The goal of final project is using the **'complaint1700'** and **'non-complaint1700'** file to train a robust classification model, which can be used afterwards to test if a new tweet is negative and help the company to separate the non-negative tweets.

I start by removing punctuation characters and stop words. And the trained three models and selected Native Bayes model which yields the highest predicting accuracy.

Lastly, I applied the model on the test dataset and resulting in accuracy around 77%.