## Assignment 1.2: Naïve Bayes Classifier

1. **Brief description of model**

Naive Bayes uses the Bayes’ Theorem and assumes that all predictors are independent. In other words, this classifier assumes that the presence of one particular feature in a class doesn’t affect the presence of another one.   
In our case of spam filtering we want to classify unseen samples as spam or non-spam. For this we use the equation given below:

**w1, w2,… wn  = all the words present in the email to be classified.**

Emails are classified as spam if:

And as non-spam otherwise.

One case that needs to be taken care of is when there is a word in the email to be tested, which is not present in the training data. We can do two things:

1. Ignore the word. But this would mean we are considering P(w|spam) and P(w|non-spam) both as 1 which would be illogical.
2. Take probability as 0 as it is not occurring in either dataset But this would mean we are considering P(w|spam) and P(w|non-spam) both as 0 which is also incorrect.

This is why we use Laplace smoothing to take care of the zero probability in Naïve Bayes. The formula is given as:

Here is the smoothing parameter(taken as 1).

1. **Accuracy of your model over each fold and the overall average accuracy.**

**FOLD 1**

Percentage of emails correctly classified: 64.78

**FOLD 2**

Percentage of emails correctly classified: 89.43

**FOLD 3**

Percentage of emails correctly classified: 66.19

**FOLD 4**

Percentage of emails correctly classified: 57.04

**FOLD 5**

Percentage of emails correctly classified: 71.12

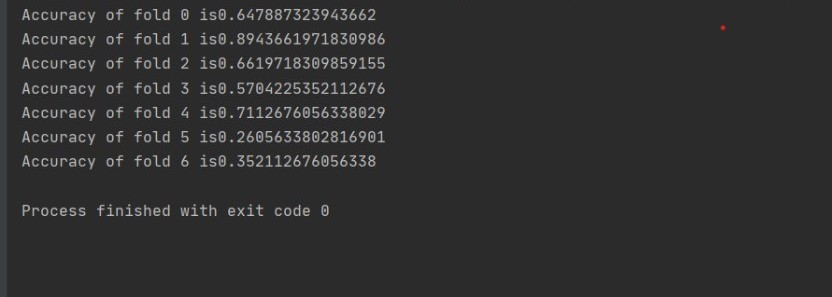
**FOLD 6**

Percentage of emails correctly classified: 26.05

**FOLD 7**

Percentage of emails correctly classified: 35.21

Overall average accuracy: 58.54



1. **Major limitations of the Naive Bayes classifier**

Naïve Bayes assumes that the conditional probabilities are independent of one another and do not affect one another: this is known as naïve conditional independence assumption. To put it another way, it means that the predictor features are all independent of one another. This, however, does not occur in real-life circumstances. This limits the algorithm's applicability in real-world scenarios.

Another drawback is that Naive Bayes model does not give any importance to the order of the words. Therefore, Naive Bayes model ignores the grammar rules of a model