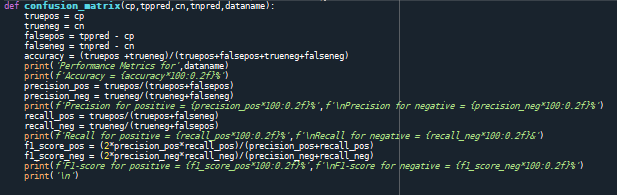
STEP 1:

The datasets used for this assignment are based on reviews from Rotten Tomatoes and Nokia. From the Blackboard website, it has been downloaded.

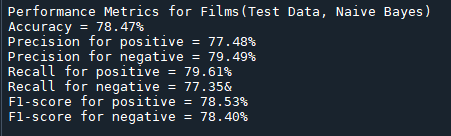
STEP 2:

In this process, files containing the dataset are read, and the dataset is divided into positive and negative sentences. These sentences are then segmented into words and stored in a sentiment dictionary.

Subsequently, the function is executed to train and test the Naive Bayes model on both the train and test datasets, which are created by partitioning the sentiment sentences. The classification results are obtained by calculating the scores of the confusion matrix using the following function:



**Fig1: Confusion Matrix Python Function**

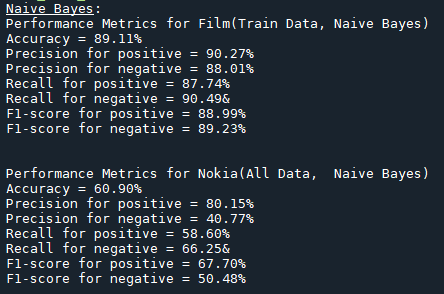
In this context, True Positive, False Positive, True Negative, and False Negative values are derived from the parameters, and various metrics such as Accuracy, Precision, Recall, and F1 Score are computed. The model underwent testing on the test data, yielding the following results:

**Fig 2: Results for Classification on test data**

Observing the results, the testing accuracy is approximately 78%. Notably, Precision, Recall, and F1-score values for both positive and negative sentence predictions exhibit similarity. This suggests that the model is demonstrating a balanced performance, achieving a satisfactory level of accuracy.

STEP 3:

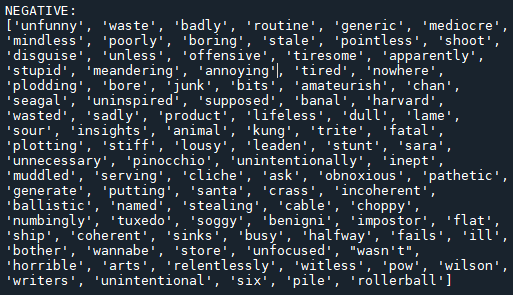
Here, the testBayes function has been invoked with the training dataset and data sourced from reviews of Nokia products. The outcomes are reported as follows:



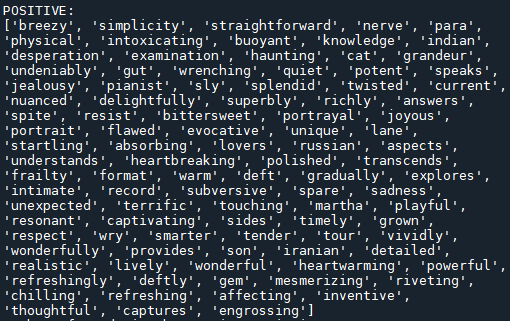
**Fig 3: Results for Classification on train data and Nokia Data**

The outcomes indicate that the Naive Bayes Classifier has proven more effective when applied to the training data than to the Nokia dataset. This is evident in the higher metric values observed for the former. Upon examining the context of both datasets, the divergence in results can be attributed to the Naive Bayes Classifier being trained on film-related data. Consequently, when tested on the same dataset, the classifier exhibited notably superior performance. In contrast, the Nokia dataset, characterised by reviews focused on Nokia products, generated relatively lower results. The distinct nature of the two datasets may explain the observed contrast in performance, with the film-related data displaying superior metric values compared to the Nokia dataset.

STEP 4:

The negative sentiment highly predictive words which are included in the sentiment dictionary are given below:

**Fig 4: Negative highly predictive words included in Sentiment Dictionary**

The positive sentiment highly predictive words which are included in the sentiment dictionary are given below:

**Fig 5: Positive highly predictive words included in Sentiment Dictionary**

Examining the list of positive and negative words provided, it is clear that neutral terms have been labeled as positive or negative and are in the Sentiment Dictionary. For instance, 'routine,' 'disguise,' and 'shoot,' although neutral, are listed as negative words. On the flip side, words like 'physical,' 'cat,' and 'answers,' which are neutral in meaning, find themselves in the positive words list. It is noteworthy that the term 'sadness,' typically associated with negative sentiment, is included in the positive sentiment terms list. Hence, we can conclude that the words selected by the model are not so much in good sentiment terms.

Counting the words in both lists reveals a total of 85 words that the model considers highly predictive and includes in the Sentiment Dictionary.