|  |  |
| --- | --- |
| **프로젝트 개발 기획서** |  |
|  |  |

문서서식포탈비즈폼

|  |  |  |  |
| --- | --- | --- | --- |
| 작 성 자 | 박상범 | 작성 일시 | 2021 년 9월 8일 |

|  |  |
| --- | --- |
| 프로젝트명 | 네이버 뉴스 카카오톡 전송 프로젝트 |

|  |  |
| --- | --- |
| 프로젝트 기간 | 2021 년 9월 7 일 ∼ 2021 년 9 월 12 일 |

|  |  |
| --- | --- |
| 프로젝트 참여 인원 | 5명 |

|  |  |
| --- | --- |
| 프로젝트 시행 목적 | * 출퇴근 또는 이동시간에 스마트폰으로 뉴스를 보려고 하면 광고들이   많아 의도치 않게 클릭하게 되거나 바로가기 앱으로 이동하게 되어  불편함을 겪는 경우가 많아 확실한 정보 획득이 필요하다.   * 집중하기 힘든 환경에 빠른 시간관리와 중요한 뉴스정보를 얻을 수 있다. * 뉴스의 필요성은 누구나 알고 있기에 쉽고 빠른 방법이 있어야 할 필요가 있다. |

|  |  |
| --- | --- |
| 프로젝트 개요 및 추진방향 | * 세상에는 많은 뉴스들이 있지만 다읽기에는 시간이 부족해졌다. 출퇴근이나 이동시간인 집중하기 힘든 환경에서는 뉴스같이 정적인 글이나 빼곡히 많은 글을 읽기에는 어려워 한눈에 빠르게 보기 쉽게 카카오에서 제공하는 오픈 API를 활용하여 뉴스 크롤링 카카오톡 메시지 전달하는 것이다. * 파이썬 언어를 이용하여 BeautifulSoup 객체 생성 후 네이버 뉴스를   각 섹션별로 URL의 구조변화와 HTML 구조를 파악하여 크롤링하고,  긴 뉴스 본문을 gensim을 통해 요약을 한다. 그리고 이 정보를  카카오톡 메시지로 전송한다. |
| 언어 & 환경 (IDE) | 파이썬 3.7 & 구글 코랩 |
| 프로젝트 리소스  및 소스 코드 | - news\_summarize.ipynb: 뉴스 요약 프로젝트 구현코드  - kakao\_utils.py: 카카오 메시지 전송을 위한 유틸리티 구현코드  - res/kakao\_message/kakao\_token.json: 카카오 access token과 refresh token이 관리되는 파일 시간 |

|  |  |
| --- | --- |
| 변수 요약 | **#0**  Genism: 뉴스 요약  BeautifulSoup4, lxml : html 구조 파싱, 관리  Requests: html을 요청 or 카카오 Open API의 REST API를 사용하기 위함  url 변화 부분 - 경제: 101, 사회: 102, IT/과학: 105  sid: 3개 섹션 받기  Input: url 받는 함수  Output: BeautifulSoup 객체 전달  default\_img: 임시 이미지 저장(이미지 없는 뉴스도 있을 수 있으니)  news.info: 뉴스 메타 정보를 가져옴(뉴스제목, 날짜, 뉴스 본문, 이미지)  **#1 네이버 뉴스 크롤링하기**  get\_soup\_obj(): BeautifulSoup 객체를 return  get\_top3\_news\_info(): for문의 내용을 함수화, 뉴스의 기본정보  sec\_url: 분야의 상위목록 주소  soup: 상위뉴스 HTML가져오기  find(): 태그 하나만 가져옴  find\_all(): 해당 태크가 여러개 있을 경우 한꺼번에 모두 가져옴.  news\_list3: 해당분야 상위뉴스 3개 가져오기  title: 뉴스 제목, news\_url: 뉴스 URL, image\_url: 이미지 URL  news\_dic: 뉴스 결과를 담아낼 dictionary  "eco": 경제  "soc": 사회  "IT": IT/과학  section\_ids: URL에 사용될 뉴스 각 부문 ID  news\_dic: '경제', '사회', ‘it/과학 분야의 상위 3개 뉴스  get\_naver\_news\_top3(): 섹션 별 뉴스 메타 정보, 뉴스 본문  **#2 뉴스 요약하기**  news\_info: 뉴스 본문 요약 정보  **#3 카카오톡 전송하기**  sections\_ko: 사용자가 선택한 카테고리를 제목에 넣기 위한 dictionary  navernews\_url: 네이버 뉴스 URL  contents: 추후 각 리스트에 들어갈 내용  template: 리스트 템플릿 형식  object\_type: text |

|  |  |
| --- | --- |
| 사전준비 | **TOKEN**    **KakaoAPI** |

|  |  |
| --- | --- |
| 결과 | 슬기로운 2조생활 |

|  |  |
| --- | --- |
| 프로젝트 시행 후  얻을 수 있는 이익 | * 바쁜 시대를 살아가는 현대인들을 위해서 짧고 간편하게 원하는 분야의 기사 스크랩을 읽을 수 있도록 하였다. |