Python
Heart Disease
Prediction
Case Study
WEEK 5

Data Science
Academy

Description: You are given a dataset of patients to be tested for heart disease. Features are given below.

- age
- sex – Male, Female
- cp - chest pain type (4 values)
- trestbps - resting blood pressure
- chol - serum cholestoral in mg/dl
- fbs - fasting blood sugar > 120 mg/dl
- restecg - resting electrocardiographic results, values: 0,1,2
- thalach - maximum heart rate achieved
- exang - exercise induced angina
- oldpeak - ST depression induced by exercise relative to rest
- slope - the slope of the peak exercise ST segment
- ca - number of major vessels colored by flourosopy, values: 0,1,2,3,4
- thal - values: 0,1,2,3
- target – is the column we want to predict.

You asked to create Support Vector Machine (with both Linear and RBF kernel) classifier, Decision Tree classifier, Random Forest classifier, Logistic Regression, Naïve Bayes classifier, KNN classifier to predict whether patients have heart disease or not and decide which algorithm performs the best on average.

Data Science
Academy