

WORD CLOUD(TAG CLOUD)

개발자 : 안 승진 (2023-11-22)

주요 기능 및 기술



❖ 주요 기능

- 온라인 상의 주요 기사 타이틀을 스크래핑하여 자주 등장하는 키워드 순으로 글자 크기를 달리하여 보여줌
- 전체 모양은 하트 이미지로 Masking 하여 보여줌

❖ 기술

- Python
- Webscraping
- Wordcloud

주요 구현 내용

```
soup = BeautifulSoup(html, "html.parser")
# 뉴스 타이틀 뽑기
titles = soup.select("a.news_tit")
# print(titles[0].get_text())
for title in titles:
    # print(title.get_text())
    print(f"---{cnt}---")
    cnt += 1
    title_list.append(title.get_text())
```

- 웹스크래핑하는 코드로서 뉴스 기사의 타이틀을 select하여 리스트에 저장

- 웹스크래핑한 기사 타이틀의 명사, 형용사를 찾아서 자주 사용된 40개의 단어를 추출

```
# twitter함수를 통해 읽어들이는 내용의 형태소를 분석한다.
sentences_tag = []
sentences_tag = twitter.pos("".join(title_list))
noun_adj_list = []
# tag가 명사이거나 형용사인 단어들만 noun_adj_list에 넣어준다.
for word, tag in sentences_tag:
    if tag in ["Noun", "Adjective"]:
        noun_adj_list.append(word)

# 가장 많이 나온 단어부터 40개를 저장한다.
counts = Counter(noun_adj_list)
tags = counts.most_common(40)
```