



HEART BEAT LEVEL PREDICTION USING MACHINE LEARNING ALGORITHMS

Nukala Naveen Teja
Department of Computer Science & Engineering
Bangalore Karnataka, India
naveentej.nukala@gmail.com

Gandhaveeti Yeswanth Reddy
Department of Computer Science & Engineering
Bangalore Karnataka, India
gandhaveetiyeswanthreddy@gmail.com

Gangireddy Gari Yoga Nanda Reddy
Department of Computer Science & Engineering
Bangalore Karnataka, India
yogareddy426@gmail.com

A. Ajil
Department of Computer Science & Engineering
Bangalore Karnataka, India
ajil.a@reva.edu.in

Abstract—Heart disease is one of the leading causes of death in the world today. Cardiovascular disease prediction is a critical challenge in clinical data analysis. Machine learning (ML) has been shown to be effective in assisting with decision-making and prediction from the vast amounts of data generated by the healthcare industry. Many lives can be saved if heart disease is detected early. Machine learning classification techniques have the potential to significantly benefit the medical field by providing accurate, unambiguous, and rapid disease diagnosis. As a result, both doctors and patients should set aside time for prediction. This paper proposes a model to compare the accuracies of applying rules to the individual results of support vector machine, decision trees, logistic regression, and Random forest on the Cleveland Heart Disease Database in order to present an accurate model of predicting heart disease.

Keywords— Machine learning, accuracy, prediction, data set, Heart disease.

I. INTRODUCTION

Machine learning is now being used in a variety of business applications, including e-commerce. One application of machine learning is prediction; our topic is predicting heart disease by processing patient datasets and data from patients to whom we need to predict the likelihood of the occurrence of a heart disease. Many studies have been conducted in an attempt to identify the most influential risk factors for heart disease and to accurately predict the overall risk. Heart disease is even referred to as a silent killer because it causes death without obvious symptoms.

Early detection of heart disease is critical for making lifestyle changes in high-risk patients and, as a result, reducing complications. This project aims to predict future heart disease by analyzing patient data and classifying whether or not they have heart disease using machine-learning algorithms.

Heart disease is a condition that impairs the function of the heart. Many aspects of health promotion and clinical medicine are dependent on an individual's risk for coronary heart disease. In this paper we have taken four

machine learning algorithms. Based on four machine learning algorithms we can be able to check that which algorithm accuracy will give best result.

II. LITERATURE SURVEY

In the Survey of Data Mining Techniques on Medical Data for Finding Frequent Diseases on a Local Level, Mohammed Abdul Khaleel presented a project.

Heart disease is a term that refers to a wide range of heart-related medical conditions. These medical conditions describe abnormal health conditions[1].

Disease Predicting System Using Data Mining Techniques, presented by M.A. Nishara Banu and B. Gomathy. MAFIA (Maximal Frequent Item Set Algorithm) and K-Means clustering are discussed in this paper. Because classification is essential for a disease's prognosis. The classification based on MAFIA and K-Means yields accurate results[2].

Costas Sideris, Nabil Alshurafa, Haik Kalantarian, and Mohammad Pourhomayoun published Remote Health Monitoring Outcome Success Prediction Using First Month and Baseline Intervention Data. RHS systems are effective in terms of cost savings and illness reduction.

In this paper, they present Wanda- CVD, an upgraded RHM framework that is cell phone based and intended to provide members with remote instruction and social assistance. CVD prevention measures are regarded as a primary priority by social insurance organisations all over the world[3].

A paper titled Prediction for Disease Similarities Using the ID3 Algorithm in Television and Mobile Phones was presented by L.Sathish Kumar and A. Padmapriya. This paper offers a pre-programmed and hidden method of dealing with hidden coronary disease designs. The provided framework makes use of data mining methods such as the ID3 algorithm. This proposed method not only assists people in learning about diseases, but it may also aid in lowering the death rate and number of disease-affected people[4].

To predict cardiovascular disease, Sharma Purushottam et al. proposed c45 rules and a partial tree technique. This paper can discover a set of rules for predicting patients' risk levels based on given health parameters. Classification accuracy can be used to assess performance.

III.SYSTEM REQUIREMENTS

Software Requirements	Hard Requirements
Operating System: Windows 7 above.	System: Pentium IV 2.4 GHz
Coding Language: Python 3.6 and above	Hard Disk: 50 GB.
IDE : Jupiter Notepad	Monitor: 15 VGA Color RAM: 512 Mb.

IV.EXISTING SYSTEMS

Heart disease is even referred to as the "silent killer" because it causes death with no visible symptoms. works with datasets from Deep Learning and Data Mining [7]. Medical diagnosis is a critical but challenging task that must be completed quickly and accurately. To help reduce the cost of performing clinical tests, appropriate computer-based information and decision support should be used. Data mining is the use of software techniques to discover patterns and consistency in large amounts of data. Furthermore, with the advent of data mining in the last two decades, there is a significant opportunity to allow computers to directly construct and classify the various attributes or classes. Learning about the risk factors associated with heart disease enables medical professionals to identify patients who are at high risk of developing heart disease. Age, blood pressure, total cholesterol, diabetes, hypertension, family history of heart disease, obesity and lack of physical activity,

fasting blood sugar, and other risk factors for heart disease have been identified through statistical analysis. All existing systems, however, have a very low accuracy.

V.IMPLEMENTATION METHOD

Decision tree:

The Decision Tree algorithm is one of the most basic yet powerful Supervised Machine Learning algorithms. Decision tree analysis is a predictive modelling tool that can be used in a variety of contexts.

Logistic Regression:

Logistic regression is a popular Machine Learning algorithm that belongs to the Supervised Learning category. Logistic regression is used to predict the outcome of a categorical variable. As a result, the outcome must be categorical. Except for how they are used, Logistic Regression and Linear Regression are very similar. Linear Regression is employed in logistic regression, instead of fitting a regression line, we fit a "S" shaped logistic function that predicts.

Random Forest:

Random forest is a classification and regression supervised learning algorithm. However, it is primarily used to solve classification problems. Ensemble learning underpins the Random Forest algorithm. Ensemble learning is a type of learning in which multiple instances of the same algorithm or different types of algorithms are combined to create a more powerful prediction model. The random forest algorithm combines multiple algorithms of the same type.

Support Vector Machine:

Support vector machines (SVMs) are supervised machine learning algorithms that are both powerful and adaptable. They're used for classification and regression. However, they are most commonly used in classification problems. SVMs were invented in the 1960s, but they weren't perfected until 1990. When compared to other machine learning algorithms, SVMs have a distinct method of implementation. They have recently gained popularity due to their ability to handle a large number of continuous and categorical variables.

The SVM model is essentially a representation of various classes in a multidimensional hyperplane.

- Import all the necessary packages like pandas, numpy, matplotlib, sklearn.
- Perform data cleaning like removing null values Nan.
- Create a dummy for categorical variables.
- Perform training in the model machine using algorithms like decision tree, Logistic

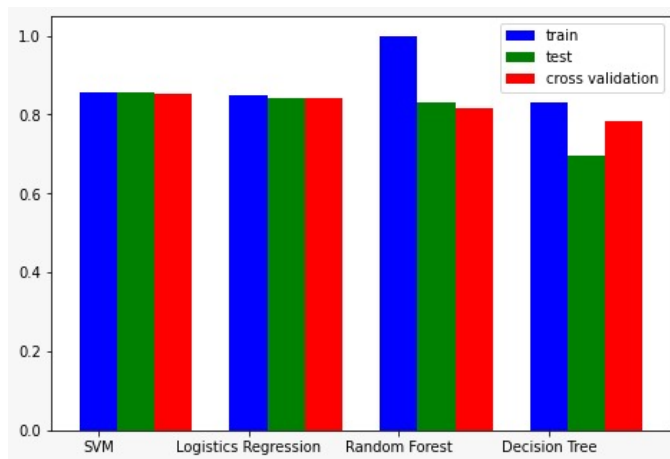
Regression, Random forest, support vector machine, decision tree.

- Finally, support vector machine gave the best result in predicting heart beat level

Here, in this research paper we are calculating the heart beat accuracy. we have taken four machine learning algorithms. In that four machine learning algorithms support vector machine will give the correct accuracy.

VI.RESULT

In this research paper we have taken four algorithms. Compared to four algorithms support vector machine will gives the correct accuracy.



Graph 1. Results Observed

VII. CONCLUSION

The early prognosis of heart disease can help as making lifestyle changes decisions, high-risk patients and as a result it reduce the complications, which can be a significant milestone in the fieldin medicineThe framework proposed in this paper is based on four algorithms: support vector machines, logistic regression, decision trees, and random forest.We conclude the support vector machine algorithm gave the best accuracy compare to other threealgorithms.

VIII. REFERENCES

- [1]. "Prediction of Heart Disease Using Machine Learning Algorithms," SonamNikhar and A.M. Karandikar, International Journal of Advanced Engineering, Management, and Science (IJAEMS), vol-2, June 2016.
- [2]. .Deeanna Kelley, "Heart Disease: Causes, Prevention, and Current Research," JCCC Honors Journal.
- [3]. "Intelligent Heart Disease Prediction System Using Data Mining Techniques," SellappanPalaniappan and RafiahAwang, IEEE Conference, 2008, pp 108-115.
- [4]. "Intelligent Syncope Disease Prediction Framework Using DM-Ensemble Techniques", FTC 2016 - Future Technologies Conference 2016. AmmarAsjad Raja, Irfan-ul-Haq, MadihaGuftar, Tamim Ahmed Khan, and DominikGreibl et al.
- [5]. Ashwinkumar.U.M and Dr. Anandakumar K.R, "Predicting Early Detection of cardiac and Diabetes symptoms using Data mining techniques", International conference on computer Design and Engineering, vol.49, 2012.