Mehmet Mert Bezirgan

Engr 421 – homework 4

**Problem**: Decision Tree Regression

**Solution:**

For this problem we have a dataset that has only 1 feature and y values are also numbers. We are going to learn split positions for all nodes and generate tree. For each node that we can split, we calculate error for each split line and pick the best splits for each node like in lab session. The only part we change is going to be the part that calculates error because our problem changed to regression problem. We implemented our loop with pre-prunning rule and we are going to stop iteration if a node has less than P data points.

First, we learn a tree with a parameter P 15, We draw data points and fit in a figure and we calculate rmse for this regressogram. We get the output 26.8777 for rmse.

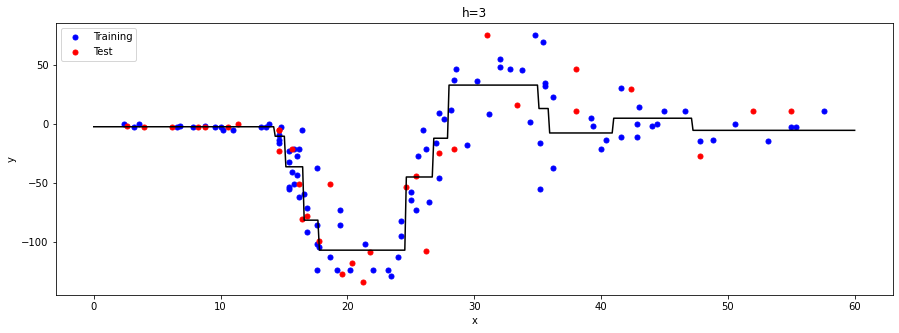


Figure 1 regressogram graph for p=15

After this step we are going to iterate over different values to test accuracy using rmse. The rmse for each p values between 5 to 50 with step size 5 and get the graph.

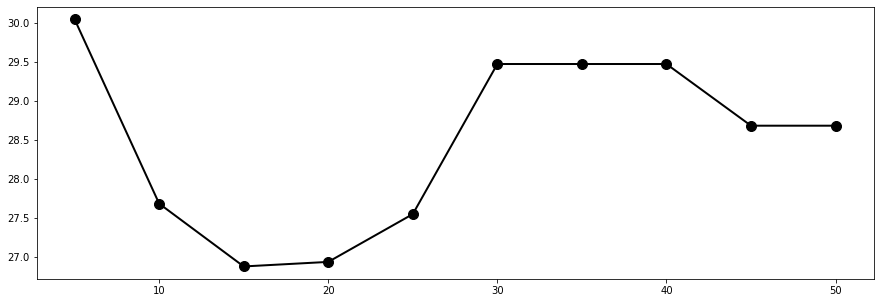


Figure 2 rmse values for each p value