

Assignment 3: Naïve Bayes algorithm for text classification

In this assignment, we will redo the task of classifying documents (assignment 2) using the same Reuter dataset. But this time, you should implement the multinomial naive Bayes algorithm instead of KNN.

Naive Bayes used to be the de-facto method for text classification.

What's the accuracy of your learner? Try various smoothing parameters for the Naive Bayes learner. Which parameters work best and why?

You have to prepare the results (i.e., accuracy for different smoothing parameters) in tabular (in excel/paper) format and bring it to the sessional class.