Problem-1

The Boston data has 13 feature and 1 target feature. First, the program splits data into train and test part. %70 of the data become train part and the rest become test part. After that, program add polynomial features to data. So, in the end of this process, program has one normal-features data and one poly-features data. Then, it moves to the processing part. In this part it creates some alpha values (50) and create different models for each alpha. It calculates accuracy for all of these models. In the last part, it picks most accurate ones. One for linear-ridge model and another one for polynomial-ridge model. You can see the results on the graph.

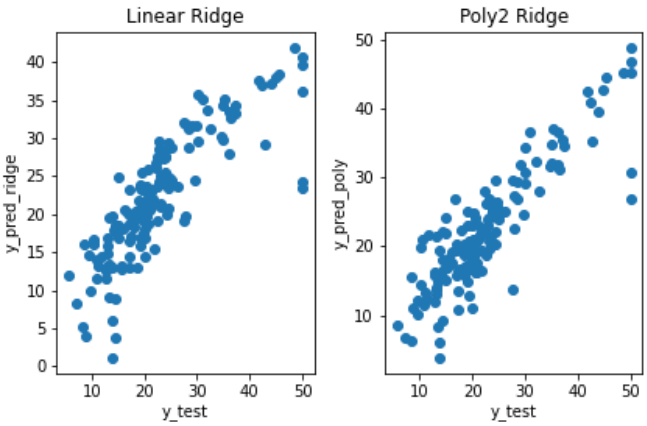

Figure 1: Linear Model vs Polynomial Model

Alpha for ridge model is = 1 * 10^ (-7)

Alpha for polynomial model is = 438.77

RMSE for ridge model = 5.322

RMSE for polynomial model = 4.574

RMSE results and graphs shows us that the polynomial model is better than the linear ridge model.

Problem-2

MNIST dataset contains had written digits' images which are 8x8 grayscale images. So, each sample has 64 features. First, the program takes data and split it train and test part. After that, it initializes the KNN model and try this model with different k values. It calculates the model score for each k value. In the end, it extracts the most efficient k value from the scores list and uses it to predict on test data. You can see the plot in figure 2. It is very efficient as you can see.

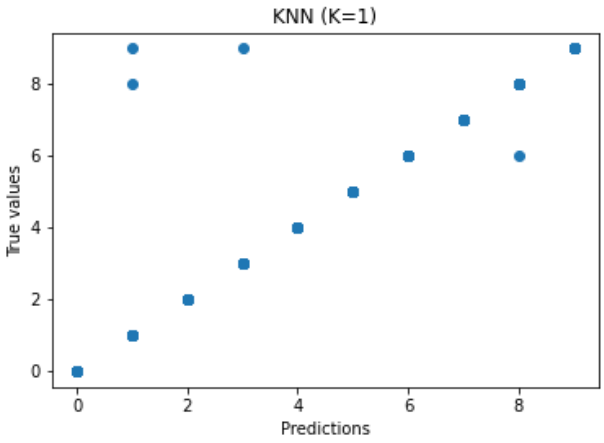

Figure 2: KNN with k=1