

You need not print out all $P_{MLE}(w_{kj}|j)$ or $P_{BE}(w_{kj}|j)$, just take a look at the values of these two types of estimators. What do you observe? Discuss what you have observed in your report.

There is a significant difference between the values of the two types of estimators: (MLE-Maximum Likelihood Estimators, BE- Bayesian Laplace estimators)

1. MLE matrix contains a lot of zeroes which correspond to the cases where the required data was not found.

BE matrix on the other hand had the minimum prior of $1/|\text{vocab}|$ in the place where there are 0s in the MLE matrix.

2. MLE matrix values are observed to be generally greater than the BE counterparts.

3. All the values in both the matrices are very low. Probably more data is needed.

Now you will evaluate your classifiers on the testing data set. First, repeat the experiments described in Section 2.2.1 on testing dataset. Compare the results obtained with the results you have obtained in Section 2.2.1. What do you observe? Discuss.

We can easily see that the performance on training data is much better than the performance on the testing data, with overall accuracies observed to be 94% and 79% for Bayesian estimators respectively. On the level of individual classes, the prediction accuracy in testing is consistently lower than the training accuracy in training.

Training with BE and testing with BE respectively:

```

Printing the predictions for all the documents
The total accuracy is
94.25841593
Class 1 accuracy is 96.666667 percent
Class 2 accuracy is 91.738382 percent
Class 3 accuracy is 88.636364 percent
Class 4 accuracy is 93.015332 percent
Class 5 accuracy is 93.739130 percent
Class 6 accuracy is 94.763514 percent
Class 7 accuracy is 78.006873 percent
Class 8 accuracy is 96.452703 percent
Class 9 accuracy is 96.140940 percent
Class 10 accuracy is 96.969697 percent
Class 11 accuracy is 97.658863 percent
Class 12 accuracy is 97.643098 percent
Class 13 accuracy is 92.554992 percent
Class 14 accuracy is 97.306397 percent
Class 15 accuracy is 97.639123 percent
Class 16 accuracy is 98.163606 percent
Class 17 accuracy is 98.532110 percent
Class 18 accuracy is 96.808511 percent
Class 19 accuracy is 96.767241 percent
Class 20 accuracy is 64.893617 percent
-----
Printing the predictions for all the documents
The total accuracy is
79.9173773987
Class 1 accuracy is 67.610063 percent
Class 2 accuracy is 78.406170 percent
Class 3 accuracy is 64.450128 percent
Class 4 accuracy is 77.806122 percent
Class 5 accuracy is 74.934726 percent
Class 6 accuracy is 74.615385 percent
Class 7 accuracy is 75.916230 percent
Class 8 accuracy is 86.582278 percent
Class 9 accuracy is 94.458438 percent
Class 10 accuracy is 90.931990 percent
Class 11 accuracy is 96.992481 percent
Class 12 accuracy is 88.860759 percent
Class 13 accuracy is 65.903308 percent
Class 14 accuracy is 83.206107 percent
Class 15 accuracy is 90.051020 percent
Class 16 accuracy is 85.175879 percent
Class 17 accuracy is 87.087912 percent
Class 18 accuracy is 88.031915 percent
Class 19 accuracy is 56.451613 percent
Class 20 accuracy is 53.784861 percent

```

Next repeat the experiments (on testing data) using your Maximum Likelihood estimator $P_{MLE}(w_{kj}!)$ instead of the Bayesian estimators. Compare your results to the results obtained using your Bayesian estimators. Can you observe the difference? Which one is better?

The results are a lot worse than the ones using Bayesian estimators, with the class accuracy for some classes going down to almost zero.

The testing data has been found to contain a lot of outliers from the train set, also it is unsorted unlike the training data, so prediction will take some time. Because it has a lot of outliers, MLE fails and has a much lesser accuracy than BE.

On the other hand, BE introduces a uniform prior on each class, which in this case is an informed and well made prior, thus it is able to overcome the difficulties faced by MLE and still give a comparable prediction.