Report on Naïve Bayes Classifier model

Result:

Number of training samples: 4459

Number of testing samples: 1115

Training Accuracy: 0.9712940121103386

Testing Accuracy: 0.9596412556053812

Training F1 Score: 0.8890814558058925

Testing F1 Score: 0.8442906574394464

Training Confusion Matrix (TP, FN, FP, TN): 513, 82, 46, 3818
Testing Confusion Matrix (TP, FN, FP, TN): 122, 30, 15, 948

Performance of the model: The model is working relatively good, did a good job in classifying spam and ham messages, with a high score of accuracy and F1 with both training and testing data sets. However, there seems to be a few misclassifications happens, which lead to F1 score and accuracy not close to each other. Overall, the model is performing well.

Challenges faced: At first the model produce just a little amount of True Negative and True Positive, hence the F1 and accuracy was off. After a little time figuring out, I was processing the data wrong, which lead to data misinterpreted, so It produce a bad result, after adjusting the data processing, the model worked out the way I want.

Possible improvements: the model is still classifying some ham as spam an vice versa, maybe that due to the way I originally process data and calculate the possibility of each word. Trying other way (than instructed in the description) would give a better result and accuracy and F1 would be closer.