ENGR 421 – HW4 – Report

Tolga Sümer

64534

Implementation:

First of all, I assigned every images in hw04\_data\_set.csv to training sets and test sets. The training set has the first 100 data and labels of each class and the test set has the remaining 33 data.

After that I set the bin width parameter to 3 and origin parameter to 0.

Later I used the data from lab 6 to assign the left borders and right borders. After that I assigned the p-hat function similar to the lab’s function. However, I used the x\_train data which is between the left and right borders inside y\_train to retrieve the y values for the p-hat function. Also, I divided the sum of y\_train of the x\_train data which is between the borders to the sum of the x\_train data which is between the borders. Later, I plotted the train and test points and used the function from lab to draw the lines. Below is the figure for regressogram.

Chart, scatter chart

Description automatically generated

Later on, I calculated the root mean squared error (RMSE) for test data points and the result is below. For using the test data for p-hat function, I divided the x\_test points to bin width to get the appropriate equivalent.

Text

Description automatically generated

After that, I used the running mean smoother from the lab. However, I changed the reciprocal like I did in the regressogram part. I divided the sum of y\_train of the x\_train data which is between the borders to the sum of the x\_train data which is between the borders. Later, I plotted the train and test points. Below is the figure for running mean smoother.

Chart, scatter chart

Description automatically generated

Later on, I calculated the RMSE for test data points and the result is below. For using the test data for p-hat function, I divided the x\_test points to data interval width to get the appropriate equivalent.

Text

Description automatically generated

Moreover, I used the kernel smoother from the lab. However, I changed the reciprocal like I did in the regressogram part. I divided the sum of y\_train multiplied with x\_train data which is between the borders to the sum of the x\_train data which is between the borders. Later, I plotted the train and test points. Below is the figure for kernel smoother.

Chart, scatter chart

Description automatically generated

Last but not least, I calculated the RMSE for test data points and the result is below. For using the test data for p-hat function, I divided the x\_test points to data interval width to get the appropriate equivalent.

Text

Description automatically generated