**비정형 데이터 활용 분석 프로젝트 기획안**

**2022년 12 월 27 일**

**과정명: [멀티잇]서비스 산업 데이터분석가 취업캠프(Python)**

|  |  |
| --- | --- |
| 프로젝트 주제 | 재활용 가능한 페트병 사진 모형을  통한 분류 예측기(다중분류 모형) -save Us(Earth) |
| 팀 명 | 비전 있는 비전공자 |
| 팀 원 | 팀 : 진세용 / 김성훈 / 김지찬 / 김소연 / 조영훈 / 문충헌 |
| 프로젝트 내용  해결하고자 하는 문제  최종 산출물의 청사진 | **탄소중립을 실천하기 위한 Save Us(Earth) 프로젝트**  -한 사람이 평생 살아가며, 배출하는 쓰레기는 약 55톤으로 심각한 환경문제를 야기하고 있는 상황이다.  -이 중에서 재활용(Recycle)될 수 있는  비율이 가장 높은 자원은 페트병이다.  페트병은 생태계오염을  가장 많이 시키는 1순위 자원이며, 플라스틱 1%만 줄여도  연간 640억 절감의 경제적 효과를 가져올 수 있다.  그렇기에, 이미지를 찍으면  스스로 플라스틱 재활용 여부를 판단하는  분류예측기를 제작하게 되었다. |
| 프로젝트 추진 일정 | ※ 일정  ● **12/14 ~ 12/15 : 주제 선정, 일정 수립 및 시장 조사**  ● **12/16 ~ 12/25 : Programming**  1) 데이터 수집: 크롤링(12/16)  2) 데이터 전처리(12/19~12/20)  3) 머신러닝 분석 :다중분류 모델 구축(12/21~12/22)  4) 머신러닝 성능 확인(12/23)  5) PPT 및 최종기획안 제작(12/23)  ● **최종 수정(12/26)**  ● **프로젝트 기획안 발표(12/27)** |
| 프로젝트 수행 방향  수행 방법/도구 | **<프로젝트 진행 순서>**  데이터수집- 데이터 전처리-모델구축-모델평가-성능검증  **1.데이터수집**  -구글 이미지 크롤링  -캐글 데이터 수집  **2.데이터 전처리**  -이미지 리사이징  -데이터 크기변환  -데이터증식  -넘파이 배열변환  -데이터정규화  **3.모델구축**  -CNN  -전이학습  -Custom CNN  **4.모델평가**  -학습 끝난 모델에서 최적 모델 찾기  **5.성능검증**  -새로운 데이터를 넣어 성능검증  사용언어: Python  사용프로그램: Jupyter Notebook  데이터 수집방법: Web Crawling  크롤링 대상페이지: google image, Kaggle data |
|  |