CSCI 5622 – Nick Ketz
Assignment 2 Analysis:

1. What is the role of the learning rate?

Learning rate or step size scales down the change in beta parameters to smooth out the process or learning the beta weights. As seen in Figure 1, a learning rate that is too large will cause beta changes that are too large and will over fit the data and potentially overshoot the optimal solution. Beta weights that are too small will take a long time to reach the optimal solution, and perhaps get stuck in non-optimal solution. This is particularly important when doing regularized regression where the over inflation of particular features will not lead to better test performance.

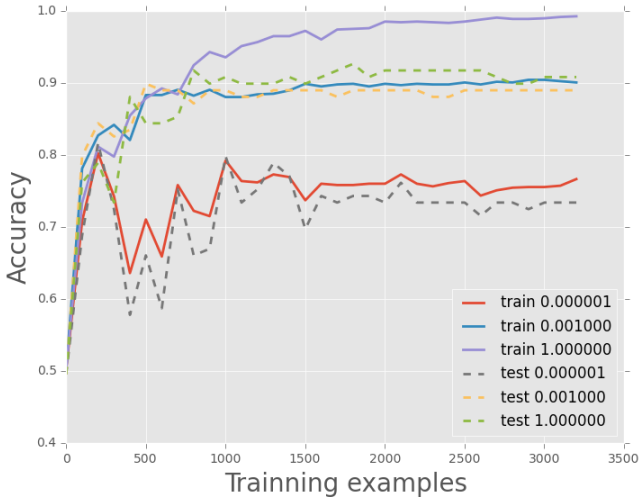

Figure 1: Accuracy for various learning rates on train and test data as a function of training examples

2. How many passes over the data do you need to complete?

This depends on the learning rate, and whether the regression is regularized or not. For Non-regularized with a well chosen learning rate, shown in Figure 1, performance plateaus at around 1000 training examples or 1 pass through the full set.

3. What words are the best predictors of each class? How (mathematically) did you find them?

See Table 1 on right, 10 best features for class 1 have the largest betas and for class 0 have the smallest betas

4. What words are the poorest predictors of classes? How (mathematically) did you find them?

See Table 1 on right, worst features have the weights that are closest to zero

Class 0	Class 1	Worst
playoffs	big	calculatior
period	luriem	elite
shots	swing	estimate
traded	better	fastball
playoff	rickert	attitudes
play	hit	breath
puck	ball	cliches
ice	baseball	essence
points	runs	fathom

Table 1 10 best and worst features