개발환경: - Pycharm(python 2.7 64bit) windows

-------------------------------------------------------------------------------------------------

사용방법:

1. Python2.7 64bit (numpy, scipy, genism, pandas)
2. word2vec\_wx, word2vec\_wx.syn1neg.npy, word2vec\_wx.syn1.npy, word2vec\_wx.wv.syn0.npy 와 sim.py를 같은 파일에 넣는다.

-------------------------------------------------------------------------------------------------

1. **vocab.py: word2vec model 내의 단어를 저장**

출력: word2vec model내의 모든 단어를 vocab.txt에 저장한다

실행 예시: python vocab.py

-------------------------------------------------------------------------------------------------

1. **noun.py: word2vec model내 없는 단어 제거**

입력: remove.txt (소설 속 중복 제거한 명사) & vocab.txt (word2vec model 단어)

출력: noun.txt

실행 예시: python noun.py

-------------------------------------------------------------------------------------------------

1. **sim.py: 소설 속 단어와 카테고리의 similarity구하기**

입력: noun.txt cat.txt

출력: result.txt 단어들의 유사도 계산

실행 예시: python sim.py

(유사도가 가장 큰 5개단어 출력하거나 유사도가 일정한 값(0.7)보다 큰 단어를 출력할 수 있다)

C:\Users\Jane\AppData\Local\Temp\1499687288(1).png

