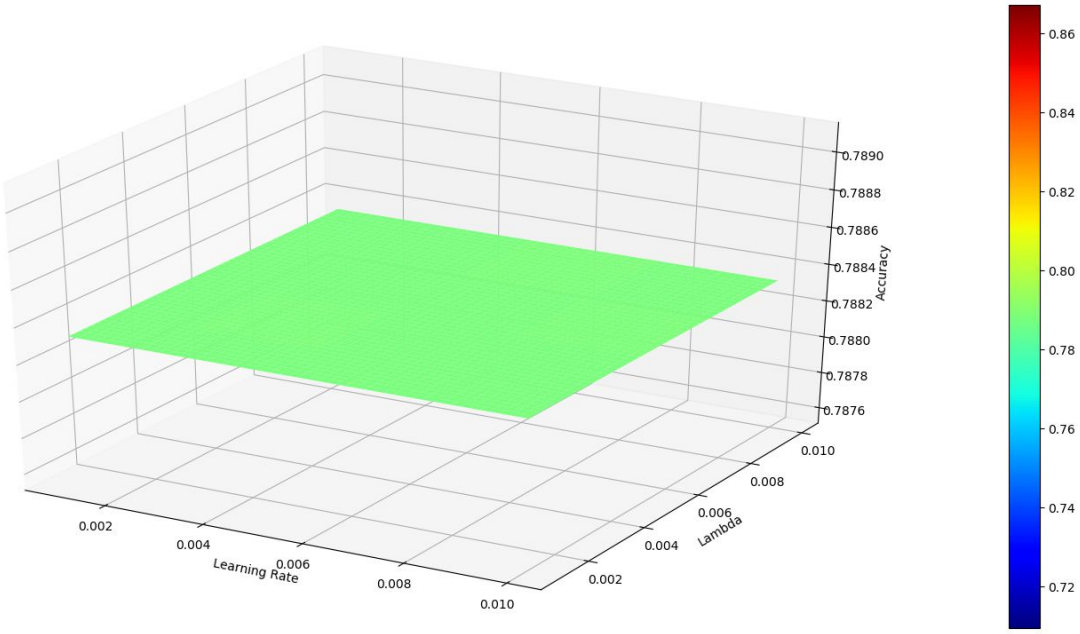
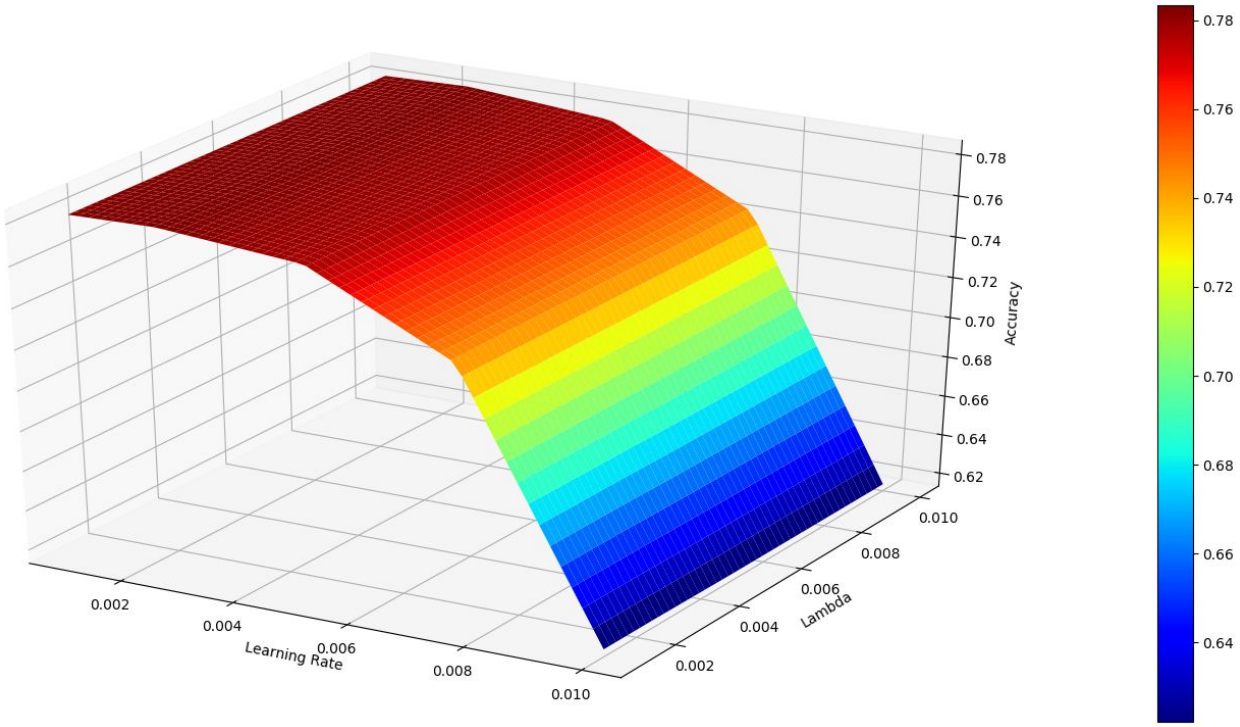
No feature selection Logistic Regression analysis

Our plots for logistic regression with no feature selection are much more interesting than the ones with feature selection, but if you look at the accuracies they are significantly lower. Some analysis on these plots will tell you that the highest accuracy is obtained at 100 iterations. By the time we get to 1000 iterations it looks like our accuracy is falling. All of the following have eta, lambda = [.001, .0025, .005, .0075, .01]

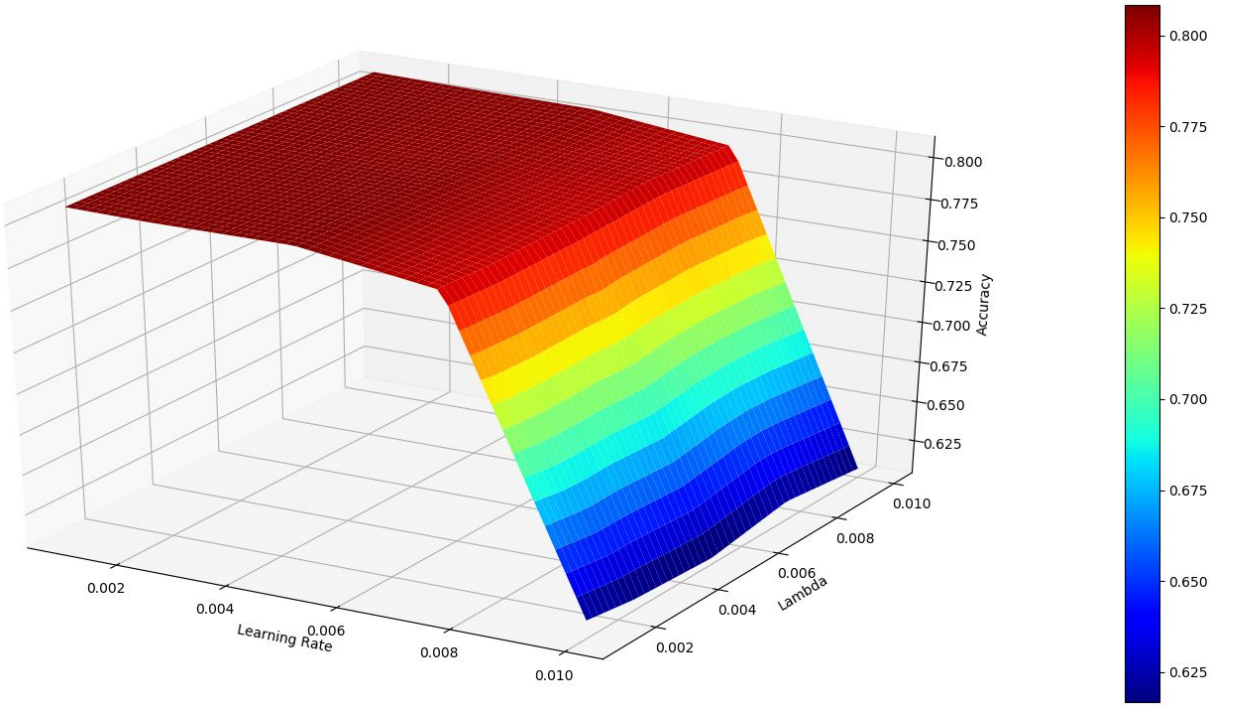
1 iteration Logistic Regression (no feature selection)



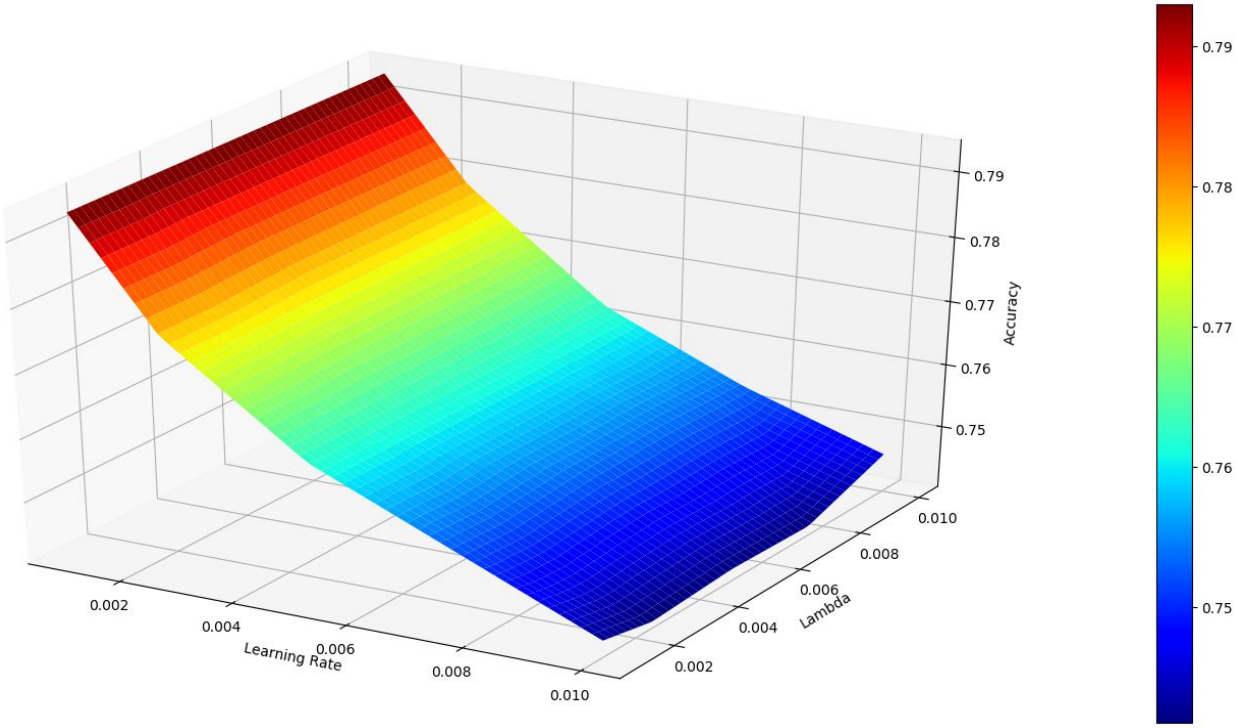
10 iteration Logistic Regression (no feature selection)



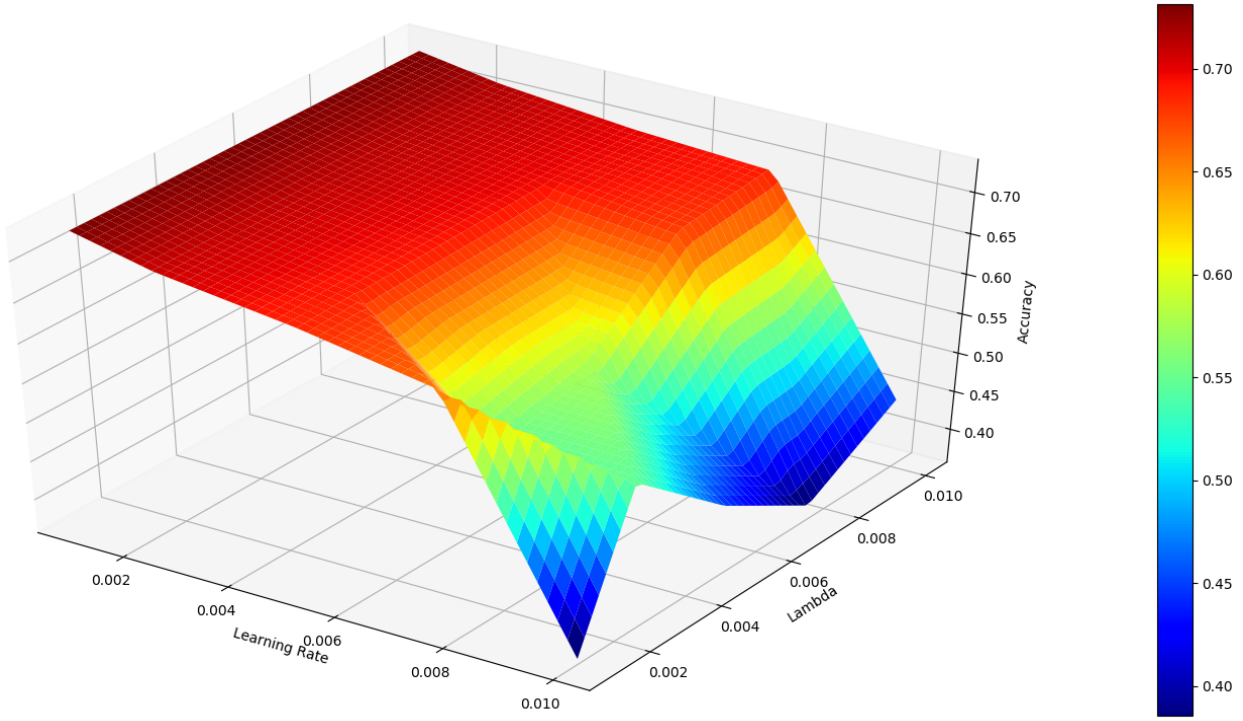
100 iteration Logistic Regression (no feature selection)



1000 iteration Logistic Regression (no feature selection)



10000 iteration Logistic Regression (no feature selection)

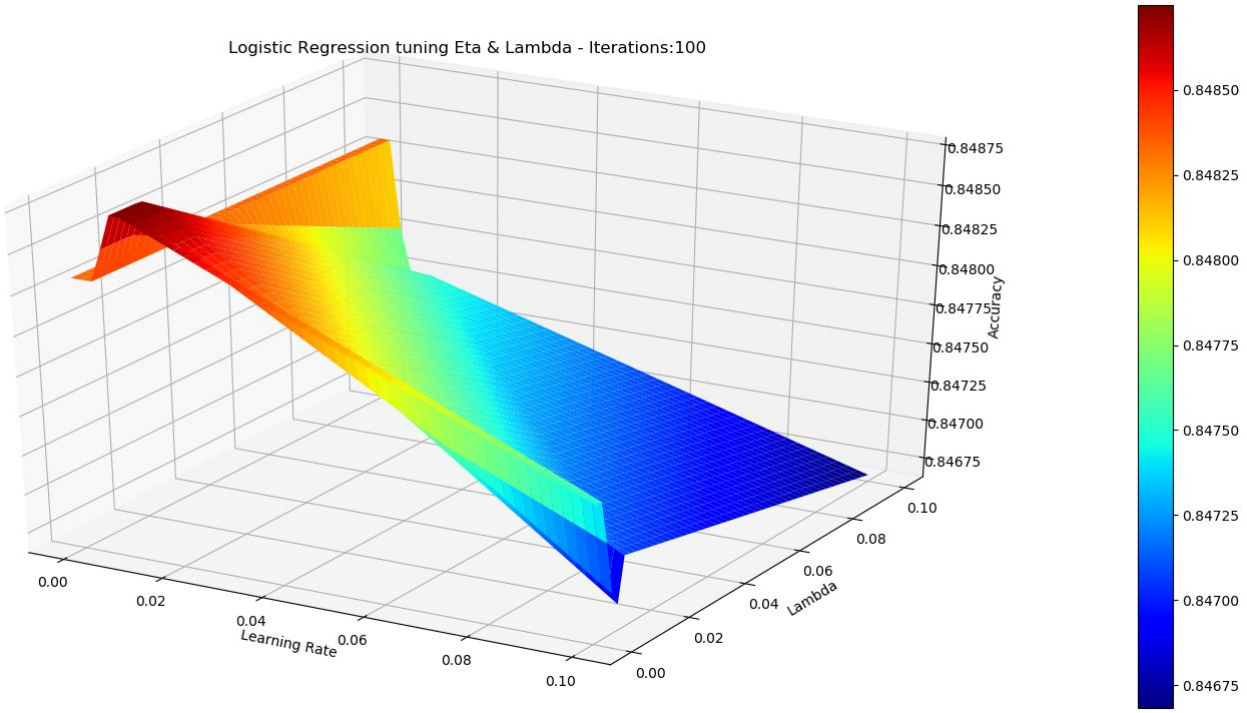


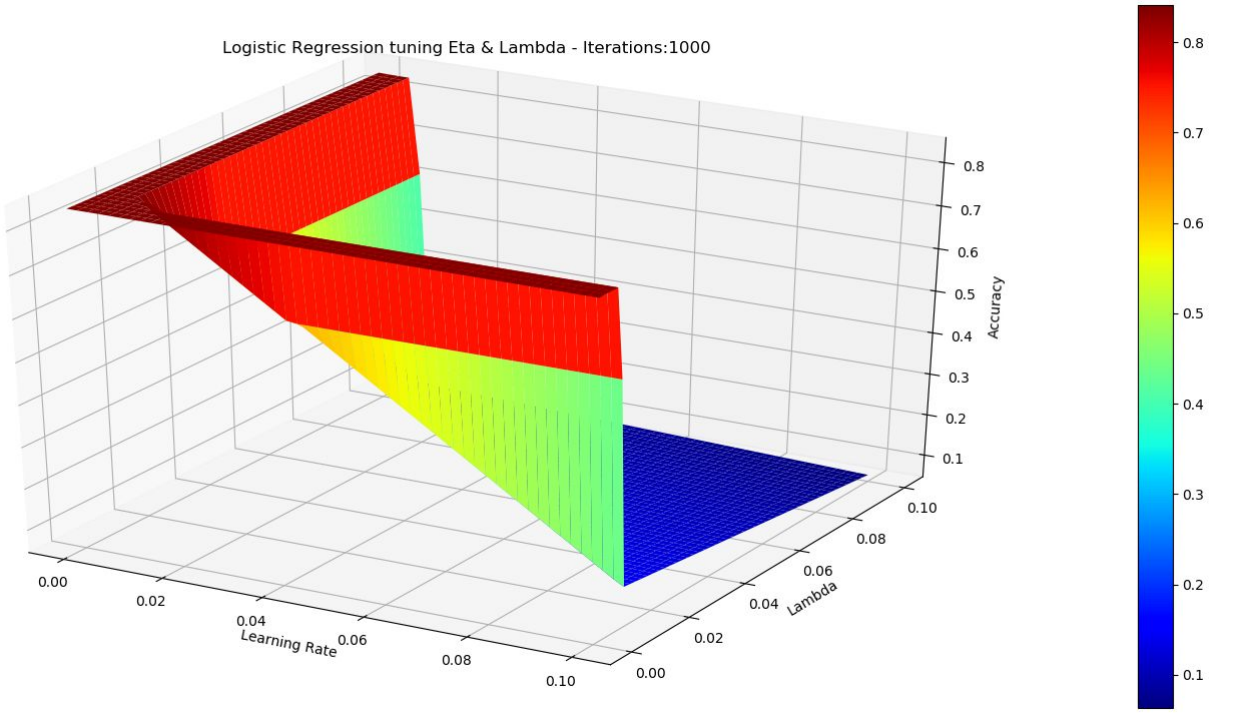
Non-mutual information feature selection

After building the likelihood matrix from Naive Bayes with Beta = 1/100, we used the likelihood probabilities to make a feature selection of the most important features. We took the maximum probability of each column, giving us the probabilities for each feature for when they impacted a class the most. We then took the top 60,000 of these probabilities, giving us 60,000 features and found better accuracy than when we don't use it. We later replaced this with mutual information right before finishing the project. Also, on these runs we were using the exact same set of validation data to keep our results consistent.

All of the following have eta, lambda = [.0001, .001, .0025, .005, .0075, .01, .1]

Additionally, note that running 1 iteration or iterations over 1000 aren't particularly interesting. 1 iteration will generally give a plane as we haven't used the hyperparameters. Over 1000 iterations and we are starting into the land of overfitting.


Logistic Regression tuning Eta & Lambda - Iterations:100
Learning Rate
Lambda
Accuracy
0.00
0.02
0.04
0.06
0.08
0.10
0.84675
0.84700
0.84725
0.84750
0.84775
0.84800
0.84825
0.84850
0.84875


Logistic Regression tuning Eta & Lambda - Iterations:1000
Learning Rate
Lambda
Accuracy
0.00
0.02
0.04
0.06
0.08
0.10
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8