|  |  |
| --- | --- |
| **[2019-데이터크롤링] 기말 과제 계획서** | |
| **작 성 자** | 산업경영시스템공학과 2017316012 노임령 |
| **프로젝트 명** | 술을 많이 마시는 학생의 성적이 더 나쁠까? |
| 프로젝트 목표 | 주중과 주말 음주량에 따른 기말 성적 예측 프로젝트 |
| 프로젝트 수행을 위한 학습내용 | 머신러닝 회귀분석 |
| 데이터 수집  계획 | kaggle의 student-alcohol-consumption 데이터를 사용한다. |
| 데이터 준비 및 탐색 작업 계획 | csv자료 읽어오기(출처 : kaggle의 student-alcohol-consumption중 student-por.csv만 사용 ) |
| 분석 모델 구축  계획 | 2개의 regression 모델을 사용한 Final Grade Prediction 실시(With G1 and G2 features)  -DecisionTreeRegressor와 LinearRegression 모델을 활용한 모델링 작업  -RandomForestClassifier  -LinearRegression  -Feature Importances |
| 결과 시각화  계획 | 데이터 구성  -조사 대상 학생들의 특성 시각화(Visualization)의 하위항목으로는 Sex, Weekly study time, Romantic Relationship, Age, Alcohol Consumption and other features가 있다.  -Alcohol Consumption and other features의 하위 항목으로는  Weekend Alcohol Consumption, Alcohol Consumption and Health, Alcohol Consumption and Final Grade가 있다. |

**※ 제출시, 파란색 글자는 삭제하고 제출할 것.**