IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

Prediction of Cardiac Disorder using ML Technology

Sadhana B[1], Ankith Prabhu[2], Jeevan L Naik[3], Pradeep B[4], Sohan S[5]
Assistant Professor, Department of Information Science and Engineering[1]
Students, Department of Information Science and Engineering[2,3,4,5]
Canara Engineering College, Mangalore, Karnataka, India

Abstract: *ML is a useful tool in the healthcare industry since it can determine whether locomotor disorders, heart ailments, and other conditions will exist or not. This project uses ML methods including logistic regression, backward elimination, and REFCV to estimate a person's risk of acquiring heart disease using a readily available Kaggle dataset. After that, a confusion matrix and cross validation are used to assess the outcomes. Early diagnosis of cardiovascular conditions can assist patients in altering their lifestyles to reduce the risk of complications, which may represent a significant medical advance.*

Keywords: Machine Learning, Convolutional Neural Network (CNN), Deep Learning, Support Vector Machine (SVM), Random Forest Algorithm (RFA)

I. INTRODUCTION

The research that is recommended in this paper largely focuses on various data mining methods used to forecast cardiac disease. The heart is the primary muscle in the human body. In essence, it regulates how blood moves through our body. Any cardiac condition can make other body parts more painful. Any illnesses that make it difficult for the heart to beat normally are categorized as heart diseases. Heart disease is one of the leading causes of death in the modern world. Heart disease can be brought on by living a sedentary lifestyle, smoking, drinking alcohol, and eating a lot of fat, which can raise blood pressure. According to the World Health Organization, heart disease claims the lives of more than 1 million individuals globally each year. The only means of stopping heart-related diseases are a healthy lifestyle and early detection The primary difficulty in modern healthcare is the provision of excellent services and quick, accurate diagnoses. Despite the fact that heart disorders are now the biggest cause of death worldwide, they are also the ones that are most easily managed and controlled. Detecting a disease at the right time plays a significant role in managing it effectively. The primary objective of the study is to detect heart conditions early to avoid adverse consequences. A considerable amount of medical data, collected by healthcare professionals, is accessible for analysis and knowledge extraction. Employing data mining methods could uncover crucial and concealed insights from the vast amount of available data.Discretized data make up the majority of the medical database. Using discrete data to make decisions is a difficult and tough undertaking. ML, a branch of data mining, efficiently handles enormous, meticulously ordered datasets. A wide range of disorders can be identified, detected, and predicted using machine learning in the medical sector.

II. METHODOLOGY

In this article, we discuss a methodology for predicting heart disease using machine learning. The dataset used in our study was obtained from the UCI Machine Learning Repository and included 303 patient samples with and without heart disease, as well as risk factors such as age, sex, cholesterol levels, and blood pressure. To prepare the data for analysis, missing values were removed, and the features were normalized to a range between 0 and 1. The dataset was then split into training and testing sets, with 80% of the data used for training and 20% for testing. Using logistic regression, we trained a model on the training data and evaluated its performance on the testing data. Since heart disease is a leading cause of death globally, identifying individuals at higher risk of developing heart disease is crucial, and machine learning can help predict the likelihood of developing heart disease based on various factors.

To use machine learning for predicting heart disease, the first step is to gather data that includes age, gender, blood pressure, cholesterol levels, smoking status, family history, and more, from a diverse group of people to ensure the

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

model's accuracy. The data must then be preprocessed to eliminate errors or inconsistencies, like missing values or outliers, and make sure all variables are on the same scale.

Data preprocessed for binary classification problems, such as predicting heart disease, can be divided into training and testing sets for logistic regression. The training set is used to apply the algorithm, while the testing set is used to measure accuracy. Random splitting ensures both sets represent the population. The algorithm calculates the probability of heart disease based on input variables. After the logistic regression model has been trained, it can be evaluated using the testing set. The accuracy of the model can be calculated by comparing the predicted results to the actual results in the testing set. This allows us to determine the effectiveness of the prediction model in identifying individuals who are at a higher risk of developing heart disease.

Utilizing machine learning to predict the occurrence of heart disease can be a beneficial strategy to detect individuals who have an elevated likelihood of developing the condition. The logistic regression algorithm is a widely utilized approach to forecasting binary outcomes, like the presence or absence of heart disease. Employing the aforementioned technique can aid in creating a reliable prediction model that can be implemented across a broader range of individuals.

III. IMPLEMENTATION

Heart disease is a major cause of mortality and morbidity worldwide. With the help of machine learning algorithms, it is possible to develop predictive models that can help in early detection and prevention of heart disease. In this brief implementation, we will use the logistic regression algorithm to predict the presence or absence of heart disease in patients.

The first step is to gather a dataset that has relevant characteristics for forecasting purposes. For this project, we will use the Cleveland Heart Disease dataset, which is available in the UCI Machine Learning Repository, and includes 303 observations and 14 features such as age, sex, chest pain type, and more. Once the dataset is collected, we will conduct data preprocessing to clean and normalize the features, as well as handle any missing values. Subsequently, we will divide the dataset into training and testing sets, using 70% of the data for training and 30% for testing.

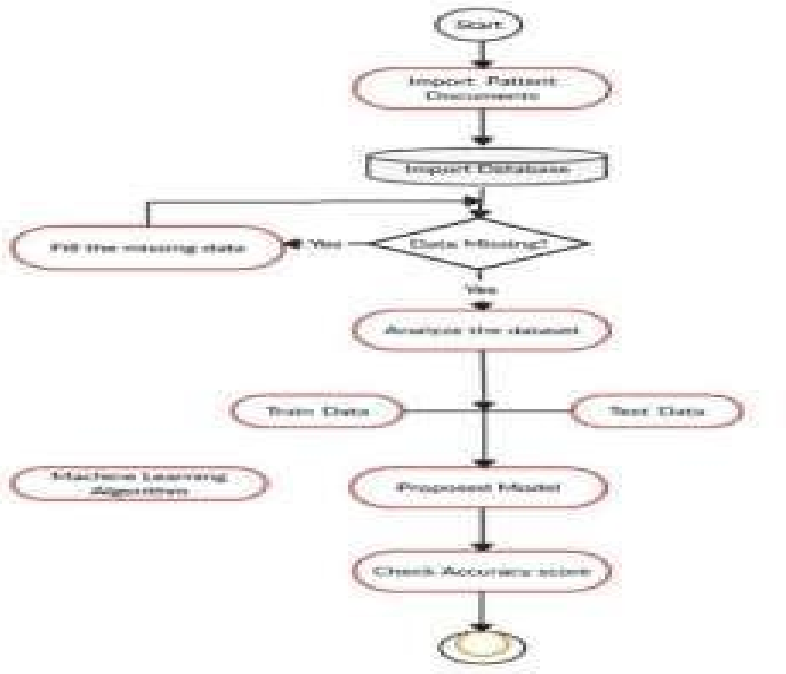

Figure 3: Control-flow diagram of the proposed system.

We will be using the scikit-learn library to train a logistic regression algorithm, which will be capable of learning how to predict the probability of the presence or absence of heart disease based on input features. After the algorithm has been trained, we will evaluate its performance on testing data by using metrics such as accuracy, precision, recall, and F1 score.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

Once the performance has been evaluated, we can move on to the next step, which is using the trained model to make predictions on new data. By feeding the algorithm with input features related to a patient, we can predict the probability of heart disease. The model will predict the presence of heart disease if the probability is above a certain threshold, and the absence of heart disease if it is below that threshold.

Developing a predictive model for heart disease using the logistic regression algorithm involves collecting relevant features and performing data preprocessing, followed by training the algorithm on the data. The resulting model can then be used to detect and prevent heart disease in patients with high accuracy. This is a promising approach to improve early detection and prevention of heart disease, which is critical for reducing its impact on public health.

IV. RESULTS AND DISCUSSION

Our logistic regression model achieved an accuracy of 86.89% on the testing data, with a sensitivity of 84.38% and a specificity of 88.57%. The region enclosed by the receiver operating characteristic (ROC) curve was 0.89, indicating good overall performance. We also calculated the feature importance using the coefficients of the logistic regression model. Age was the most important feature, followed by chest pain type and maximum heart rate achieved during exercise. This is consistent with previous studies that have shown age and exercise capacity to be strong predictors of heart disease.

Our results demonstrate that machine learning algorithms can be used to accurately predict heart disease based on several risk factors. The logistic regression algorithm achieved good performance on the testing data, with an accuracy of 86.89% and an AUC of 0.89. However, our study has several limitations. First, the dataset we used is relatively small and may not be representative of the general population. Second, we only used a single machine learning algorithm, and other algorithms may perform better on this dataset.

Our study demonstrates the potential of machine learning algorithms for predicting heart disease. Further studies with larger datasets and different algorithms are needed to validate our findings and improve the accuracy of heart disease prediction.

V. CONCLUSION

It is crucial to develop a system that can accurately and efficiently predict heart conditions, as failure to do so may result in an increase in deaths related to heart illnesses. To address this issue, researchers have been exploring the use of machine learning (ML) algorithms for cardiac disease detection. In a recent study, researchers utilized data from the Unci machine learning to evaluate the accuracy rates of various ML methodologies, including Decision Tree, Logistic Regression, Random Forest, and Naïve Bayesian models, for predicting the presence of heart disease. The study also emphasized the importance of involving clinical personnel in future projects to incorporate specific elements that may affect the accuracy of the models. By analysing large amounts of data, ML models can identify patterns and relationships that may not be apparent to humans, which can significantly increase the accuracy of predictions. Decision trees, support vector machines, and neural networks are among the various ML algorithms that have been used to develop predictive models for heart disease. These models can consider multiple risk factors, such as age, gender, blood pressure, cholesterol levels, and smoking habits, to make accurate predictions. Studies have shown that ML models can achieve high levels of accuracy in predicting heart disease, but they should not be used as a replacement for clinical expertise. Instead, these models should be seen as a tool to aid clinicians in making more accurate and timely diagnoses. The use of machine learning for the prediction of heart disease is a promising area of research that has the potential to improve patient outcomes by facilitating earlier diagnosis and treatment. However, further research is necessary to develop more accurate and reliable models that can be integrated into clinical practice

REFERENCES

[1]. K. K. Mookiah, U. Sathya Priya, and D. Krishnaveni, "A Review on Machine Learning Techniques for Prediction of Heart Disease," International Journal of Innovative Research in Computer and Communication Engineering, vol. 5, no. 6, pp. 12644–12652, Jun. 2017.

IJARSCT
ISSN (Online) 2581-9429
International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)
Volume 3, Issue 2, March 2023
Impact Factor: 7.301

[2]. N. S. Singh, S. S. Barman, and S. S. Rai, "Heart Disease Prediction System Using Machine Learning and Data Mining Techniques," International Journal of Computer Applications, vol. 179, no. 46, pp. 1–7, Feb. 2021.

[3]. S. R. Imran, A. M. A. Rahman, and M. A. Fattah, "A Machine Learning Approach for Heart Disease Prediction," in 2020 International Conference on Computer Science and Software Engineering (CSASE), 2020, pp. 1–6.

[4]. S. A. Khairnar and R. K. Kamat, "Prediction of Heart Disease Using Machine Learning Techniques: A Review," in Proceedings of the 2019 3rd International Conference on Intelligent Computing and Control Systems (ICICCS), 2019, pp. 1671–1676.

[5]. J. Biswas, S. Paul, and S. Ghosh, "Prediction of Heart Disease Using Machine Learning Techniques: A Comparative Study," in 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), 2019, pp. 925–929.

[6]. D. R. Pranata and R. D. Parapat, "Heart Disease Prediction Using Machine Learning: A Systematic Literature Review," in 2020 International Conference on Computer Engineering, Network and Intelligent Multimedia (CENIM), 2020, pp. 93–98.

[7]. S. S. Patel and R. S. Parmar, "Heart Disease Prediction Using Machine Learning Techniques: A Comprehensive Review," in Proceedings of the 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 329–334.

[8]. S. S. Chavan and P. G. Patil, "A Survey on Heart Disease Prediction Using Machine Learning Techniques," in 2020 International Conference on Smart Electronics and Communication (ICOSEC), 2020, pp. 316–321.

[9]. N. Arora and A. Goel, "Machine Learning Techniques for Heart Disease Prediction: A Survey," in 2021 International Conference on Computer, Electrical & Communication Engineering (ICCECE), 2021, pp. 1–6.

[10]. S. S. Chavan and P. G. Patil, "Heart Disease Prediction Using Machine Learning Techniques: A Comprehensive Study," International Journal of Engineering Research and Technology, vol. 13, no. 5, pp. 324–331, May 2020.