Mike Roylance
Ling 572
Homework 6

Q1)

The command I used is also included in the tar file under q1.sh:

#!/bin/sh
rm -rf q1
mkdir q1
mallet import-svmlight --input examples/train2.vectors.txt examples/test2.vectors.txt --output q1/train2.vectors q1/test2.vectors
mallet train-classifier --input q1/train2.vectors --trainer MaxEnt --output-classifier q1/m1
mallet classify-svmlight --input examples/test2.vectors.txt --output q1/classified.txt --classifier q1/m1
classifier2info --classifier q1/m1 *> q1/m1.txt*
vectors2classify --training-file q1/train2.vectors --testing-file q1/test2.vectors --trainer MaxEnt --report test:raw test:accuracy test:confusion train:confusion train:accuracy*> q1/out.stdout 2>q1/out.stderr*

Q2)

What is the test accuracy?	82.6667%
Is this the same test accuracy as Q1?	Yes
Why or why not?	It is the same because the algorithm that mallet uses for MaxEnt is the same as the one used for this homework problem.

Q3 & Q4) The files have been included in the tar file for this. I didn't see any specific questions to answer on the homework assignment, but if you'd like to look at my code you can see the src/MainQX for each question (as well as the corresponding bash file).

Additional Notes:

I haven't included the example directory, my bash scripts by default will reference it. Please include an alias to the directory if you would like to call q1.sh or q4.sh to see those results.

Also, I wrote this in Java using Maven to compile. The Maven version on patas is 3.0, it will display a debug message to the console (but will not affect the file output).

I have tested these scripts on patas