CS771 - Assignment 1

Vikas Jain, 13788

February 3, 2016

Overview of results is present here, for complete details see output files for each part in their respective directories in the submission.

1) **Naivebayes**

   10-fold cross validation of total 2893 emails(2412 nonspam + 481 spam)
   positive = spam, negative = nonspam

Method	1a	1b	1c
True Positives	475	470	470
True Negatives	2394	2402	2402
False Positives	18	10	10
False Negatives	6	11	11
Total	2893	2893	2893
Accuracy	0.9917	0.9927	0.9927
Precision	0.9635	0.9792	0.9792

2) **Linear Discriminator (Perceptron)**

   10-fold cross validation of total 2893 emails(2412 nonspam + 481 spam)
   positive = spam, negative = nonspam

Method	2a	2b	2c
True Positives	462	472	458
True Negatives	2402	2402	2392
False Positives	10	10	20
False Negatives	19	9	23
Total	2893	2893	2893
Accuracy	0.9900	0.9934	0.9851
Precision	0.9788	0.9792	0.9582

3) K - Nearest Neighbors

60,000 Training images

10,000 Test Images

Below Matrix representing no. of images labelled correctly out of 10,000 against value of ‘k’ used and distance metric used

	k = 1	**k = 2**	**k = 3**	**k = 4**
L1/Manhattan	9631	9599	9642	9621
L2/Euclidean	9691	9653	9714	9683
Minkowski	9691	9653	9714	9683