Read Me

Note : I have pickled various intermediate results which can be used for faster execution and evaluation of code.

How to Run

- There are 4 code files for the following four combinations of classifier and features selection techniques:
  - TF-IDF and KNN
  - TF-IDF and Naïve Byes
  - MI and KNN
  - MI and Naïve

- Each code file is well documented with the function of specific piece of code is commented inside the code cell itself. All the code cell are properly numbered (##cell 0 ## ,##cell 1## etc).

- All pickled files are enclosed with the submission inside a folder. Please find the same.

- Cell 0 of each code file contains piece of code to fetch various pickled results. Please uncomment the code and give path to the pickled files needed.

- Last comment line of cell 0 shows the execution sequence of code cells inside the python file using the pickled values. For example ## Cell 0 ## → ##Cell 1 ## -→ ##Cell 5## →##Cell 7## .Here execution of Cell 2,3,4,6 are skipped because the produced results (which was highly time consuming) are already fetched in cell 0.