5. RESULTS AND DISCUSSION

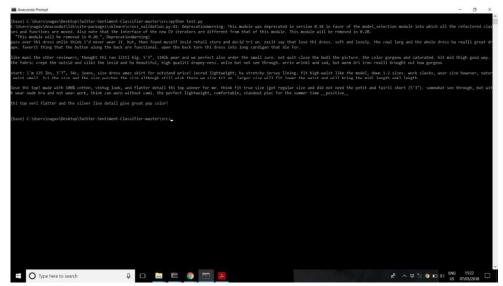


Figure 5.1 Stemming

The figure 5.1 shows the result regarding the stemming work. The training dataset will have reviews that includes many verb forms. For training the model the root word should be selected which can be done using porter stemmer.

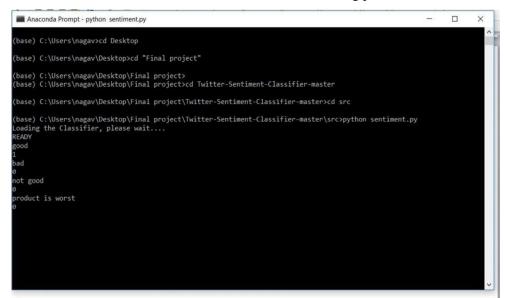


Figure 5.2 Output for a review

The figure 5.2 is regarding verification of the model accuracy by giving the review directly. This doesn't mean the test data but only by giving the words

that having sentiment the model will be able to give the results by classifying that sentiment into positive or negative correspondingly.

```
Training.txt - Notepad
File Edit Format View Help
        The Da Vinci Code book is just awesome.
        this was the first clive cussler i've ever read, but even books like Relic, and Da Vinci code were more plausible than this.
        i liked the Da Vinci Code a lot.
        i liked the Da Vinci Code a lot.
        I liked the Da Vinci Code but it ultimatly didn't seem to hold it's own.
        that's not even an exaggeration ) and at midnight we went to Wal-Mart to buy the Da Vinci Code, which is amazing of course.
        I loved the Da Vinci Code, but now I want something better and different!..
        i thought da vinci code was great, same with kite runner.
        The Da Vinci Code is actually a good movie...
        I thought the Da Vinci Code was a pretty good book.
        The Da Vinci Code is one of the most beautiful movies ive ever seen.
        The Da Vinci Code is an * amazing * book, do not get me wrong. then I turn on the light and the radio and enjoy my Da Vinci Code.
        The Da Vinci Code was REALLY good.
        i love da vinci code....
        i loved da vinci code..
```

Figure 5.3 Dataset used for training

The figure 5.3 shows the dataset used for training the classifier.

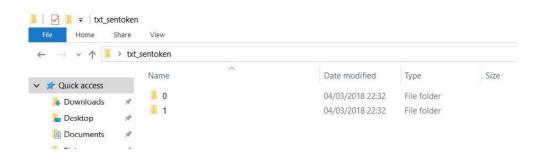


Figure 5.4 Folder containing class labels

The figure 5.4 shows the folder containing class labels as folders in which in turn contains the text documents which contains the reviews

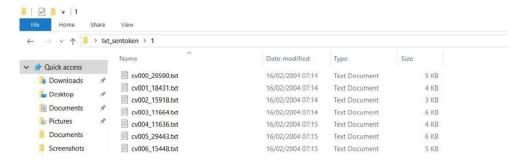


Figure 5.5 Dataset containing txt files which is used for training

The figure 5.5 is the dataset in which the text files are present in which the class label and corresponding reviews are present

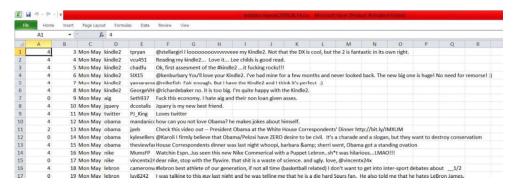


Figure 5.6 Dataset used for testing

The figure 5.6 shows the dataset which is used for testing, in which the reviews are present at the sixth column.

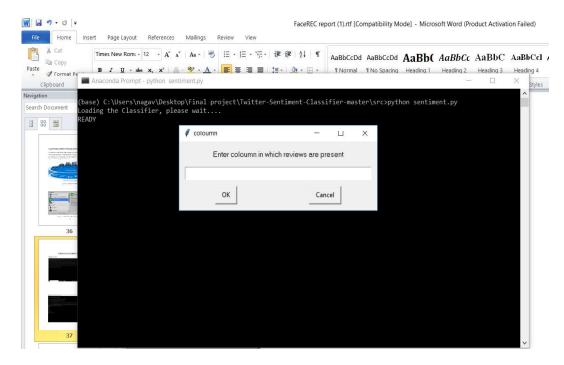


Figure 5.7 Dialog box for review selection in test data

This figure 5.7 is the first part of the actual test data result. When the test data i.e. the dataset for prediction is given to the model this will shows a dialogue box asking for in which column of the csv file the actual review regarding product is present and by entering the value (number in which the review is present) the final prediction will be given.

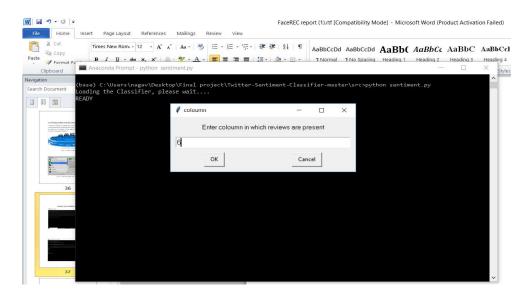


Figure 5.8 Enter the column number

As shown figure 5.8 the column number of the dataset should be given in which the reviews are present.

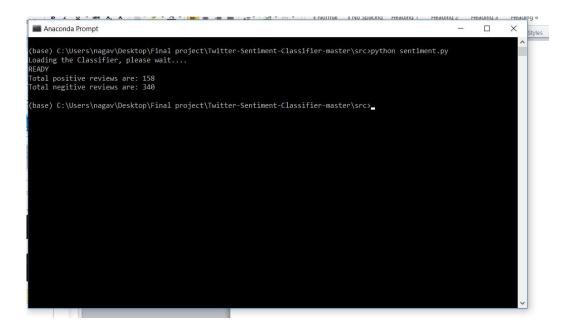


Figure 5.9 Predicting how many positives and negatives

The shown figure 5.9 is the result obtained when the dataset to be predicted is given as input to the classifier.

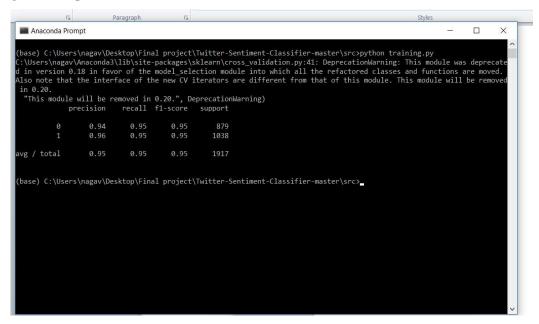


Figure 5.10 Accuracy of the model

The figure 5.10 shown is the accuracy that has been obtained from the model.