

Final exam - Email Classification

Sofia Huang
due 12/20/2022

1. Create two datasets, Testing and Training

Create two datasets, Training and Testing. You may choose a topic to classify your emails on (but choose only 1 topic). This can be spam, shopping emails, school emails, etc. Make sure that these are plain-text documents and that they do not include HTML tags. The documents in the Testing set should be different from the documents in the Training set. Upload your datasets to your GitHub repo. Please do not include emails that contain sensitive information.

A: What topic did you decide to classify on?

I chose to classify on ads vs non-ad emails. A lot of my emails that I immediately delete are subscription-based shopping advertisements so I thought it would be a good topic to classify on versus emails that contain important information or need a response.

For the results I have selected emails that are not advertisements to be relevant and advertisement emails to be non-relevant. Ideally, the classifier would filter my emails to show only the relevant, or non-advertisement emails.

2. Naive Bayes classifier

Use the example code in the class Colab notebook to train and test the Naive Bayes classifier. Use your Training dataset to train the Naive Bayes classifier. Use your Testing dataset to test the Naive Bayes classifier. Create a table to report the classification results for each email message in the Testing dataset. The table should include what the classifier reported (relevant or non-relevant) and the actual classification.

A. For those emails that the classifier got wrong, what factors might have caused the classifier to be incorrect? You will need to look at the text of the email to determine this.

The classifier classified two emails incorrectly. Both were relevant (not ads) emails, classified as non-relevant (ads). The first email that was classified incorrectly did not have a salutation or a closing line like many of the relevant emails from the training set which could have caused the classifier to think it was non-relevant. The second email was a club meeting recap that included future events and encouraged members to attend them. It contained phrases like "click this link" and "please sign up" which might have been a reason for the classifier to mark it as an advertisement.

Table 1: Classification Results

Email	Classifier Output	Actual Classification
1	Non-relevant	Non-relevant
2	Non-relevant	Non-relevant
3	Non-relevant	Non-relevant
4	Non-relevant	Non-relevant
5	Non-relevant	Non-relevant
6	Relevant	Relevant
7	Non-relevant	Relevant
8	Non-relevant	Relevant
9	Relevant	Relevant
10	Relevant	Relevant

3. Confusion Matrix

Draw a confusion matrix for your classification results (see Module 13, slides 42). This should be a table in Markdown or LaTeX, NOT a screenshot of output or image generated by another program. There's an example of a LaTeX confusion matrix in the Overleaf report template.

For the confusion matrix, I categorized relevant (not ads) as positive and non-relevant (ads) as negative.

Table 2: Confusion Matrix

		Actual	
		Relevant	Non-relevant
Predicted	Relevant	3 (TP)	0 (FP)
	Non-relevant	2 (FN)	5 (TN)

A. Based on the results in the confusion matrix, how well did the classifier perform?

The classifier performed relatively well as it did not have any false positives, only 2 false negatives. Larger training and testing sets would be able to show more accurate results for the classifier to assess its performance.

B. Would you prefer an email classifier to have more false positives or more false negatives? Why?

I would rather have an email classifier have more false positives (classifying a non-relevant email as relevant) so that there is less chance of missing a relevant email if the inbox was filtered by the classifier.

Extra Credit

Report the precision and recall scores of your classification results. Include the formulas you used to compute these values.

$$Precision = \frac{TP}{TP + FP} = \frac{3}{3 + 0} = \frac{3}{3} = 1$$

$$Recall = \frac{TP}{TP + FN} = \frac{3}{3 + 2} = \frac{3}{5} = 0.6$$

References

- **Confusion Matrix** <https://towardsdatascience.com/understanding-confusion-matrix-a9ad42dcfd62>
- **Latex Fraction** <https://tex.stackexchange.com/questions/106700/text-mode-in-fractions>