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 files:

main\_rocchio.py

main\_knn3.py

Dataset folder should be in the same directory with main\_rocchio.py and main\_knn3.py files.

For Rocchio algorithm, you should run main\_rocchio.py .

For kNN algorithm with k = 3, you should run main\_knn3.py. If you want to change the value of k, you should edit the file and set the parameter “knn\_k” equals to 5 for example.

The python files creates 2 files during the execution. Those files are written to the same directory.

Top\_20\_legitimate\_features.txt

Top\_20\_spam\_features.txt