Ömer Faruk Tuna

**README**

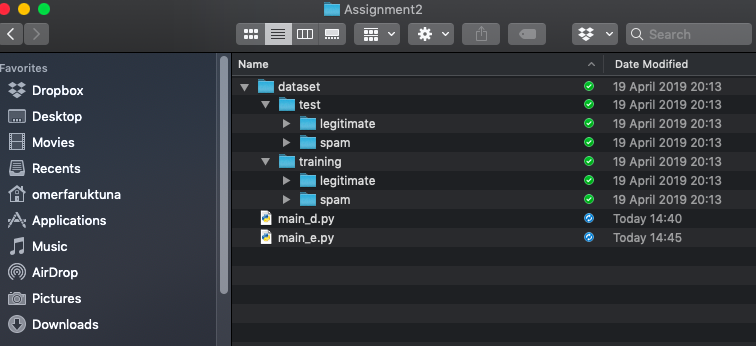
The code is written and tested using Python 3.7. Programs take a couple of minutes to run. Pleae wait untill it finishes and shows the output.

There are 2 python file:

main\_d.py

main\_e.py

The structure of the folder is as below below. That is: dataset folder should be in the same directory with main\_e.py and main\_d.py files.



For section “e” of Assignment 2, you should run main\_e.py. It uses Laplace smoothing with alpha equals 1 and calculates all the perforance values for Naive Bayes based spam mail classifier.

For section “d” of Assignment 2, you should run main\_d.py. It calculates the performance values without using smoothing.

The python files creates 3 files during the execution. Those files are written to the same directory.

legitimate\_features.txt

spam\_features.txt

common\_features.txt