Homework 3 CMPS242 - Fall 2015

Ehsan Amid Andrei Ignat Sabina Tomkins

1 Problem One

1.1 Part a

Save the data to an aff file and run the linear regression algorithm in Weka on the full training set. Report the model and Root mean squared error". Note that testing on the training set means we will have a relatively small error.

Root Mean Square Error
0.1897

1.2 Part b

Suppose you had an unlabeled instance x = 3,3,5. What t prediction for t would the model from part (a) give?

Should probably write the coefficients out. -0.1343 * x1 +1.8477 * x2 +-0.8966 * x3 +4.3608
Answer:5.0180

1.3 Part c

1.616 * x2 + -1.0388 * x3 + 5.7513 Is the change qualitatively what you would expect. Yes. RMSE:0.4473

1.4 Part d

Will come back to this.

1.5 Part e

If the examples are re-ordered (so the rows of X and elements of t are permuted), what happens to the learned w vector and why?

2 Problem Two

	Nearest Neighbor	Naive Bayes	Logistic Regression
Accuracy	100%	76.3021%	78.2552%