

\*본 문서는 데이터 캠퍼스 경기대팀의 시연용 코드 설명을 위한 문서입니다.

### <유의 사항>

1. 본 프로젝트는 크롬 85버전 환경에서 진행되었습니다. 버전이 호환되지 않는다면 <https://chromedriver.chromium.org/downloads> 에서 크롬 버전과 동일한 chrome webdriver.exe를 다운로드 해서 chromedriver.exe 파일과 교체해주세요.
2. Kyonggi\_univ 폴더를 각각 파이참 프로젝트와 주피터 노트북 디폴트 경로에 넣어주세요. .py 파일은 파이참, 러닝 모델 코드(ipython 파일)는 **아나콘다 프롬프트를 반드시 관리자 권한으로 실행시켜서 주피터 노트북을 구동해주시고 가상환경이 아닌 기본 base 환경에서 진행해주시기 바랍니다.** (모듈 설치시에 컴파일러 때문에 그렇습니다.)

### <폴더 설명>

1. original\_code는 원본 코드가 들어있는 폴더입니다. 해당 폴더에 있는 코드는 길게는 수십 시간이 소요됩니다. 해당 폴더 안에는 각 채널별로 데이터 전처리 및 분석을 위한 추가적인 코드가 포함된 폴더가 존재합니다. 산업체 측에서 제공받은 데이터는 요청에 의하여 첨부하지 못하는 점 양해 부탁드립니다.
2. test\_code 폴더는 시연을 위해 작성한 코드가 포함되어 있습니다. 시연을 위해 수정한 코드로 데이터의 건수를 제한하여 수집합니다. (약 1~2 게시글만 수집하도록 수정했습니다.)

### <코드 설명>

1. insta\_main.py 파일은 인스타그램 크롤러입니다.
2. amazon\_main.py 파일은 아마존 크롤러입니다.
3. walmart\_main.py 파일은 월마트 크롤러입니다.
4. sentimental\_analysis 폴더는 러닝 모델 코드가 있는 폴더입니다. ipython 파일입니다. 주피터 노트북으로 실행 해주시면 감사하겠습니다.
5. keyword\_analysis 폴더는 키워드 분석 코드로 ipython 파일입니다.
6. additional\_analysis 폴더는 분석 코드로 ipython 파일입니다.
7. doc2vec.py 파일은 산업체 측에서 제공받은 국내 데이터를 활용하여 학습 시킨 딥러닝 모델입니다.