

Heart Disease Prediction

Abstract:

Heart disease is easier to treat when it is detected in the early stages. Machine learning techniques may aid a more efficient analysis in the prediction of the disease. Moreover, this prediction is one of the most central problems in medicine, as it is one of the leading diseases related to an unhealthy lifestyle. So, an early prediction of this disease will be useful for a cure or aversion.

Problem Statement:

Analyze the heart disease dataset to explore the machine learning algorithms and build decision tree model to predict the disease.

Dataset Information:

Each attribute in the heart disease dataset is a medical risk factor.

Variable Description:

Column	Description
age	Age of the patient
gender	Gender of the patient - (0,1) - (Male, Female)
chest_pain	It refers to the chest pain experienced by the patient -(0,1,2,3)
rest_bps	Blood pressure of the patient while resting(in mm/Hg)
cholesterol	Patient's cholesterol level (in mg/dl)
fasting_blood_sugar	The blood sugar of the patient while fasting
rest_ecg	Potassium level (0,1,2)
thalach	The patient's maximum heart rate
exer_angina	It refers to exercise-induced angina - (1=Yes, 0=No)

PGAA

Problem Statement – Decision Tree

old_peak	It is the ST depression induced by exercise relative to rest(ST relates to the position on ECG plots)
slope	It refers to the slope of the peak of the exercise ST-Segment-(0,1,2)
ca	Number of major vessels - (0,1,2,3,4)
thalassemia	It refers to thalassemia which is a blood disorder - (0,1,2,3)
target	The patient has heart disease or not - (1=Yes, 0=No)

Scope:

- Understand data by performing exploratory data analysis
- Training and building classification algorithms to predict heart disease
- Understand various model performance metrics and measure the performance of each model

Learning Outcome:

The students should be able to predict heart disease from medical records with the help of classification models. They should also be able to perform EDA and re-build the model and check if there is any significant change in the predictive scores.