개발환경: - Pycharm(python 2.7) windows

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

1. **txt\_pre.py : 형태소 분석 결과 텍스트 파일 속 불필요한 ‘ ’, ‘\n’등을 제거****(실행시간 32분)**

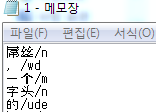
입력: 형태소 분석 후 결과 파일 /books\_result/1.txt, 2.txt, ...

출력: 처리된 형태소 분석결과 /text/1.txt, 2.txt, ...

실행 예시: python txt\_pre.py 800 (windows cmd에서, 소설 개수가 800개일때)

(의미 : text폴더의 1.txt부터 800.txt까지 순차적으로 처리하여 Txt 폴더에 출력파일을 생성함)

C:\Users\Jane\AppData\Local\Temp\1499968221(1).png



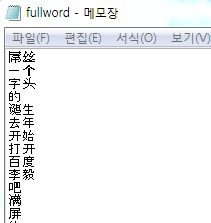
-------------------------------------------------------------------------------------------------

1. **fullword.py : 형태소 분석 결과 텍스트 파일 속 의미 있는 단어만 저장 (실행시간 1시간10분)**

입력: 처리된 형태소 분석결과 /text/1.txt, 2.txt, ...

출력: 의미 단위로 처리된 단어들을 하나의 텍스트 파일로 통합 fullword.txt

C:\Users\Jane\AppData\Local\Temp\1499970361(1).png



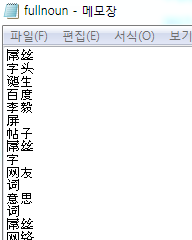
-------------------------------------------------------------------------------------------------

1. **fullnoun.py: 형태소 분석 결과 텍스트 파일 속 명사만 저장 (실행시간 50분)**

입력: 처리된 형태소 분석결과 /text/1.txt, 2.txt, ...

출력: 추출된 명사들을 모두 통합된 하나의 텍스트 파일 fullnoun.txt

**C:\Users\Jane\AppData\Local\Temp\1499977630(1).png**

****

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