

To run the programs:

It contains 4 main scripts. Each script performs different tasks. The tasks of scripts are explained below:

baseline.py: Basic baseline data of class frequencies

DataInspection.py: Toxic class distribution

Models:

naivebayes.py: Naive bayes classifier

svc.py: Support vector classifier