10-601: Homework 4

Due: 18 October 2014 11:59pm (Autolab)

TAs: Qihui Li, Siping Ji

Name: Alvin Chou

Andrew ID: alvinch o

Please answer to the point, and do not spend time/space giving irrelevant details. Please state any additional assumptions you make while answering the questions. For Questions in this assignment, you need to submit your answers in a single PDF file on autolab, either a scanned handwritten version or a LaTeXpdf file. Please make sure you write legibly for grading. For Implementations, submit your m-files on autolab.

You can work in groups. However, no written notes can be shared, or taken during group discussions. You may ask clarifying questions on Piazza. However, under no circumstances should you reveal any part of the answer publicly on Piazza or any other public website. The intention of this policy is to facilitate learning, not circumvent it. Any incidents of plagiarism will be handled in accordance with CMU's Policy on Academic Integrity.

⋆: Code of Conduct Declaration

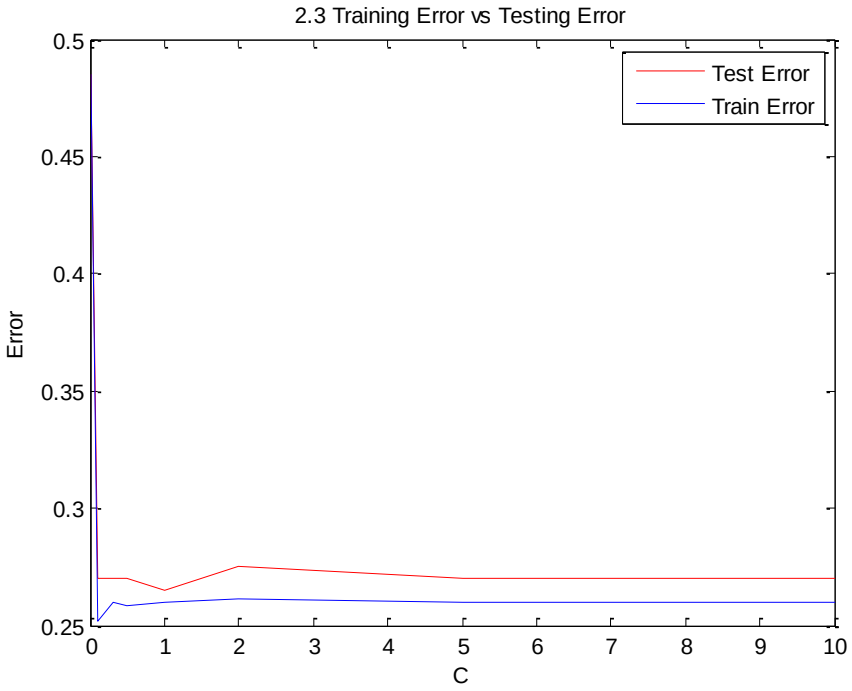
- Did you receive any help whatsoever from anyone in solving this assignment? Yes / No.
- If you answered *yes*, give full details: _____ (e.g. *Jane explained to me what is asked in Question 3.4*)
- Did you give any help whatsoever to anyone in solving this assignment? Yes / No.
- If you answered *yes*, give full details: _____ (e.g. *I pointed Joe to section 2.3 to help him with Question 2*).

Alvin Chou
alvincho
10601 HW4

2.2) Confidence Interval increaseas with higher partitions and lower significance percentage (ie 95% yields higher confidence interval than 99%).

	2 Partitions		10 Partitions	
	95%	99%	95%	99%
Accuracy	0.75	0.75	0.85	0.85
Lower Interval	0.6651	0.6385	0.6935	0.6443
Upper Interval	0.8349	0.8615	1.0065	1.0557

2.3)



Optimal C occurs when C = 0.1

2.4) From t-test analysis, the P-value and CI of all tests (left, right, and both) are within the 95% significance range, meaning that the null hypothesis CANNOT be reject as the difference is not deemed to be significant. As such, there cannot be a determination made as to which algorithm is better.

k	1	2	3	4	5	6	7	8	9	10	avg
NN Test Accuracy	92.67	96.00	95.33	92.67	92.33	92.00	93.33	94.67	91.33	93.00	93.33
NN Train Accuracy	99.30	99.26	99.48	99.04	99.48	99.33	99.44	99.59	99.30	99.26	99.35
LR Test Accuracy	96.97	96.97	98.48	98.48	96.97	98.48	98.48	98.48	96.97	96.97	97.73
LR Train Accuracy	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	98.99	100.00	99.90

	Confidence Interval		Hypothesis	P-Value
Left	-INF	7.375	0	0.5
Right	5.7563	INF	0	0.5
Both	5.5669	7.5644	0	1