

Name:
Caroline Otten

Project:
CA 02 - Spam eMail Detection using Naive Bayes Classification Algorithm

Link to GitHub:
<https://github.com/cotten1/CA02>

Link to Google Drive Folder:
<https://drive.google.com/drive/folders/1QsyAMHl6zpQheZLNfIA0y-CIs6w4fUZH?usp=sharing>

Results:

```
Creating dictionary from input
->extracting each word from emails and adding them to a dictionary
->removing items from dictionary that include non alphabetical characters
    or contains more than one word
-> returning dictionary containing the 3000 most common words in emails
    reading and processing emails from TRAIN and TEST folders
    extract email features and put in matrix
->iterating over each email to create matrices containing the words and
    word counts
    extract email features and put in matrix
->iterating over each email to create matrices containing the words and
    word counts
Training Model using Gaussian Naive Bayes algorithm .....
Training completed
testing trained model to predict Test Data labels
Completed classification of the Test Data .... now printing Accuracy Score
by comparing the Predicted Labels with the Test Labels:
0.9615384615384616
```