Assignment 4 (Vector Space Retrieval)

Title: Assignment 4 (Vector Space Retrieval)
Author: Chenfeng Fan (fanc@brandeis.edu)
Data: Wed, 20 March 2019, 11:59 PM

Dependencies:
    Python 3.7 (https://www.python.org/downloads/)
    NLTK 3.4 (https://www.nltk.org/)
    Flask 1.0.2 (http://flask.pocoo.org/)
    Bootstrap 4.3 (https://getbootstrap.com/)
    PyCharm (https://www.jetbrains.com/pycharm/download/
#section=mac)
    MacOS 10.14 (https://www.apple.com/mac/)

Build Instructions:
$ pip install -U Flask
$ pip install -U nltk

Run Instructions:
To start the application:
   >>> python vs_query.py
To terminate the application, use control-c
To use the application within a browser, use the url: http://127.0.0.1:5000/

Modules: NLTK, Flask
Stemmer: SnowballStemmer from NLTK
Folders: shelve folder stores the shelve files and temp json file, corpus folder stores films corpus json file and test corpus json file
Testing: Tested by searching manually.
Query examples: good day, king of sweden, hello, hello world xxx, xxx dsarwe cfrfwe, 西虹市首富, and blank
Benchmark: Generate the shelve file cost about 50sec