# **2021 데이터사이언스 ‘나만의 프로젝트’ 계획서**



|  |  |  |
| --- | --- | --- |
| 항목 | 내용 | 비고 |
| 팀이름 | 교수님 소고기 사주세요,,,, |  |
| 팀원이름 | 지은미  장희권 | 202001555  202001549 |
| 제목 | Bike Sharing Demand |  |
| 목적 (프로젝트로 해결하고 싶은 문제) | 날씨, 휴일 여부, 계절에 따른 자전거 대여량을 예측하는 모델을 설계 |  |
| 프로젝트 요약 | 데이터를 전처리와 시각화하고  결정 트리로 모델링하여  외부요인에 따른 자전거 수요 예측 |  |
| 모델 / 입력 / 출력 | Model : 결정 트리(Decision Tree)  Input : 시간, 휴일, 계절, 일하는 날, 날씨,  온도, 체감, 온도, 습도, 풍속  Output : 수요량 |  |
| 장점 / 독창성 | 결정 트리를 한 개만 쓰기에는 과적합이 되어 퍼포먼스가 낮게 나올 수 있다. 그러므로 여러 개의 결정 트리를 이용하여 예측하는 모델을 설계할 것이다. |  |
| 구현 기술 | 결정 트리(Decision Tree)로 모델 구성 |  |
| 데이터 확보 방법 / 데이터 특징 / 데이터 전처리 | 데이터 확보 방법 : Kaggle 대회 데이터셋 활용  데이터 특징 : 1년간 워싱턴 D.C.애서 날씨 및 시간, 기온, 습도, 풍속 등의 정보를 기반으로 1시간 간격의 자전거 대여 횟수를 기록하 데이터  데이터 전처리 : datetime을 연,월,일,요일,시간으로  분리해 새로운 feature로 만듦 |  |