**Machine Learning: Assignment-3**

**REPORT**

**1. Submitted by:**

2013A7PS152H- Simran Kapur

2013A7PS162H- Kinjal Jain

2013AAPS263H- Arth Chetan Patel

2013A7PS387H- Aditya Chivukula Satya Krishna

**2. Language used:** Java for part 1 and R for part 2

**3. Strategy used for training data in Part1:**

We first created 10 arraylists, each for every digit from 0-9 and then took input line by line from traininglabel file corresponding to each of which we read 28 lines from trainingimages file and stored it in the form of a 2 dimensional array with values 0 or 1 as per the symbol encountered. Depending on the actual label of that array, we added it to the arraylist belonging to that label.

We then found the priors value for each digit. And then for each array in that digit’s array list we counted the number of zeroes corresponding to each pixel and stored this result in a three dimensional array which we finally used to classify our testimages data. Finally we calculated the accuracy by comparing our classification with the actual testlabel data.

**4. Smoothening factor used, k:**

After checking and implementing the code using various values of k from 1 to 50, we realized that the accuracy was highest for the value of k being 1 which is 77.100006%

Therefore, we decided to take k=1.

**5. Accuracy for Part1 :** 77.100006%