

READ ME:

Explore Enron Dataset – a jupyter file that contains my exploration of the Enron dataset, feature selection and feature scores using select k best, and playing around the decision tree classifier, naïve bayes, and k-nearest neighbors (prints out accuracy, precision, and recall time). This includes parameter tuning with GridSearchCV.

Enron Dataset Report – a pdf report of my answers to the questions from Udacity.

Poi_id.py - a python file where I used my final algorithm.

Tester.py – a python file for the tester.

4 pkl files – final_project_dataset(provided by Udacity), my_classifier.pkl, my_dataset.pkl, my_feature_list.pkl (which were extracted from tester.py file)