

PROJECT

Topic: Human Heart Disease Prediction Using Random Forest Classification

Submitted by : Neethu Subramanian

Guide: Amal Ganesh

About project:

Heart disease also called as coronary artery disease is a condition that affects the heart. Heart disease is a leading cause of death worldwide. Physicians generally make decisions by evaluating current test results of the patients. Previous decisions taken by other patients with the same conditions are also examined. So diagnosing heart disease requires experience and highly skilled physicians. Heart disease will become a leading cause of death by 2020. Heart disease diagnosis is an important yet complicated task. Today many hospitals collect patient data to manage health care of patients. This information is in different format like numbers, charts, text and images. But this database contains rich information but poorly used for clinical decision making.

The automation or decision support system would be extremely advantageous. Data mining can be used to automatically infer diagnostic rules and help specialists to make diagnosis process more reliable. The purpose of predictions in data mining is to discover trends in patient data in order to improve their health care .Knowledge discovery in data bases is applied to extract useful patterns from the medical data sets using various data mining techniques. Data mining have shown a good result in prediction of heart disease and is widely applied for prediction of heart disease. Due to shortage of doctors and experts in medical field to predict heart disease, and because of neglecting the patient's symptoms, data mining is emerged as an analysis tool.

Several data mining techniques are used by researchers to help the health care professionals to predict the heart disease. Random forest is an ensemble and most accurate learning algorithm, suitable for the medical application. Random forest is an ensemble classifier which combines bagging and random selection of features. Random forest can handle data without preprocessing. Random forest algorithm has been used in prediction and probability estimation. Random forest consists of many decision trees and outputs the class, which is the mode of individual trees class .It is one of the most accurate classifier. It produces a highly accurate classification for many data sets especially for heart disease data sets.

PUBLICATIONS

- **Paper** – Prognosis of Heart Disease Using Data Mining Techniques
Conference - ICICT 2018
Journal – IEEE Xplorer
Status - Presented

- **Paper** – Methods of Large Scale Text Classification In Natural Language Processing
Conference – ICICT 2018
Journal - IEEE Xplorer
Status -Presented