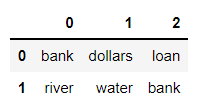
# PROGRAMMING PROJECT 4

TASK1:



The answer for the artificial data set, it contains top 3 values because while concluding for the smaller data set we noticed only three significant values for each topic

A picture containing graphical user interface

Description automatically generated

These are top 5 values for each topic for the newsletter data set. Few of the topics above make sense, like bmw lights change service oil, can be considered a part of 1 topic, but few topics have words that don’t have any relationship.

Task2:

Text

Description automatically generated

Chart, line chart

Description automatically generated

From the graphs topic representation graph has more error rate and bag of words has a lesser error rate.