İrem Şahin

ENGR421

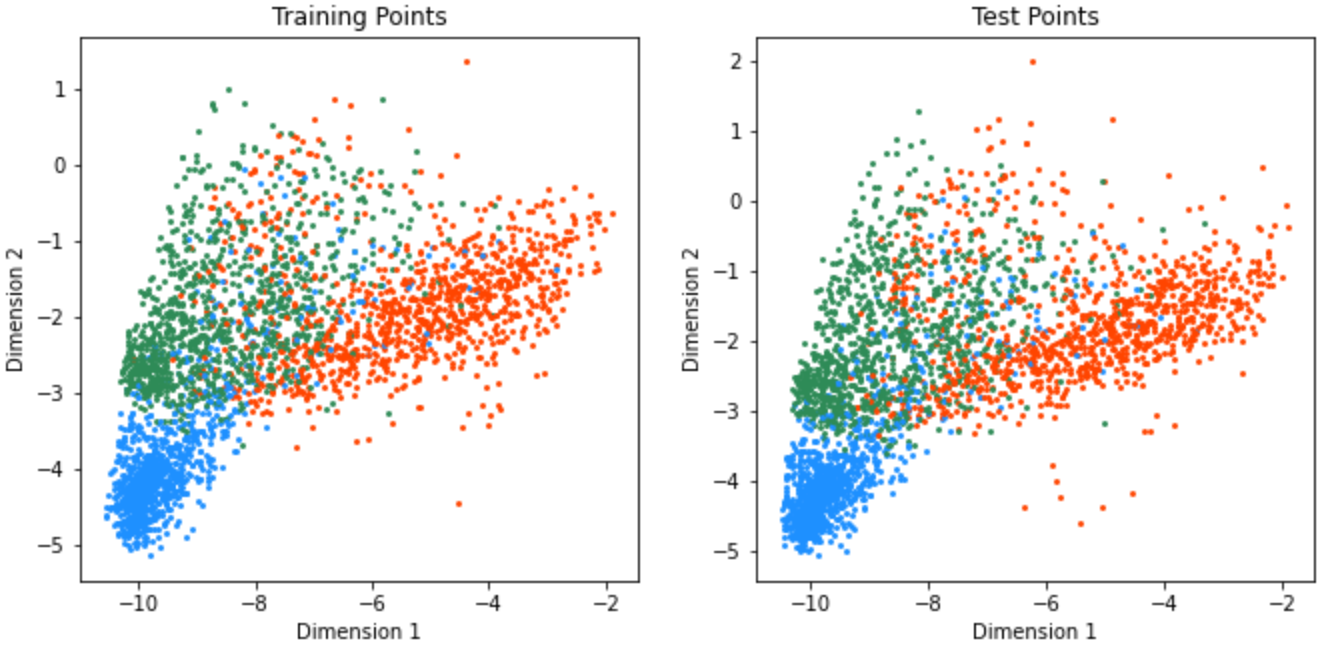
Mehmet Gönen

HW7-Report

For this project, I have written my code in a Jupyter Notebook .ipynb file. I have started my code by importing necessary libraries and writing our parameters down which were given in the pdf file. I have started my code by importing the data, assigning them to arrays and derive necessary parameters. I have derived class number, number of samples and features(784 in this case).

Next, I have derived some parameters from the data, such as class mean and point mean(total mean of every point). Then, I have implemented the algorithms we have seen in our class. First, to derive eigenvectors, I have calculated the within class matrix and between class matrices in a loop. Then, I have added an epsilon value of 1e-10 to prevent the singular matrix error in case our within matrix is singular. Then, by using scipy's linalg library, I have derived the eigenvalues and vectors.

After this, I have continued my dimension reduce. In the pdf, we were asked to do linear discriminant analysis with R=2. By getting the first two eigenvectors and multiplying them with our X and X_test, I have done the analysis. Then, I have scatter plotted them like this:



For the last part, I have implemented the kNN algorithm. By calculating the euclidean distances between the points and assigning the point to the mode of the nearest 5 points, I have predicted the data's classes for both my train data and test data. My confusion matrices were as below:

1: For y-train and y-predicted

```
[[846  14 152]
 [ 39 816 119]
 [141  67 806]]
```

2: For y-test and y-predicted

```
[[835   8 157]
 [ 49 851 100]
 [134  60 806]]
```