

CS6140 Machine Learning Fall 2014 Homework 3, Wei Luo

PROBLEM 1

Using Gaussian Discriminant Analysis, I got:

average train error rate: 0.091212 average test error rate: 0.096498

Since the average accuracy is over 0.9, it seems to me that the gaussian assumption holds for this data set.

PROBLEM 2

The Error Table for Naïve Bayes Classifier, Model with Bernoulli Random Variables:

fold#	false positive rate	false negative rate	overall error rate
1	0.142322	0.067010	0.110629
2	0.134021	0.112426	0.126087
3	0.123288	0.089286	0.110870
4	0.104651	0.084158	0.095652
5	0.119718	0.062500	0.097826
6	0.127208	0.056497	0.100000
7	0.118467	0.127168	0.121739
8	0.114695	0.049724	0.089130
9	0.128571	0.083333	0.110870
10	0.123596	0.062176	0.097826
average	0.12365375	0.07942785	0.10606291

The Error Table for Naïve Bayes Classifier, Model with Gaussian Random Variables:

fold#	false positive rate	false negative rate	overall error rate
1	0.186207	0.140351	0.169197
2	0.121622	0.091463	0.110870
3	0.102273	0.056122	0.082609
4	0.138996	0.084577	0.115217
5	0.175000	0.077778	0.136957
6	0.178832	0.112903	0.152174
7	0.136842	0.085714	0.117391
8	0.138686	0.096774	0.121739
9	0.199301	0.137931	0.176087
10	0.157143	0.094444	0.132609
average	0.15349013	0.09780588	0.13148496

The Error Table for Naïve Bayes Classifier (4-bins Histogram):

fold#	false positive rate	false negative rate	overall error rate
1	0.086957	0.221622	0.140998
2	0.097222	0.151163	0.117391
3	0.076642	0.139785	0.102174
4	0.103321	0.190476	0.139130
5	0.064982	0.218579	0.126087
6	0.084559	0.196809	0.130435
7	0.078292	0.195531	0.123913
8	0.081851	0.201117	0.128261
9	0.094737	0.194286	0.132609
10	0.063604	0.225989	0.126087
average	0.08321663	0.19353558	0.12670848

The Error Table for Naïve Bayes Classifier (9-bins Histogram):

fold#	false positive rate	false negative rate	overall error rate
1	0.032374	0.251366	0.119306
2	0.025180	0.340659	0.150000
3	0.025271	0.371585	0.163043
4	0.007692	0.485000	0.215217
5	0.021818	0.356757	0.156522
6	0.027119	0.400000	0.160870
7	0.027027	0.358209	0.171739
8	0.049123	0.291429	0.141304
9	0.023649	0.329268	0.132609
10	0.017544	0.365714	0.150000
average	0.02567962	0.3549987	0.15606102

The ROC curves and AUC for each classifier is:

