10-601 Machine Learning - Homework 4 Report

Dawei Wang

daweiwan@andrew.cmu.edu

October 18, 2014

Table 1: 95% and 99% Confidence Interval for $k = 2$ and $k = 10$

	$k = 2$			$k = 10$	
Acc	.95%	.99%	Acc	.95%	.99%
64.0%	[0.546, 0.734]	[0.516, 0.764]	70.0%	[0.499, 0.901]	[0.436, 0.964]
73.0%	[0.643, 0.817]	[0.616, 0.844]	80.0%	[0.625, 0.975]	[0.570, 1.030]
70.0%	[0.610, 0.790]	[0.582, 0.818]	55.0%	[0.332, 0.768]	[0.263, 0.837]
70.0%	[0.610, 0.790]	[0.582, 0.818]	65.0%	[0.441, 0.859]	[0.375, 0.925]
68.0%	[0.589, 0.771]	[0.560, 0.800]	55.0%	[0.332, 0.768]	[0.263, 0.837]
64.0%	[0.546, 0.734]	[0.516, 0.764]	65.0%	[0.441, 0.859]	[0.375, 0.925]
70.0%	[0.610, 0.790]	[0.582, 0.818]	80.0%	[0.625, 0.975]	[0.570, 1.030]
73.0%	[0.643, 0.817]	[0.616, 0.844]	70.0%	[0.499, 0.901]	[0.436, 0.964]
78.0%	[0.699, 0.861]	[0.673, 0.887]	80.0%	[0.625, 0.975]	[0.570, 1.030]
76.0%	[0.676, 0.844]	[0.650, 0.870]	65.0%	[0.441, 0.859]	[0.375, 0.925]

Held-out Set Test The test was carried out for ten times. See Table 1 for the result.
Observation Larger k results in wider confidence intervals and more fluctuating accuracies.
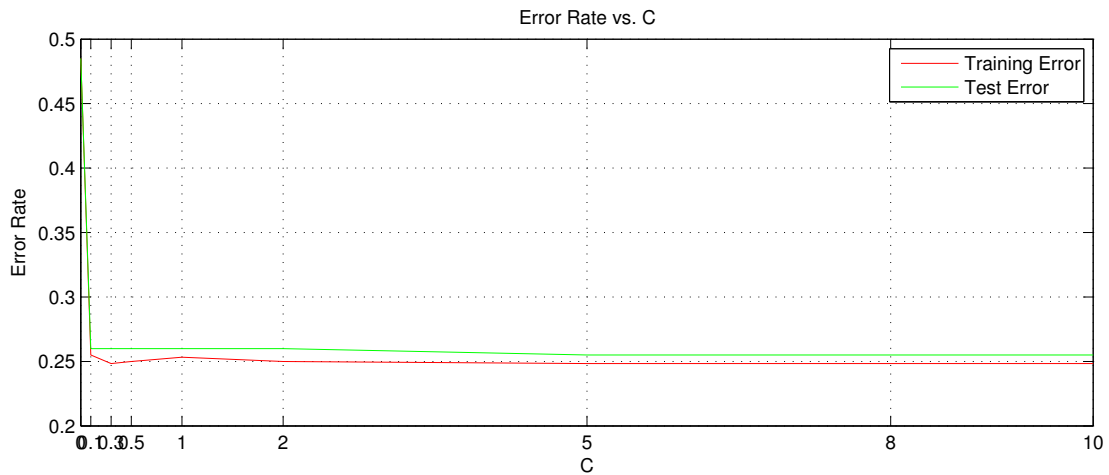

Figure 1: Error Rate versus C

Find C by Cross-Validation The test was carried out for five times. See Figure 1 for the result. The red polylines correspond to the training errors, and the green to the testing errors.
Observation $C = 0.1$ yields the best training and testing errors for this particular set of data.

Table 2: LR and NN Comparison: Error Rate

AL	1	2	3	4	5	6	7	8	9	10
LR	0.000	0.000	0.000	0.091	0.000	0.000	0.045	0.000	0.000	0.000
NN	0.000	0.000	0.000	0.091	0.000	0.000	0.045	0.000	0.000	0.000
LR	0.045	0.000	0.000	0.000	0.000	0.000	0.091	0.000	0.000	0.000
NN	0.045	0.000	0.000	0.000	0.000	0.000	0.091	0.000	0.000	0.000
LR	0.000	0.045	0.045	0.000	0.000	0.000	0.000	0.045	0.000	0.000
NN	0.000	0.045	0.045	0.000	0.000	0.000	0.000	0.045	0.000	0.000
LR	0.045	0.000	0.000	0.045	0.000	0.000	0.045	0.045	0.000	0.000
NN	0.000	0.000	0.000	0.045	0.000	0.000	0.045	0.045	0.000	0.000
LR	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.045	0.087
NN	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.045	0.087

LR and NN Comparison Only the instances labeled with 4 and 7 are used in this test. Both classifiers were first trained on the filtered `Xtrain` and `Ytrain`, then tested against 10 disjoint subsets obtained by partitioning the filtered `Xtest` and `Ytest`. See Table 2 for the result. Only the fourth group shows difference. **Observation** For the group that shows difference: $p = 0.343$ under two-tailed test, $p = 0.171$ under one-tailed test; but it's not significant enough to assert which one is better. For this particular data-set, they have performed equally well.