

## README HW1 CS 189

Name: Tyler Freund

SID: 3036171683

Reproduction of results instructions:

First, you need to run "featurize.py" in order to update the features used for the spam dataset.

Next, you must run the .ipynb file in this folder titled "HW#1 code - CS 189.ipynb" - as a jupyter lab/jupyter notebook then the results should be recreated. (Be sure to run each cell in order from top to bottom otherwise variables may be mis assigned and skew results)

Important Note:

In question 5 I used the approach of a grid search so there is a grid search cell for each of the datasets. These will print out the accuracy of the grid search NOT the accuracy of my predictions (Also I put the grid search input size down to 100 for efficiency of runtimes and since it should not impact my predictions). However, the actual predictions (variable name `y_test`) output and saved as csv files should be reflective of my scores found on kaggle.