ML Assignment 1: Naïve-Bayes

Anirudh Agrawal	Aviral Agarwal	Vikramjeet Das
2018A7PS0099H	2018A7PS0192H	2018A7PS0280H

1 Naïve-Bayes

Naive Bayes is a probabilistic classifier that makes classifications using the A Posteriori decision rule in Bayesian setting.
According to Bayes' Rule

$$P(S|V) = \frac{P(V|S)P(S)}{P(V|S)P(S) + P(V|S')P(S')}$$

For the given spam filtering problem-

$$P(S|w_1...w_n) = \frac{P(S)\prod_{i=1}^{n} P(w_i|S)}{P(S)\prod_{i=1}^{n} P(w_i|S) + P(H)\prod_{i=1}^{n} P(w_i|H)}$$

2 Implementation

The Algorithm can be roughly outlined by the following steps-

1. Find all the unique terms present in the emails by tokenizing.
2. Remove the stop words to obtain better accuracy.
3. For each term w_i present in an email find the spamsy value of that word by using the formula-

$$P(w_i|S) = \frac{count_{w_i}|spam}{count_{w_i}}$$

4. For each term w_i present in an email find the hamsy value of that word by using the formula-

$$P(w_i|H) = \frac{count_{w_i}|ham}{count_{wi}}$$

These are the parameters required for building the model

5. To predict the class of an unknown email, tokenize and remove stop words.
6. If

$$P(S|w_1...w_n) > P(H|w_1...w_n)$$

then the email is classified as spam, otherwise ham

3 Results

Cross Validation Method was used to test the algorithm wherein the dataset was divided into 7 equal parts, one of which was used to test the model while others were used to train the model in each iteration. The accuracy in individual folds ranged from 0.74 to 0.86 with the mean accuracy approximately being 0.80

4 Limitations

- The Naïve Bayes algorithm assumes conditional independence of the words i.e the words occurring in the sentence do not depend on each other nor their ordering matters. In real-life scenarios this may not always be the case.
- Naïve Bayes will have a difficult time making predictions for observations it has never seen before. For example, in our example, it cannot make sense of a new word it has never seen in the dataset.