

Code States AIB(AI Bootcamp)

수강생 프로젝트 기획서

수강생 이름: 박동주, 박진수, 김영진,

주보국, 권병현

시작일 : 2022. 08.30 프로젝트 : 프로젝트1

1. 프로젝트 주제

SNS에서 유료광고 표시를 텍스트가 아닌 이미지에 광고를 표시하여 태그 필터링을 피하는 악성 광고에 강제로 광고 태그 입력

2. 프로젝트 목적



2020년 9월 이후로 SNS 유료광고 표기법이 시행되어 인스타그램 같은 SNS에서 원치 않는 광고를 해쉬태그를 이용하여 필터링 할 수 있게 되었다. 법의 허점을 이용하여 해쉬태그로 광고임을 표시하지 않고이미지에 유료광고 표시를 하여 필터링을 회피하는 사례가 늘어나고 있다. 하여 글 업로드 시 딥러닝을 이용하여 이미지에서 유료광고 표시를 찾아 강제적으로 해쉬태그를 부여하는 기능을 구현하는 것이 본프로젝트의 목적이다.

- 3. 프로젝트 구현 기술 및 담당자 선정
 - 데이터 수집
 - 담당자 (정/부) : 박진수 / 김영진
 - 데이터 전처리

.....

- 담당자 (정/부) : 김영진 / 박동주
- 모델 설계 및 구현
 - 담당자 (정/부) : 권병현 / 주보국
- 배포
 - 담당자 (정/부): 주보국 / 권병현
- 발표 자료 제작(시각화)
 - 담당자 (정/부): 박동주 / 박진수

4. 프로젝트에 필요한 강의 컨텐츠

- 김영진 Section3, 데이터베이스설계 관련(youtube,강의)
- 주보국 도커 교과서/ 핸즈온 머신러닝/ Section3/ VSC에서 git으로 협업하기/ FASTAPI
- 박동주 시각화(태블로 인강[클래스101]/ 전처리관련 섹션공부 및 youtube 강의
- 박진수 Section 3, Section 4, Data 다루는 강의 (youtube)
- 권병현 Section 3, Section 4, OCR-ViT 관련 서적, BERT, pytorch 사용법, FLASK, html

5. 프로젝트 디자인

Enter URL: https://www.instagram.com/p/Ch8Y3AthGhp/

Submit URL



Output : 유료광고 이미지 검출 (or 유료광고 이미지 미검출)

6. 프로젝트 개발 일정

2022년 8월
30 31 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17

□ 프로젝트 선정

전처리

모델 설계 및 구현

추석

보고서 작성

보고서 작성

7. 프로젝트 활용 데이터

- 인스타그램 게시글
 - 데이터 설명 : 인스타그램 게시글
 - 데이터 수집 방법: instaloader 라이브러리를 사용하여 데이터 스크래핑 하려 했지만 클라이언트와 서버 문제, 수집할수 있는 데이터의 제한으로 인해 insta API를 사용 불가능
 - 데이터 타입(종류) : CSV / TXT / JSON / JPG
 - 데이터 크기 : 인스타그램 게시글 3만 개
 - 데이터 활용 방안(예정): 게시글 안에는 여러개의 이미지 사진이 저장되어 있습니다. 첫번째 사진과 마지막 사진의 연관성이 없을 경우 광고라고 판단을 내리는 알고리즘을 사용할것입니다.

8. 팀 및 개인 Github 링크 작성

• 주보국 : https://github.com/amitis94/cp1

박동주: https://github.com/copperstick/project_cp1

박진수: https://github.com/PJinss/CP1.git
 권병현: https://github.com/bhkwon89/cp1

• 김영진 : https://github.com/rladudwls615/CP1.git