**Natural Language Processing (Spring 2017)**

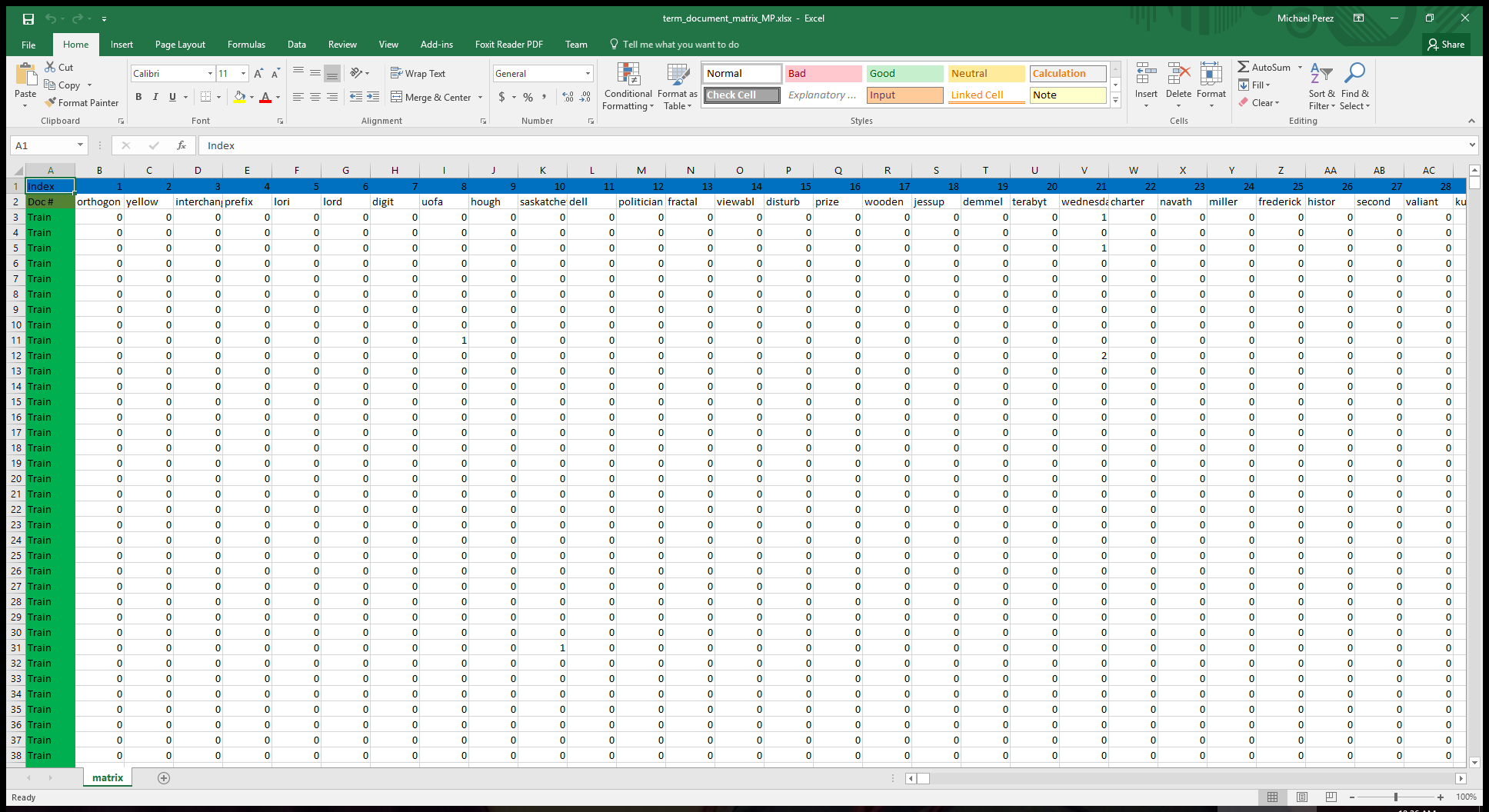
**Homework # 2**

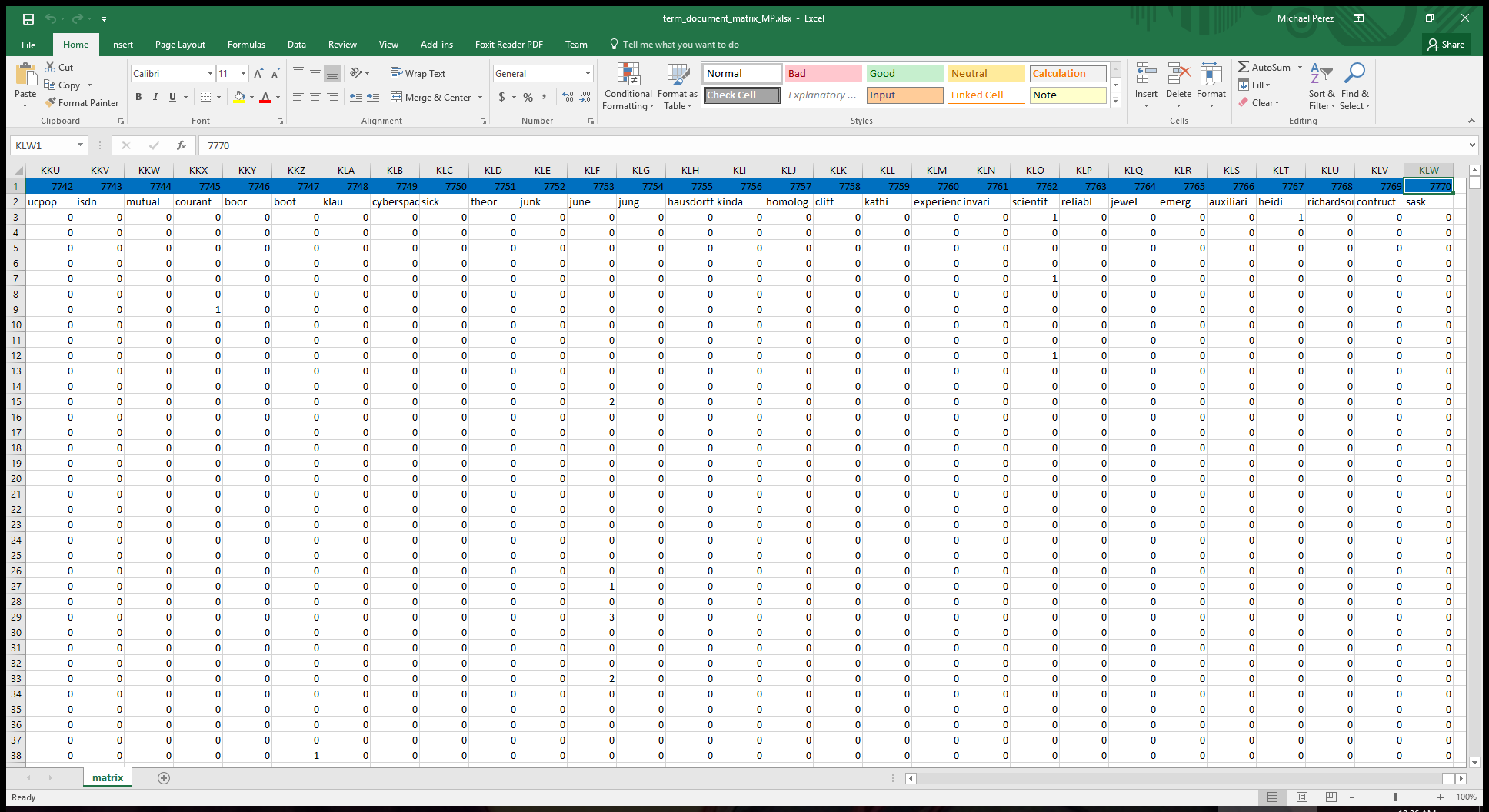
**DUE: 2/26/2017 11:59pm EST Total: 10 points**

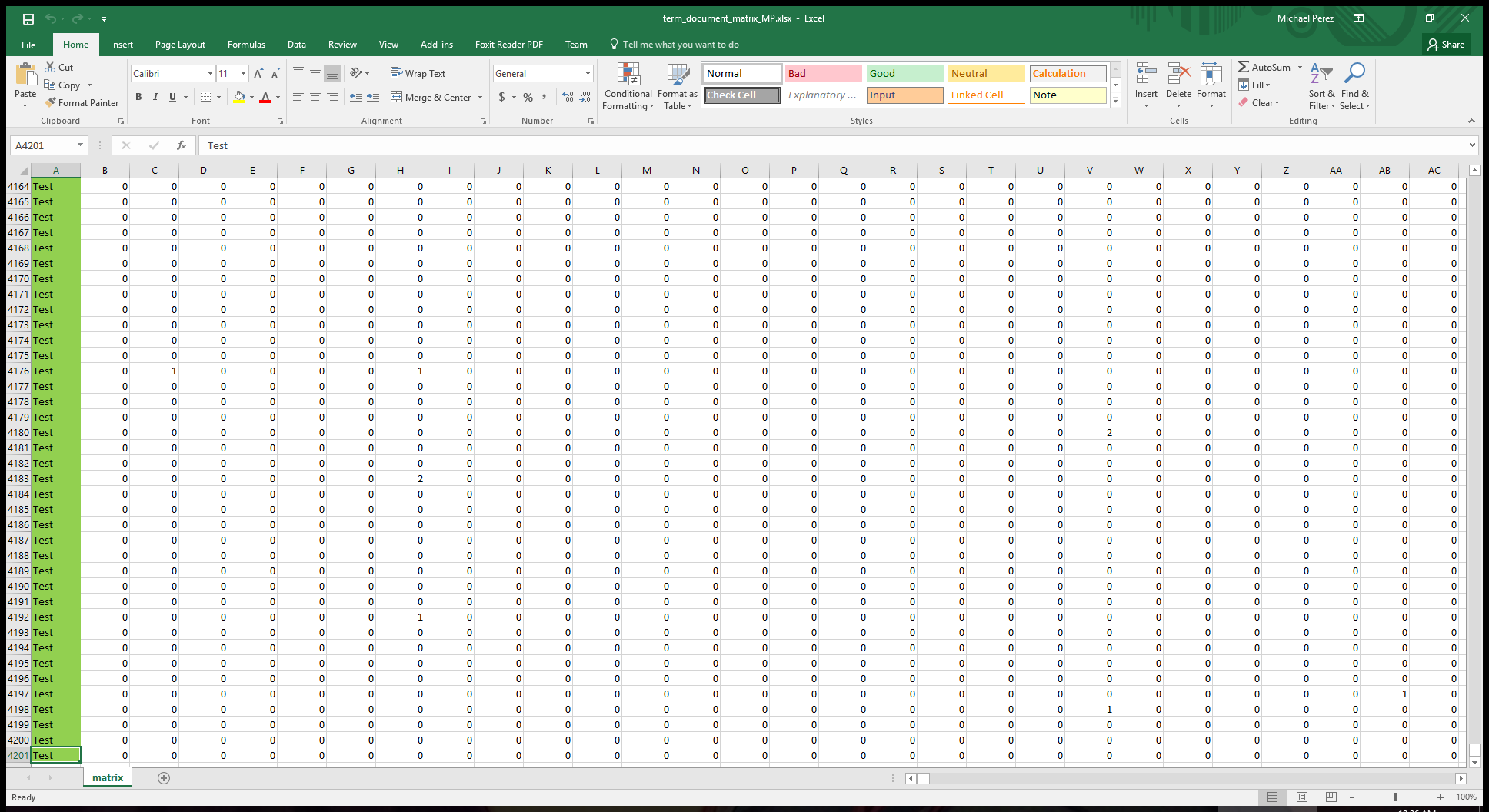
# Term-Document Matrix

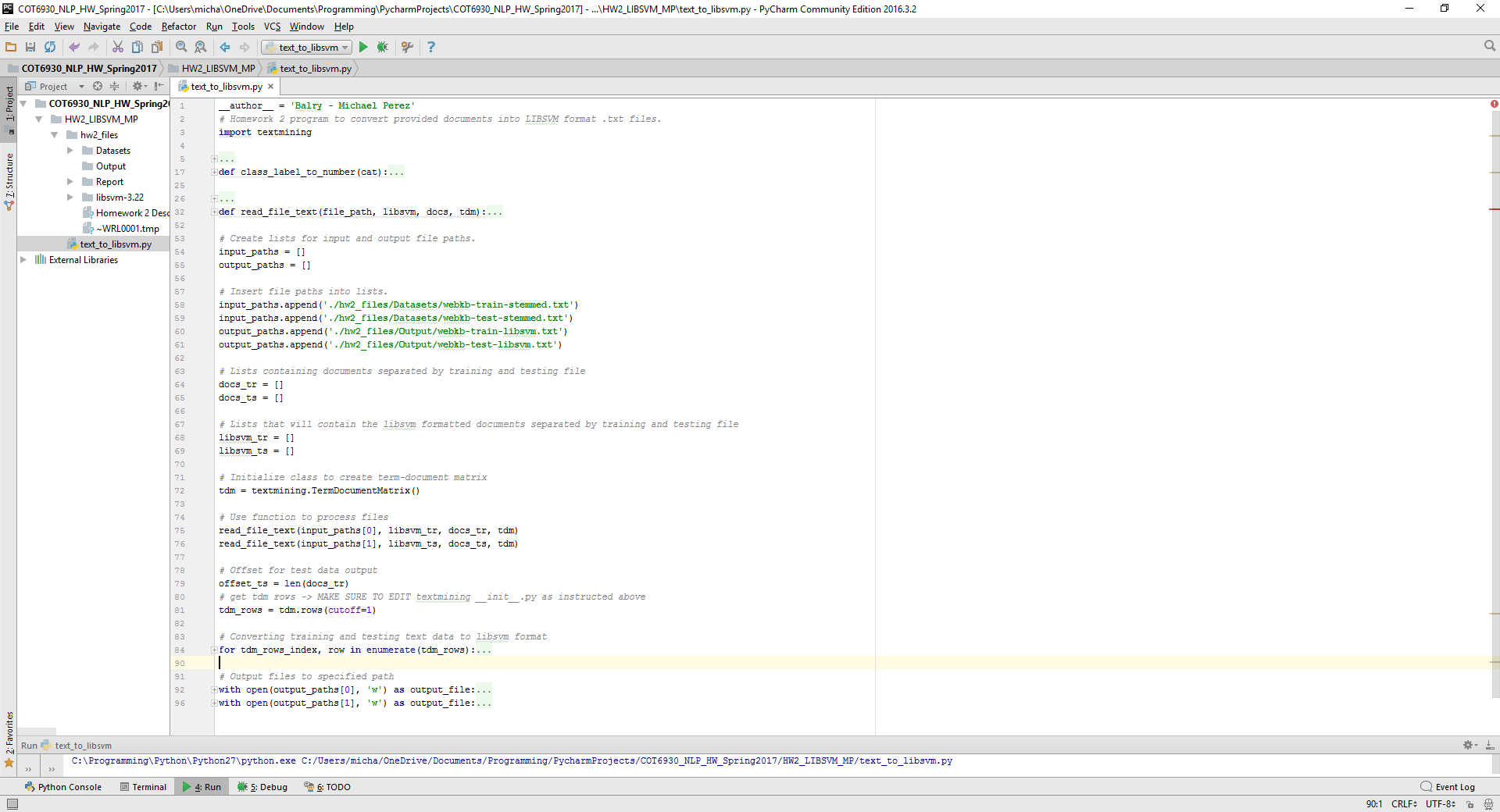
This turned out to be a large matrix when outputted into .csv and it is largely useless because of its sparsity. See screenshots below for a sample. Matrix will be attached, but it is 86.2 MB.

The code for making the conversion will be attached (made using python). Limited screenshots available below as well.

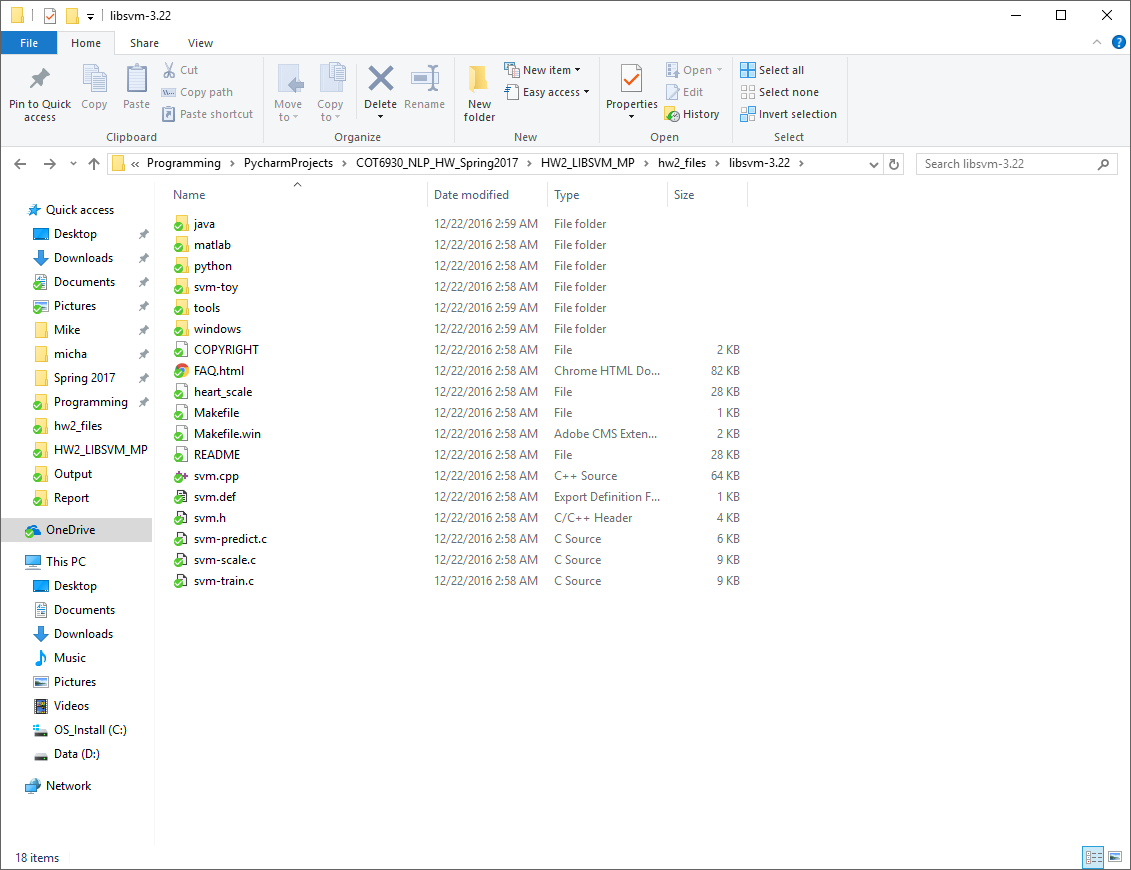








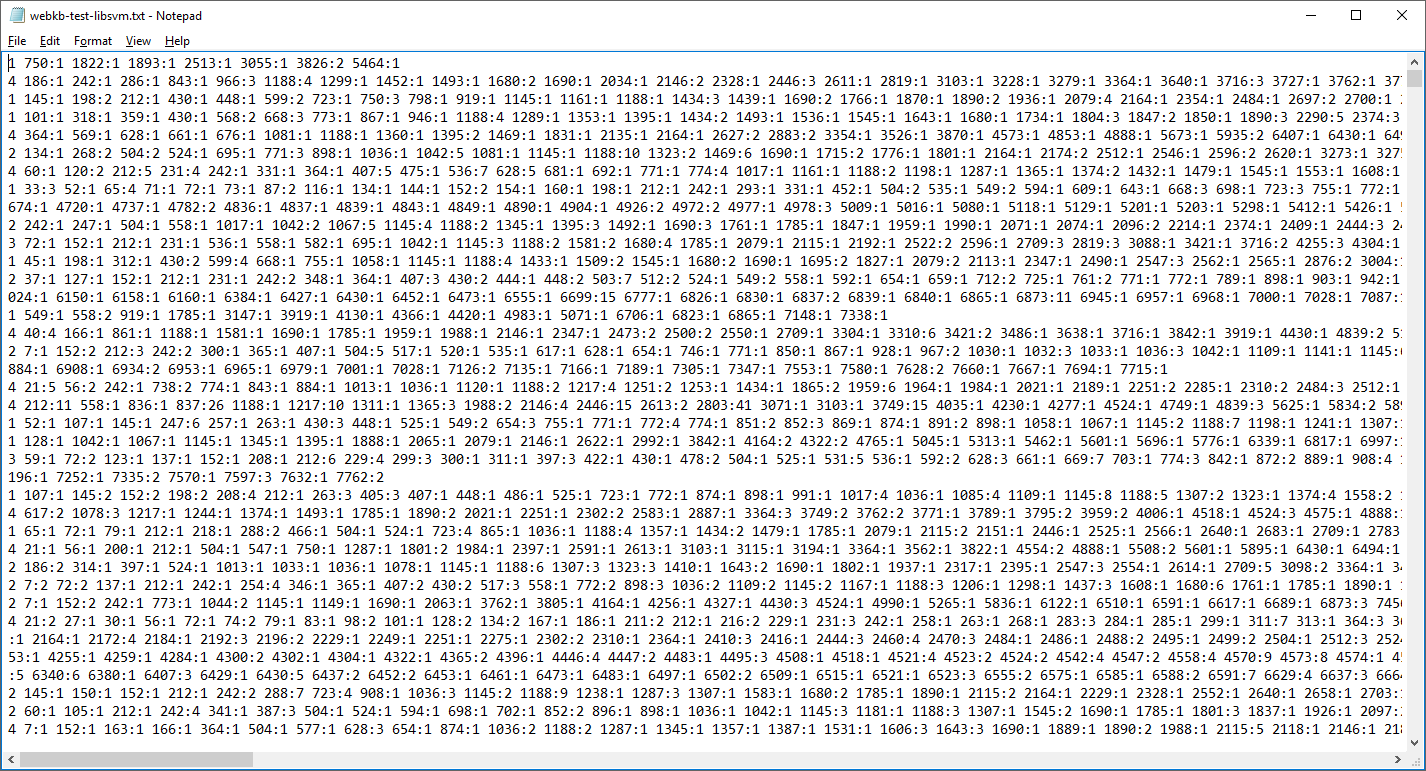
# LIBSVM Download



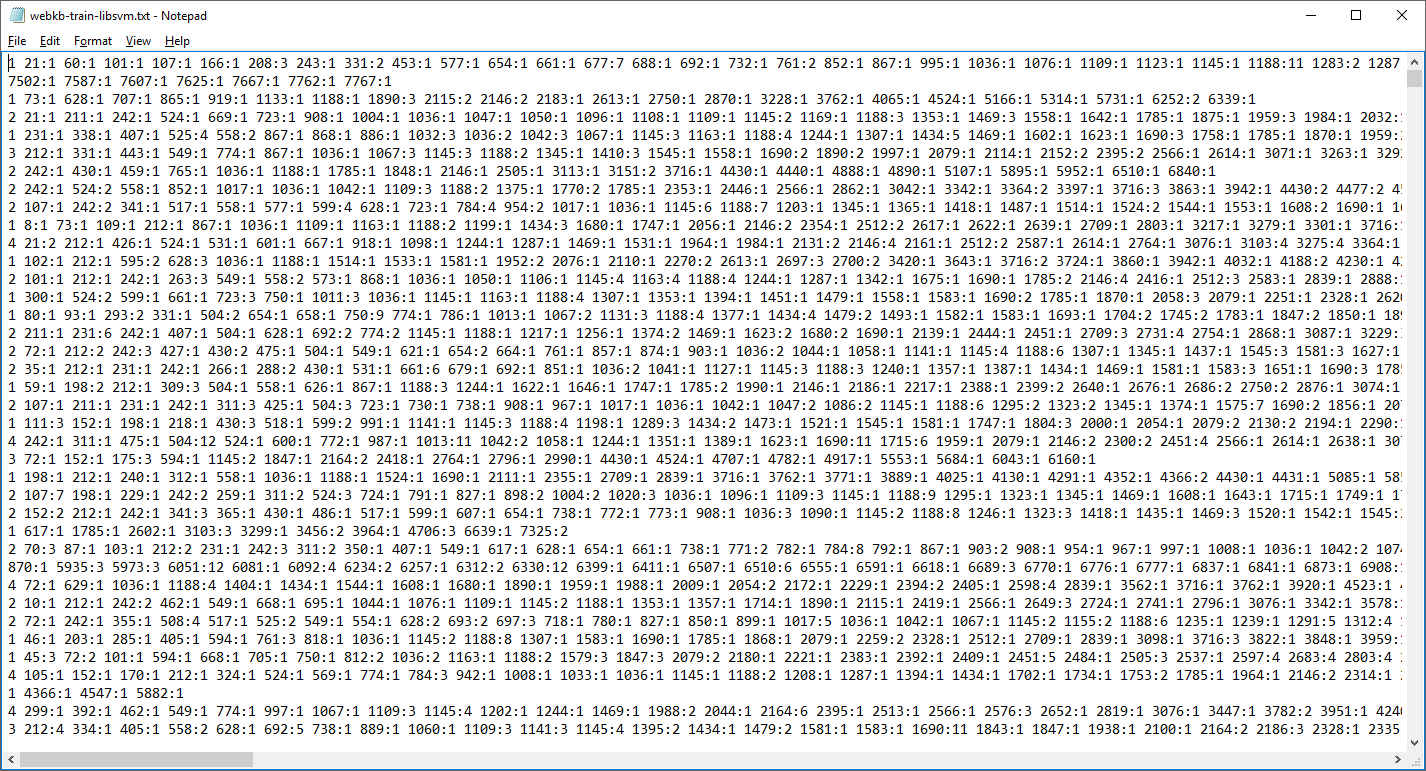
# LIBSVM Formatted Text

Attached files under the following names “webkb-test-libsvm.txt” and “webkb-train-libsvm.txt”. Screenshots can be found below as well:

webkb-test-libsvm.txt

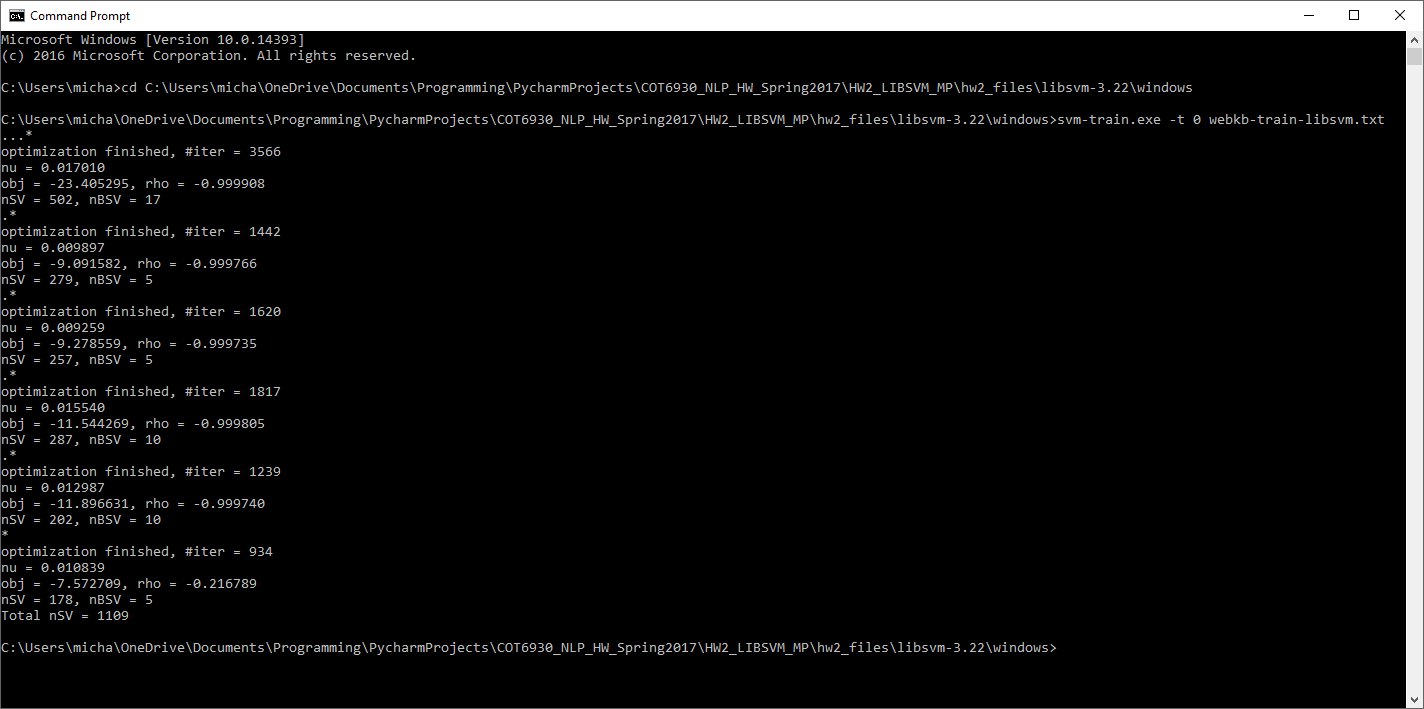


webkb-train-libsvm.txt



# LIBSVM Training and Testing Data Demo

svm-train.exe -t 0 webkb-train-libsvm.txt



svm-predict.exe webkb-test-libsvm.txt webkb-train-libsvm.txt.model webkb\_output\_file\_MP

