**INSTRUCTION SET**

**Code can be downloaded from :-**

[**https://github.com/debasisdwivedy/Sentiment-Analysis-From-Product-Review**](https://github.com/debasisdwivedy/Sentiment-Analysis-From-Product-Review)

**DATASET**

The larger dataset can be made available upon request. Current I have added “Professor Damir Cavar “ and “Zeeshan” with access to the list. The person needs to have a drop box account.

The project folder has two smaller dataset which one can use to run the experiments in their local machine. If you want to run the larger dataset , make sure you run it on Big Red 2 or any other super computer.

***Requirements***

* Python 3.5.2
* Cython (optional)
* NLTK
* Gensim
* Scikit-learn 0.19
* pandas 0.19
* matplotlib
* Numpy
* Scipy

***Procedure***

The program files can be imported to your favorite python editor( in my case it was pycharm) and can be run.

*Singleton Approach*

Each file can be run individually as a main method has been provided for that purpose. Files that can be run individually are mentioned below: -

* *FetchData.py* :- This takes a text file as input and converts it into CSV file for further processing. This file needs to be run first.
* *NaiveBayes.py* :- This file provides us with default naïve bayes and optimized naïve bayes results.
* *SVM.py* :- This file run the SVM classifier and return back the results.
* *BagOfWords.py* :- This file provides us with the result of random forest classifier combined with bag of words model.
* *Word2Vecor.py* :- This file applies the deep learning neural net provided by gensim to out dataset and provides us the result..
* *Doc2Vecor.py* :- This file applies the improved word2vector model where the context is preserved.
* *Create\_TIMBL.py* :- This file creates the input TIMBL file. The output is created in two folders named “*TIMBL without Subjectivity Lexicon*” and “*TIMBL with Subjectivity Lexicon*”. Each folder will create a test and a train file for out experiment. This file is then taken and run on KARST (IU’s supercomputer that has TIMBL installed). The command used to run it on KARST is :-

*>Timbl -f trainfile -t testfile*

*>* *default run*

*Combined Approach*

* *Predict.py* :- This file runs all the algorithm one after the other and displays the result. This file can be run to view the results of all at one go.