

Heart Disease Prediction using Supervised Machine Learning Algorithms

G.Anuhya(AP19110010462), K.Varshini Reddy(AP19110010403) ,
M.Tanushq(AP19110010416), Sripada Srikar(AP19110010474), Souvik
Mandal(AP19110010485)

forecast the early risk of cardiovascular disease using machine learning algorithms.

Abstract: Cardiovascular disease (CVD) is a challenging health problem in today's world. It is reported that four out of every five CVD deaths are caused by heart-attacks. CVD claims the lives of 17.9 million people each year, accounting for almost 32% of all deaths worldwide. We need to put in place a system that can detect heart-attack symptoms early and hence can be prevented. It is infeasible for a common person to examine expensive tests like the ECG at a regular interval. It is vital to recognise and treat such disorders as soon as possible by developing a system that is convenient and reliable in estimating the likelihood of a heart disease. To address this challenge, machine learning (ML) can be used to anticipate diseases in a variety of fields, including healthcare. The purpose of this research is to apply various ML algorithms (e.g., logistic regression, Random Forest, etc.) to find the most important predictors of heart disease and to forecast total risks as a binary classification task. The main objective is to determine whether the patient has a 10-year risk of heart disease or not. To implement and validate our model, the data collection, preparation, classification, and results are analysed using the Python language in Google Collab environment. We can predict the risk of a heart attack in a person based on information such as age, blood pressure, artery thickness, and so on. Our method can be used to tackle the possibility of heart disease.

Keywords: CVD, ECG: Electrocardiography, ML: Machine Learning.

I. INTRODUCTION

The heart is essential to the proper functioning of the human body. According to the World Health Organization, heart illnesses cause 12 million deaths worldwide each year. Cardiovascular disease account for half of all deaths.

Health is wealth. In today's society, it is more vital than ever to live a healthy lifestyle, but this is not achievable for everyone due to a variety of circumstances such as poverty, stress, unhealthy diets, and so on. According to the World Health Organization, 50 million deaths occur each year, with heart disease accounting for 12 million deaths. Majority of all deaths are caused by cardiovascular disease.

This occurs because people will not attend to regular check-ups because they cannot afford it, and because diagnostic labs and skilled doctors are very few in most countries. In this scenario, machine learning comes into play, as we can

The risk of death from heart disease can be reduced if it is detected early.

Machine learning algorithms are capable of dealing with enormous data sets, pre-processing the data and identifying patterns. Data sets are divided into testing and training data, with the training set being trained using various machine learning algorithms to predict the target variable. To obtain the appropriate accuracy, the model's accuracy is compared to the accuracy of other models.

II. Literature Survey

Many studies have been conducted to evaluate the classification accuracy of various machine learning algorithms using various datasets.

Singh and Choudary et al. [1] created a robust classifier using the Ada-Boost Algorithm. It creates a model using training data and then creates a second model to fix the inaccuracy in the first model. Ada-boost, a binary classification technique, was the first effective boosting algorithm. Their model was overfitting when they used the decision tree approach on 13 attributes from their data set. For the optimization of the output generated by the Decision tree, they used the Ada-Boost algorithm. They achieved an accuracy of 89% using the Ada-Boost algorithm.

Sharma, Yadav, and Gupta et al.[2] have looked into a variety of applications that demonstrated the importance of machine learning approaches in a variety of areas. They have used a benchmark dataset of UCI Heart disease prediction for this research work, which consists of 14 different parameters related to heart disease. They used a variety of machine learning techniques, such as random forest and decision trees, to achieve accuracy of 99 and 88 percent, respectively. They discovered that SVM, with a 98 percent accuracy, and random forest, with a 99 percent accuracy, produced the greatest results when compared to all other ML algorithms.

Archana and Kumar et al[3] used a variety of machine learning algorithms, which include linear regression, decision tree, support vector machine, and k-nearest-neighbour, to train the

UCI machine learning data set, which contains 303 samples and 14 input measures, with 73 percent of the data used for training and 37 percent for testing. When using a decision tree, they found that the number of nodes is unbalanced, resulting in overfitting and lower accuracy. After experimenting with various algorithms, they discovered that Knn has the best accuracy with 87%.

Bertsimas et al. [4] proposed a novel methodology to extract ECG-related features and predict the type of ECG recorded. Our models leverage a collection of almost 40 thousand ECGs labelled by expert cardiologists across different hospitals and countries and are able to detect 7 types of signals: Normal, AF, Tachycardia, Bradycardia, Arrhythmia, Other or Noisy. We exploit the XGBoost algorithm, a leading machine learning method, to train models achieving sample F1 Scores in the range 0.93 – 0.99.

References

1. Choudhary, G., & Singh, S. N. (2020, October). Prediction of heart disease using machine learning algorithms. In *2020 International Conference on Smart Technologies in Computing, Electrical and Electronics (ICSTCEE)* (pp. 197-202). IEEE.
2. Sharma, V., Yadav, S., & Gupta, M. (2020, December). Heart disease prediction using machine learning techniques. In *2020 2nd International Conference on Advances in Computing, Communication Control and Networking (ICACCCN)* (pp. 177-181). IEEE.
3. Singh, A., & Kumar, R. (2020, February). Heart disease prediction using machine learning algorithms. In *2020 international conference on electrical and electronics engineering (ICE3)* (pp. 452-457). IEEE.
4. Bertsimas, D., Mingardi, L., & Stellato, B. (2021). Machine learning for real-time heart disease prediction. *IEEE Journal of Biomedical and Health Informatics*, 25(9), 3627-3637.