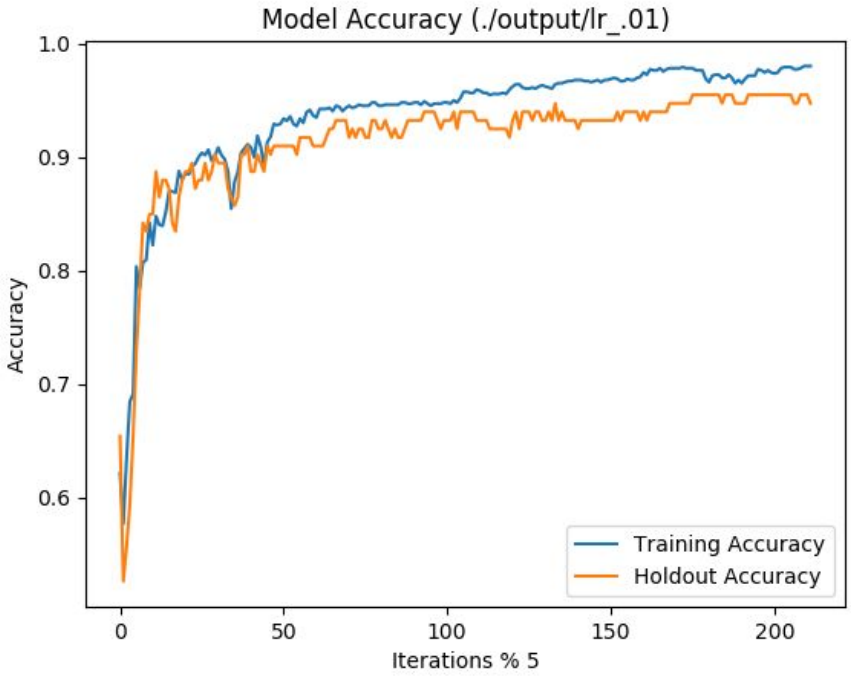
Raymond Duncan
CSCI 5622
Logreg Writeup

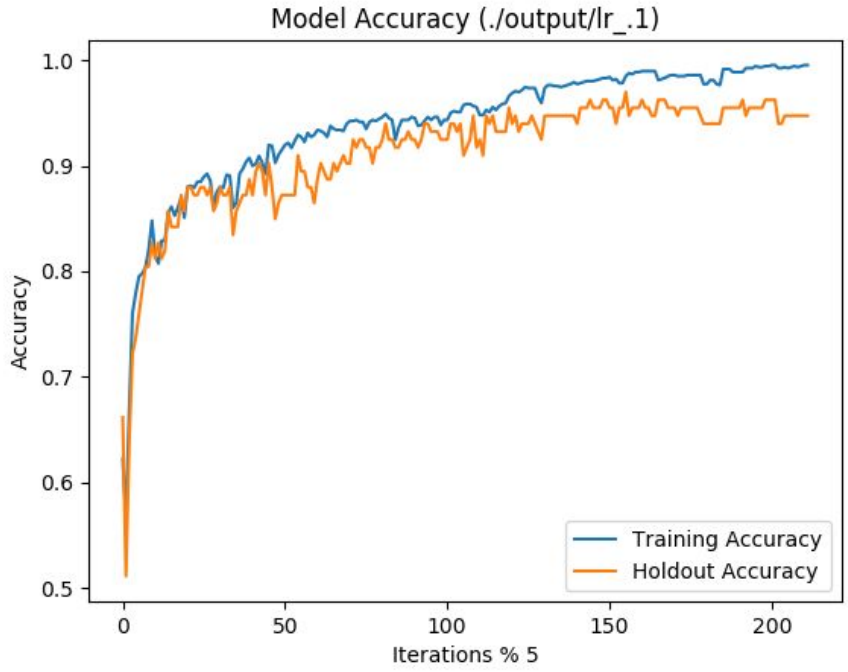
1) How did the learning rate affect the convergence of your SGA implementation?
With eta as a constant, it seemed that the larger eta was the more examples were required to reach maximum accuracy. (Graphs are those with eta in the title)

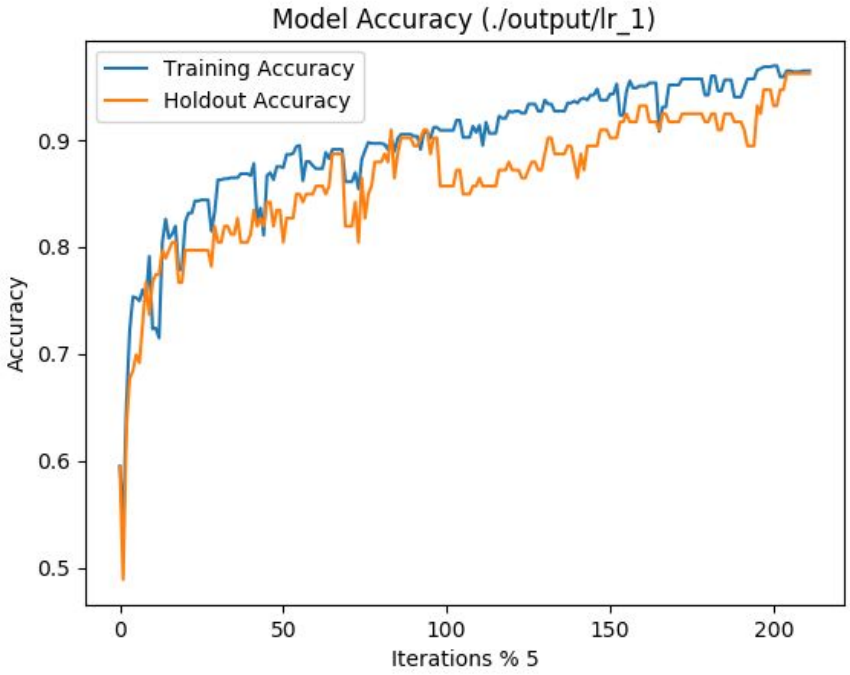
2) What was your stopping criterion and how many passes over the data did you need to complete before stopping?
I tried running the model with several different inputs for the passes argument. I decided the stopping criterion was when the holdout accuracy was consistently above 90% (It didn't decrease to below 90% for the remainder of the accuracy measurements.. With this dataset, it was possible after only one pass, but in order to ensure as little variance in the end accuracies, I decided to run the final model with 2 passes. (Graphs are those with passes in the title) (Final graph has final in the name)

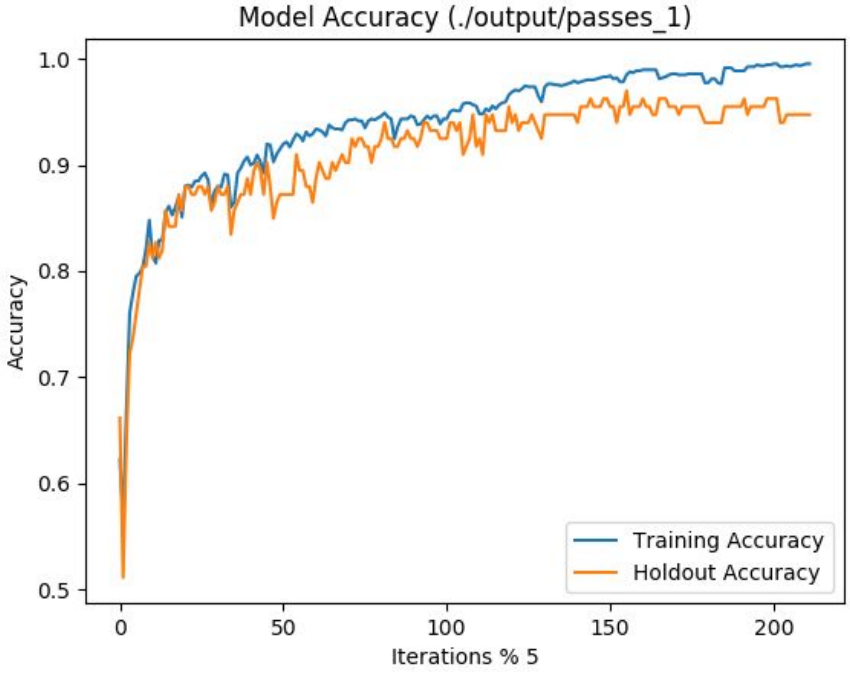
3) What words are the best predictors of each class? How (mathematically) did you find them?
The 5 best predictors of each class were:
Cars: car, cars, ford, toyota, engine
Motorcycles: bike, did, bikes, ride, riding
I found these by finding the features with the highest (motorcycle) and lowest (car) weights

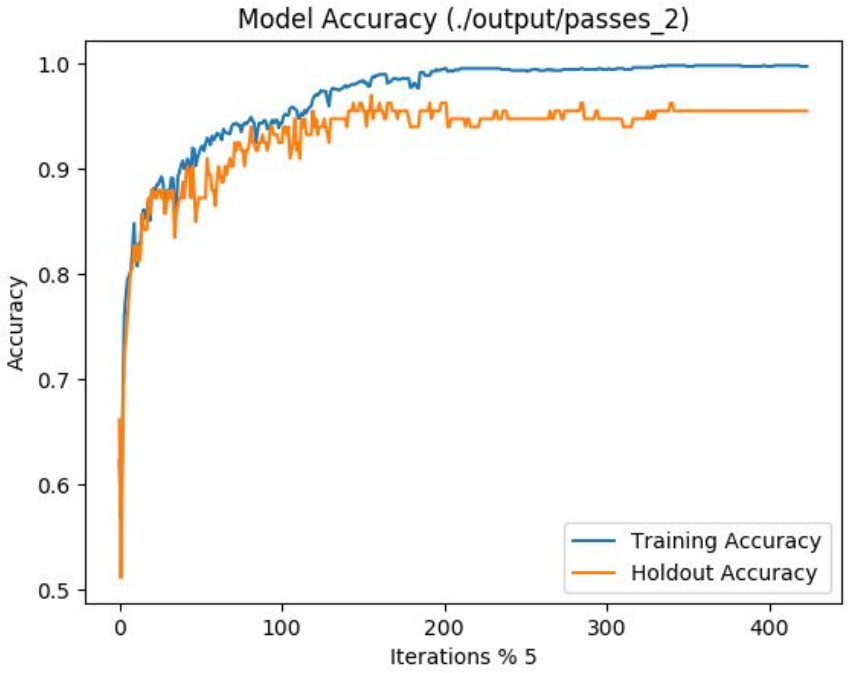
4) What words are the poorest predictors of classes? How (mathematically) did you find them?
5 of the worst predictors for either class are:
highest, horribly, arrive, medium, heavily
I found these by find the features with weights closest to 0.

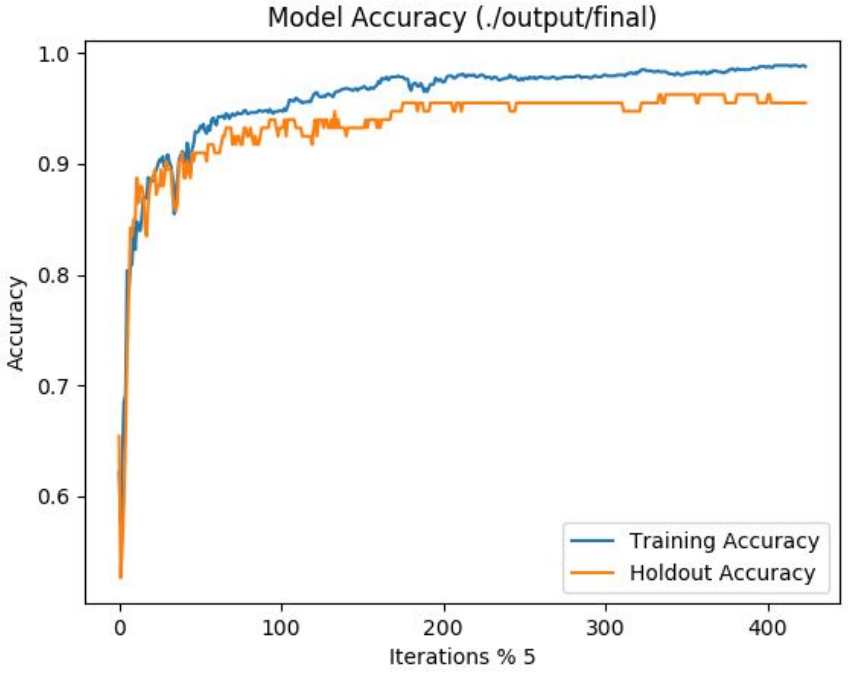

Model Accuracy (./output/lr_.01)
Accuracy
Iterations % 5
Training Accuracy
Holdout Accuracy
1.0
0.9
0.8
0.7
0.6
0
50
100
150
200


Model Accuracy (./output/lr_.1)
Accuracy
Iterations % 5
Training Accuracy
Holdout Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0
50
100
150
200


Model Accuracy (./output/lr_1)
Training Accuracy
Holdout Accuracy
Accuracy
0.9
0.8
0.7
0.6
0.5
0
50
100
150
200
Iterations % 5


Model Accuracy (./output/passes_1)
Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
Training Accuracy
Holdout Accuracy
0
50
100
150
200
Iterations % 5


Model Accuracy (./output/passes_2)
Accuracy
Iterations % 5
Training Accuracy
Holdout Accuracy
1.0
0.9
0.8
0.7
0.6
0.5
0
100
200
300
400


Model Accuracy (./output/final)
Accuracy
Iterations % 5
Training Accuracy
Holdout Accuracy
1.0
0.9
0.8
0.7
0.6
0
100
200
300
400