

Assignment 2: Naive Bayes Text classification

Building a text classifier in python for movie review corpus in 2000 NLTK movie review corpus, you should write a 600(+/- 50) words of short report to summarize:

- 1: Your proposed text classification method (you can use relevant NLTK text classification functions or other text classification functions)
- 2: Pre-processing methods such as removing stop words, punctuations.
- 3: Feature selection methods such as selection the 3000 most important words.
- 4: You should evaluate your model on 10 percent of movie review model, please report the precision, recall, accuracy, and F-score of each class.

My evaluation will based on the whether you can clearly demonstrate your method in text classification, pre-processing, feature selection, and perform evaluation.

Please using the same dataset as we used in class 3: NLTK and sentiment analysis

You should submit both code and report