Assignment 3

1). Binary classification tree:

a. Train a fully grown binary classification tree based on Gini impurity using the dataset A4_test.csvand visualize it.

b. Compute the Sum of Squared Errors (SSE) on the test dataset (A4_test.csv) at each depth and plot the variation of SSE with depth.

c. Determine the optimal pruning depth by selecting the depth where the SSE change is minimal.

d. Visualize the pruned tree.

2). Write a code to implement regression models using dataset A2_P2.csv .Divide the dataset into training and testing sets (80:20). Implement the following models using the training dataset and compute MSE on the test dataset:

a. Linear regression.

You can use inbuilt functions from the sklearn package. Use bar plots to compare MSE and feature coefficients (weights) for the three methods