

2주차 과제 출제

뉴스 기사 텍스트 데이터 분석

뉴스 기사 텍스트 데이터를 크롤링하고 분석하는 파이썬 프로그램을 작성하세요.

1. 원하는 기사를 찾은 후 해당 페이지 URL에서 뉴스 기사 텍스트 데이터를 크롤링하세요:
2. 크롤링한 텍스트 데이터를 토큰화하고, 각 단어의 빈도를 계산하세요.
3. 가장 빈도가 높은 단어 10개를 찾아 출력하세요.
4. 뉴스 기사의 길이 (단어 수)를 계산하세요.
5. 문장의 평균 길이를 계산하세요.
6. 기사의 감정을 분석하기 위해 긍정적인 단어와 부정적인 단어의 빈도를 계산하세요. (긍정적인 단어 목록과 부정적인 단어 목록을 정의하세요)
7. 뉴스 기사의 주제나 주요 토픽을 분석하기 위해 가장 많이 등장하는 명사를 찾아 출력하세요.
8. 크롤링한 기사를 요약하는 알고리즘을 개발하고, 기사의 요약된 내용을 출력하세요.(딥/머신러닝팀만 해당)

사용 라이브러리

웹 크롤링: **BeautifulSoup or Scrapy or ChromeDriver etc.**

자연어 처리: **NLTK or SpaCy etc.**