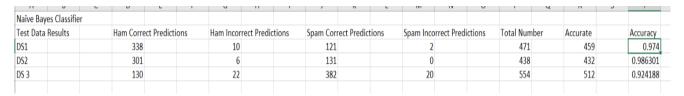
## Homework 2 Report

## Naïve Bayes Classifier

For the three given data sets(hw2\_train/test, enron1\_train/test, enron4\_train/test), an accuracy displayed in the following picture was obtained.



The program is a straightforward Python implementation of the Multinomial Naïve Bayes classifier from the given link in the HW description.

## Logistic Regression

## Perceptron Algorithm

For Perceptron algorithm, it was found that for approximately 750 iterations, the accuracy on the validation set is maximized. Results for all the 3 datasets are available in the Excel sheet in the Java Project folder and one can observe that the accuracy on the validation set tends to maximize around 750 iterations.

Perceptro	n:	
DS 1		
For Valida	tion Set:	
No Of Iterations		Accuracy
1		0.79
50		0.84
100		0.8
250		0.80
500		0.80
750		0.8
1000		0.8
DS2		
For Valida	tion Set:	
No Of Iterations		Accuracy
1		0.822
50		0.81481
100		0.83
250		0.844
500		0.86666
750		0.8444
1000		0.82962

DS3		
For Validation Se	t	
No Of Iterations	Accuracy	
1	0.7826	
50	0.9192	
100	0.8944	
250	0.92	
500	0.88	
750	0.937	
1000	0.826	
For Test Set		
DS1		
No Of Iterations	Accuracy of Ham	Accuracy of Spam
750	0.91954023	0.9
DS2		
750	0.912052117	0.88590604
DS3		
750	0.881578947	0.938618926

Validation was performed by training the Perceptron only on Training data and testing was performed by training the perceptron on both training and validation data.