

## CS6140 Machine Learning Fall 2014 Homework 2, Wei Luo

### PROBLEM 1

**B)** Training and testing performance across all four learning algorithms is:

	Decision or Regression Tree	Linear Regression (Normal Equations)	Linear Regression (Gradient Descent)	LogisticRegression (Gradeint Descent)
Spambase	Train ACC: 0.91 Test ACC: 0.90	Train ACC: 0.92 Test ACC: 0.88	Train ACC: 0.91 Test ACC: 0.90	Train ACC: 0.92 Test ACC: 0.91
Housing	Train MSE: 26.49 Test MSE: 24.28	Train MSE: 22.08 Test MSE: 22.63	Train MSE: 22.13 Test MSE: 22.58	N/A.

**C)**

Confusion Matrix for Decision Tree:

		Prediction	
		1	0
Label	1	152	29
	0	14	264

Confusion Matrix for Linear Regression:

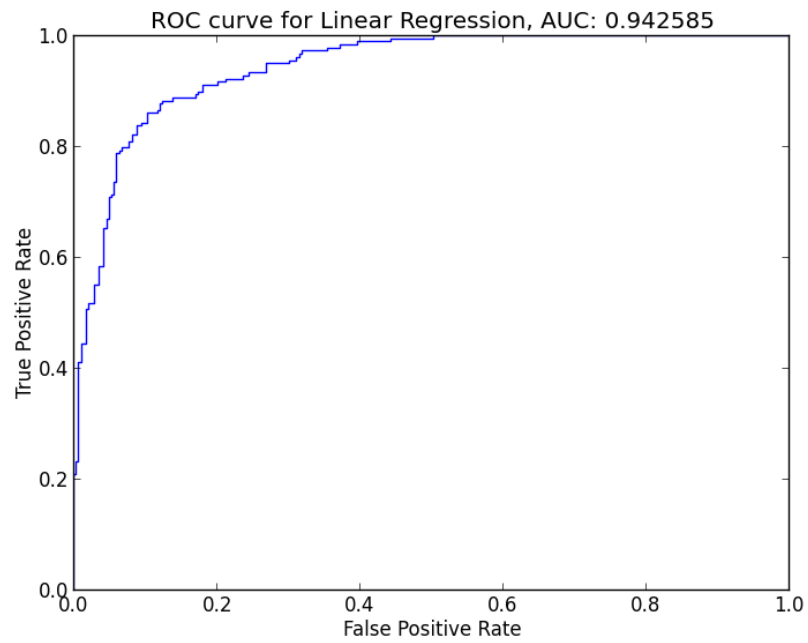
		Prediction	
		1	0
Label	1	162	19
	0	22	256

Confusion Matrix for Logistic Regression:

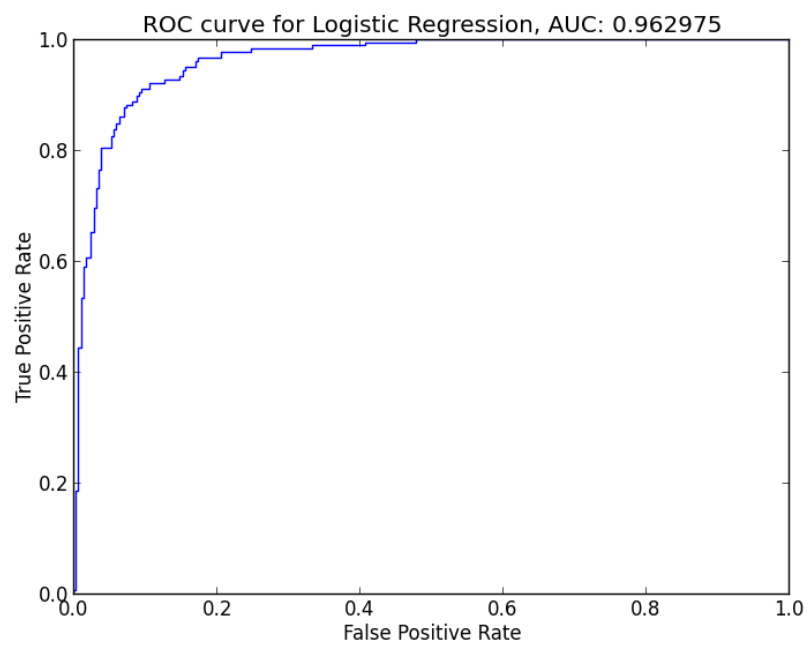
		Prediction	
		1	0
Label	1	154	27
	0	10	268

**D)**

ROC curve for Linear Regression:



ROC curve for Logistic Regression:



## PROBLEM 2

Iteration 1, total\_mistake 136  
Iteration 2, total\_mistake 68  
Iteration 3, total\_mistake 50  
Iteration 4, total\_mistake 22  
Iteration 5, total\_mistake 21  
Iteration 6, total\_mistake 34  
Iteration 7, total\_mistake 25  
Iteration 8, total\_mistake 0

Classifier weights:   -14.    2.52873259    5.70717051    8.52231457   11.32560723

Normalized with threshold:    0.18062376    0.40765504    0.60873676    0.80897195

## PROBLEM 3