For this assignment, I started off with the same code base I used from the last assignment. I did a little bit of exploration, but after running a base model (with no feature engineering) I found I was getting 100% train accuracy and around 90% test accuracy, so I figured that I would have to combat overfitting and feature engineering wouldn’t be too important. I tried PCA, but didn’t get that good results so I figured feature selection wouldn’t be that useful. I then tried some different feature engineering pipelines to find the best one. It ultimately made sqrt, log1p, and polynomial features and standard scaled them. Next I used imbalanced learn to see if that would help the accuracy since the default dataset wasn’t that balanced and it had some good results. Finally, I spent a lot of time on hyperparameter testing which turned out to be amazing. I did a whole bunch of grid and random searches and wrote down a list of observations of some of the best performing parameters which I used for a final grid search. I used google colab, ran a bunch of different grid searches and just ended up keeping one example. I ran it in the background and came back when it was done. The final grid search improved my test results by about 7-8% to make my accuracy around 99% so I called it there.