

# CSCI 5622 - Logistic Regression Assignment

Byron Becker

## Analysis

1. How did the learning rate affect the convergence of your SGA implementation?

I graphed the holdout accuracy and test accuracy as a function of training example updates with eta ranging from 0.01 – 5 (through the first pass of training examples), and we can see that the learning rate has an affect on how fast the accuracy initially converges, as well as determining a limit on how accurate our feature vector can become if you give it more training data. One can see that the hold out accuracy seems to converge faster initially at smaller eta (relative to the tested etas), with the fastest initial convergence around 0.05. We can see that at large eta (2-5), the accuracy in the hold out set is very volatile through the update process, which corresponds with taking larger jumps. Additionally, I found that etas of 0.05-1.0 gave very high hold out and test accuracies, both being over 0.95 -> 95% within the first pass, whereas larger etas were not able to reach that accuracy in the same amount of updates.

2. What was your stopping criterion, and how many passes over the data did you need to complete before stopping?

My stopping criterion was essentially to find a point (number of passes) where the accuracy was not improving or decreasing visibly (being within 0.03 of the previous pass). As in the previous question, I found that etas ranging from 0.01 to 1.0 converged the quickest and produced the highest accuracies. After trying several pass variations, I found that the accuracies only needed 1 pass to converge, and using an eta of 0.05, I received a test accuracy of 0.994334 -> 99.4%, and a hold out accuracy of 0.962 -> 96.2%.

3. What words are best predictors of each class? What are the worst predictors of each class?

### Best Predictors:

Automobiles: car, cars, ford, toyota, usa, engine, warning, parts, also, keys, oil, dealer, concepts, wagon, acs

Motorcycles: moa, battery, shaft, mike, ca, exhaust, sun, dog, riding, helmet, rider, ride, bikes, bike, dod

### Poorest Predictors of Both Classes:

agate, winds, tower, character, hired, labor, requests, anyhow, respectable, 5th, lemon, inaccurate, becomes, goal, presume, verify, prone, helping, hardly, tight

To find the best predictors of each class, I sorted the feature vector and took the features at the beginning of the vector (most negative), which were indicative of automobiles the "0" class. Likewise, I took the features at the end of the sorted vector (most positive), which I then interpreted as indicative of the motorcycle or "1" class. To get the poorest predictors of either class, I looked at twenty features smack dab in the middle of the sorted feature vector (close to zero), meaning they had little effect at all in classifying the document, and therefore were poor predictors of document classification.











