Lab 2 – Handwriting Recognition Report

This report will cover what I have learnt about classifiers, algorithms, feature extractions and more over the time period I performed this lab exercise.

KNN Classifier

The KNN classifier, while computationally demanding, is one of the most accurate classifiers. This classifier also doesn't have a learning algorithm. For a 2-value classification, if the K value is even, then a tie occurs and so this third data point is given one of the two labels at random.

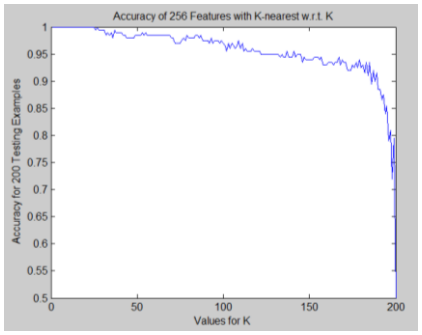

Figure 1: Above Graph for values 3 & 6

```
confusionMatrix =

    99     1
     8    92


labels =

     3
     6
```

Figure 2: Confusion Matrix for K = 120

Features

I made many features and I will only list the most creative features.
One feature used the Fibonacci sequence to select the data points.

Another transformed the original matrix into a more compact matrix with only every fourth column of the original matrix becoming part of the new matrix.

My fifth feature found the percentage of the total sum of all the values located in the top half of the matrix and the right half of the matrix and stored this as two features

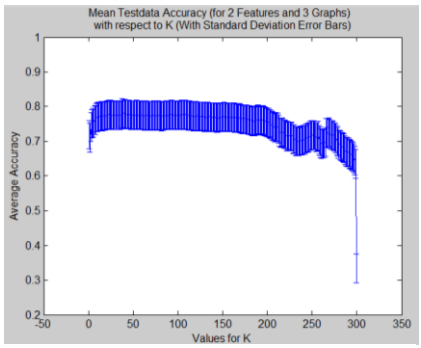

Figure 3: Above Graph for values 3,6, and 8

```
Average Accuracy is: 74.7758%
Feature Score: 37.3879
```

Figure 4: Score & Accuracy for Figure 3

Perceptron

A perceptron has a learning algorithm which updates its randomly assigned weights and threshold through the use of its learning rate variable. The perceptron cycles through a training set for a finite number of iterations to get a good value for the individual weights and the threshold.

The perceptron is then tested using testing data and from there, an accuracy can be calculated. Mistakes can be shown through the use of the showdata() function and a confusion matrix.

The perceptron, while slower at training data in comparison to the KNN classifier, is more memory efficient and can handle non-linearly separable data.

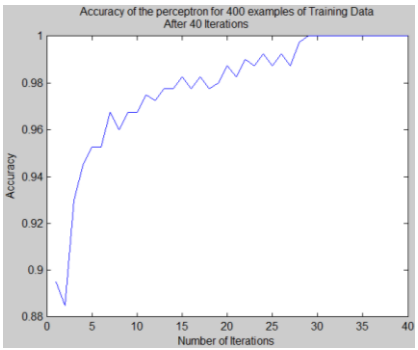

Figure 5: Accuracy for Values 3 & 6

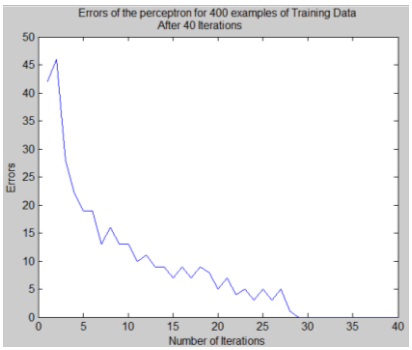

Figure 6: Errors for Values 3 & 6