CSC411: Assignment 2

Due on Sunday, November 12^{th}, 2017

Student Name: **Gokul K. Kaushik**

Student Number: **999878191**

Table of Contents

1 - Class Conditional Gaussians

1 - Using Bayes rule to derive an expression for $p(y = kx, \mu, \sigma)$

2 - Expression for the negative likelihood function (NLL)

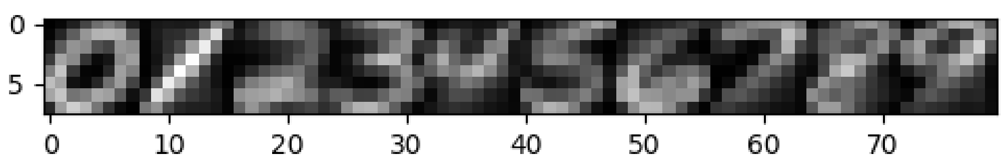
3 - Partial Derivatives of the Likelihood

4 - Find the maximum likelihood estimates for μ and σ

2 - Handwritten Digit Classification

0 - Loading the data and Plotting the Feature Means

The means (from 700 samples per digit) for each feature (64 features in total for an 8-by-8 pixel image) for 10 digits (digit 0 to digit 9) are plotted below:



1 - K-NN Classifier

Train and Test Classification Accuracy for K=1 and K=15

	Accuracy	
K Value	Train Classification	Test Classification
K = 1		
K = 15		

Tie Breaker Method

There are cases in K-Nearest Neighbours where there isn't one most frequent neighbours (there might be two neighbours that occur equally frequently).

Therefore, in such cases a tie breaking decision needs to be made. I have chosen to reduce the number of nearest neighbours by one **- effectively removing the last neighbour and repeating the check until a decision can be made**.

Data: K Nearest Neighbours Array with a tie
Result: Nearest Neighbours Decided
initialization;
while *not at end of this document* **do**
  read current;
  if *understand* **then**
    go to next section;
    current section becomes this one;
  else
    go back to the beginning of current section;
  end
end

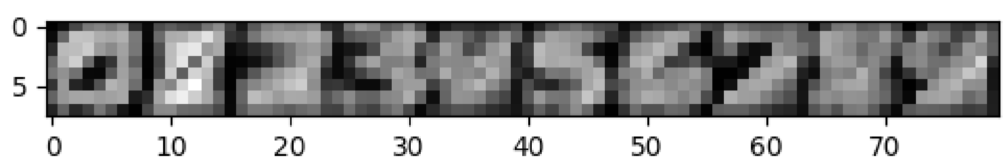
Algorithm 1: Tie Breaker Pseudo code

This method was chosen because:

1. The tie remains until the *most distant* neighbour from one of the most frequent neighbours is removed. This rewards the neighbour with the closest value in such cases.
2. This decision making method is intuitive to understand and easy to implement.

2 - Conditional Gaussian Classifier Training

Plot of the Diagonal Elements of the Covariance Matrix



Average Conditional (Log) Likelihoods

The average conditional (log) likelihoods were compute for the train set and the test set are:

Train Data Set: -0.124624436669
Test Data Set: -0.196673203255

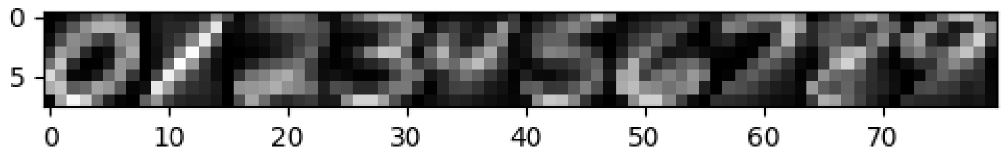
Accuracy for the Most Likely Posterior Class

For the most likely posterior class, the training and test set accuracies are:

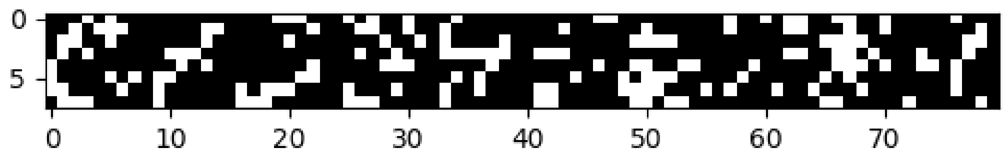
Train Data Set: 0.9814
Test Data Set: 0.9727

3 - Naive Bayes Classifier Training

Plot of Eta



Plot of Generated Binarized Data (Using a Binomial Disribution)



Average Conditional (Log) Likelihoods

The average conditional (log) likelihoods were compute for the train set and the test set are:

Train Data Set: -0.9437538618
Test Data Set: -0.987270433725

Accuracy for the Most Likely Posterior Class

For the most likely posterior class, the training and test set accuracies are:

Train Data Set: 0.7741
Test Data Set: 0.7642

4 - Model Comparison