



Homework #2: My Word Cloud

Sunglok Choi, Assistant Professor, Ph.D.
Computer Science and Engineering Department, SeoulTech
sunglok@seoultech.ac.kr | <https://mint-lab.github.io/>

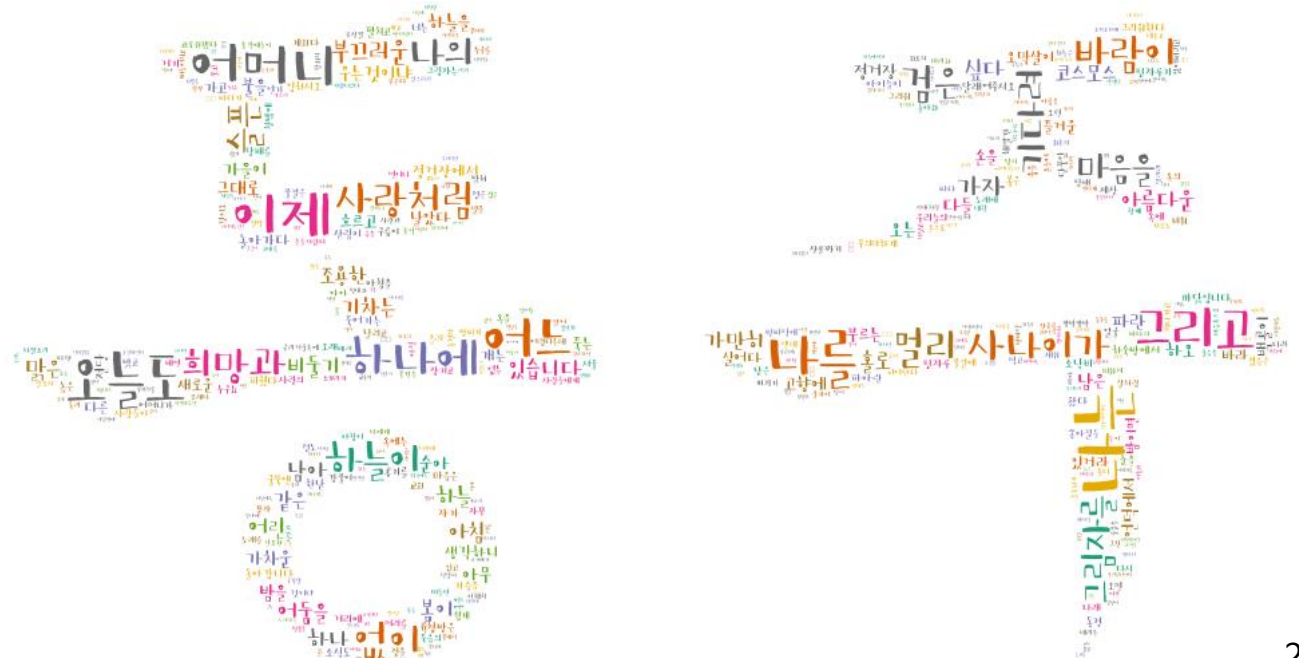
Review) Word Cloud

■ 대략적인 절차

- 글에서 POS tagging을 통해 **명사(noun)** 단어를 추출 및 **불용어(stopword)** 제거
- Python의 [WordCloud](#) 패키지를 통해 추출된 단어들 렌더링
 - 불용어(stopword)를 설정 가능 (참고: [한국어 불용어](#))
 - generate() 또는 generate_from_text() 함수를 통해 직접 단어의 빈도수 계산
또는 generate_from_frequency() 함수를 통해 직접 계산한 빈도수 입력 가능
 - 주의) 한글 사용을 위해서는 사용할 한글 폰트를 직접 설정하여야 함

■ 참고

- [WordCloud Official Document](#)
- [고적고적 cODE 블로그](#), [Github 예제](#) (김영진)
- [온라인 워드클라우드 생성기](#) (나름수익모델연구소)



My Word Cloud

- Task
 - 나만의 word cloud (a.k.a. [tag cloud](#)) 만들기
 - 언어: 한국어(또는 모국어)로 된 자신이 선택한 텍스트
 - 품사: 명사+ α (적어도 조사는 제외)
 - 도전적 요소: 새로운 패키지 설치 ([KoNLPy](#), [WordCloud](#)), skeleton code가 주어지지 않음
- Guideline
 - 제출 파일 (hw02_YourStudentID.zip)
 - 소스코드
 - 작품 파일 (**작품명.png**; 가능한 고해상도로 제작)
 - 부가 파일 (text file, mask image file, ...)
 - 제출 장소: e-Class > 숙제
- Point: 20
 - 심미적인 요소는 고려하지 않음
 - 감점 요소: 언어 (-5), 품사 (-5)
- Deadline: **2022/03/30 23:59**