CS5100J – Data Analysis : Assignment 2

Task 1 :

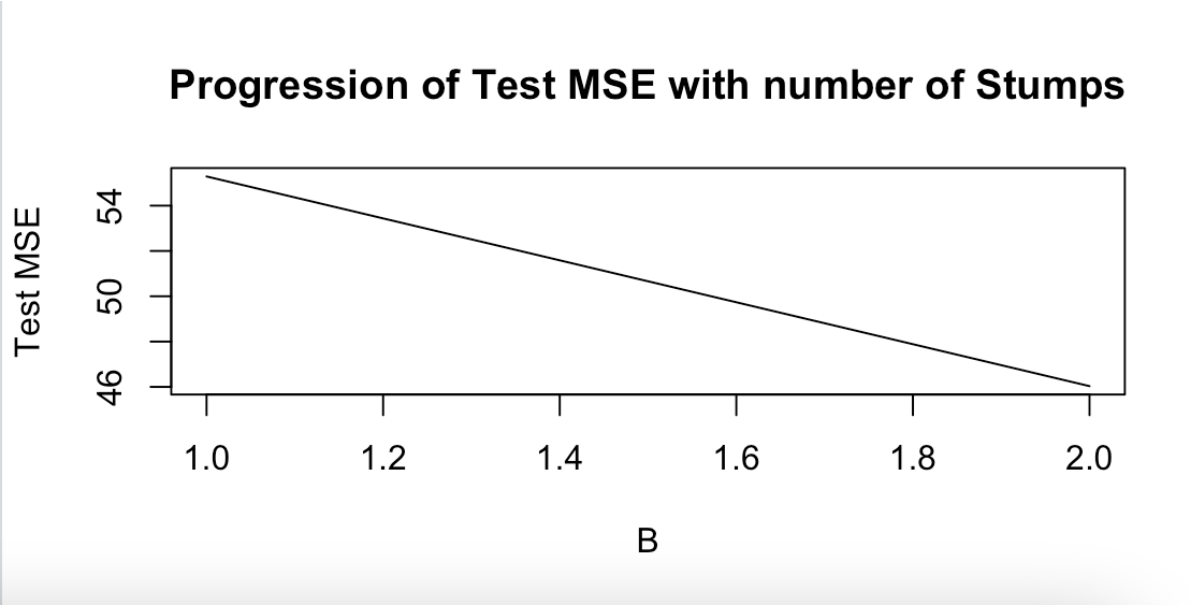
The Test MSE for decision stumps on the test set: 52.635319

Task 2:

"The Test MSE for B=1000 for boosted decision stumps (BDS) ,n = 0.1 on the test set: 539.508857"

Task 3:

Plot the test MSE for a fixed value of η as a function of B ∈ [1,B0] (the number of trees) for as large B0 as possible :



Did you observe overfitting : Yes the graph overfits for a large value of B. As the number of trees increases, the Test MSE for the BDS decreases, and becomes constant after a certain number of trees.

Observation : For Learning rate = 1, the MSE for BDS is 56.866842. As we change learning rate to 0.01, the MSE increases. As the B increases, MSE decreases. Then, it becomes constant after a certain number of trees.