COMP 421 – HOMEWORK 05

REPORT

Özgür Taylan Özcan

First, I read X and Y data from the given file and divided them into two by assigning the first 100 data points to the training set and 33 data points to the test set.

Then I implemented the decision tree regression function. In the lab session, we created a decision tree for classification. So, I followed a similar way for implementing the decision tree with some modifications to achieve regression. This function uses pre-pruning with a given pre-pruning parameter and it returns a list containing the following learned tree structures: node_splits, node_means, is_terminal.

After implementing the decision tree regression function, I called it with pruning parameter 10 to learn the tree structures using the training data.

I also implemented a prediction function that takes the tree structures and predict the labels for a given set of data points. Then I plotted a graph containing training and test data points as well as the fit on this graph. I got the same figure with the homework description.

I calculated the RMSE value for P = 10 and printed it. Then I used a loop to calculate RMSE values for P = 1 to 20, and then plotted a P vs RMSE graph. I got the same results with the homework description.