

Description: In these projects we will apply what we've learned in class to real-world datasets for analysis.

Tasks:

1. Using the **spam.csv** file provided in the datasets zip file (or the **emails.csv**), perform an NLP analysis to create a classification model of the data. Use appropriate test/train splits or other methods of validation (such as k-fold cross validation or LOOCV). Create appropriate metrics and a confusion matrix.

Submission: Write up a work document describing your code and the results obtained. Include with your submission your R code file.