Detecting Spam Emails

Sample Emails:

'Hey man can we get together to watch the basketball game tomorrow?',

'Up to 20% discount on parking, exclusive offer just for you. Don't miss this reward!',

'Free offer available now! Do not miss this opportunity to win!',

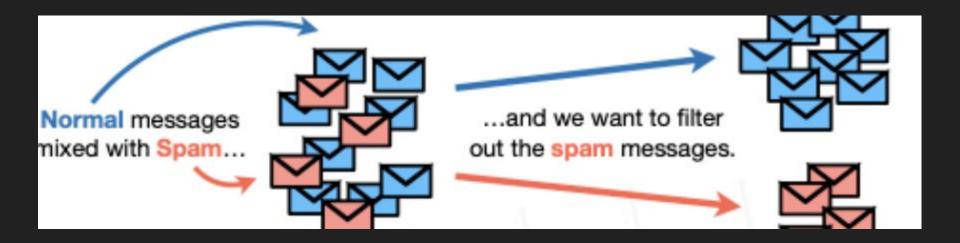'Can we go to the gym to workout?'

CountVectorizer

Counts the number of words in the sentences while checking for duplicates. It pays attention to sales type of words such as free, buy, reward, etc.

free	miss	we	to
0	0	1	1
0	1	0	0
1	1	0	0
0	0	1	1

Naive Bayes Classifier

- Text Classification
- Helps with spam filtering
- Fast Run-Time

Normal messages mixed with Spam... ...and we want to filter out the spam messages.

Check out the Code Here