

Instructions

Dependencies

- Python3
- Tensorflow 1.11.0
- Keras 2.2.4
- Jupyter Notebook
- Scikit Learn

How to use the folder

- There are two folders inside the main folder one for "Binary_Classification" and another for "Six_Class_Classification" in each of these folders there is a pre-processing file and classifier file run the pre-processing file by providing the right path for the dataset components. ("train.tsv" for train_news,"test.tsv" for test_news and "valid.tsv" path for valid_news) and generate preprocessed data ("preprocessed.npz" in the first case and "preprocessed6class.npz" in the second case)
- For opening the classifier files make sure to run them using jupyter notebook and run them line by line and obtain the results

The highest accuracy for both the binary classification and 6 class classification is given by Logistic Regression which can be obtained by running the logistic regression block in both the Jupyter Notebooks

Binary Classification – Maximum Accuracy was obtained through Logistic Regression

- Test Accuracy-60.8%
- Valid Accuracy-62.4%

Classifier	Test Accuracy in %	Valid Accuracy in %
Multinomial Naive bayes	60.7	61.4
Logistic Regression	60.8	62.1
Xgboost	59.3	58.1
Random Forest	60.8	61.4
SVM	56.4	52.0
Neural Network	56.6	52.3

Six Class Classification - Maximum Accuracy was obtained through Logistic Regression

- Accuracy-23.5%
- Valid Accuracy-24.1%

Classifier	Test Accuracy in %	Valid Accuracy in %
Multinomial Naive bayes	21.9	23.4
Logistic Regression	23.5	24.1
Random Forest	21.2	20.1
SVM	21.2	20.1
Neural Network	23.0	23.13

The code can also be found at my github repository - <https://github.com/themechanicalcoder/Fake-News-Detection>