Assignment 1

Name: Apurv Upasani **USC-ID:** 4839310230

1) What are the precision, recall and F-score on the development data for your classifier in part I for each of the two datasets. Report precision, recall and F-score for each label.

Soln: Output for Classifier in Part –I

Spam Filter

Label	Precision	Recall	F-Score
HAM	0.9889	0.984	0.9864
SPAM	0.9565	0.9696	0.963

Sentiment Analysis

Label	Precision	Recall	F-Score
POS	0.856	0.781	0.8148
NEG	0.799	0.864	0.8295

2) What are the precision, recall and F-score for your classifier in part II for each of the two datasets. Report precision, recall and F-score for each label.

Soln: Output for Classifier in Part –II

Spam Filter

Label	Precision	Recall	F-Score
HAM	0.9477	0.98	0.9637
SPAM	0.9392	0.8512	0.893

Sentiment Analysis

Label	Precision	Recall	F-Score
POS	0.745	0.774	0.7599
NEG	0.7648	0.735	0.7499

3) What happens exactly to precision, recall and F-score in each of the two tasks (on the development data) when only 10% of the training data is used to train the classifiers in part I and part II? Why do you think that is?

Soln: Output for Classifier –I with 10% training data

Label	Precision	Recall	F-Score
HAM	0.982	0.98	0.9809
SPAM	0.945	0.95	0.9478

Label	Precision	Recall	F-Score
POS	0.8558	0.801	0.8274
NEG	0.8129	0.865	0.8381

Output for Classifier – II with 10% training data

Label	Precision	Recall	F-Score
HAM	0.9267	0.98	0.9526
SPAM	0.93	0.7851	0.8514

Label	Precision	Recall	F-Score
POS	0.7071	0.526	0.632
NEG	0.6229	0.783	0.6938

The Precision, Recall and F-Score decreases when we consider only 10% of the training data. We can observe that the fall in Spam dataset is significantly less than Sentiment dataset. This is mainly due to the distribution of the training data in each case. Also, the content of Spam dataset has specific features for a particular label. Hence, even with 10% data, it is possible to separate the two classes. In case of sentiment analysis, many features overlap thereby making it more difficult to classify the labels with 10% training data and hence we see a sharp decrease in precision and F-Score.