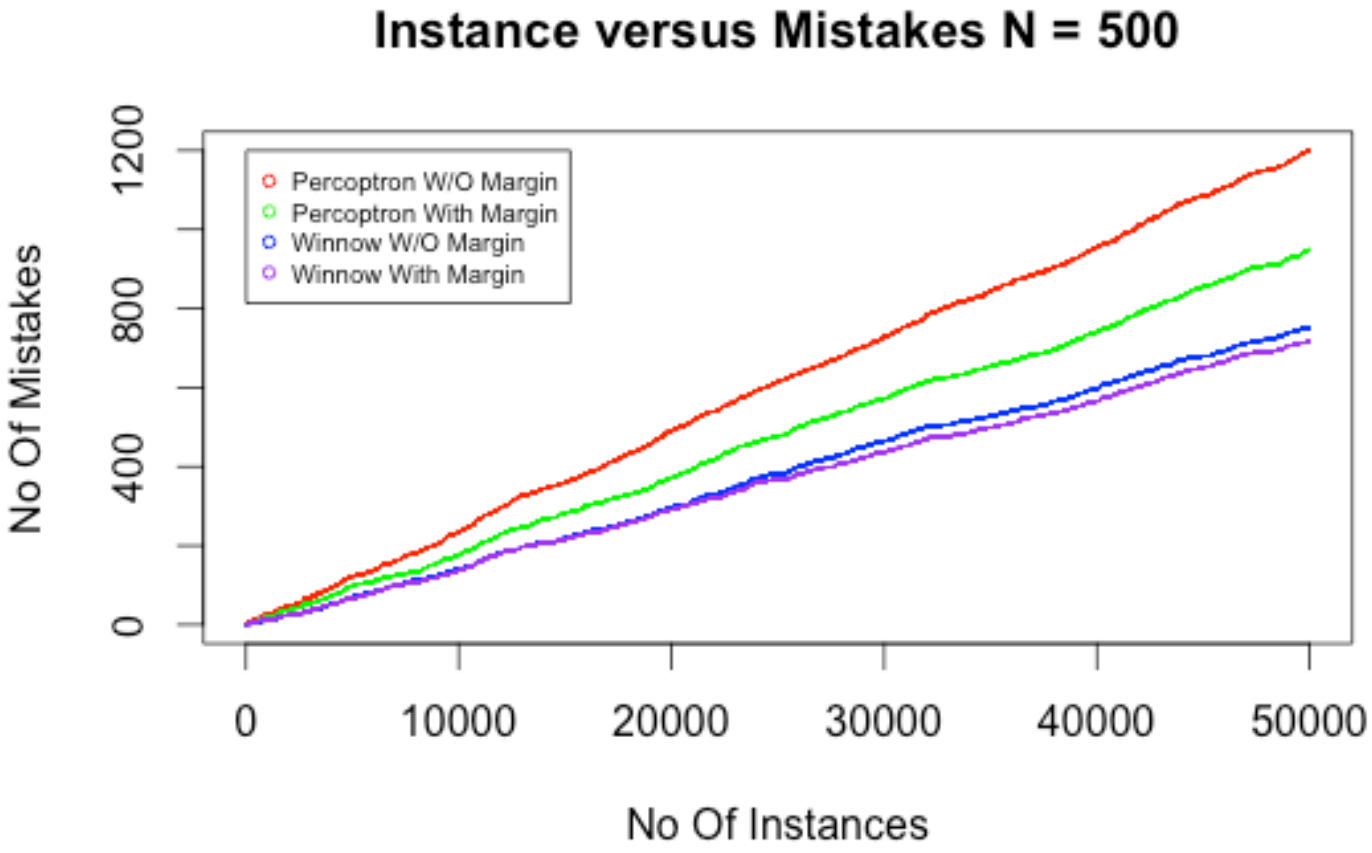
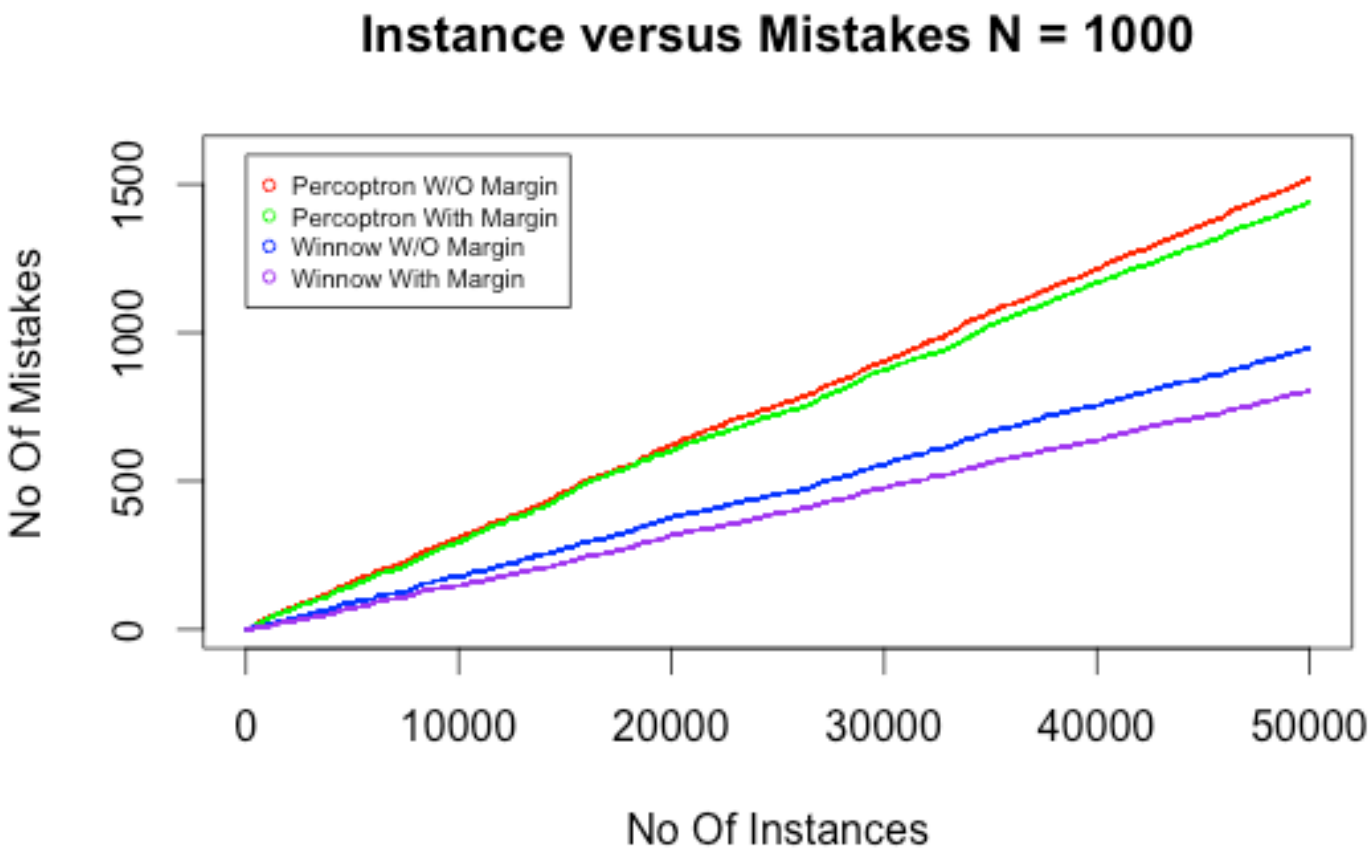
CS6140: Machine Learning – Homework 4
Name: Vishal Rajpal

Problem 1:

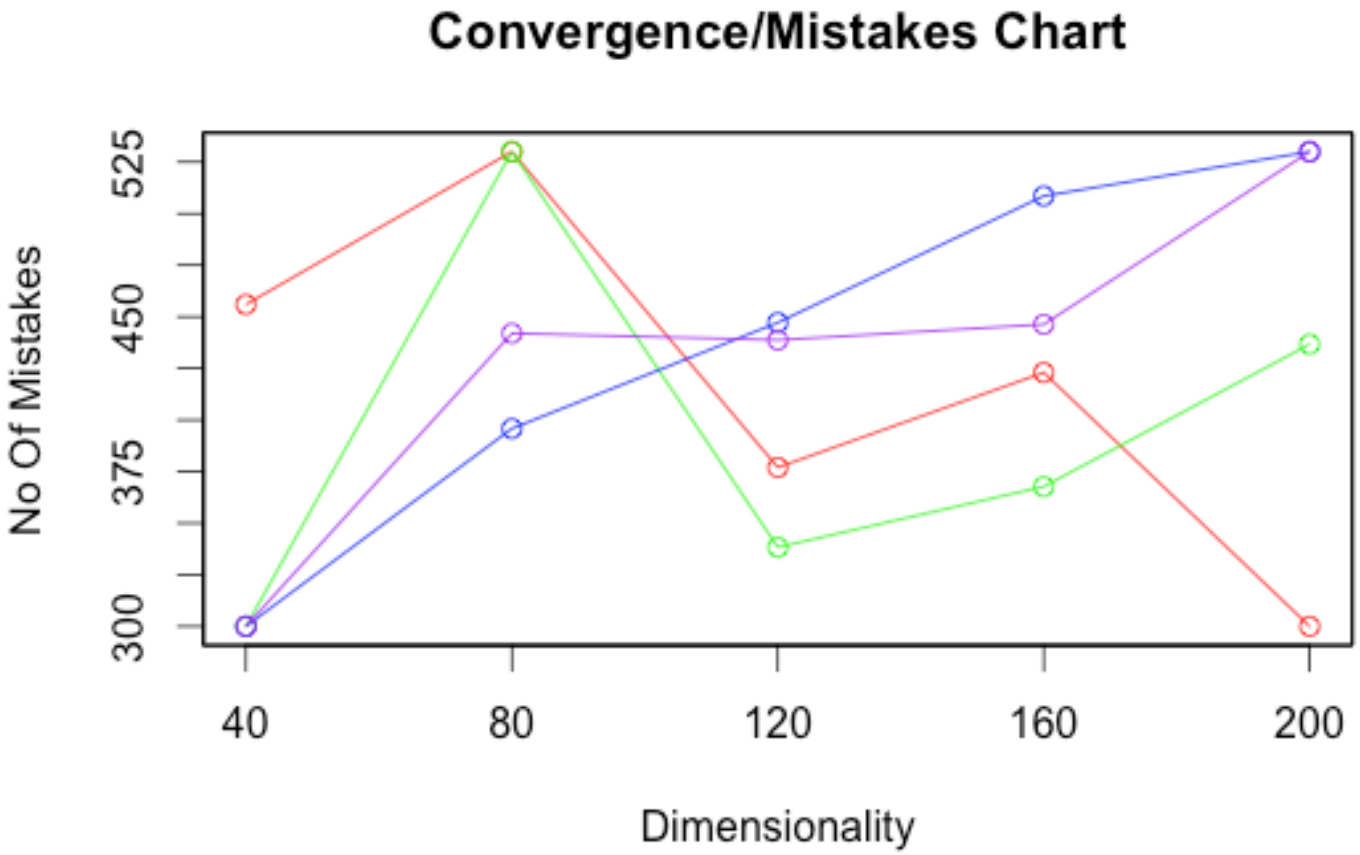
Algorithm	n=500			n=1000		
	Params	Acc($\mathcal{D}_2$)	M	Params	Acc($\mathcal{D}_2$)	M
Perceptron ($\gamma = 0$)	$\gamma = 0, \eta = 1.0$	97.74%	1200	$\gamma = 0, \eta = 1.0$	96.85%	1521
Perceptron ($\gamma > 0$)	$\gamma = 1.0, \eta = 0.001$	98.24%	949	$\gamma = 1.0, \eta = 0.25$	97.42%	1442
Winnow ($\gamma = 0$)	$\gamma = 0, \eta = 1.1$	98.97%	751	$\gamma = 0, \eta = 1.1$	98%	950
Winnow ($\gamma > 0$)	$\gamma = 2.0, \eta = 1.1$	99.05%	716	$\gamma = 0.3, \eta = 1.1$	98.38%	807




Instance versus Mistakes N = 1000
Percoptron W/O Margin
Percoptron With Margin
Winnow W/O Margin
Winnow With Margin
No Of Mistakes
0
500
1000
1500
0
10000
20000
30000
40000
50000
No Of Instances

Problem 2:

		Perceptron	Perceptron (γ)	Winnow	Winnow (γ)
$n = 40$	Params	$\gamma = 0, \eta = 1.0$	$\gamma = 1.0, \eta = 0.25$	$\gamma = 0, \eta = 1.1$	$\gamma = 2.0, \eta = 1.01$
	Acc($\mathcal{D}_2$)	99.65%	99.65%	99.98%	100%
	M	456	317	228	94
$n = 80$	Params	$\gamma = 0, \eta = 1.0$	$\gamma = 1.0, \eta = 0.25$	$\gamma = 0, \eta = 1.1$	$\gamma = 2.0, \eta = 1.01$
	Acc($\mathcal{D}_2$)	98.47%	98.72%	99.87%	99.95%
	M	530	677	353	486
$n = 120$	Params	$\gamma = 0, \eta = 1.0$	$\gamma = 1.0, \eta = 1.5$	$\gamma = 0, \eta = 1.1$	$\gamma = 0.3, \eta = 1.1$
	Acc($\mathcal{D}_2$)	97.97%	97.97%	99.95%	99.98%
	M	377	377	420	477
$n = 160$	Params	$\gamma = 0, \eta = 1.0$	$\gamma = 1.0, \eta = 1.5$	$\gamma = 0, \eta = 1.1$	$\gamma = 0.04, \eta = 1.1$
	Acc($\mathcal{D}_2$)	97.39%	97.39%	99.77%	99.98%
	M	423	423	500	498
$n = 200$	Params	$\gamma = 0, \eta = 1.0$	$\gamma = 1.0, \eta = 0.25$	$\gamma = 0, \eta = 1.1$	$\gamma = 2.0, \eta = 1.1$
	Acc($\mathcal{D}_2$)	95.74%	96.74%	100%	100%
	M	300	531	528	729



Problem 3:

Algorithm	$m = 100$			$m = 500$			$m = 1000$		
	Params	Acc($\mathcal{D}_2$)	M	Params	Acc($\mathcal{D}_2$)	M	Params	Acc($\mathcal{D}_2$)	M
Perc.	$\gamma = 0$ $\eta = 1.0$	81.69%	2286	$\gamma = 0$ $\eta = 1.0$	78.74%	4311	$\gamma = 0$ $\eta = 1.0$	83.66%	860
Perc. (γ)	$\gamma = 1.0$ $\eta = 0.001$	86.44%	2291	$\gamma = 1.0$ $\eta = 0.001$	88.05%	4791	$\gamma = 1.0$ $\eta = 0.001$	88.26%	445
Winnow	$\gamma = 0$ $\eta = 1.1$	83.35%	1743	$\gamma = 0$ $\eta = 1.0001$	87.87%	1086	$\gamma = 0$ $\eta = 1.01$	85.75%	1026
Winnow(γ)	$\gamma = 0.001$ $\eta = 1.1$	83.35%	2074	$\gamma = 2.0$ $\eta = 1.0001$	87.98%	1065	$\gamma = 0.001$ $\eta = 1.01$	85.75%	1026

Problem 4: Naïve Bayes

Algorithm	$m = 100$		$m = 500$		$m = 1000$	
	Accuracy on Testing	M	Accuracy on Testing	M	Accuracy on Testing	M
Naïve Bayes	93.28%	673	93.61%	672	62.79%	3721

Findings:

Performance Discussion: Winnow with Margin performed best in all the cases and Perceptron Without Margin performed worst. The results were consistent in all the experiments. Winnow generally performed better than Perceptron as it ignores the irrelevant features, on the other hand Perceptron cannot ignore the irrelevant features due to the presence of parameter bias in the calculation, which leads it to be prone to noise and make larger mistakes, compared to Winnow which makes logarithmic mistakes in the number of irrelevant features.

Convergence: All the algorithms were converging with S=1000. The only case i.e. Winnow with Margin was not able to converge when n=40. It just stopped learning and the best predictions it could make 791. Even I produce a different data set, this number stayed same. In all the other values of n, Winnow with Margin was able to converge with S=1000.