

BEFORE FILTERING THE STOP WORDS

Before filtering the stopwords, the accuracy that we got from the set is as shown below:

The accuracy of the TRAIN data is **0.9374**

The accuracy of the TEST data is **0.8309**

AFTER FILTERING THE STOP WORDS

After filtering the stopwords, the accuracy that we got from the set is as shown below:

The accuracy of the TRAIN data is **0.9503**

The accuracy of the TEST data is **0.9019**

We could see that the performance has been improved after filtering out the stop words. This is because removing stop words does not only reduce the vector space but also improve the performance by increasing the execution speed, calculations, and also the accuracy.

Training models takes time and if you have less tokens to be trained, the training time should decrease thereby it means if the stopwords are removed, training time should decrease by increasing the performance.

The accuracy results are written in results.txt everytime my code is run.

.....