The code gives output of two algorithms Naïve Bayes and Logistic Regression with and without stop words. The input required for the code is the path to train and test data set, the number of iterations for finding the right weights, the lambda value, and the eta value.

Naïve Bayes: The accuracy of Naïve Bayes without removing the stop words is 92.19% and after removing the stop words the accuracy is 92.31% which is nearly similar values. It shows that Naïve Bayes output does not change so much with removing the stop words. We are removing the numbers, symbols and all other characters that are not alphabet for the first part and then we remove the stop words.

Naïve Bayes Accuracy after removing stop words: 0.9231868643222761

Naïve Bayes Accuracy without removing stop words: 0.9219214600635702

Logistic Regression: The logistic regression algorithm is executed on 500 iteration with eta value = 0.01 and three lambda values 0.1, 0.01, 0.001.

Without Removing the Stop Words

Logistic Regression with lambda = 0.1: 0.9037656903765691

Logistic Regression with lambda = 0.01: 0.8744769874476988

Logistic Regression with lambda = 0.001: 0.8263598326359832

After Removing the Stop Words

Logistic Regression with lambda = 0.1: 0.8514644351464435

Logistic Regression with lambda = 0.01: 0.8640167364016736

Logistic Regression with lambda = 0.001: 0.8640167364016736