**NAÏVE BAYES**

**How to execute the script in command prompt:**

The **naiveBayes.py** takes 3 command line parameters (please use double quotes to surround string parameters):

1. Path to training dataset up till folder name that contains ham and spam folders. Do NOT include the ending backward slash.
2. Path to test dataset up till folder name that contains ham and spam folders. Do NOT include the ending backward slash.
3. Path to stopWords.txt including file name

$ python naiveBayes.py <path-to-training-dataset-up-till-folder-name-that-contains-ham-and-spam-folders> <path-to-test-datatest-set-up-till-folder-name-that-contains-ham-and-spam-folders> <path-to-stop-words-text-file-including-file-name>

Example:

$ python naiveBayes.py “C:\ML\hw2\_train\train” “C:\ML\hw2\_test\test” “C:\ML\stopWords.txt”

**NOTE:** The stopWords.txt can be found in the project submission zip.

**LOGISTIC REGRESSION**

**How to execute the script in command prompt:**

The **logisticRegression.py** takes 3 command line parameters (please use double quotes to surround string parameters):

1. Path to training dataset up till folder name that contains ham and spam folders. Please include the ending backward slash, and add another one. The last slash is added to escape the previous one. If your path happens to have a directory beginning with lowercase n or t, then this might result in \n or \t, giving a ‘The filename, directory name, or volume label syntax is incorrect’. View example below for clarity.
2. Path to test dataset up till folder name that contains ham and spam folders. Please include the ending backward slash, and add another one. The last slash is added to escape the previous one. If your path happens to have a directory beginning with lowercase n or t, then this might result in \n or \t, giving a ‘The filename, directory name, or volume label syntax is incorrect’. View example below for clarity.
3. Path to stopWords.txt including file name

$ python logisticRegression.py <path-to-training-dataset-up-till-folder-name-that-contains-ham-and-spam-folders> <path-to-test-datatest-set-up-till-folder-name-that-contains-ham-and-spam-folders> <path-to-stop-words-text-file-including-file-name>

Example: $ python logisticRegression.py “C:\ML\hw2\_train\train\\” “C:\ML\hw2\_test\test\\” “C:\ML\stopWords.txt”

**NOTE:** The stopWords.txt can be found in the project submission zip.

End of document