Homework #4 Report
Written by: mhtay

P1

(gaussian_kernel.m, polynomial_kernel.m, p1.m)
Testing on the data using p1.m, the accuracy in octave was found to be 0.5 and 0.675 for the Gaussian and polynomial kernels respectively. The accuracy in matlab was 0.775 and 0.7250 for the Gaussian and polynomial kernels respectively.

P2.1

(ConstructInterval.m)

P2.2

(TrainHeldOut.m, PartitionHeldOut.m (for testing only), p2_2.m, nb_test.m, nb_train.m)

PartitionHeldOut was used to partition the data into K=2 or K=10 partitions. C=0.5.

	K =2	K=10
Accuracy	0.74	0.75
Lower Interval (0.95 confidence)	0.654	0.560
Upper Interval (.95 confidence)	0.826	0.940
Lower Interval (0.99 confidence)	0.627	0.501
Upper Interval (0.99 confidence)	0.853	0.999

P2.2-Question

The larger the number of partitions, the higher the accuracy, due to more data available for training, but the confidence intervals increase due to less testing data, and the higher the confidence percentage, the larger the confidence interval.

P2.3

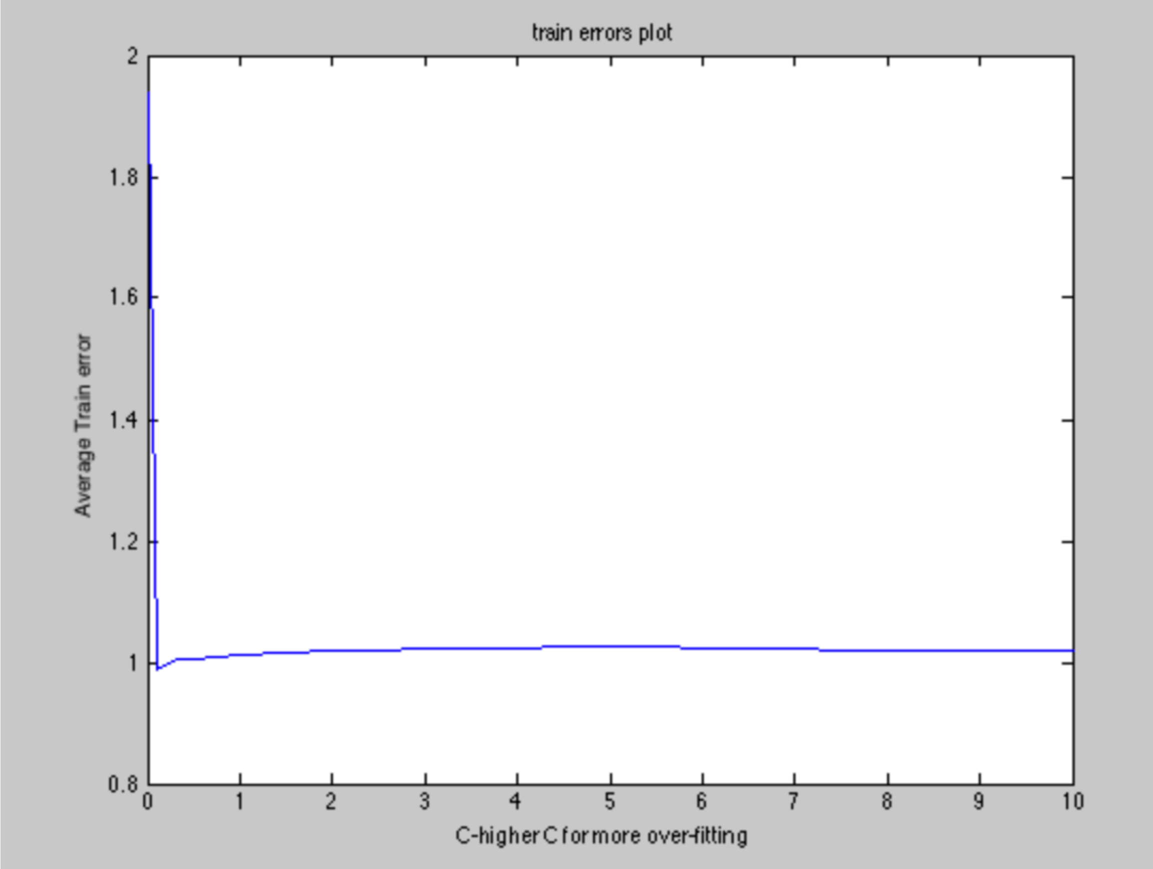
(PartitionCrossSet.m (for testing), TrainCrossSet.m)

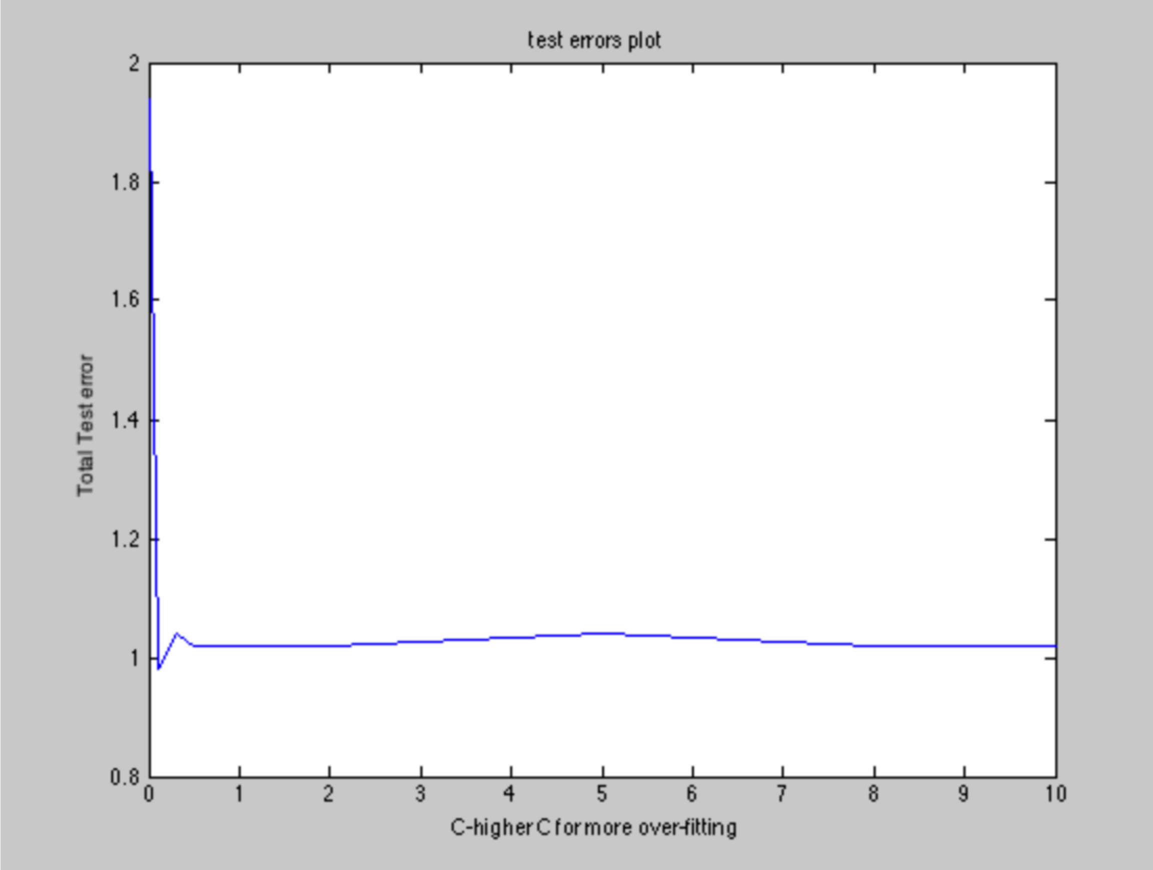
P2.3-Question

Using 4-fold cross validation-Using no kernel, the errors are:

C-penalty on errors	Training Error	Testing Error
0	1.94	1.94
0.1	0.9867	0.98
0.3	1.0067	1.04
0.5	1.0067	1.02
1	1.0133	1.02
2	1.0200	1.02
5	1.0267	1.04
8	1.0200	1.02
10	1.0200	1.02

Plot of graph:





Optimal value of C for no kernel SVM is when C= 0.1 as this produces the minimal training and testing error.

P2.4
Best Parameters for Logistic Regression:
lambda = 1; (regularization term)
Best Parameters for NN (for digits classification):
Hidden layers size: 64;
Lambda = 0.1;

Code:
compare_classifiers.m

Partition Number	Accuracy for Logistic Regression	Accuracy for Neural Network
1	0.5 = 0.5	1.0
2	1-0.545 = 0.455	1.0
3	1-0.409 = 0.591	1.0
4	1-0.318 = 0.682	1.0
5	1-0.591 = 0.409	1-0.045 = 0.955

6	1-0.455 = 0.545	1.0
7	0.5	1-0.045 = 0.955
8	1-0.591 = 0.409	1-0.045 = 0.955
9	0.5	1.0
10	1-0.364 = 0.636	1.0

Partition Number	Error for Logistic Regression	Error for Neural Network
1	0.5 = 0.5	0
2	0.545	0
3	0.409	0
4	0.318	0
5	0.591	0.045
6	0.455	0
7	0.5	0.045
8	0.591	0.045
9	0.5	0
10	0.364	0

Random Variable Y computed as difference of errors:

```
Y_mean =  0.4636;
Y_std_dev = 0.0796;
t-value = 18.4190;
```

Doing t-test for 1 sample t-test: Where Y_i is the vector containing the values for the random variable Y.

For two-tailed test:

```
[h, p, ci,stats] = ttest(Y_i)
p = 1.8747e-08
```

Reasoning:
Probability of observing data given null hypothesis is true is 1.8747e-08. Reject the null hypothesis and that the two methods have the same error rate at 5% significance level in favor of the hypothesis that they are different.

For one-tailed test (null hypothesis being that the expectation of Y is 0):

```
[h, p, ci,stats] = ttest(Y_i,zeros(1,size(Y_i,1)),'Tail','Right')
p = 9.3737e-09
```

Reasoning:
Probability of observing data given null hypothesis is true is 9.3737e-09. Reject the null hypothesis and that the two methods have the same error rate at 5%

significance level in favor of the hypothesis that they are error rate of logistic regression is higher than error rate of neural network.