



e-ISSN: 2319-8753 | p-ISSN: 2347-6710

IJRSET

International Journal of Innovative Research in
SCIENCE | ENGINEERING | TECHNOLOGY



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

Volume 13, Issue 4, April 2024

ISSN INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA

Impact Factor: 8.423

9940 572 462

6381 907 438

ijirset@gmail.com

www.ijirset.com

Precision Cardiovascular Health Prediction using ML Algorithm.

Harshitha M¹, Karuna C Kolar¹, Kavana N S¹, Uma Halli¹, Dr. Arun Kumar G Hiremath²,
Prof. Priya V³

U G Students, Department of CSE, Bapuji Institute of Engineering and Technology, Davangere, India¹

Associate Professor, Department of CSE, Bapuji Institute of Engineering and Technology, Davangere, India²

Assistant Professor, Department of CSE, Bapuji Institute of Engineering and Technology, Davangere, India³

ABSTRACT: Cardiovascular diseases (CVDs) present a global health challenge, demanding accurate prediction models for prevention and management. This study employs machine learning algorithms to forecast cardiovascular health outcomes, leveraging diverse datasets such as electronic health records and lifestyle factors. Through rigorous validation, our models demonstrate superior predictive accuracy, facilitating personalized risk assessments. Integration of multiple data modalities ensures transparent predictions, holding promise in alleviating the burden of CVDs and improving patient outcomes globally. Continued research and implementation efforts are crucial for realizing the full potential of precision cardiovascular health prediction in clinical practice.

KEYWORDS: predictive frameworks, lifestyle factors, personalized risk assessments, feature selection, interpretability methods, precision medicine strategies, healthcare interventions.

I. INTRODUCTION

Globally, Cardio Vascular Disease (CVDs) is the primary cause of morbidity and mortality, accounting for more than 70% of all fatalities. According to the 2017 Global Burden of Disease research, cardiovascular disease is responsible for about 43% of all fatalities. Common risk factors for heart disease in high-income nations include lousy diet, cigarette use, excessive sugar consumption, and obesity or excess body fat. However, low- and middle-income nations also see a rise in chronic illness prevalence. Between 2010 and 2015, the global economic burden of cardiovascular diseases was expected to reach roughly USD 3.7 trillion.

In addition, technologies such as electrocardiograms and CT scans, critical for diagnosing coronary heart disease, are sometimes too costly and impractical for consumers. The reason mentioned above has resulted in the deaths of 17 million people. Twenty-five to thirty percent of firms' annual medical expenses were attributable to employees with cardiovascular disease. Therefore, early detection of heart disease is essential to lessen its physical and monetary cost to people and institutions. According to the WHO estimate, the overall number of deaths from CVDs would rise to 23.6 million by 2030, with heart disease and stroke being the leading causes. To save lives and decrease the cost burden on society, it is vital to apply data mining and machine learning methods to anticipate the chance of having heart disease.

Heart disease, specifically (CVDs), is a leading cause of morbidity and mortality worldwide, accounting for over 70% of all global deaths. According to the Global Burden of Disease Study 2017, CVD accounts for more than 43% of all deaths. Common risk factors associated with heart disease include unhealthy food, tobacco, excessive sugar, and overweight or extra body fat, often found in high-income countries. However, low- and middle-income countries are also seeing an increase in the prevalence of chronic diseases. The economic burden of CVDs worldwide has been estimated to be approximately USD 3.7 trillion between 2010 and 2015.

Furthermore, devices such as electrocardiograms and CT scans, essential for detecting coronary heart disease, are often too expensive and infeasible for many low- and middle-income countries. Therefore, early determination of heart disease is crucial to decrease its physical and financial burden on individuals and organizations. According to a WHO report, by 2030, the total number of deaths due to CVDs will increase to 23.6 million, mainly from heart disease and stroke. Therefore, it is crucial to use data mining and machine learning techniques to predict the likelihood of developing heart disease in order to save lives and reduce the economic burden on society.

In the medical field, a vast amount of data is generated daily using data mining techniques, and we can find hidden patterns that can be used for clinical diagnosis. Therefore, data mining plays a vital role in the medical field, which can be proved by the work conducted in the past few decades. Many factors, such as diabetes, high blood pressure, high cholesterol, and abnormal pulse rate, need to be considered when predicting heart disease. Often, the medical data available need to be completed, affecting the results in predicting heart disease.

Machine learning plays a crucial role in the medical field. Using machine learning, we can diagnose, detect, and predict various diseases. Recently, there has been a growing interest in using data mining and machine learning techniques to predict the likelihood of developing certain diseases. Although some studies have attempted to predict the future risk of the progression of the disease, they have yet to find accurate results. The main goal of this work is to accurately predict the possibility of heart disease in the human body.

In this work, we aim to investigate the effectiveness of various machine learning algorithms in predicting heart disease. The project's success hinges on rigorous testing and validation, ensuring its accuracy and safeguarding patient data. As this technology continues to evolve and access to data grows, the Heart Attack Prediction System using Machine Learning has the potential to revolutionize cardiac care. By providing early warnings and enabling timely interventions, it not only saves lives but also alleviates the burden of heart disease on healthcare systems, offering a proactive approach to combat this prevalent and deadly condition

II. RELATED WORK

In [1] the purpose of this study was to develop a machine-learning model based on routine, quantitative, and easily measured variables to predict the presence and extent of subclinical atherosclerosis (SA) in young, asymptomatic individuals. The risk of having SA estimated by this model could be used to refine risk estimation and optimize the use of imaging for risk assessment. Evaluation of ML algorithm effectiveness surveyed machine learning approaches for early detection of CVD.

In [2] this work aims to gain knowledge from a heart disease dataset and analyses several data mining classification techniques seeking improved accuracy and a lesser error rate in the results. The dataset is analyzed using two different data mining tools, i.e., WEKA and Tanagra. The results show that the Naive Bayes algorithm and the C-PLS algorithm outperform others with an accuracy of 83.71% and 84.44% respectively.

In [3] this work suggests that growing healthcare industry is generating a large volume of useful data on patient demographics, treatment plans, payment, and insurance coverage—attracting the attention of clinicians and scientists alike. In recent years, a number of peer-reviewed articles have addressed different dimensions of data mining application in healthcare. In this paper, a review of the literature on healthcare analytics using data mining and big data.

In [4] this work describes the main aim to discuss and compare the similarities and differences between the conventional risk prediction approach and AI risk prediction approach in the CVD studies. Strengths and limitations of both approaches are outlined. Besides, the research progress and the usage of biomarkers in the risk assessment of CVD are reviewed. Furthermore, problems and challenges faced in the current CVD risk prediction studies are discussed and lastly, future prospects of CVD prediction models are proposed. Before the advancement of information technology, most researchers and scientists used basic statistics to make inferences and conclusions of their findings.

III. EXISTING SYSTEM

Research Studies and Academic Projects:

Researchers and data scientists have been leveraging machine learning algorithms such as logistic regression, decision trees, random forests, and neural networks in analyzing historical patient data to forecast the likelihood of a heart attack. These models rely on manually provided data points like blood pressure, age, gender, and smoking habits to generate predictions, aiming to improve early detection and prevention strategies for cardiovascular diseases.

Healthcare Applications and Startups:

Within the realm of health monitoring apps, numerous platforms enable users to input their health data manually, including crucial metrics like blood pressure, age, gender, and lifestyle habits. These applications often incorporate

algorithms to offer users insights into their heart health risks, including the potential for heart attacks. Additionally, wearable devices like smartwatches gather data on physical activity, heart rate, and sometimes blood pressure, which some advanced wearables utilize, along with user-provided information, to evaluate heart health risks, albeit with varying degrees of accuracy.

Telemedicine Platforms:

Telemedicine platforms have evolved to integrate predictive algorithms that utilize manual input data from patients to provide initial assessments of heart health risks. These platforms typically prioritize secure communication channels to safeguard patient privacy while delivering valuable preliminary insights into potential cardiovascular concerns.

Hospital Information Systems:

Hospitals and healthcare providers rely on Electronic Health Records (EHR) systems, often equipped with modules designed for predicting and managing chronic diseases. These systems leverage vast amounts of patient data to aid in the early detection, management, and prevention of cardiovascular diseases, contributing to more effective patient care and treatment strategies.

IV. METHODOLOGY

1. Data Collection:

- **Medical History:** Information about previous heart-related issues, surgeries, or family history of heart diseases.
- **Lifestyle Factors:** Data on habits such as smoking, physical activity, diet, stress levels, and alcohol consumption.
- **Clinical Measurements:** Includes various health metrics such as blood pressure, cholesterol levels, blood sugar levels, BMI (Body Mass Index), and other relevant clinical data.

2. Data Processing and Feature Selection:

- The collected data is processed to remove inconsistencies and outliers.
- Feature selection techniques are employed to identify the most relevant variables for predicting heart attacks. These might include specific medical conditions, genetic predispositions, or lifestyle choices known to influence cardiovascular health.

3. Machine Learning Model Development:

- **Algorithm Selection:** Choosing appropriate machine learning algorithms, like logistic regression, random forests, or neural networks, based on the nature and complexity of the data.
- **Training the Model:** Using a subset of the collected data to train the machine learning model. During this phase, the model learns patterns and relationships within the data that are indicative of heart attack risks.

4. Prediction and Risk Assessment:

- **Input Patient Data:** Medical professionals can input a patient's information, including medical history, lifestyle factors, and clinical measurements, into the trained model.
- **Risk Assessment:** The machine learning model processes the input data and calculates the individual's risk of experiencing a heart attack. This risk assessment is based on the patterns learned during the training phase.
- **Machine Learning Models:** Utilize machine learning algorithms (such as Logistic Regression, Random Forest, or Neural Networks) to create predictive models. Train these models using historical patient data, considering various risk factors, and use them to predict the patient's risk.

5. Timely Warnings and Interventions:

- **Tailored Interventions:** Based on the risk assessment, tailored interventions and preventive measures can be recommended. These could include lifestyle changes, medication, or further diagnostic tests.



System Design:

The system design for precision cardiovascular health prediction using ML algorithms begins with comprehensive data collection and integration processes to gather diverse datasets, including electronic health records, genetic profiles, and lifestyle factors. Feature engineering and selection methods are then employed to extract pertinent predictors from the integrated data. Supervised, unsupervised, and deep learning models are developed to construct predictive frameworks capable of accurately forecasting cardiovascular health outcomes based on the chosen features. Rigorous model evaluation and validation procedures are carried out to assess predictive accuracy and generalizability, with a focus on interpretability and transparency to enhance clinical relevance. The system is designed to be scalable, allowing for integration with existing healthcare infrastructure and standards.

System Development:

The system development phase of precision cardiovascular health prediction using ML algorithms involves various steps. Initially, data preprocessing tasks are undertaken to clean and preprocess the collected data, including handling missing values and outliers. Subsequently, machine learning models are trained using the pre-processed data, with experimentation on various algorithms and hyperparameters to optimize performance. Validation and hyperparameter tuning procedures are then performed to enhance predictive accuracy and generalization. Interpretability methods are implemented to ensure that healthcare professionals can understand and trust the predictions made by the system. Finally, the developed models are integrated into existing healthcare systems or deployed as standalone applications/interfaces, with ongoing maintenance and support provided to ensure system reliability and effectiveness.

4.1 SYSTEM ARCHITECTURE

The architecture for precision cardiovascular health prediction utilizing ML algorithms, specifically logistic regression, is designed to efficiently process and analyse diverse datasets while ensuring accurate predictions and clinical relevance. The system architecture comprises several key components: data ingestion and preprocessing, feature engineering, model training and evaluation, and prediction generation. Initially, electronic health records, genetic profiles, and lifestyle factors are ingested and pre-processed to handle missing values, outliers, and inconsistencies. Feature engineering techniques are then applied to extract relevant predictors, followed by logistic regression model training using the pre-processed data. Rigorous evaluation procedures, including cross-validation, assess model performance and generalizability, ensuring robust predictions.

User Layer:

In the user layer, individuals interact with the system by providing essential input data related to their cardiovascular health. This input includes personal medical history, lifestyle habits, demographic information, and any relevant symptoms or risk factors. Users engage with the system to receive predictions or risk assessments regarding their cardiovascular health status, influencing the data used for prediction and interpretation of results.

GUI Layer:

The graphical user interface (GUI) presents users with user-friendly input forms and data entry fields for inputting their information for cardiovascular health prediction. It also displays prediction results in a visually understandable format, such as charts or textual summaries. Acting as a bridge between the user layer and the system layer, the GUI translates user actions into system commands and vice versa, enhancing user experience and accessibility.

System Layer:

At the system layer, user input provided through the GUI is processed to make predictions about cardiovascular health. This layer houses the logistic regression model, trained using historical patient data to predict the likelihood of cardiovascular events based on input variables. Data preprocessing tasks, such as handling missing values and normalizing input features, are conducted to prepare the data for model training and prediction. Generated predictions are then conveyed back to the GUI layer for user presentation, completing the feedback loop of the system.

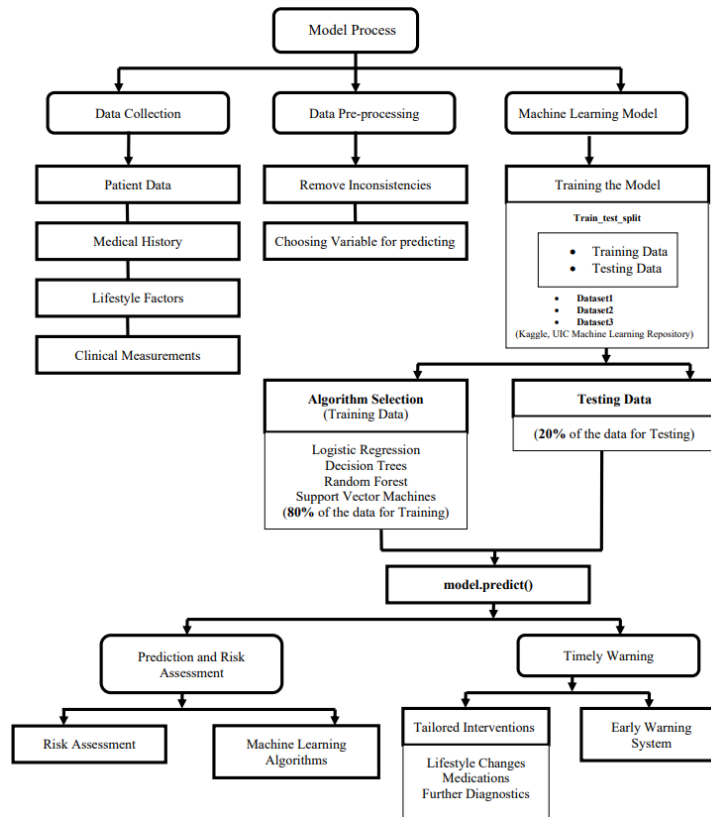


Fig 4.1: Block Diagram

4.2 FLOWCHART

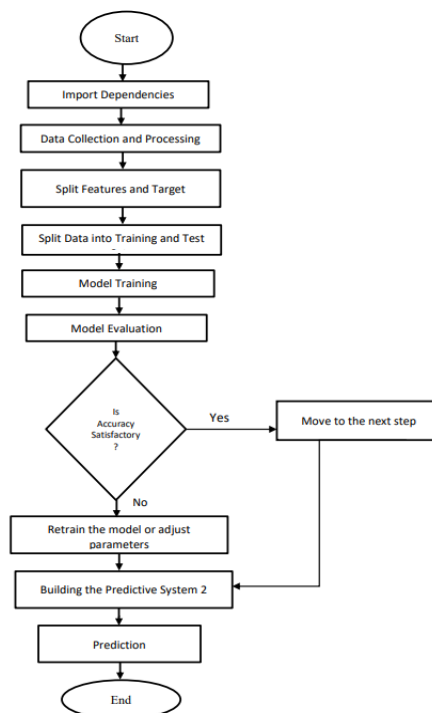


Fig 4.2: Flowchart



V. FUTURE SCOPE

The future of heart attack prediction models holds vast potential for advancements in precision medicine, incorporation of social determinants of health, enhanced predictive analytics, longitudinal monitoring, interoperability with healthcare systems, and population health management. By tailoring models to individual genetic profiles and lifestyle factors, integrating data on social determinants, and leveraging advanced machine learning techniques, these models can provide more accurate and personalized risk assessments. Seamless integration into electronic health records and telehealth platforms facilitates widespread adoption, while population-level utilization informs public health initiatives. Ultimately, these innovations aim to revolutionize cardiovascular risk assessment and preventive care, leading to improved outcomes and population health worldwide.

VI. CONCLUSION

In conclusion, the development and integration of machine learning (ML) models for heart attack prediction represent a significant leap forward in cardiovascular healthcare. By leveraging advanced algorithms to analyse diverse patient data, these models offer personalized risk assessments, enabling proactive interventions and targeted preventive measures. With the potential to identify individuals at heightened risk of heart attacks and guide clinical decision-making, ML-based approaches hold promise for improving patient outcomes and reducing the burden of heart disease on public health. By embracing innovation and collaboration between healthcare professionals, researchers, and technology experts, we can harness the power of data-driven insights to transform cardiovascular care, ultimately saving lives and promoting heart health for all.

REFERENCES

- [1] Halah A. Al-Alshaikh, Prabu P, Abeer A. AlSanad , “Halah A. Al-Alshaikh, Prabu P, Abeer A. AlSanad ”, 3 April 2024
- [2] Sara Retal, Brahim EL Bhiri, “Towards revolutionizing precision healthcare: A systematic literature review of artificial intelligence methods in precision medicine”, Informatics in Medicine, 2024
- [3] Elias Dritsas, Maria Trigka, "Efficient Data-Driven Machine Learning Models for Cardiovascular Diseases Risk Prediction", 19 January 2023
- [4] Aizatul Shafiqah Mohd Faizal a, T. Malathi Thevarajah b, Sook Mei Khor c, Siow-Wee Chang a. "A review of risk prediction models in cardiovascular disease: conventional approach vs. artificial intelligent approach", August 2021



INTERNATIONAL
STANDARD
SERIAL
NUMBER
INDIA



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN SCIENCE | ENGINEERING | TECHNOLOGY

 9940 572 462  6381 907 438  ijirset@gmail.com



www.ijirset.com

Scan to save the contact details