On the previous presentation, the researchers show the sample on how they compute and get the accuracy result for the system. The researchers use 50 data set from there corpus and this sample data set undergo unto the process on how the researchers will going to make the system (e.g cleaning the data, annotation and etc.). The Bag of Words of this 50 data was containing 378 attribute and 3 classification. These said data was divide into two for testing and training (10 data was for testing and remaining 40 for training) As the result, the system got a accuracy of 10% of accuracy for the testing and 54% for the training this means the system already learning given this sentences.

To make the result more efficient, using the 100% percent corpus, the researchers manually undergo the process for the system like cleaning data, annotating and etc. After that the researcher change the type of method for the result unlike the methods that the researchers do on the previous sample data, the researchers uses the Cross Validation fold-10 in WEKA but still the purpose is the same to get the higher accuracy of the system as much as possible and as the Bag of Words got bigger the attribute contained 7062 and still 3 classification. 2000 thousand data was not inserted all in the Weka because the time was not enough. The researcher splices the data hundreds by hundreds to determine the changes and observe the data that inserted that can decrease or damage the accuracy of the system since the main goal of the proponents is to increase the accuracy to make the system efficient.

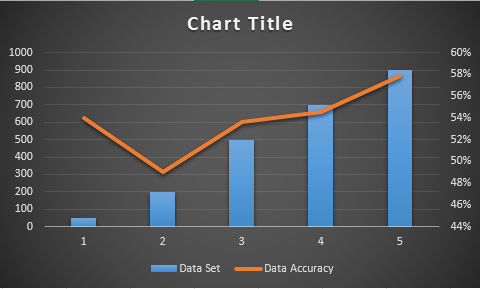


Figure 1 A graph that describe the more data inputting to Weka the possible of increasing its accuracy will high.

Using the formula describe on Theoretical Framework the researcher make some additional computation for the data observation and more result.

