Assignment 2 (Machine Learning)

Please submit your Assignment 1 and Assignment 2 (compile into one file) to the folder \Assignments\Assignment(Machine Learning Part) by 29 Oct 2023 (folder will be open for submission from 15 Oct).

Question:

The training data is given in TrainingData.xlsx, where columns A-D are the features, and column E gives the class label. The test data is given in TestData.xlsx (30 samples, no label is provided).

(1) Are there any missing values and outliers in the training data? If there are, describe how you address the issues and explain the rationale.

(2) Train a binary classification tree using the training data.

(3) Predict the class label of the test data.