

포트폴리오 개요

No	프로젝트 이름	기간		담당 역할	주요 기술	주요 내용 요약
		시작	종료			
1	거주자 우선 주차장 공유 및 관리	23년 1월	23년 2월	팀원	데이터 분석 / AI모델 학습	융합 프로젝트 (최우수상 수상)
2	서울시 상권 & 고객 분석 서비스	22년 12월	22년 12월	리더	데이터 분석 / 머신러닝(다중분류)	빅데이터 프로젝트
3	전국시티투어 전문 플랫폼	22년 10월	22년 11월	리더	사이트 개발	인터페이스 개발
4	따릉이 핫 스팟	22년 9월	21년 12월	팀원	데이터 분석 / 사이트 개발	산학협력캡스톤설계
5	프로 야구팀 득점 예측 / 심장병 예측	22년 4월	22년 5월	리더	선형회귀 / 로지스틱회귀	데이터 마이닝 프로젝트
6	공사 입찰공고별 낙찰 여부 예측	22년 1월	22년 2월	팀원	머신러닝(이중분류)	공공조달 빅데이터 경진대회
7	축구선수 시장 가치 평가	21년 11월	21년 12월	팀원	데이터 분석 / 해석	빅데이터 언어 프로젝트
8	하계 인턴십	21년 7월	21년 8월	팀원	실시간 데이터 송수신 및 관리	휴먼디지털
9	오픈소스가 뭐죠	21년 10월	21년 11월	리더	Git, Github 활용	오픈소스 프로젝트
10	백신 예약 프로그램	21년 10월	21년 12월	리더	링크드리스트	C언어 프로젝트

프로젝트 명 : 거주자 우선 주차장 공유 및 관리

기간	'23년 1월 ~ '23년 2월(2개월)
담당 역할	<p>팀원</p> <p>분석</p> <ul style="list-style-type: none"> - 데이터 수집 - 데이터 전처리 - 데이터 시각화 <p>AI모델</p> <ul style="list-style-type: none"> - 이미지 학습 - 텍스트 데이터 추출
사용언어 / 프로그램	Python, R, Yolo, pandas, Numpy, OpenCV, Selenium, matplotlib, Colab, roboflow, AWS
주요 기술	- 데이터 분석 / AI모델 학습
요약 내용	<p>융합 프로젝트</p> <p>- '공유를 통해 주차난을 해결하고 불법 주차 차량 단속'</p>
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none"> - 거주자 우선 주차자 관리, 공유 시스템 <p>상세 내용</p> <ul style="list-style-type: none"> - 데이터를 분석을 통해 서울 근교인 인천, 경기도 통근통학 인구수와 주차공간을 비교해 서울시 주차공간의 공유가 필요함과 단속CCTV위치와 주차장 위치를 비교해 주차공간 관리의 필요성을 나타냄. - 라즈베리에 PIR센서, 조명센서, 카메라를 설치하여 차량이 들어올 때 찍은 사진을 Request방식으로 AWS서버로 전송되면 직접 학습시킨 AI모델이 번호판을 인식 후 텍스트 형태로 출력. - 모델 1 (정확도 95%) : 차량과 번호판 라벨링 / 번호판을 잘라내어 숫자 판독 모듈인 Easy OCR로 텍스트 추출 - 모델 2 (정확도 98.4%, mAP 98.3%) : 표지판의 숫자를 하나하나 라벨링 / 인식을 낮은 숫자 라벨 삭제, x축 좌표기준으로 정렬 후 텍스트 추출 <p>이슈 및 해결 방안</p> <ul style="list-style-type: none"> - 모델 1에서 자린 이미지를 모델 2를 실행시켜 텍스트 출력을 하려했으나 모델2 학습시 차량 전체 이미지를 넣어 주었기 때문에 인식 실패 => 모델 2만 사용하기로 결정 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none"> - DB에 있는 정보와 비교하여 주차장 사용자와 소유자에게 문자 알림을 보내는 서비스를 개발. - 최우수상 수상.

프로젝트 명 : 서울시 상권 & 고객 분석 서비스

기간	'22년 12월 ~ '22년 12월(1개월)
담당 역할	<p>리더</p> <p>분석</p> <ul style="list-style-type: none">- 데이터 수집- 데이터 전처리- 데이터 시각화 <p>머신러닝(다중분류)</p> <ul style="list-style-type: none">- 분류모델 사용하여 최적의 모델 선정
사용언어 / 프로그램	Python, Colab
주요 기술	<ul style="list-style-type: none">- 데이터 분석 / 머신러닝(다중분류)
요약 내용	<p>빅데이터 프로젝트</p> <ul style="list-style-type: none">- '외식업종별 자영업자에게 창업 지역 추천, 고객 특성에 맞는 마케팅 전략 수립을 목표'
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none">- 서울시 상권 & 고객 분석 서비스 <p>상세 내용</p> <ul style="list-style-type: none">- 상관분석을 통해 다른 컬럼들이 외식업종의 총 매출에 어느 정도 영향을 미치는지 분석.- 외식 업종별로 매출이 우세한 지역을 파악해 지도에 나타내고 해당 업종의 주요 고객층을 분석해 파이그래프와 해시태그로 나타냄.- 요일, 성별, 시간대, 연령대를 독립변수로 하고 하이퍼 파라미터(K값, max_depth, min_samples_split, eval_metric)를 조정하며 KNN, Decision Tree, Random Forest, GridSerachCV, XGB, LGBM 모델을 사용해 상권(골목상권, 발달상권, 전통시장, 관광특구)을 예측. <p>이슈 및 해결 방안</p> <ul style="list-style-type: none">- 업종에 초점을 두느냐, 지역에 초점을 두느냐 따라 해석이 달라짐. => 지역에 초점을 두기로 결정 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none">- 팀원 수가 적음에도 각자 맡은 일을 적극적으로 수행하였다.- 해석이 계속 달라져 의견을 정리하는데 어려움이 있었다

프로젝트 명 : 전국시티투어 전문 플랫폼

기간	'22년 10월 ~ '22년 11월(2개월)
담당 역할	<p>리더</p> <p>개발</p> <ul style="list-style-type: none"> - 데이터 전처리 - 웹페이지 틀/구성/디자인 - 게시판 구현 - 데이터베이스 관리 - 로그인/회원가입 구현
사용언어 / 프로그램	Python, HTML, CSS, JavaScript, AJAX, Django
주요 기술	<ul style="list-style-type: none"> - 사이트 개발
요약 내용	<p>인터페이스 개발</p> <ul style="list-style-type: none"> - "전국시티투어" 사이트 개발
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none"> - 전국시티투어 전문 플랫폼 <p>상세 내용</p> <ul style="list-style-type: none"> - 흩어져 있는 시티투어 사이트를 모아 한 곳에서 시티투어를 검색할 수 있는 사이트 - K-pop 인기로 늘어나는 한국여행 패키지 증가와 분산된 플랫폼, 코로나로 정체되어있던 여행산업 활성화를 위해 개발 <p>이슈 및 해결 방안</p> <ul style="list-style-type: none"> - 1차 검색 필터 후 2차 검색 필터를 원함 => AJAX사용 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none"> - HTML, CSS, JavaScript를 사용해 실제로 사이트를 개발해보는 경험을 할 수 있었다.

프로젝트 명 : 따릉이 핫 스팟

기간	'22년 3월 ~ '22년 6월(3개월)
담당 역할	<p>팀원</p> <p>분석</p> <ul style="list-style-type: none">- 데이터 전처리 <p>개발</p> <ul style="list-style-type: none">- 따릉이 대여소 위치 지도 표시- 스트리트 뷰 연동
사용언어 / 프로그램	Python, JavaScript, Django, Tableau
주요 기술	<ul style="list-style-type: none">- 데이터 전처리, 시각화, 사이트 개발
요약 내용	<p>산학협력캡스톤설계</p> <ul style="list-style-type: none">- 따릉이 GPS 데이터 분석을 통한 POI 시각화
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none">- 따릉이 핫 스팟 <p>상세 내용</p> <ul style="list-style-type: none">- 따릉이 GPS 데이터로 point of interest 찾기- Uber H3로 육각형 공간 분할하여 GPS 오류 데이터 삭제- Tableau를 사용해 자주 지나가는 지역을 시각적으로 표현- 따릉이 대여소 추가 설치 및 위치 변경 / 모임장소 추천 / 자전거 거도로&가로등 설치 등 정책 연구 활용 목적 <p>이슈 및 해결 방안</p> <ul style="list-style-type: none">- GPS 오류 데이터 삭제 => 속도 계산 후 25km 이상 삭제 / seq 0 or 1 & 120이상 & 마이너스 & 5000이상 삭제 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none">- Tableau를 활용해 다양한 시각화를 해볼 수 있었고 사이트를 개발하여 POI지점을 보여주었다.

프로젝트 명 : 프로 야구팀 득점 예측 / 심장병 예측

기간	'22년 4월 ~ '22년 5월(2개월)
담당 역할	리더 R언어 사용한 머신러닝 - 데이터 수집 - 데이터 전처리 - 데이터 시각화 - 회귀/분류모델 사용하여 예측 - 변수 선택법
사용언어 / 프로그램	R, R studio
주요 기술	- 분류/회귀모델 사용하여 결과 예측
요약 내용	데이터마이닝 프로젝트 - '선형회귀 / 로지스틱 회귀'
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none"> - 프로야구팀 득점예측 / 심장병 예측 <p>상세 내용</p> <ul style="list-style-type: none"> - 프로야구 팀 타격 기록으로 득점을 예측하기 위해 데이터를 수집하여 스케이링&정규화를 진행하고 box plot 등으로 가시화하여 이상치를 알아보며 summary를 사용해 해석해 보았음. MSE와 MAE를 이해해보고 변수선택법을 활용해 정확도 향상. - 담배핀 횟수, 고혈압기록, 당뇨병 기록 등으로 심장병에 걸릴지 예측하기 위해 데이터를 수집하여 범주형 데이터는 더미변수로 변환하고 스케이링&정규화 과정을 거쳐 correlation matrix등으로 가시화 함. Train모델/test모델 분류 후 Summary를 사용해 해석해 봄. Confusion matrix등으로 결과 해석. 변수선택법을 통해 정확도 향상 <p>이슈 및 해결 방안</p> <ul style="list-style-type: none"> - Confusion matrix를 제대로 이해하지 못해 나온 값이 숫자만 나와있어 무엇인지 몰랐음 => 다른 코딩방법으로 라벨 기록 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none"> - 프로야구팀 득점을 예측해보고 심장병을 머신러닝을 통해 예측해 볼 수 있는 시간이었다.

프로젝트 명 : 공사 입찰공고별 낙찰 여부 예측

기간	'22년 1월 ~ '22년 2월(2개월)
담당 역할	<p>팀원</p> <p>머신러닝(이진분류)</p> <ul style="list-style-type: none"> - 데이터 수집 - 데이터 전처리 - 분류모델 사용하여 최적의 모델 선정
사용언어 / 프로그램	Python, Colab
주요 기술	<ul style="list-style-type: none"> - 머신러닝(이진분류)
요약 내용	<p>공공조달 빅데이터 경진대회</p> <p>- '공사 입찰공고별 입찰업체 내역'자료를 이용해 낙찰 여부를 예측하는 모델 개발'</p>
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none"> - 공사 입찰공고별 낙찰 여부 예측 <p>상세 내용</p> <ul style="list-style-type: none"> - 학습에 필요한 데이터만 추출하여 포괄적인 데이터 분류를 위해 소분류→대분류로 컬럼을 수정. - 소수 레이블 중심으로 오버샘플링을 진행 후 SVM, Logistic 회귀, random forest, XG Boost로 학습. - 모델평가를 하기 위해 정확도, 오차 행렬, 정밀도, 재현율, F값, ROC 곡선을 나타냄. <p>이슈 및 해결 방안</p> <ul style="list-style-type: none"> - 소수 레이블 처리 => 오버샘플링 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none"> - 여러 학습모델과 평가를 사용해 보여 이해해보는 시간이 되었다.

프로젝트 명 : 축구선수 시장가치 평가

기간	'21년 11월 ~ '21년 12월(2개월)
담당 역할	팀원 분석 - 데이터 수집 - 데이터 전처리 - 그래프 해석
사용언어 / 프로그램	Python, Colab
주요 기술	- 데이터 분석 / 해석
요약 내용	빅데이터 언어 프로젝트 - '프리미어리그 경기 데이터를 활용해 축구선수 시장가치 평가'
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none">- 20/21 프리미어리그 경기 데이터를 활용해 축구선수 시장가치 평가 <p>상세 내용</p> <ul style="list-style-type: none">- 기본적으로 필요한 '시장가치', '골', '어시스트' 등의 데이터는 CSV 파일로 내려받음.- '슈팅' 데이터는 크롤링으로 가져옴.- 시장가치 변수를 y축으로 하고 다른 경기 기록 변수를 x축으로 하여 R 언어로 Linear Regression 그래프를 그려 보았고 OLS 회귀분석 결과를 보여 해석. <p>이슈 및 해결 방안</p> <ul style="list-style-type: none">- X축과 y축의 변수에 따라 그래프 모양과 원하는 해석이 달라져 고민 => 팀원들과 회의와 교수님 질의를 통해 해결 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none">- '득점 + 2*어시스트'의 R-squared값이 가장 높게 나와 이를 토대로 선수들의 시장가치가 고평가, 적정평가, 저평가.

프로젝트 명 : 하계 인턴십

기간	'21년 7월 ~ '21년 8월(2개월)
담당 역할	<p>팀원</p> <ul style="list-style-type: none"> - VBA코드 분석 - IoT 디바이스(드론) 동작 - 실시간 데이터 송수신(postman, Mendix)
사용언어 / 프로그램	Pyhon, Json, MindSphere, Mandix, mqtt, node-red
주요 기술	<ul style="list-style-type: none"> - 실시간 데이터 송수신
요약 내용	<p>하계 인턴십(휴먼디지털)</p> <ul style="list-style-type: none"> - 실시간 데이터 송수신
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none"> - 하계 인턴십 <p>상세 내용</p> <ul style="list-style-type: none"> - MindSphere를 사용해 VBA코드를 분석해보고 물리 asset, 논리 asset을 생성하여 fleet manager를 통해 가상 데이터 생성 & 출력되는지 확인 - Postman을 활용하여 post, get, put 등의 HTTP 통신 메시지 유형을 익혀 mindSphere에 데이터 송신 및 fleet manager 확인 - Mendix를 활용해 IoT와 application-고객관리, 생산관리 APP을 만들어 보고 MindSphere에 접근해 데이터 송신 확인 - IoT 디바이스(자율 주행 자동차, 드론)을 동작 제어해보고 mqtt와 node-red사용법을 배워 얻은 데이터를 송신하여 시각화 <p>이슈 및 해결 방안</p> <ul style="list-style-type: none"> - IoT 동작 오류 => Python 버전에 따른 오류 파악 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none"> - 실시간으로 데이터를 송수신한 것을 시각화하여 기록하는 경험을 할 수 있었다.

프로젝트 명 : 오픈소스가 뭐조?

기간	'21년 10월 ~ '21년 11월(2개월)
담당 역할	<p>리더</p> <ul style="list-style-type: none"> - Git 사용 - Github 기여 - 자료수집
사용언어 / 프로그램	Python, Git, Github
주요 기술	- Github 타 프로젝트에 기여
요약 내용	<p>오픈소스 프로젝트</p> <ul style="list-style-type: none"> - awesome-sushi, AN-EN-Tags, just-news에 기여
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none"> - 오픈소스가 뭐조? <p>상세 내용</p> <ul style="list-style-type: none"> - Awesome-sushi는 오마카세 맛집 정보 업로드 사이트로 기여 방법에 맞춰 Git으로 추가 - AN-EN-Tags는 IOS 사파리환경에서 언어 선택을 누를 수 있는 버그 발견해 버그 리포트 & 해결책 제시 - Just-news는 뉴스를 볼 때 불필요한 광고를 제거해주는 프로젝트로 업데이트 봇이 다운되어 현황표와 다르게 제대로 작동하지 않는 버그 발견 / just-news봇을 통해 스크립트 동작여부를 확인할 수 있는 테스트페이지 332줄 소스코드 추가 하여 이슈 폐쇄 <p>이슈 및 해결 방안</p> <ul style="list-style-type: none"> - 기존에 기여하려던 프로젝트 삭제 => 타 프로젝트로 변경 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none"> - Git과 Github를 사용해보며 장점과 단점을 파악하였고, 쉽게 타 프로젝트에 원하는 기여할 수 있었다.

프로젝트 명 : 백신 예약 프로그램

기간	'21년 10월 ~ '21년 12월(2개월)
담당 역할	리더 개발 - 예약 정보 찾기 - 예약 취소
사용언어 / 프로그램	C언어
주요 기술	- C언어로 링크드리스트를 활용
요약 내용	C언어 프로젝트 - 링크드리스트를 사용해 백신예약프로그램 개발
프로젝트 내용	<p>프로젝트 주제</p> <ul style="list-style-type: none"> - 백신 예약 프로그램 개발 <p>상세 내용</p> <ul style="list-style-type: none"> - 예약자 추가 / 예약자 정보 찾기 / 예약자 삭제 / 예약 일정 변경 / 예약 리스트 보기 기능 개발 <p>이슈 및 해결 방안</p> <ul style="list-style-type: none"> - 프로그램이 꺼지면 정보 리셋 => 이전 데이터를 파일로 저장해 불러옴 <p>프로젝트 결과 및 Lessons Learned</p> <ul style="list-style-type: none"> - C언어에 대해 공부하며 링크드리스트에 대해 이해해 볼 수 있는 시간이었다.