Umang Chaturvedi
17CS60R69

- All models is processed on processed_data.txt kept in "/data" folder.

- Run extract_data.py for extracting tar file in text file.
  Run preprocess_module.py for removing urls and language words other than english.

- Laplace smoothing - ngram_model_laplace.py
  Good Turing smoothing - ngram_model_goodturing.py
  Kneserney smoothing – ngram_model_kneserney

- Run rak_graph to get plot of freqency distribution of bigram and list of top 10 bigrams.

- Processed_data.txt is need to be filled before running codes. (kept empty due to storage constraint on moodle)