iJRASET
International Journal For Research in Applied Science and Engineering Technology



INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 12 **Issue:** XII **Month of publication:** December 2024

DOI: https://doi.org/10.22214/ijraset.2024.65889

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 12 Issue XII Dec 2024- Available at www.ijraset.com

Predicting Heart Attack Risk Using Machine Learning: A Review

Radhika S K[1], Deeksha M V[2], Dhanushri M[3], Gayana K O[4], Inchara D[5]
Computer Science Department, Visvesvaraya Technological University

Abstract: ***Machine learning in the healthcare field involves developing predictive models for early disease detection, such as analyzing patient data and identify individuals at high risk of conditions like heart attacks, diabetes, or cancer. This paper explores the use of machine learning to predict the risk of heart attacks, helping to assess and classify the likelihood of occurrence. The study highlights its potential to improve early detection and support better healthcare decisions.***
Keywords: Heart Attack Prediction, Machine Learning Algorithms, Risk Assessment, Supervised Learning, Healthcare AI

I. INTRODUCTION

In recent times with the extensive use of advanced technology, addressing the need for early detection and prevention of heart attacks is crucial as heart attack cases are rising. Traditional diagnostic methods, often manual and reliant on limited datasets, can be time consuming in identifying risk patterns. Machine learning provides a powerful solution by analyzing complex medical data to provide majorly accurate automated risk assessments. This paper explores the potential of machine learning in predicting heart attack risk, highlighting its benefits in improving healthcare outcomes and its transformative impact on preventive care.

II. LITERATURE SURVEY

In this section, several authors have presented various machine learning techniques and methodologies for predicting heart attack risks.

[1] The paper "Enhancing Heart Attack Prediction with Machine Learning: A Study at Jordan University Hospital", authored by Mohammad Alshraideh and colleagues, presents a comprehensive approach to predicting heart attack risk using machine learning techniques. The study utilized a dataset of 486 patient records from Jordan University Hospital, incorporating 58 initial features such as age, blood pressure, cholesterol levels, and other clinical indicators. To improve model performance, the authors applied Particle Swarm Optimization (PSO) for feature selection, reducing the number of features to 19 key predictors. The study evaluated several machine learning algorithms, including SVM, Random Forest, Decision Tree, Naive Bayes, and KNN. Among these, the SVM classifier achieved the highest accuracy of 94.3%, highlighting the significant role of PSO in optimizing feature selection and enhancing prediction accuracy. The research underscores the potential of machine learning to support early heart attack diagnosis and assist healthcare professionals in clinical decision-making. By improving prediction performance, these methods can help identify high-risk patients more accurately, leading to timely interventions. This approach is particularly valuable in regions with high cardiovascular mortality rates, where early detection and prevention are crucial for reducing the burden of heart disease. The study demonstrates how advanced machine learning techniques can play a critical role in enhancing healthcare outcomes and guiding medical practice.

[2] The paper titled "Heart Attack Prediction Using Machine Learning," authored by Dr. M. Manoj Prabu, S. Ramprakash, G. Ajith, S. Sethuragavan, and R. Sathyan, presents a novel approach to predicting heart attack risks through a wearable device that monitors various body parameters. The device tracks crucial indicators such as heart rate, blood oxygen levels (SpO2), body temperature, and humidity, using sensors like DHT11, KY039, and MAX30100. These sensors collect real-time data, which is then transmitted to an IoT platform (ThingSpeak) for processing and analysis. Machine learning classifiers such as Naive Bayes, Support Vector Machine (SVM), and K-Nearest Neighbors (KNN) are applied to the data, enabling the system to predict the likelihood of a heart attack. The system provides real-time alerts to the patient, as well as notifications to healthcare providers, offering a proactive approach to heart disease management. It serves as a portable and cost effective alternative to traditional heart monitoring devices, making it particularly beneficial for patients in remote areas with limited access to healthcare facilities. These advancements aim to enhance user convenience, accuracy, and accessibility, ultimately contributing to better early detection and prevention of heart attacks, particularly for individuals who are at high risk or live in underserved areas.

©IJRASET: All Rights are Reserved

[3] The paper titled "Predicting the Chance of Heart Attack with a Machine Learning Approach – Supervised Learning" by Lanting Zhu and Guillermo Goldsztein explores the application of machine learning, specifically supervised learning, to predict the risk of heart attacks using a dataset of 303 patient records. The dataset includes 13 features, such as age, cholesterol levels, blood pressure, and chest pain type, all of which are critical indicators of heart health. Logistic regression was used as the primary model, with the sigmoid activation function to convert raw outputs into probabilities of a heart attack. The data was split into training and validation sets, and the model was trained over 400 epochs to optimize its performance. The resulting model achieved a high validation accuracy of 90%. Binary cross entropy was used as the loss function, ensuring effective error evaluation. The research highlights the potential of machine learning in healthcare, particularly in predicting heart attack risks. It emphasizes that expanding the dataset with more diverse patient information and additional features could further enhance the model's accuracy and generalizability. The study showcases the growing role of AI in medical diagnostics, suggesting that such models could assist healthcare providers in early detection, leading to better patient outcomes and timely interventions.

[4] The paper titled "A Machine Learning Approach for Heart Attack Prediction," published in the International Journal of Engineering and Advanced Technology (IJEAT), Volume-10 Issue-6, August 2021, is authored by Suraj Kumar Gupta, Aditya Shrivastava, Satya Prakash Upadhyay, and Pawan Kumar Chaurasia. The research focuses on predicting heart attack risks using various supervised machine learning classifiers, including Gradient Boosting, Decision Tree, Random Forest, and Logistic Regression. The researchers utilized two key datasets for their analysis: the Framingham dataset and the UCI Heart Disease dataset. The methodology involved several steps: data collection, data preprocessing to handle missing values using imputation methods, model iteration with pipeline enhancements such as hyper parameter optimization and feature engineering, and the selection of the best-performing model for deployment. The iterative approach aimed to improve model accuracy through continuous refinement. The study found that the Gradient Boosting classifier achieved the highest accuracy, with 85.5% on the Framingham dataset and 85.6% on the UCI dataset. The Random Forest classifier also performed well, achieving 84.7% accuracy on the Framingham dataset and 85.6% on the UCI dataset. The Decision Tree classifier achieved 77.1% accuracy on the Framingham dataset and 80.2% on the UCI dataset, while the Logistic Regression classifier achieved 67.7% accuracy on the Framingham dataset and 84.5% on the UCI dataset. The study's findings highlight the effectiveness of machine learning techniques in predicting heart attack risks. The research underscores the potential for using these advanced algorithms to provide early detection and intervention, ultimately improving patient outcomes and healthcare efficiency.

[5] The paper titled "Heart Attack Prediction by Using Machine Learning Techniques" by Sangya Ware, Shanu K. Rakesh, and Bharat Choudhary investigates the application of six machine learning algorithms—Support Vector Machine (SVM), Random Forest, Naive Bayes, Logistic Regression, K-Nearest Neighbors (KNN), and Decision Tree—for predicting heart disease. Using the Cleveland dataset from the UCI Machine Learning Repository, which contains 303 patient records with 14 features (e.g., age, cholesterol, chest pain type), the study preprocessed the data to remove noise and handle missing values. The dataset was split into training (60%) and testing (40%) subsets, and the models were evaluated based on metrics like accuracy, precision, recall, and F1 score. SVM emerged as the top-performing algorithm with an accuracy of 89.34%, followed by Random Forest (86.07%) and Naive Bayes (85.25%), while Decision Tree had the lowest accuracy (74.59%). The study underscores the effectiveness of machine learning in providing accessible, cost-efficient tools for early heart disease detection, especially in regions with limited access to medical specialists. It highlights that different algorithms perform better under specific conditions, suggesting that a combination of methods or advanced techniques like deep learning could further improve outcomes. Future directions include expanding the dataset to include a more diverse population, adding more predictive features, and integrating advanced classification methods for greater accuracy and generalizability in real world applications.

[6] The paper titled "Predicting the Risk of Heart Disease Using Advanced Machine Learning Approach", authored by Dr. Rakhi Waigi, Dr. Sonali Choudhary, Dr. Punit Fulzele, and Dr. Gaurav Mishra, presents a machine learning-based method for predicting heart disease risk using a large dataset from Kaggle containing many records. The dataset included 11 parameters, such as age, gender, height, weight, systolic and diastolic blood pressure, cholesterol and glucose levels, smoking and alcohol habits, and physical activity status. The study implemented and tested two classifiers, Decision Tree and Naive Bayes, to classify individuals into risk and non-risk categories for heart disease. The Decision Tree model achieved a prediction accuracy of 72.77%, with further analysis using heat maps revealing weak correlations between parameters like cholesterol and smoking with heart disease risk. The authors acknowledged the moderate accuracy of the Decision Tree and proposed Naive Bayes as a potentially better alternative. They also recommended integrating real-time monitoring parameters, such as pulse rate, body temperature, and blood flow rate, using wearable devices to improve accuracy in future work.

This study emphasizes the importance of advanced machine learning techniques in healthcare for early detection and prevention of cardiovascular diseases and highlights the need for continuous improvement in prediction models through better data and additional features.

[7] The paper titled "Prediction of Heart Attack Using Machine Learning," published in the IITM Journal of Management and IT, is authored by Akshit Bhardwaj, Ayush Kundra, Bhavya Gandhi, Sumit Kumar, Arvind Rehalia, and Manoj Gupta. The research focuses on predicting heart attack risks using machine learning, specifically utilizing Logistic Regression. The study used the Framingham Heart Study dataset, a long-term cardiovascular cohort study, and employed various preprocessing techniques to handle missing values and outliers. The dataset was split into training (80%) and testing (20%) sets, with K-fold cross validation applied to improve the model's accuracy. The model achieved an impressive accuracy of 87% in predicting heart attack risks. The study identified critical risk factors such as age, sex, smoking habits, cholesterol levels, blood pressure, BMI, and heart rate. It found that men are more susceptible to heart disease than women, and increasing age, smoking, and high blood pressure contribute significantly to heart disease risk. The research highlights the potential of machine learning in early diagnosis, which can help in lifestyle changes and reduce the likelihood of heart problems in the future. It also envisions the development of an AI-based system that can assist medical professionals in making timely treatment decisions, offering quicker predictions with minimal processing delays.

[8] The paper titled "Heart Attack Mortality Prediction: An Application of Machine Learning Methods", authored by Issam Salman, applies machine learning techniques to predict hospital mortality in acute myocardial infarction (AMI) patients using a dataset of 787 records from the Czech Republic and Syria. Addressing challenges like incomplete and imbalanced data, the authors preprocess the dataset by standardizing measurements, handling missing values, and encoding attributes as ordinal, discrete, or binary. Several machine learning models, including C4.5 Decision Tree, Logistic Regression, Naive Bayes, and Tree-Augmented Naive Bayesian (TAN), were evaluated using metrics like accuracy, precision, recall, F-measure, and specificity. The C4.5 Decision Tree achieved the highest accuracy of 94.2% with discrete attributes, while the enhanced TANI algorithm with binary attributes outperformed others, achieving an AUC of 0.953 and a log-likelihood of -2744.43. By analyzing Bayesian Networks, the study also uncovered significant variable relationships, such as the influence of LDL cholesterol levels and hospital type on mortality. The findings highlight the need for diverse models to gain comprehensive insights and emphasize the potential of tailored algorithms for healthcare data. Future research is recommended to explore larger datasets and advanced methods like deep learning to enhance accuracy and generalizability.

Reference	Dataset	Tools used	ML algorithms used	Outcome
Mohammad Alshraideh [1], 2024	Jordan University Hospital	Particle Swarm Optimization (PSO) for feature selection	SVM, Random Forest, Decision Tree, Naive Bayes, KNN	SVM (94.3% accuracy)
Dr. M. Manoj Prabu [2], 2023	Real-time data from wearable device (heart rate, SpO2, body temperature, humidity)	IoT platform (ThingSpeak)	Naive Bayes, Support Vector Machine (SVM), K-Nearest Neighbors (KNN)	Real-time alerts, proactive heart disease management
Lanting Zhu [3], 2022	Kaggle datasets	Logistic Regression, sigmoid activation function, binary cross-entropy	Logistic Regression	Validation accuracy (90%)
Suraj Kumar Gupta [4] , 2021	Framingham , UCI Heart Disease	Data collection, preprocessing (imputation), hyperparameter optimization, feature engineering	Gradient Boosting, Decision Tree, Random Forest, Logistic Regression	Gradient Boosting 85.5%, Random Forest 85.6%, Decision Tree 77.1% , Logistic Regression 67.7% .
Sangya Ware [5], 2020	Cleveland dataset	Data preprocessing (noise removal, handling missing values), accuracy,	Support Vector Machine (SVM), Random Forest, Naive Bayes, Logistic	SVM (89.34%), Random Forest (86.07%), Naive Bayes (85.25%), Decision

		precision, recall, F1 score evaluation	Regression, K Nearest Neighbors (KNN), Decision Tree	Tree (74.59%)
Dr. Rakhi Waigi [6], 2020	Kaggle dataset	Data preprocessing, heat maps, prediction accuracy evaluation	Decision Tree, Naive Bayes	Decision Tree (72.77% accuracy)
Akshit Bhardwaj [7], 2019	Framingham Heart Study	Data preprocessing, K fold cross validation	Logistic Regression	Accuracy (87%)
Issam Salman [8], 2019	Czech Republic and Syria (AMI patients)	Data preprocessing (standardization, handling missing values, encoding attributes), accuracy, precision, recall, F-measure, specificity evaluation	C4.5 Decision Tree, Logistic Regression, Naive Bayes, Tree Augmented Naive Bayesian (TAN)	C4.5 Decision Tree (94.2% accuracy), TANI (AUC 0.953, log likelihood 2744.43)

Table I: Summarization of Various Authors

III. CONCLUSION

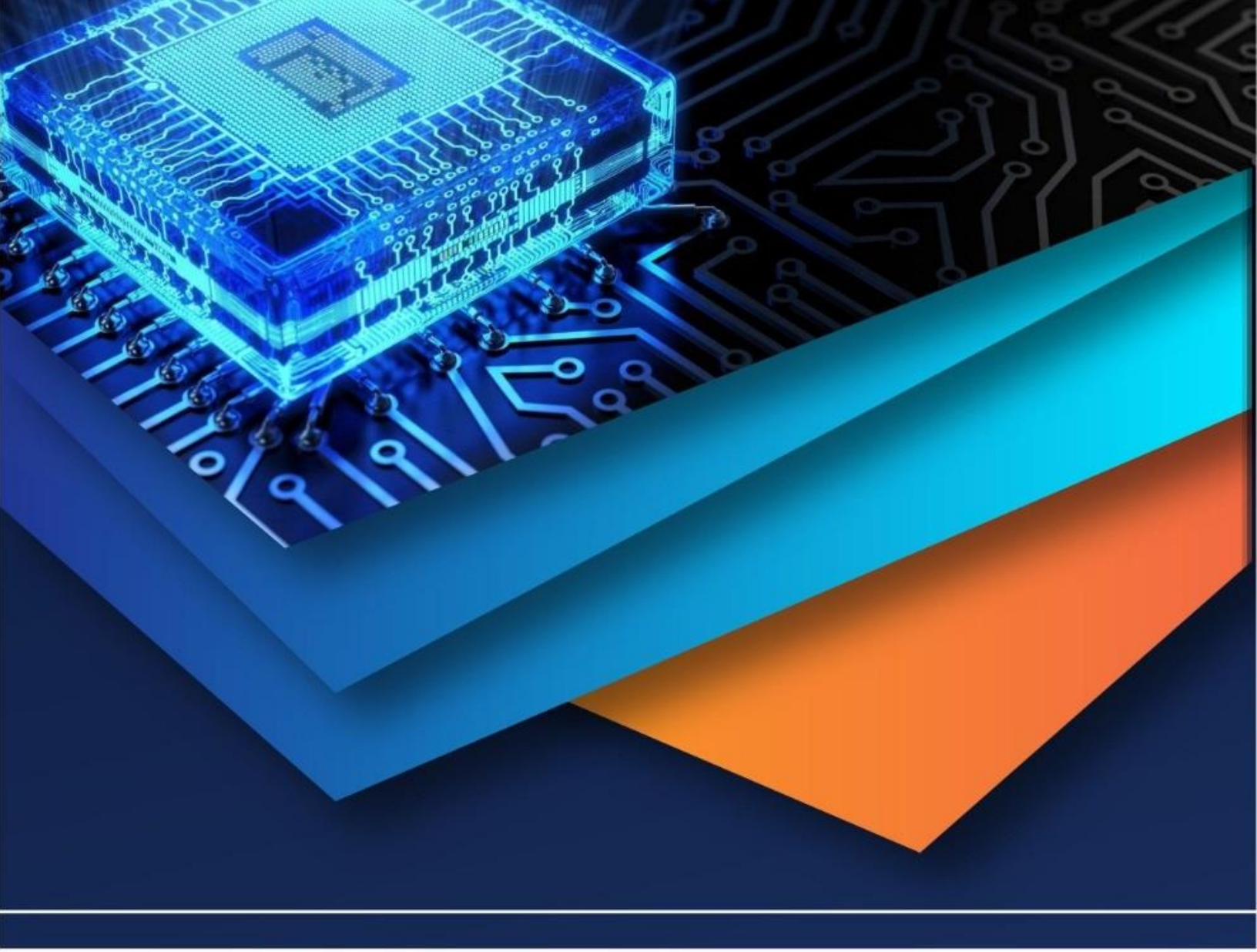
The integration of machine learning in heart attack risk prediction offers a better solution to the challenges of early diagnosis and accurate risk assessment. By analyzing complex medical data, ML models can provide timely and reliable predictions, improving patient outcomes. Future research should focus on enhancing model accuracy, incorporating diverse datasets, and integrating real-time monitoring to further improve prediction capabilities and enable widespread application in healthcare systems.

IV. ACKNOWLEDGMENT

We express our sincere gratitude to our guide, Mrs. Radhika S K for her suggestions and support during every stage of this work. We also convey our deep sense of gratitude to Dr. Jalesh Kumar, Head of the department. we would like to express our sincere gratitude to all those who have contributed to this paper submission to IJRASET. we would like to thank co-authors for their valuable insights and contributions to this work. we would also like to extend our appreciation to the reviewers and editors for their time and effort in reviewing and improving this manuscript. Finally, we are grateful to my institution for providing me with the necessary resources to carry out this research.

REFERENCES

[1] Mohammad Alshraideh, Najwan Alshraideh, Abedalrahman Alshraideh ,Yara Alkayed, Yasmin Al Trabsheh and Bahaaldeen Alshraideh , “Enhancing Heart Attack Prediction with Machine Learning: A Study at Jordan University Hospital” Volume 2024, Article ID 5080332.
[2] Dr.M. Manoj Prabu, S. Ramprakash, G. Ajith, S. Sethuragavan, R. Sathyan, Sri Shakthi Institute of Engineering and Technology, Coimbatore,”Heart Attack Prediction using Machine Learning” India Volume: 07 Issue: 03 | March – 2023.
[3] Lanting Zhu1 and Guillermo Goldsztein, “Predicting the Chance of Heart Attack with a Machine Learning Approach – Supervised Learning” Volume 11 Issue 3 (2022)Centennial High School, USA.
[4] Suraj Kumar Gupta, Aditya Shrivastava, S. P. Upadhyay, Pawan Kumar Chaurasia, " A Machine Learning Approach for Heart Attack Prediction,", International Journal of Engineering and Advanced Technology (IJEAT) ISSN: 2249-8958 (Online), Volume 10 Issue-6, August 2021.
[5] Sangya Ware, Shanu k Rakesh, Bharat Choudhary, ” Heart Attack Prediction by using Machine Learning Techniques”, International Journal of Recent Technology and Engineering (IJRTE) ISSN: 2277-3878, Volume-8 Issue-5, January 2020.
[6] Dr Rakhi Waigi, Dr Sonali Choudhary, Dr Punit Fulzele, Dr Gaurav Mishra, “Predicting The Risk Of Heart Disease Using Advanced Machine Learning Approach”, European Journal of Molecular & Clinical Medicine, ISSN 2515-8260 Volume 7, Issue 07, 2020.
[7] Akshit Bhardwaj, Ayush Kundra,Bhavya Gandhi, Sumit Kumar4, Arvind Rehalia, Manoj Gupta, ”Prediction of Heart Attack Using Machine Learning”, 2019.
[8] Issam Salman, ”Heart attack mortality prediction: an application of machine learning methods”, Department of Software Engineering, Faculty of Nuclear Sciences and Physical Engineering, Czech Technical University, The Czech Republic Volume 27 Number 6 Article 24



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)