International Journal of Recent Development in Engineering and Technology
Website: www.ijrdet.com (ISSN 2347-6435 (Online) Volume 13, Issue 3, March 2024)

# Effective Heart Disease Prediction Using Hybrid Machine Learning Techniques

**Anjali Pandey, Prof. Saurabh Sharma, Prof. Vishal Paranjape**

*Global Nature Care Group Sangathan Institution, Jabalpur, Madhya Pradesh*

***Abstract:*** Type 2 diabetes is one of the most common diseases in the world. Diabetes has been diagnosed for a long time, but just recently. As Deep Learning (DL) grows increasingly difficult, researchers are trying to solve the complicated puzzles. Thus far, the accuracy of the model is 0.011%. This problem uses a range of machine learning techniques, many of which have never been applied previously. It is intriguing to look into their ability to predict diabetes. There isn't currently a comparison or review of all the combined modeling and technique concepts. a duration of six years This article addressed all ML and DL prediction strategies. To make the Pima Indians more effective, they were also assigned peculiar ML classifications. This article proposes employing the classifiers for a more comprehensive diabetes prediction, even though they have a rating of 68–74 percent.

***Keywords:*** Cardiovascular Diseases; Machine Learning; Support Vector Machines; Decision Tree Classifier; KNN Classifier;LogisticRegression.

## I. INTRODUCTION

People have hectic schedules and routines that lead to stress and anxiety. Furthermore, there has been a noticeable rise in the population of obese, agitated, and cigarette addicts [13]. This poses a significant danger of heart disease development. Cardiovascular diseases (CVDs) are the major cause of death worldwide, accounting for 17.9 million deaths annually or 31% of all deaths [14]. Four out of every five deaths from CVD are caused by heart attacks and strokes, and one-third of these deaths occur prematurely in people under the age of 70[14]. This work can be used to predict a possible cardiac condition by comparing the accuracy of many algorithms and selecting the most effective one. Heart failure is a frequent consequence of CVDs. Heart disease is affecting more men and women than ever before, regardless of age[13]. For those who have cardiovascular disease or are at high risk for developing it, early detection is critical, and this is where an AI model can be very beneficial.

## II. PROBLEM STATEMENT

During these moments of stress and anxiety, persistent diseases are not uncommon. The only means to lower death rates are through early disease detection and efficient treatment. Because of this, we are trying to predict and analyze heart disease at an early stage by considering many characteristics such as age, gender, blood pressure, heart rate, diabetes, and more. Because there are so many other variables involved, the forecast is somewhat of a test.

## III. LITERATURE REVIEW

Many earlier studies that employed machine learning to diagnose cardiac issues served as the basis for this work. Below is a summary of the relevant literature:

• Vijeta Sharma et al.'s study, "Heart Disease Prediction using Machine Learning Techniques,"

The primary goal of the author's research was to evaluate various machine learning techniques for the prediction of heart disease and determine which one worked best. It has been found that naive Bayes and decision trees are not as good at predicting cardiac illness as random forests and support vector machines. Moreover, this study does not cover several important parameters that should be considered during the model-development process. Various medical histories are not included. This study uses a limited number of data.

The work by Aditi Gavhane et al. titled "Prediction of Heart Disease Using a Machine Learning" This paper is primarily concerned with the investigation of neural network methods for efficient analysis of cardiac disease prediction. This work also makes use of the Multi-Layer Perceptron Algorithm to forecast cardiac disease. Age, sex, blood pressure, diabetes, cholesterol, and heart rate are among the variables it considers. We can also improve the accuracy of the model by adding a few other features.

• Farzan Tasnim et al. carried out a comparative study on feature selection and data mining techniques for the prediction of heart disease. In order to identify the most efficient technique, we provide a comparison analysis of many machine learning algorithms for cardiac disease prediction in this paper.

The best machine learning algorithms for predicting heart disease have been found to be random forest and support vector machines. Principal component analysis is also used to improve accuracy. This study does not address other important issues that need to be considered.

• Prof. (Dr.) Kanak Saxena and colleagues created an effective decision tree-based system for predicting heart disease. This study aims to demonstrate the predictive power of machine learning algorithms and health markers in the context of cardiac disease. The decision tree and computer-aided illness diagnostic algorithm employed in this paper's core research are used to forecast heart disease in order to assess whether or not a consumer has it. This approach is also not very successful and has low precision. We can also draw the conclusion that it is not a good idea to use decision trees to forecast cardiac disease.
Rashmi G. Sabojiet al. conducted A Scalable Solution for Heart Disease Prediction Using Classification Mining Technique.

This article's primary focus is on the comparative comparison of naïve basis and random forest for heart disease prediction. In this study, random forest was found to be a better method than naive bayes, using data from Hadoop disturbed files. There are just two approaches used in the inquiry. A few parameters were gathered during the feature engineering process. The result has been observed to exhibit overfitting.

## IV. GENEARLIZED METHODOLOGY

After analyzing the outcomes of the previous methods, we used Python and pandas operations to perform heart disease classification for the data received from the UCI repository. It provides an easily comprehensible visual depiction of the dataset, workspace, and advancements in predictive analytics. The machine learning (ML) process begins with pre-processing data and proceeds to feature selection based on data cleaning, categorization, and performance assessment. The random forest approach is utilized to improve the outcome's accuracy.

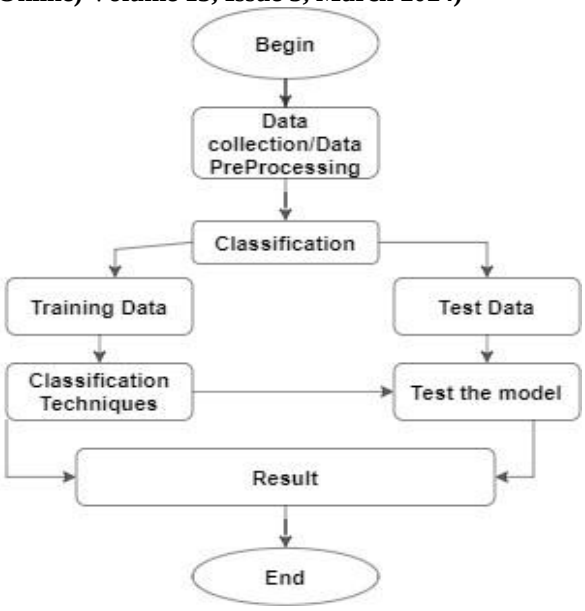


Fig. 1: Flowchart of the Heart disease Prediction

- Comparative study of Machine learning algorithms for heart disease prediction.
- Data gathering from different sources.
- Feature engineering and parameter tuning.
- Finding out the best Machine learning algorithm for heart disease prediction and use of other Machine learning features for boosting the performance and increasing the accuracy.
- Model Implementation
- Using Machine learning matrices for performance evaluation.

## V. ALGORITHM

K-Nearest Neighbor is a classification algorithm. When ki is a small positive integer, the class that is most common among a data point's closest neighbors defines the class of that particular data point.

One algorithm utilized in machine learning approaches for regression and classification is the support vector machine. Because it performs better than other algorithms, it is commonly used as a classification approach. Using an attribute's everlasting coordinate from the dataset, this method displays a hyperplane.

The binary predictive analysis technique known as logistic regression is utilized when the target variable is dichotomous. The logistic regression model explains the relationship between one dependent binary variable and one or more independent variables.

Decision Tree Classifier: It groups the characteristics to infer the intended value. Classification trees are a type of tree model where the objective parameter might have only a few possible values. In these, the traits that coincide to generate the class labels are defined by the branches, and the class labels are represented by the leaves. Regression trees are decision trees in which the objective border can have infinite values.

## VI. ADVANTAGES

- Early and online forecasting can be particularly helpful in the event of a medical emergency.
- A free web platform offers patients a diagnosis at a reasonable price.

## VII. CONCLUSION

Predicting cardiac disease is a critical and labor-intensive task in the medical field. The mortality rate can be lowered if the illness is discovered in its early stages and preventative measures and suitable treatment are feasible. The algorithms are tested using multiple features. The recommended approach seeks to predict illnesses with accuracy. The decision classifier approach has shown to be highly successful in illness prediction when parameters such as age, BMI, cholesterol, and more are used. The feature of BMI improved prediction accuracy. Thus, by assessing the results, the proposed method generates a more precise prediction of cardiovascular diseases. Our project is focused on assessing and developing a system that would enable the processing and analysis of patients' real-time data in relation to their current and previous symptoms for a variety of illnesses.

The finest algorithms for forecasting and analysis, with the highest accuracy rates, include KNN, Support Vector, Random Forest, and Decision Tree. KNN can be easily implemented in a web-based system because it requires less computational power and is easy to use.

## REFERENCES

[1] Vijeta Sharma, Shrinkhala Yadav, Manjari Gupta,"Heart Disease Prediction using Machine learning Techniques ", C.M.R Institute of Technology Bangalore, India ,2020.

[2] Md. Touhidul Islam1, Sanjida Reza Rafa2, Md. Golam Kibria3,"Early Prediction of Heart Disease Using PCA and Hybrid Genetic Algorithm with £-Means", department of Materials and Metallurgical Engineering, Bangladesh University of Engineering and Technology, Dhaka,2020.

[3]Aditi Gavhane, Isha Pandya, Gouthami Kokkula, Prof. Kailas Devadkar (PhD), " Prediction of Heart Disease Using a Machine Learning ", Sardar Patel Institute of Technology Mumbai, India,2018.

[4] Farzan Tasnim, Sultana Umme Habiba, “A Comparative Study on Heart Disease Prediction Using Data Mining Techniques and Feature Selection",International Islamic University Chittagong Kumira, Chittagong-4318, Bangladesh,2021.

[5] Purushottam, Prof. (Dr.) Kanak Saxena, Richa Sharma," Efficient Heart Disease PredictionSystem using Decision Tree ", Amity University Uttar Pradesh Noida, India,2015.

[6] Rashmi G Saboji, Prem Kumar Ramesh,"A Scalable Solution for Heart Disease Prediction using Classification Mining Technique",C.M.RInstitute of Technology Bangalore,India,2017.

[7]SanjayKumarSen”PredictingandDiagnosingofHeartDiseaseUsingMachineLearningAlgorithms”-InternationalJournalofEngineeringandComputerScienceJune2017.

[8] Th Ritika Chadha, Shubhankar Mayank, Anurag Vardhan and Tribikram Pradhan," Application of Data Mining Techniques on Heart Disease Prediction: A Survey”, Manipal University, Manipal 576104, India,2016

[9]SonamNikhar,A.M.Karandikar,"PredictionofHeartDiseaseUsingMachineLearningAlgorithms"-Vol-2Issue-6,June2016.

[10] Jaymin Patel, Prof. Tejal Upadhyay, Dr. Samir Patel, “Heart Disease Prediction using Machine learning and Data Mining Technique" Sept 2015-March2016.

[11]SanaBharti,ShailendraNarayanSingh"Analyticalstudyofheartdiseasecomparingwithdifferentalgorithms":Computing,Communication& Automation (ICCCA), 2015 International Conference.

[12]S.Chaitrali,DangareandS.ApteSulabha,"ImprovedStudyofHeartDiseasePredictionSystemusingDataMiningClassificationTechniques", IJCA,June2012.

[13]VRamalingam,AyantanDandapath,MKarthikRaja,“Heartdiseasepredictionusingmachinelearningtechniques:Asurvey”,March2018 International Journal of Engineering &Technology (IJET).

[14]HongzuLi,PierreBoulanger,“A Survey of Heart Anomaly Detection Using Ambulatory Electrocar diagram (ECG)”,May 2020.