

A Comparative study of Machine Learning Algorithms in prediction of Heart Disease

Ms. C. Keerthana,
Research Scholar,
Department of Computer Science,
Nallamuthu Gounder Mahalingam College,
Coimbatore, India.
kirthy6@gmail.com

Dr.B.Azhagusundari ,
Associate Professor,
Department of Computer Science,
Nallamuthu Gounder Mahalingam College,
Coimbatore, India.
azhagumithra78@gmail.com

Abstract—Heart disease is the one among the foremost common disease. This disease is sort of common now days, we used different attributes which may relate to the present heart diseases well to locate the higher method to predict and that we additionally used algorithms for prediction. This study explores a different models based on such algorithms and techniques and analyse their performance. Models based on supervised learning and unsupervised learning algorithms are Support Vector Machines (SVM), K-Nearest Neighbour (KNN), NaïveBayes, Decision Trees (DT), Neural Networks, Apriori algorithm, Random Forest (RF) and ensemble models are discovered very mainstream among the researchers

Keywords— Machine Learning algorithm, Naïve Bayes, Decision Tree, Random Forest, Heart disease Prediction)

I. INTRODUCTION

Heart disease is one among the many causes of death during the planet. It cannot be easily predicted by the medical practitioners because it may be a troublesome task which demands expertise and better knowledge for prediction[7]. Cardiovascular Diseases (CVDs) are the principle reason for an enormous number of deaths on the earth over the foremost recent few decades and has emerged because the most life-threatening disease, in India also as within the whole world. During this way, there's a requirement of reliable, accurate and feasible system to diagnose such diseases in time for correct treatment[3]. Machine Learning algorithms and techniques are applied to different medical datasets to automate the examination of huge and sophisticated data. an automatic system in diagnosis would enhance medical efficiency and furthermore reduce costs. The prediction of heart condition requires an enormous size of knowledge which is excessively complex and large to process and analyze by conventional techniques[6].

Machine Learning

Machine learning is an emerging region of AI. It's essential center is to style systems, permit them to find out and make predictions support the experience. It trains machine learning algorithms utilizing a training dataset to

make a model[2]. The model uses the new information data to predict heart condition. Utilizing machine learning, it detects hidden patterns within the info dataset to manufacture models. It makes accurate predictions for brand spanking new datasets. The dataset is cleaned and missing values are filled. The model uses the new info data to predict heart condition and afterward tested for accuracy. Machine learning techniques are categorized as:

i) Supervised Learning

The model is trained on a dataset that's labeled. It's input file and its outcomes. Data are classified and split into training and test dataset. Training dataset trains our model while testing dataset capacities as new data to urge accuracy of the model. The dataset exists with models and its yield. The classification and regression are its example.

ii) Unsupervised Learning

Data prepared aren't classified or labeled within the dataset. Point is to get hidden patterns within the data. The model is trained to develop patterns. It can easily predict hidden patterns for any new info dataset, yet after exploring data; it makes determination from datasets to explain hidden patterns[8]. During this technique, no responses within the dataset are seen. The clustering method is an example of an unsupervised learning technique.

In this survey, extract hidden patterns by applying data processing techniques, which are noteworthy to heart conditions and to predict the presence of heart condition in patients where the presence is valued on a scale and is to get the acceptable machine learning technique that's computationally efficient even as accurate for the prediction of heart disease. data processing combines Statistical investigation machine learning and database technology to extract hidden patterns and relationships from large databases. The implementation of labor is completed on Cleveland heart diseases data set from the University of California Irvine (UCI) machine learning repository to check on different data processing techniques. This paper explores different models supported such algorithms and techniques and analyze their performance. Models supported supervised learning algorithms, for