

INSTRUCTIONS FOR FILES ACCESS:

We have 2 Baseline models

1. Train-test split model
2. K-fold model (Used Stratified K-fold as well)

Baseline model code are in files

Baseline_TrainTestSplit.ipynb

Baseline_Kfold.ipynb

All improved models in below files

Improved_Model_Trigram.ipynb

Improved_NaiveBayes_SVM_RF.ipynb

Improved_NoGrams_Tfidf_CountVectorizer.ipynb

Improved_parameter_tuning.ipynb

Improved_Adding Lexical Features and K-Fold Cross Validation.ipynb

Improved_LSTM.ipynb

Best model is evaluated on Essays dataset and the file is under below name

Improved_TrainTestSplit_Evaluate.ipynb