

In order to run this code, firstly we need to add to PATH variable the location of “stanford-ner.jar” which is located at .\Stanford_NER\stanford-ner-2016-10-31. Now, the code uses the following python packages:

1. Pandas
2. Nltk
3. BeautifulSoup
4. Pickle
5. Sklearn from scikit-learn
6. Sumy
7. Numpy

Once the environment is ready, the main file to run is “run.py”. It accepts the following optional arguments:

- t – If provided, it will run the code for training the data. If not provided, code will test the data.
- d – If provided with date in YYYY-MM-DD format of the match day, it will scrap the data from web corresponding to that date. It is useful for both training on new data and providing new test data as well. If not provided, it will use the existing training and testing data.
- f – If provided with argument, the argument will act as k in k-fold cross validation(k=10 default). Used only while training the data.
- c – If provided with argument, the argument specify the type of classifier to train. 1= “Action- Not just Action” classifier, 2=Event classifier, 3= Combined Classifier(default). Used only while training the data.
- m – If provided with argument, the argument specify the model to be used for training “Action-Not just Action” classifier. Different argument value that can be used are : Logistic Regression(default), Multinomial, Gaussian, Bernoulli. Used only while training the data.
- n – If provided with argument, the argument specify the model to be used for training event classifier. Different argument value that can be used are : Logistic Regression, Multinomial, Gaussian, Bernoulli. Used only while training the data.
- s – If provided with argument, the argument value specifies the type of summarizer to be used. Different argument value that can be used are : LexRank(default), TextRank. Used only while testing the data.

Example to run the code to train the combined classifier with new match data of 04/15/2017 with Multinomial model for both the classifier and LexRank for text summarization and perform 2-fold cross validation :

```
python run.py -t -d 2017-04-15 -f 2 -c 3 -m Multinomial -n Multinomial -s LexRank
```