ENGR 421 – HW5 – Report

Tolga Sümer

64534

Implementation:

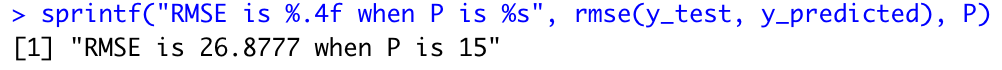
First of all, I assigned every images in hw04\_data\_set.csv to training sets and test sets. Because when I downloaded hw05\_data\_set.csv in Blackboard, hw04\_data\_set.csv was downloaded. However, it seems in the pdf that the data was same in the last homework so I thought that it wouldn’t be a problem. The training set has the first 100 data and labels of each class and the test set has the remaining 33 data.

Then, I implemented a decision tree regression algorithm using the following pre-pruning rule: If a node has P or fewer data points, convert this node into a terminal node and do not split further, where P is a user-defined parameter. Here I used the code in Lab 7.

Chart, scatter chart

Description automatically generatedAfter that, I set P to 15 and used the decision tree to draw the data. The result can be seen below. I also used the prediction function in Lab 7 to predict the left borders and right borders. After that I created lines between the actual left and right borders and the predicted left and right borders.

Moreover, I calculated RMSE of the regressogram for the test data points. The output can be seen below. I calculated the predicted values using the code in Lab 7.



Chart, line chart

Description automatically generatedFurthermore, I set P to 5, 10, 15 all the way to 50 and calculated the RMSE of each regressogram for test data. I calculated the decision tree using the algorithm in the Lab 7 and calculated the predicted values using the code in the lab. The output can be seen below.