

A Hybrid Random Forest Linear Model approach to predict the Heart Disease

U. Sivaji¹, N V Krishna Rao², Chinnareddy Srivani³, Tharuna Sree⁴, Mamatha Singh⁵

^{1,3,4,5}Department of Information Technology, Institute of Aeronautical Engineering, JNTU, Hyderabad, INDIA.

²Department of CSE, Institute of Aeronautical Engineering, JNTU, Hyderabad, INDIA.

ABSTRACT

The project focuses mainly on cardiovascular disease prediction in the real world. Heart disease prediction involves various risk factors. AI (ML) is showing that it contributes to making decisions and based on the huge amount of information provided by the medical services industry. A brief look at predicting heart infection with ML procedures is provided by various exams. We propose a new strategy in this article that aims to identify significant characteristics by applying AI techniques to improve the cardiovascular expectation accuracy. The preliminary model offers a range of highlights and a number of known methods of classification. By the coronary disease expectation model with a hybrid random forest linear model that is blend of two different algorithms. We produce an improved presentation level with an exactness level of 88.7 percent. Dataset is from the UCI repository. The proposed HRFLM algorithm will assist doctors in diagnosis heart patients.

Key Words: Hybrid Random Forest Linear Model, Support Vector Machine.

1. Introduction

Heart diseases are increasing day by day. In the present day, medical field doctors use many techniques to treat the patients like CT scan, ECG, X-Ray. However, it requires the doctor to look at the reports. So this is beneficial to both the doctors and patients. The doctors can check whether their prediction is correct or not. Patients can check if the doctor is not available and also can take a second opinion. Heart problem was a type of disease that influences cardiovascular capacity. The source of mortality is cardiovascular disappointment in the present age. WHO has assessed that Twelve million people experience the ill effects of heart issues. Numerous respiratory conditions, stroke, renal confusion additionally hit. Thump was a sort of cardiovascular condition that happens attributable to extending, blockage, in any case, decreasing of platelets courses that pass across the cerebrum or caused using raised platelets pressure. The fundamental issue facing the energy area today was foundation strength. Low treatment has obliterating and unaccepted outcomes. So this is beneficial to both the doctors and patients. There are two broad categories such as supervised learning algorithms and unsupervised learning algorithms for heart disease prediction. In an existing system for finding the accuracy of the heart disease prediction, they have used methods such as Support Vector Machine, Decision trees, etc. This project aims to detect the accuracy of Heart Disease by using Random Forest Linear Model Algorithm.

II. Literature survey

[1] Effective Cardiovascular Disease Prediction using Hybrid Machine Learning Techniques. “MadhaviVeeranki, JayanagBayana”.

This research paper focuses on cardiovascular disease prediction by processing raw data by providing a modern and novel discernment of heart disease. This paper shows the accuracy comparison of six different algorithms. They used hybrid machine learning algorithms for improved predictive strategies.

[2] Hybrid Intelligent System Framework for the Prediction of Heart Disease using Machine Learning Algorithms. “Amin UIHaq, Jian Ping Li, Muhammad HammadMemon, Shah Nazir, and Ruinan Sun” .

This paper proposes a hybrid machine learning-based algorithm for the diagnosing the cardiovascular infection. The dataset used here is Cleveland heart disease from UCI Repository. Classification algorithms used are K-nearest neighbor, SVM algorithm, Naïve Bayes, Decision Tree, ANN, RF, and Selection Algorithms like Relief, LASSO, and mRMR. Validation is has done by using the K-fold method. The logistic regression gave an accuracy of 89%.

[3] Heart Disease Prediction System using Hybrid Technique of Data Mining Algorithms. “Navdip Singh, Sonika Jindal”.

This article, 14 attributes used here are to get more precise outcomes. From outcomes, it's has been seen that the projected model gives the right results in contrast with existing models. For effective cardiovascular prediction. They have selected algorithms like Naïve Bayes and Genetic Algorithm for effective prediction of Heart Disease. Different algorithms are used here for the prediction and analysis.

[4] Prediction of Cardiovascular disease using Hybrid Machine Learning Algorithm . “Ramkumar P, Thanusha K, Soumya U, Sahana K, Sushma M” .

Decision tree, Neural Networks, Naïve Bayes, neural network, K-NN are the presented methods used to predict Heart Disease. In this proposed system Hybrid Random Forest Linear Model produces accurate results in the prediction of cardiovascular infection. It also considers all the features without any feature restriction. In comparison with other algorithms, HRFLM gave the best accuracy.

[5] Hybrid Machine Learning Techniques for Heart Disease. “S Sharanya, S Lavanya, R. Chandhini, R. Bharathi, K. Madhulekha”.

Identifying the treatment of raw health will help save human lives and prevent abnormalities in the heart conditions at an early stage. The proposed method is the hybrid method of detecting cardiovascular infection. The proposed hybrid approach used here is fuzzy logic and the k-nearest neighbor algorithm that provides 94% accuracy. This method proved the accuracy of the highest prediction rate.

[6] Effective Heart Disease Prediction using Hybrid Machine learning Techniques. “Senthilkumar Mohan, ChandrasegarThirumalai, and Gautham Srivastava”.

In this study, classification models are on Hybrid Learning Algorithms. This Hybrid Algorithm used for prediction gives the best accuracy. This study is one of the first to compare SVM, decision tree, and other cardiac disease detection methods performance with actual data set.

III. Problem definition

Millions of people die due to heart disease each day for several reasons like lifestyle, Food. In the world of today the heart is one of the causes of death. It affects the physical life and mental life of the individual. One person dies every 36 seconds from cardiovascular disease in the United States.

Therefore, there's an urge to unravel the matter of heart disease detection. The latest technologies provide additional ways for accurate prediction of results. High BP, Excess Cholesterol, diabetes, depression, obesity are among the risk factors of heart disease. Other causes of heart disease include smoking, poor diet, alcohol consumption. This proposed methodology focuses on finding the better accuracy of occurrence of heart disease. With this methodology, we will be able to detect heart infection at an early stage. The purpose of this paper is to provide a hybrid algorithm that gives the highest accuracy of 88.7% .

IV. Proposed methodology

We use support vector machine, Extension extreme machine learning algorithm, Hybrid Random Forest Linear Model, Naïve Bayes, and deep Learning ANN algorithms in the proposed method to detect the heart disease from the raw dataset collected. The voting classifier is used internally for Random Forest and Linear Model. The algorithm which gives better results considered for prediction. Pre-Processing of data is using 297 patient records. The feature selection is using the attributes of the patient. Classification is by using the proposed algorithms. Here the results of all algorithms are compared. Splitting of data is using decision trees. This paper compares the accuracy, precision, and error rate of all the algorithms. HRFLM algorithm gives the highest accuracy of 88.7%.

V. System architecture

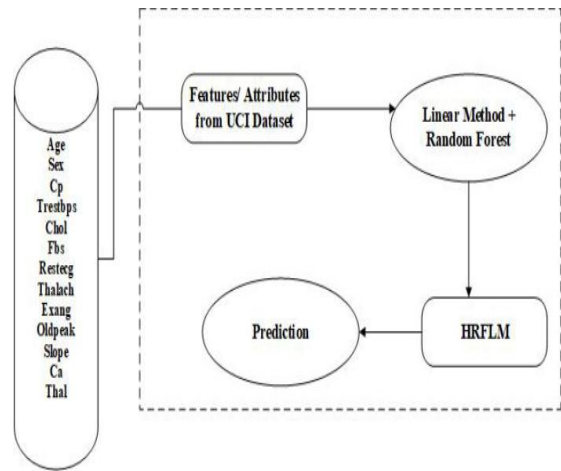


Fig:1 Proposed Methodology using HRFLM algorithm

VI. Results



Fig:2 Home Screen

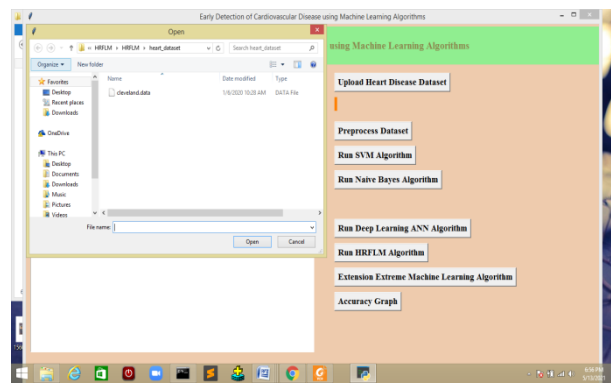


Fig:3 Uploading Cleveland dataset

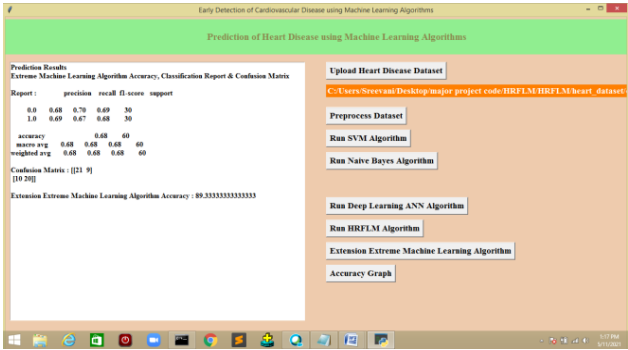


Fig:4 Extension Extreme Machine Learning Algorithm Screen



Fig:5 Proposed HRFLM Screen

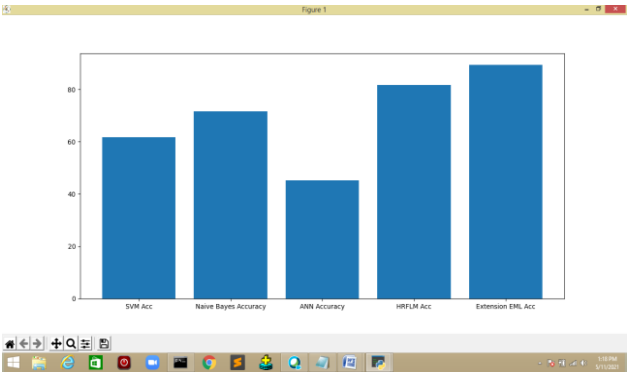


Fig:4 Graph showing accuracy of algorithms

Accuracy of Hybrid Random Forest Linear Model algorithm after testing: 88.7%

VII. Conclusion

The results show that a Hybrid Random Forest Linear Model is appropriate for the prediction of cardiovascular disorders.

VIII. References

- [1] MadhaviVeeranki, JayanagBayana: Effective Cardiovascular Disease Prediction using Hybrid Machine Learning Techniques vol. 9, no.-4, pp 2249-8958, 2020.
- [2] Amin UlHaq, Jian Ping Li, Muhammad HammadMemon, Shah Nazir, and Ruinan Sun: Hybrid Intelligent System Framework for the Prediction of Heart Disease using Machine Learning Algorithms vol. 2018, article id 3860146, <https://doi.org/10.1155/2018/3860146>.
- [3] Navdip Singh, Sonika Jindal: Heart Disease Prediction System using Hybrid Technique of Data Mining Algorithms, vol.4, pp 2414-532X, 2018.
- [4] Ramkumar P, Thanusha K, Soumya U, Sahana K, Sushma M: Prediction of Cardiovascular disease using Hybrid Machine Learning Algorithm, vol. 7, no. 14 pp. 2394-5125,2020.
- [5] S Sharanya, S Lavanya, R Chandhini: Hybrid Machine Learning Techniques for Heart Disease, vol 7, no 3, pp 2349-6495(P)|2456-1908(O).
- [6] Data Mining Algorithms, vol.4, pp 2414-532X, 2018.