Applied Artificial Intelligence

COMP 6721

Project Assignment 2 Report

Submitted By:
Group – FL-G04
Gursimran Singh - 40080981
Aravind Ashoka Reddy - 40103248
Darshan Dhananjay - 40079241

Submitted To: Dr. René Witte

Analysis

(a) How we generated this table using our program:

We calculated the FP (False positive) and FN (False Negative), TP (True Positive) and TN (True Negative) values for each class namely: **spam (Positive Class)** and **ham (Negative Class)**. Then using the formulas in the **Table 1**, we calculated the metrics for each class. For **Table 2**, we used the values that we calculated to generated the Confusion Matrix for 800 test files in which 400 were labelled as spam and the rest were ham.

(b) Discussion of the results

For individual classes- Spam and Ham:

As per the results, for Ham class it was 98.75% accurate and for Spam class, the accuracy was 83%. For both the classes, precision was 100% as expected which means all of the predicted positives were correctly classified. The recall was 83% for Spam class which means that 17% of the spam emails find a way into the inbox and are not filtered whereas

For the Model:

The model has 90.87% accuracy and the precision is 98.51% which means that out of 100 emails classified as spam, the classifier incorrectly classifies around only 2 ham emails as spam emails which looks good. However, the recall for the model is 83% which means that 17% emails still find a way to the inbox, which is not a good score for a spam detector. The F1 score is 90.09 for the classifier which is a harmonic mean of precision and recall.

Table 1 *Results for each class- Spam and Ham*

	Ham (%)	Spam (%)	MODEL (%)	Formula
Accuracy	**98.75**	**83**	**90.87**	**(TP+TN) / (TP+FP+TN+FN)**
Precision	**100**	**100**	**98.51**	**TP / (TP+FP)**
Recall	**98.75**	**83**	**83**	**TP / (TP+FN)**
F1-measure	**99.37**	**90.71**	**90.09**	**2PR / (P+R)**

Table 2 *Confusion Matrix*

		Predicted		
	N = 800	**SPAM**	**HAM**	TOTAL
Actual	**SPAM**	**TP = 332**	**FN = 68**	400
	HAM	**FP = 5**	**TN = 395**	400
	TOTAL	**337**	**463**	**800**

References

[1] "Russell, S., & Norvig, P. (2002). Artificial intelligence: a modern approach.,".

[2] R. Witte, "Lecture Slides Naive Bayes Classiifier," Concordia University, January 2020. [Online].

[3] "Evaluating a Spam filter Classifier," 2020. [Online]. Available: https://freecontent.manning.com/evaluating-a-classification-model-with-a-spam-filter/.