# NLP Assignment #2 Report

Submitted by:

Jigna Reshamwala [425008356]

This report contains details about NLP Assignment#2. In this project about sentiment analyzer, I have implemented term frequency based Naïve Bayes classifier, binarized Naïve Bayes and Perceptron to classify documents based on the sentiments ‘positive’ or ‘negative’. I could achieve following results.

***Naïve Bayes:***

Frequency Based(With Stop Words):

Average Accuracy:0.8165

Frequency Based(Without Stop Words):

Average Accuracy:0.8110

Binarized:

Average Accuracy:0.7385

***Perceptron***

#of Iterations -1:

Average Accuracy:0.6265

#of iterations -50

Average Accuracy:0.823

1. ***How to compile and run the code***

My code is tested for Python 2.7.12. It might give errors with Python 3.

To compile and run the code following are the two methods:

Using IDE

* Import the project in any Python IDE
* Edit Run Configurations and set script parameters for NaiveBayes.py to any of the following commands:

python NaiveBayes.py ../data/imdb1

python NaiveBayes.py -f ../data/imdb1

python NaiveBayes.py -b ../data/imdb1

* Run the script to get the output
* To Run Perceptron.py first set the script parameters to Perceptron.py ../data/imdb1/ 50

From Terminal

* Open Terminal
* Go the the folder NLP\_HW2/python
* Type in the following command:

python NaiveBayes.py ../data/imdb1

python NaiveBayes.py -f ../data/imdb1

python NaiveBayes.py -b ../data/imdb1

python Perceptron.py ../data/imdb1/ 50

***Testing another test set:***

You can use python NaiveBayes.py ../data/imdb1 ../data/test\_data to test the test data. Here I have used the same data from train data to test data. You can replace the files in test\_data. Similarly python Perceptron.py ../data/imdb1/ ../data/test\_data 50 can be used to test perceptron.

1. ***Results and Analysis***

Results of the following commands:

python NaiveBayes.py ../data/imdb1



python NaiveBayes.py -f ../data/imdb1



python NaiveBayes.py -b ../data/imdb1



python Perceptron.py ../data/imdb1/ 1 (run for 1 iteration)



python Perceptron.py ../data/imdb1/ 50 (run for 50 iteration)



We can observe that perceptron performance gets better with increase in number of iterations. This improvement can be seen only till 50 number of iterations, then the accuracy remains almost constant even with the increase in number of iterations.

1. ***Any known bugs, problems, or limitations of the program***

There are no known bugs.