iJRASET
International Journal For Research in
Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH
IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 Issue: XI Month of publication: November 2023

DOI: https://doi.org/10.22214/ijraset.2023.57048

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

Review on Heart Disease Prediction Using Machine Learning Algorithms

Prof. Sushmita Khalane[1], Sahil R. Chikhale[2], Nikita S. Sawakare[3], Akshay Y. Kurkunde[4], Mayur S. Autade[5]
Computer Engineering, Sandip Institute of Technology and Research Centre Nashik, Maharashtra, India

Abstract: Heart disease remains a pressing global health challenge, accounting for a significant portion of morbidity and mortality worldwide. Timely identification of individuals at risk and the early prediction of heart disease plays a pivotal role in reducing its impact. Machine learning, with its capacity to discern intricate patterns within vast datasets, offers a promising avenue for enhancing the accuracy of risk assessment. This project focuses on the application of machine learning algorithms to predict heart disease. By leveraging features such as patient demographics, medical history, and clinical measurements, we aim to develop a predictive model that can assist healthcare professionals in identifying individuals susceptible to heart disease well before symptoms manifest. The utilization of machine learning algorithms, including logistic regression, Random Forest, and K Nearest Mean, allows us to extract meaningful insights from heterogeneous healthcare data. These algorithms can analyze and weigh the significance of various risk factors, thus enabling accurate risk assessment. This endeavor holds profound implications for public health, as early detection can pave the way for timely interventions and lifestyle modifications. The potential to reduce the burden of heart disease on healthcare systems and improve patient outcomes underscores the significance of this research. The ability to predict heart disease with high precision is of paramount importance, as it can aid healthcare professionals in identifying individuals at risk and tailoring preventative measures accordingly. Early diagnosis and intervention can significantly reduce the burden of heart disease on individuals and healthcare systems alike.In this project, we harness the potential of machine learning algorithms to analyze a comprehensive dataset comprising various patient attributes, including demographic information, medical history, and clinical measurements such as blood pressure, cholesterol levels, and electrocardiogram (ECG) data. By training and validating our model on this diverse dataset, we aim to create a predictive tool capable of discerning patterns and relationships that may elude traditional diagnostic methods.

Index Terms: Heart disease, machine learning, predictive modeling, risk assessment, healthcare, early detection, blood pressure, cholesterol levels, electrocardiogram (EGG).

I. INTRODUCTION

Heart disease, encompassing conditions such as coronary artery disease, heart failure, and arrhythmias, remains a formidable global health challenge. According to the World Health Organization, cardiovascular diseases are the leading cause of death worldwide, claiming nearly 18 million lives annually. Early detection and accurate risk assessment are pivotal in combating this pervasive health issue. Machine learning algorithms can analyze vast and intricate healthcare datasets with remarkable precision, offering insights that were once elusive. These algorithms have demonstrated tremendous potential in diagnosing diseases, predicting outcomes, and guiding treatment decisions.

II. MACHINE LEARNING

Machine learning for heart disease prediction leverages computational algorithms to analyze patient data, identify- ing patterns and relationships crucial for assessing the risk of heart-related conditions. The process involves collecting diverse datasets, selecting informative features, and choosing appropriate algorithms like decision trees or neural networks. By training on historical data and testing for generalization, these models are evaluated using metrics such as accuracy and precision. The optimization of hyperparameters and considera- tion of interpretability enhance the model's predictive capabilities. Ultimately, machine learning contributes to more effective and data-driven heart disease risk assessment, supporting early intervention and personalized healthcare strategies.

A. Supervised Learning

Supervised learning is employed as a key machine learning paradigm. Supervised learning algorithms are trained on a labeled dataset, where historical patient data is paired with cor- responding outcomes (presence or absence of heart disease). The algorithm learns to make predictions by generalizing from this labeled training data.

Features such as demographic information, medical history, and diagnostic test results serve as input, and the algorithm adjusts its parameters during training to optimize predictive accuracy. Once trained, the model can predict the likelihood of heart disease in new, unseen data. This supervised approach enables the algorithm to learn from known outcomes, making it a valuable tool for creating predictive models in the healthcare domain. Here are some commonly used supervised learning algorithms that can be applied to heart disease prediction:

1) Logistic Regression
2) Decision Trees
3) Random Forest
4) Support Vector Machines (SVM)
5) K-Nearest Neighbors (KNN)

The process of supervised learning involves several key steps,

a) Data Collection
b) Data Preprocessing
c) Splitting the Data
d) Choosing a Model
e) Training the Model
f) Evaluation
g) Prediction

B. Unsupervised Learning

Unsupervised learning plays a pivotal role in uncovering intrinsic patterns within patient data without relying on ex- plicit outcome labels. Through techniques like clustering or dimensionality reduction, unsupervised learning algorithms autonomously identify inherent structures within the dataset, potentially revealing nuanced relationships or patient sub- groups. By exploring these latent patterns, the approach contributes to a more comprehensive understanding of the diverse factors influencing heart health. This unsupervised exploration aids in personalized risk assessment, allowing for the identification of distinct patient profiles or characteristics that may be indicative of varying degrees of cardiovascular risk. Overall, unsupervised learning enriches the predictive modeling process by providing valuable insights into the complex interplay of factors associated with heart disease. Unsupervised learning encompasses various techniques and algorithms. Here are some commonly used methods:

1) K-Means Clustering
2) Hierarchical Clustering
3) PCA (Principal Component Analysis)
4) Apriori Algorithm

C. Reinforcement Learning

Reinforcement learning, although less commonly applied in traditional medical diagnostics, holds promise for optimizing personalized interventions in the context of heart disease prevention. Unlike supervised learning, reinforcement learning involves an agent making sequential decisions in an environment, learning through trial and error to maximize a cumulative reward signal. In the realm of heart disease, reinforcement learning could be employed to develop adaptive treatment plans or recommend lifestyle adjustments. The algorithm learns by interacting with patient data, refining its decision-making strategy over time based on the observed outcomes. While the application of reinforcement learning in healthcare is an evolving field, its potential lies in tailoring interventions to individual patient responses, contributing to more effective and patient-centric approaches to heart disease management.

III. RELATED WORK

In recent times, many health monitoring systems have been developed to monitor the health condition of patients. We are reviewing some recent works developed in this field. The study of heart disease prediction has witnessed a surge in research efforts leveraging machine learning techniques. These investigations delve into the selection of appropriate features, the comparison of various machine learning models, and the integration of real-time prediction systems. Recent studies by Ramya G. Franklin, Shaik Farzana, and others introduce innovative approaches like encryption techniques, dynamic prediction using multi-machine learning approaches, and cognitive strategies to elevate prediction accuracy.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

Collectively, these studies underscore the evolving landscape of heart disease prediction, with a focus on improving diagnostic precision, reducing costs, and optimizing the delivery of healthcare services. The literature survey reveals a comprehensive exploration of machine learning applications in heart disease prediction, showcasing various methodologies and algorithms. Halima El Hamdaoui's study [3] underscores the effective- ness of the Na¨ıve Bayes algorithm, achieving the highest accuracy in heart disease prediction, with a preference for cross-validation despite a slight decrease in accuracy. Rahul Katarya's survey [1] provides a valuable overview of expert automated systems, emphasizing the importance of feature selection for effective heart disease prediction. P. Sujatha and
K. Mahalakshmi's research [4] compares the performance of diverse supervised machine learning algorithms, identifying the random forest classifier as the most accurate on the Kaggle heart disease dataset. Simran Verma and Abhishek Gupta's review [2] highlights the impact of dataset quality on algorithm accuracy, calling attention to the need for a comprehensive exploration of data cleaning and pruning methods. Garima Choudhary and Shailendra Narayan Singh's work [12] op- timizes decision tree output using the Ada-Boost algorithm, achieving a commendable accuracy of 0.89. M. Kavitha and
G. Gnaneswar's hybrid model [5] combines Decision Tree and Random Forest for heart disease prediction, implemented in a TkInter Python application. M. Snehith Raja and M. Anurag [17] advocate for the efficiency of the Random Forest algorithm in early detection of heart disease, emphasizing its potential for reducing long-term mortality.

IV. METHODOLOGY OF SYSTEM

The methodology for our heart disease prediction project involves a systematic process to develop an accurate and reliable predictive model. We commence with the collection of a comprehensive dataset comprising diverse patient attributes. Subsequently, we meticulously preprocess the data, addressing missing values and outliers to ensure data quality. Feature selection follows, focusing on identifying the most pertinent variables that contribute significantly to heart disease predic- tion. The algorithm selection phase involves choosing appro- priate machine learning techniques, and the model undergoes intensive training on the prepared dataset. Hyperparameter tuning refines the model's parameters, optimizing its predictive capabilities. Rigorous evaluation using metrics like accuracy and precision ensures the model's effectiveness. Upon satisfac- tory performance, the model is deployed, making it accessible for real-world predictions

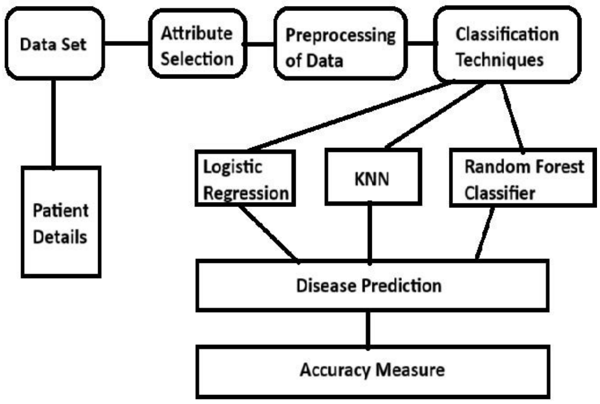


Fig. 1. Architecture Diagram of Prediction System.

1) Data Collection

- Gather a comprehensive dataset containing relevant patient information, such as demographics, medical history, and diagnostic test results. A reliable dataset is crucial for training and evaluating the machine learning model.

2) Data Preprocessing

- Clean and preprocess the dataset to handle missing values, outliers, and ensure consistency. This step may involve feature scaling, normalization, and en- coding categorical variables to prepare the data for the machine learning model.

3) Feature Selection

- Identify and select the most informative features that significantly contribute to heart disease predic- tion. This process ensures that the model focuses on relevant variables, improving its efficiency and interpretability.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

S. No.	Attribute	Description	Type
1	Age	Patient's age (29 to 77)	Numaric
2	Sex	Gender of patient(male-0 female-1)	Nominal
3	Cp	Chest pain type	Nominal
4	Trestbps	Resting blood pressure(in mm Hg on admission to hospital ,values from 94 to 200)	Numerical
5	Chol	Serum cholesterol in mg/dl, values from 126 to 564)	Numerical
6	Fbs	Fasting blood sugar>120 mg/dl, true-1 false-0)	Nominal
7	Resting	Resting electrocardiographics result (0 to 1)	Nominal
8	Thali	Maximum heart rate achieved(71 to 202)	Numerical
9	Exang	Exercise included agina(1-yes 0-no)	Nominal
10	Oldpeak	ST depression introduced by exercise relative to rest (0 to .2)	Numerical
11	Slope	The slop of the peak exercise ST segment (0 to 1)	Nominal
12	Ca	Number of major vessels (0-3)	Numerical
13	Thal	3-normal	Nominal
14	Targets	1 or 0	Nominal

Fig. 2. Attributes of Prediction System.

4) Algorithm Selection
- To choose appropriate machine learning algorithms based on the nature of the problem, we consider al- gorithms like Na¨ıve Bayes, Decision Trees, Random Forest, or others commonly used in heart disease prediction tasks.

5) Model Training
- Train the selected machine learning model on the preprocessed dataset. This involves feeding the algo- rithm with labeled data to enable it to learn patterns and relationships between features and heart disease outcomes.

6) Hyperparameter Tuning
- Fine-tune the hyperparameters of the chosen algo- rithm to optimize its performance. This step involves experimenting with different parameter configura- tions to achieve the best results.

7) Model Evaluation
- Assess the trained model's performance using evalu- ation metrics such as accuracy, precision, recall, and F1 score. This step ensures that the model provides reliable predictions and generalizes well to new, unseen data.

8) Deployment
- Integrate the trained and optimized model into the heart disease prediction system for real-world use. This may involve developing a user interface or in- tegrating the model into existing healthcare systems.

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

V. CONCLUSION AND FUTURE SCOPE

In conclusion, the heart disease prediction project utilizing machine learning has will achieve significant milestones with the development of a robust predictive model for assessing heart disease risk. This project will not only demonstrate the potential of machine learning in healthcare but also show its immediate benefits in clinical practice. The accurate risk assessment, early detection capabilities, and improved clinical decision support provided by the model have the potential to transform cardiovascular healthcare. The heart disease pre- diction project, driven by machine learning, stands on the threshold of a promising future. While it has already achieved significant strides in risk assessment and early detection, its potential for growth and impact remains vast. Looking ahead, the project can evolve in several key directions. The integration of multi-modal data, encompassing genomics and wearable device data, promises a more comprehensive risk assessment. Continuous learning mechanisms can ensure that the model remains adaptive and responsive to evolving health trends. In the future, we consider enhancing our heart disease prediction project by improving feature engineering, optimizing machine learning models, enabling real-time monitoring, integrating with electronic health records, conducting clinical validation, and focusing on interpretability. Ensure data security and privacy, expand the system to predict long-term outcomes, and collaborate with healthcare providers for feedback. Design a user-friendly interface and plan for scalability to accommodate a broader user base while staying aligned with evolving healthcare practices and technological advancements.

REFERENCES

[1] Rahul Katarya; Polipireddy Srinivas (2020). Predicting Heart Disease at Early Stages using Machine Learning: A Survey. International Confer- ence on Electronics and Sustainable Communication Systems (ICESC).

[2] Simran Verma; Abhishek Gupta (2021). Effective Prediction of Heart Disease Using Data Mining and Machine Learning: A Review. Interna- tional Conference on Artificial Intelligence and Smart Systems (ICAIS).

[3] Halima El Hamdaoui; Sa¨ıd Boujraf; Nour El Houda Chaoui; Mustapha Maaroufi (2020). A Clinical support system for Prediction of Heart Dis- ease using Machine Learning Techniques 5th International Conference on Advanced Technologies for Signal and Image Processing (ATSIP).

[4] P. Sujatha; K. Mahalakshmi (2020). Performance Evaluation of Su- pervised Machine Learning Algorithms in Prediction of Heart Dis- ease. IEEE International Conference for Innovation in Technology (INOCON).

[5] M. Kavitha; G. Gnaneswar; R. Dinesh;Y. Rohith Sai; R. Sai Suraj (2021). Heart Disease Prediction using Hybrid machine Learning Model. 6th International Conference on Inventive Computation Technologies (ICICT).

[6] Amin Ul Haq; Jianping Li; Muhammad Hammad Memon; Muhammad Hunain Memon (2019). Heart Disease Prediction System Using Model Of Machine Learning and Sequential Backward Selection Algorithm for Features Selection. IEEE 5th International Conference for Convergence in Technology (I2CT).

[7] Ramya G. Franklin; B. Muthukumar (2020). Survey of Heart Disease Prediction and Identification using Machine Learning Approaches. 3rd International Conference on Intelligent Sustainable Systems (ICISS) Heart Disease Prediction

[8] Pranav Motarwar; Ankita Duraphe; G Suganya; M Premalatha (2020). Cognitive Approach for Heart Disease Prediction using Machine Learn- ing. International Conference on Emerging Trends in Information Tech- nology and Engineering (ic-ETITE)

[9] Sanchayita Dhar; Krishna Roy; Tanusree Dey; Pritha Datta; Ankur Biswas (2018). A Hybrid Machine Learning Approach for Prediction of Heart Diseases. 4th International Conference on Computing Commu- nication and Automation (ICCCA)

[10] Shaik Farzana; Duggineni Veeraiah (2020). Dynamic Heart Disease Prediction using Multi-Machine Learning Techniques. 5th International Conference on Computing, Communication and Security (ICCCS)

[11] Abderrahmane Ed-Daoudy; Khalil Maalmi (2019). Real-time machine learning for early detection of heart disease using big data approach. International Conference on Wireless Technologies, Embedded and Intelligent Systems (WITS)

[12] Garima Choudhary; Shailendra Narayan Singh (2020). Prediction of Heart Disease using Machine Learning Algorithms. International Con- ference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)

[13] Kunal Rao; Pawan Raj Gopal; Kusum Lata (2021). Computational Analysis of Machine Learning Algorithms to Predict Heart Disease. 11th International Conference on Cloud Computing, Data Science Engineer- ing (Confluence)

[14] Lama A. Alqahtani; Hanadi M. Alotaibi; Irfan Ullah Khan; Nida Aslam (2020). Automated prediction of heart disease using optimized machine learning techniques. 11th IEEE Annual Ubiquitous Computing, Electronics Mobile Communication Conference (UEMCON)

[15] Farzana Tasnim; Sultana Umme Habiba (2021). A Comparative Study on Heart Disease Prediction Using Data Mining Techniques and Feature Selection. 2nd International Conference on Robotics, Electrical and Signal Processing Techniques (ICREST)

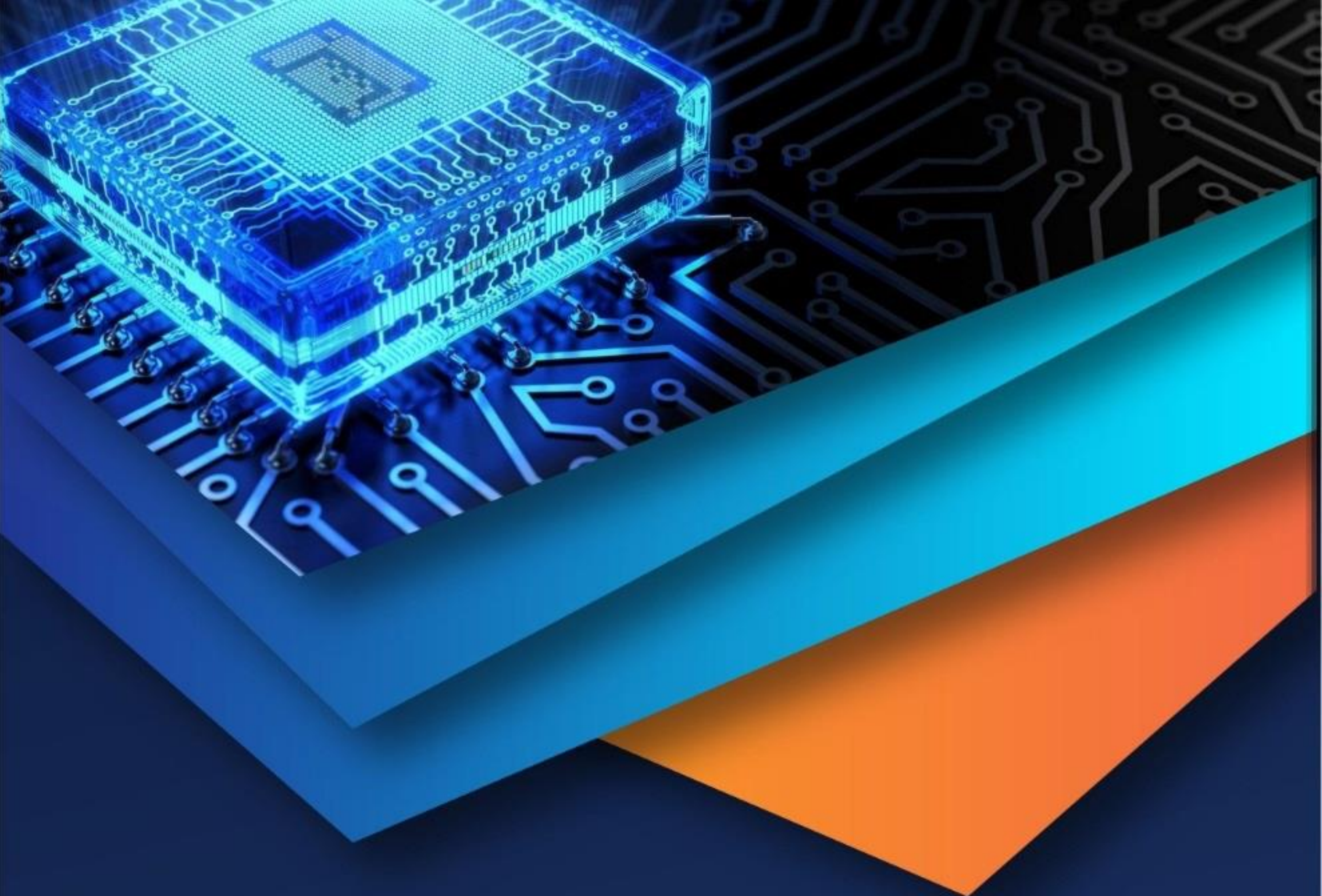
[16] Rahma Atallah; Amjed Al-Mousa (2019). Heart Disease Detection Using Machine Learning Majority Voting Ensemble Method 2nd International Conference on new Trends in Computing Sciences (ICTCS)

[17] M. Snehith Raja; M. Anurag; Ch. Prachetan Reddy; Nageswara Rao Sirisala (2021). Machine Learning Based Heart Disease Prediction System. International Conference on Computer Communication and Informatics (ICCCI) Heart Disease Prediction

[18] Muhammad Affan Alim; Shamsheela Habib; Yumna Farooq; Abdul Rafay (2020). Robust Heart Disease Prediction: A Novel Approach based on Significant Feature and Ensemble learning Model. 3rd In- ternational Conference on Computing, Mathematics and Engineering Technologies (iCoMET)

[19] Chu-Hsing Lin; Po-Kai Yang; Yu-Chiao Lin; Pin-Kuei Fu (2020). On Machine Learning Models for Heart Disease Diagnosis. IEEE 2nd Eurasia Conference on Biomedical Engineering, Healthcare and Sustainability (ECBIOS).

[20] Ruhani Chowdhury; Ezaz Ahmed; Md. Abu Dayan Siddik; Akhlak Uz Zaman (2021). Heart Disease Prognosis Using Machine Learning Classification Techniques. 6th International Conference for Convergence in Technology (I2CT)



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089 (24*7 Support on Whatsapp)