2] aimed to evaluate the performance of different data analysis methods for predicting heart stroke. The study uses Random Forest and Decision Tree as machine-learning strategies, and a hybrid model (combined RF and DT) integrating both methods to predict heart attacks using the heart disease dataset. The data obtained from their work The Decision Tree achieved 79% accuracy, Random Forest 81%, and the hybrid model combining both achieved 88%.

As we delve into the methodologies used in [3], we can understand the importance of the topic and compare the results of different algorithms. The World Health Organization's survey indicated that approximately 10 million lives are claimed by heart disease. One major challenge in healthcare today is the early prediction of heart disease following an individual's diagnosis. The research aims to create an effective heart disease prediction model. Different machine learning algorithms are applied to the training set: K-Nearest Neighbors, Random Forest, and Decision Trees. In terms of accuracy, Random Forest achieved 98.27%, followed closely by Decision Trees at 97.67%, and K-Nearest Neighbors at 97.6%.

The performance evaluation of a large data processing system for heart disease prediction based on machine learning was provided by the authors in [4]. Due to a lack of knowledge and variables related to lifestyle, the number of heart disease patients is constantly increasing. Consequently, there is a requirement to identify and prevent heart disease effectively. The study used various methods: SVM, DT, RF, and LR. Specifically, the researchers worked with the processed Cleveland data from the heart disease database, which consisted of 303 records, each with fourteen attributes. These studies suggest that Random Forest achieves the best classification accuracy, with an average accuracy of 87.50%. Accuracy scores for other models are as follows: Support Vector Machine attained 85.82%, Decision Tree yielded 82.80%, and Logistic Regression achieved 85.70%.

In [6], authors developed a Time Efficient IOS Application For CardioVascular Disease

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 640-647 www.ijaem.net ISSN: 2395-5252

Prediction Using Machine Learning. The study has integrated functions into the development of a mobile-based iOS application, allowing users to input personal details and receive timely and accurate predictions. It recognizes the importance of optimizing system efficiency within the iOS environment. In the process of data analysis, several methods are employed. The classification models utilized in this analysis include Logistic Regression, Decision Tree, Random Forest, and XGBoost. The datasets used contain 70,000 records of patient data each containing 11+ features. The application demonstrates significant predictive capabilities for cardiovascular disease, with XGBoost achieving the highest accuracy. The iOS version excels in time efficiency due to its intricate mathematical and computational methods, while its performance on other Android devices is comparatively lower. Logistic Regression and XGBoost showed similar performance, with accuracies of 72.39% and 72.70% and F1-Scores of 71.00% and 72.00%, respectively. Decision Tree lagged slightly behind, scoring 62.87% accuracy and 63.00% F1-Score. Random Forest achieved an accuracy of 69.18% with an F1-Score of 69.00%.

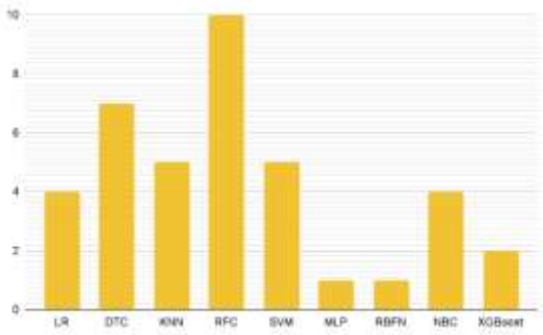
Research [7] aimed to develop a machine-learning model for predicting cardiac diseases based on pertinent characteristics. The ultimate objective is to give a beneficial decision support system to medical practitioners. In pursuit of the research objective, researchers employed various machine learning algorithms, each known for its distinct capabilities. These algorithms included RFC, SVM, DTC and Naive Bayes (NB). Using a benchmark dataset from UCI containing 14 different heart disease-related parameters, the findings revealed that Support Vector Machine (SVM) and Random Forest outperformed Gaussian Naive Bayes and Decision Tree in predicting heart diseases. Random Forest demonstrated high performance, achieving accuracy and recall of 99.70%. Then, SVM achieved an accuracy and recall of 99.50%. Decision Tree and Naive Bayes Classifier lagged with lower scores: Decision Tree had 85.10% accuracy and 84.80% recall, while Naive Bayes Classifier scored 90.40% accuracy and recall.

Research [8] will look at how machine learning may help perceive people susceptible to heart sickness by using assessing elements inclusive of but not restricted to chest soreness, cholesterol levels, and age. Machine learning methods are used in this study to predict whether a person would develop cardiovascular disease. Two supervised learning algorithms, the K-NN algorithm, and the Random Forest algorithm are applied to this dataset to build predictive models. K-Nearest Neighbor achieved 86.885% accuracy, while Random Forest attained 81.967%.

Authors of the study [9] created a clinical assistance system that uses a machine learning technique to predict heart disease. Using machine learning algorithms and pertinent risk factor data from medical records, the project tries to develop a useful support system. To achieve this goal, the study employs various machine learning algorithms, including NB, SVM, K-NN, RF, and DT. Initially, a train-test split was employed, but this technique tends to lead to overfitting due to data division. A 10-fold cross-validation approach was used to overcome this issue, effectively reducing bias and variance. For accuracy with cross-validation, Naive Bayes achieved 82.17%, K-Nearest Neighbor scored 76.56%, Support Vector Machine reached 79.20%, Random Forest obtained 69.30%, and Decision Tree 75.57%. In terms of accuracy with split data, Naive Bayes achieved 84.28%, K-Nearest Neighbor scored 81.31%, Support Vector Machine reached 81.42%, Random Forest obtained 77.14%, and Decision Tree 82.28%.

Research [10] focuses on the application of machine learning methods for heart disease prediction. This research focuses on leveraging the power of Data Science to process vast amounts of medical data, employing Data Mining and Machine Learning Techniques to facilitate precise predictions of heart diseases. The algorithms investigated include LR, NB, SVM, DT, K-NN, RF and XGBoost, using Python. The research findings indicate that various machine learning algorithms exhibit varying levels of accuracy in predicting heart diseases. Among the algorithms studied, Random Forest stands out as the most accurate, and K-Nearest Neighbor performs less effectively, yielding the lowest accuracy.

The author cannot compare the achievements because various models worked on distinct data sets. However, the chart below shows which methods are most applied.



Most applied methods

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 640-647 www.ijaem.net ISSN: 2395-5252

III. DATA AND FINDINGS

Numerous datasets are available for research in a similar vein, and they tend to share common factors that are pertinent to heart attacks. These datasets exhibit minimal variations, which makes it possible to combine them.

Our dataset consists of 1190 records of patients from Cleveland, Hungarian, Switzerland, and Long Beach VA, and the Stalog (Heart) Data Set. It includes 11 features and 1 target variable. This dataset is presented as the most comprehensive among commonly used heart disease datasets.
Attributes from our dataset along with their corresponding values:

Attribute	Description
age	Integer value representing the patient's age.
sex	0 = Male, 1 = Female. Represents the patient's gender.
chest pain type	0 = Typical Angina, 1 = Atypical Angina, 2 = Non-anginal Pain, 3 = Asymptomatic. Chest pain classification.
resting bp s	Value from 94 to 200. Resting blood pressure (in mm Hg).
cholesterol	Value from 126 to 564. Cholesterol level (in mg/dl).
fasting blood sugar	1 = True, 0 = False. Indicates whether fasting blood sugar is > 120 mg/dl.
resting ecg	0 = Normal, 1 = ST-T wave abnormality, 2 = Left ventricular hypertrophy. Resting electrocardiographic results.
max heart rate	Value from 71 to 202. Maximum heart rate achieved during exercise.
exercise angina	1 = Yes, 0 = No. Indicates if the patient has exercise-induced angina.
oldpeak	Value from 0 to 6.2. ST depression induced by exercise relative to rest.
ST slope	0 to 2. Describes the slope of the peak exercise ST segment.
target	0 = Less chance of heart attack, 1 = More chance of heart attack. Prediction result.

IV. DATA ANALYSIS, DATA CLEANING, AND FEATURE SELECTION

The complete research code, including data analysis, models, and performance evaluation, is available at this link: code.

Several libraries are necessary to effectively manipulate and analyze tabular data. Library pandas are required to work with data in tabular representation, while NumPy is used to round the data in the correlation matrix. Library missingno provides support to visualize missing values in the data and matplotlib, seaborn, and plotly are required for data visualization. Extra libraries were included in addition to the primary ones when needed.

The author loaded the dataset and then executed several lines of code to gain deeper insight into the data. Also, gained a better knowledge of the dataset's structure, data types, statistical summaries, dimensions, and unique values by using commands like dtypes, info, describe, shape, and nunique. df.isnull().sum() calculates the sum of missing values for each column in the DataFrame.
It was observed that there are no missing values in the dataset.

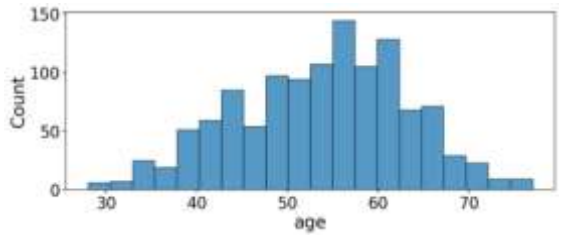
In the dataset, 272 duplicate entries were found. To ensure data integrity, the process of removing these duplicates, preserving only unique records for analysis, is being proceeded with. After the dataset analysis, a reduction in the number of rows was noticed, indicating that duplicates were successfully removed. Currently, it is (918, 12).
As the journey of data analysis begins, the focus shifts to visual representation through graph building.

Negative Correlation of "cholesterol" and "max heart rate" with "target"

It can be observed that two features ("cholesterol" and "max heart rate") are negatively correlated with the "target".

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 640-647 www.ijaem.net ISSN: 2395-5252

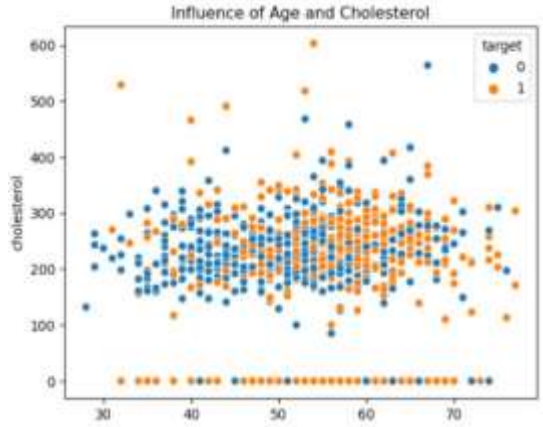
Other features are positively correlated with the "target". Positive correlation implies that changes in feature values tend to align with changes in the "target" variable in the same direction. In essence, as the feature values increase or decrease, the "target" variable follows a similar pattern.



Features Positively Correlated with "target"

V. ANALYTICAL OBSERVATIONS

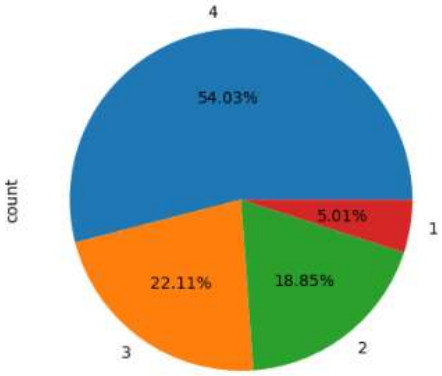
When looking at the age distribution, the graphical representation shows a concentration predominantly between the ages of 50 and 65, indicating a significant peak within this age range.



Age and Cholesterol Impact on Heart Attack Risk

This graph uses age and cholesterol features.
0 = less chance of heart attack
1 = more chance of heart attack

x-axis represents age, and the y-axis cholesterol. Datapoint in this system gives the following results: elderly people are the most affected by heart disease and young ones are the least affected, an increase in cholesterol level will have a higher risk of heart attack.



Insightful Visualization of Chest Pain Types Distribution

By displaying the frequency of each category in the 'chest pain type' column, this visualization helps in understanding the distribution of different types of chest pain. Asymptotic Chest Pain (ASY) is most common followed by Non-Anginal (22%), Atypical (19%), and Typical (5%) Chest Pain.

After a detailed investigation, it is concluded that all features available in customized dataset are significant and are likely to contribute meaningfully to the outcomes. Removing duplicates and null values, and there is no feature with constant or near-constant values.

Next, using the train_test_split function from the sklearn.model_selection package to split the dataset into training and testing sets. The input features (X_train and X_test) are separated from the target features (y_train and y_test) via this process. The test_size=0.3 option specifies that 30% of the data is used for testing and 70% is utilized to train the model. The generated sets allow for the training and evaluation of machine learning models.

X_train.shape: (642, 11)- 642 samples with 11 features for training.
X_test.shape: (276, 11) - 276 samples with 11 features for testing.
y_train.shape: (642,) - 642 samples for trainingtarget.
y_test.shape: (276,) - 276 samples for testing target.

VI. RESEARCH METHODOLOGY

In evaluating the model's performance, that is, evaluating its effectiveness, author utilized various crucial metrics to provide a comprehensive assessment.

These metrics are defined as follows:
Accuracy: Measure how often our model gets things right. It tells us the percentage of instances

International Journal of Advances in Engineering and Management (IJAEM)
Volume 7, Issue 05 May 2025, pp: 640-647 www.ijaem.net ISSN: 2395-5252

that our model correctly classifies among all the instances it encounters.

Precision: Precision is all about the exactness of our model. It helps us understand how well our model identifies positive cases among all the cases it predicts as positive.

Recall: Analyze how well the model was able to identify every positive instance among all the positive cases that are provided. F1 Score: The accuracy metric computes how many times a model made a correct prediction across the entire dataset.

Confusion Matrix: A confusion matrix is a table that displays a model's performance. It makes it easier to see how well the model is classifying occurrences into different categories [16].

Random Forest Classifier (RFC)

Random Forest is an ensemble learning method that combines the power of multiple decision trees to make more accurate predictions, where there is a labeled target variable. This technique has demonstrated its effectiveness in healthcare-related predictive modeling tasks [11]. This implementation imports and instantiates the Random Forest Classifier using scikit-learn. The fit() function is then used to train the forest model using the training data (X_train) and matching labels (y_train). Lastly, the score() function is used to assess the trained model's accuracy using test data (X_test) and labels (y_test) that have not yet been seen.

Logistic Regression (LR)

For heart stroke prediction, using Logistic Regression, a fundamental classification algorithm. LR models the logistic function to make binary or multiclass decisions when the dependent variable(target) is categorical [12].

Author replicated the methodology employed for generating predictions with Random Forest for Logistic Regression as well. This involved utilizing the predict() method to generate predictions based on the trained Logistic Regression model's learned patterns from the test data.

Decision Tree Classifier (DTC)

Our heart stroke prediction model will utilize a DTC, a hierarchical tree-based method that aids decision-making based on a set of rules and attributes.A decision tree is a type of supervised machine learning that is used to categorize or forecast depending on the answers to a previous set of questions [13]. The methodology employed for generating predictions with Random Forest was replicated for Decision Tree as well.

K-Nearest Neighbor (KNN)

For heart stroke prediction, author employ KNN, an algorithm based on finding the nearest neighbors in attribute space to classify new data points. It relies on the idea that similar data points tend to have similar labels or values [14]. The KNN algorithm, which identifies the class of a data point by considering the majority class among its nearest neighbors. With this method, predictions are made based on how close together data points are in the feature space, which sheds light on how unseen data is classified.

Support Vector Machine (SVM)

The SVM algorithm, known for its effectiveness in finding the optimal hyperplane that best separates data points of different classes, was employed to predict the labels of the test data. Our prediction model will be based on a method for separating data into different classes by defining a decision boundary. This algorithm's goal is to locate a hyperplane that, as much as feasible, separates the data points of one class from those of another class [15].

VII. PERFORMANCE ANALYSIS

The table below summarizes the results presented above. It shows accuracy, precision, recall and F1 score for each model used within this study.

Model	Accuracy	Precision	Recall	F1 Score
RFC	86,96%	85%	91,89%	88,31%
LR	84,06%	83,77%	87,16%	85,43%
DTC	79,71%	79,87%	83,11%	81,46%
KNN	65,22%	69,23%	69,23%	69,23%
SV	65,57%	73,38%	72,44%	72,90%

Summary of Model Performance Metrics

VIII. KEY CONTRIBUTIONS AND CHALLENGES

This study emphasizes the importance of algorithm selection and dataset features in molding model efficiency by comparing how different algorithms perform across diverse datasets. Evaluating algorithms' performance across several datasets reveals the effect of algorithmic complexity and data heterogeneity on predicted accuracy. Such comparison analyses provide useful insights for both researchers and healthcare practitioners, assisting in the prudent selection of algorithms for predictive modeling efforts in healthcare.

While the app offers helpful insights, it's crucial to be aware of its limitations and not rely just on it when making medical decisions. In medical settings, nevertheless, it can be a useful tool. First, medical professionals might utilize it to record observations and carefully analyze findings. The reliability of the application could be increased by subsequent iterations which will identify patterns and using the app to identify other improvements.

IX. CONCLUSION

The use of machine learning algorithms in predicting heart attacks is critical for enhancing preventative healthcare efforts. The goal of this work was to create effective predictive models for the early diagnosis of heart attacks utilizing various machine learning methods. Our analysis includes the combination of various datasets encompassing patient demographics, clinical measures, and past medical information.

The author's findings affirm the importance of feature selection in optimizing machine-learning models for heart stroke prediction. Careful selection resulted in enhanced model accuracy by focusing on the most important factors.

Data cleaning and preprocessing improved model precision and quality greatly. These critical procedures ensured high-quality data representation, which improved the models' capacity to reliably discover underlying patterns.

The accuracy and reliability of heart stroke prediction models were improved by the use of machine learning techniques. Complex algorithms, such as Random Forest, succeeded at capturing intricate data associations, but simpler algorithms, such as Logistic Regression, prioritized interpretability over accuracy. Choosing a machine learning algorithm has an impact on accuracy and reliability.

The use of machine learning algorithms for heart stroke prediction offers significant potential for improving patient outcomes. The author expects a decrease in heart stroke by identifying high-risk patients and enabling prompt interventions. This proactive strategy, which includes lifestyle changes and targeted monitoring, has the potential to greatly affect cardiovascular health.

In conclusion, this research highlights the potential of machine learning-driven approaches in detecting and preventing early heart attacks. The combination of multiple datasets, combined with careful data preprocessing, different algorithmic techniques, and complete model evaluations, has resulted in a robust prediction framework.

Future studies in this area should explore deeper ethical concerns, improve data quality metrics, and validate these models in clinical applications. This validation will increase their applicability and impact on healthcare outcomes.

While the app delivers useful information, it is critical to recognize its limitations and avoid relying completely on it for medical decisions. Nonetheless, in medical contexts, it can be a useful tool for documenting observations and reviewing results, with the possibility of incremental improvements.

REFERENCES

[1]. Mohamed Wed Eladham, Ali Bou Nassif, Mohammad AlShabi, 2023, "Heart Attack Prediction Using Machine Learning," Volume 14.
[2]. Dr. M. Kavitha, G. Gnaneswar, R. Dinesh, Y. Rohith Sai, R. Sai Sura, 2021, "Heart Disease Prediction using Hybrid Machine Learning Model," Volume 01.
[3]. M.Snehith Raja, M.Anurag, Ch.Prachetan Reddy, Nageswara Rao Sirisala, 2021, "Machine Learning Based Heart Disease Prediction System," Volume 01.
[4]. Abderrahmane Ed-daoudy, Khalil Maalmi, 2019, "Performance Evaluation of Machine Learning Based Big Data Processing Framework for Prediction of Heart Disease," Volume 12.
[5]. Thankgod Obasi, M. Omair Shafiq, 2020, "Towards Comparing and Using Machine Learning Techniques for Detecting and Predicting Heart Attack and Diseases," Volume 05.
[6]. Vansh Kedia, Aman Bhatia, Swesh Raj Regmi, Siddhant Dugar, Khushi Jha, Bickey Kumar Shah, 2021, "Time Efficient IOS Application for Cardiovascular Disease Prediction Using Machine Learning," Volume 06.
[7]. Vijeta Sharma, Shrinkhala Yadav, Manjari Gupta, 2020, "Heart Disease Prediction using Machine Learning Techniques," Volume 05.
[8]. Apurv Garg, Bhartendu Sharma, Rijwan Khan, 2021, "Heart Disease Prediction using Machine Learning Techniques," Volume 10.
[9]. Halima El Hamdaoui, Saïd Boujraf, Nour El Houda Chaoui, Mustapha Maaroufi, 2020, "A Clinical Support System for Prediction of Heart Disease using Machine Learning Techniques," Volume 09.
[10]. Pooja Anbuselvan, 2020, "Heart Disease Prediction using Machine Learning Techniques," Volume 12.

[11]. https://www.datacamp.com/tutorial/random-forests-classifier-python, last accessed 5/25/2025.
[12]. https://medium.com/data-science/logistic-regression-detailed-overview-46c4da4303bc, last accessed 5/25/2025.
[13]. https://www.mastersindatascience.org/learning/machine-learning-algorithms/decision-tree/, last accessed 5/25/2025.
[14]. https://www.analyticsvidhya.com/blog/2018/03/introduction-k-neighbours-algorithm-clustering/#h-what-is-knn-k-nearest-neighbor-algorithm, last accessed 5/25/2025.
[15]. https://se.mathworks.com/discovery/support-vector-machine.html, last accessed 5/25/2025.
[16]. https://www.fiddler.ai/model-evaluation-in-model-monitoring/what-is-model-performance-evaluation, last accessed 5/25/2025.