

Results Folder

- Classification.ipynb: Contains all the classification results for Q4.2
- EnsembleLearning.ipynb, Bi-LSTM Polarity.ipynb and Bi-LSTM Subjectivity.ipynb: Contains all the innovation results for Q4.3

CSV Folder ⇒ Original Folder

- tweets_data.csv: original crawled tweets
- reddit_data.csv: original crawled reddit posts

CSV Folder ⇒ Cleaned Folder

- final_labelled.csv: data with both polarity and subjectivity labels used for the training and testing of classification models
- non_labelled.csv: remaining data not labelled
- labelled_rater2.csv: 2 raters labelled data for inter-annotator agreement
- automatic_classification_labelled_9k.csv: data labelled using our best performing classification model
- labelled_pol.csv: data with only the polarity label derived from the final_labelled.csv
- labelled_sub.csv: data with only the subjectivity label derived from the final_labelled.csv

CSV Folder ⇒ Uncleaned_Data Folder

- unclean_labelled.csv: labelled data that requires preprocessing
- unclean_others.csv: non-labelled data that requires preprocessing