Supervised Learning Assignment

Problem Statement:

Consider yourself to be Sam who is a data scientist. He has been approached by a medical center to get more insights into the reasons for a person having heart disease

Tasks to be performed:

1. Sam has to build a tree-based model on top of the 'heart' dataset:
    a. To find out if the person has 'heart disease' on the basis of his 'age'.
    b. To find out if the person has 'heart' disease on the basis of his 'resting blood pressure'
    c. To find out if the person has 'heart' disease on the basis of 'maximum heart rate achieved'
    d. To understand which factor influences the patient having heart disease the most -> 'age', 'resting blood pressure' or 'maximum heart rate achieved'

2. Sam has to build a probabilistic classification model on top of the 'heart' dataset:
    a. To find out if the person has 'heart disease' on the basis of the 'type of chest pain' the patient has
    b. To find out if the person has 'heart disease' on the basis of the patient's 'serum cholesterol'
    c. To find out if the person has 'heart disease' on the basis of the 'number of major vessels colored by fluoroscopy'
    d. To understand which factor influences the patient having heart disease the most -> 'type of chest pain', 'serum cholesterol' or the 'number of major vessels colored by fluoroscopy'
