**Name – PRBHAKAR KUMAR**

**Roll – IRM2017008**

**Assignment 7 –KNN Regression + Classification**

In the question we first created the test dataset by slicing the given dataset from the indices 31 to 60, as was given in the objectives, and also separated our testing dataset. Thereafter, for each of the testing dataset, we calculate the Euclidean distance of the point from each of the training dataset, and only consider the K nearest point while considering the output.

In regression we take mean of the classes of K nearest points, while in classification the mean is replaced by the mode operation. Now one of the major hyper-parameter that plays a significant role is the K value, which we checked for two values:

**K = 3**

Classification Accuracy – 80%

Regression Squared Error – 1.666666666666667

**K = 5**

Classification Accuracy – 80%

Regression Squared Error – 1.3200000000000003

Here, we can see that the classification accuracy though didn’t increase but the value of squared error in case of regression decreases, as higher the value of K, more the neighbouring point are considered in accounting the output.