

경진대회 기안서

제목	오일 상태 판단 모델 개발 (정상, 이상의 이진분류)
프로젝트 날짜	2022년 11월 14일 ~ 11월 25일
팀 이 름 / 구 성 원	알프스 팀 강보원 : 프로젝트 기획 및 데이터 분석 김동률 : 코드 및 데이터 분석 박혜빈 : 기획 및 데이터 분석 송승우 : 코드 및 데이터 분석
프 로 젝 트 개 요	DACON에서 개최하는 '건설기계 오일 상태 분류 AI 경진대회'에 참가하여 건설 장비 내부 기계 부품의 마모 상태 및 윤활 성능을 오일 데이터 분석을 통해 확인
프로젝트 선정 이유	1) 기존에 참여하는 프로젝트에만 그치지 않고, 공모전을 통해 데이터 분석 능력 향상 증진을 추구함. 2) 전처리와 분류를 기반으로 알고리즘의 실제 적용을 통해 데이터 분석을 하고자 함. 적절한 데이터를 찾던 도중 DACON이 실시하는 공모전에 참가함.
개발 환경	언어: Python 3.9.7 데이터 분석: numpy 1.20.3, pandas 1.3.4 OS (운영체제): Window 11, Window 10 IDE: Anaconda3 22.9.0, Jupter lab 3.9.7 System: Memory 20G, Hard 256GB
개발 모델 및 알고리즘	1. 데이터 셋 : 오일 Train 데이터, 오일 Test 데이터, oil_info(오일정보) 데이터, sample_submission(타겟값 저장용) 2. 알고리즘 앙상블 - Voting, RandomForest, lightGBM, XGBoost 교차검증 - K-fold, stratified K-fold 파라미터 조정 - GridSearch, hyperOPT 회귀(이진분류) - LogisticRegression 3. 평가지표 : macro F1 score 4. 사용라이브러리 : numpy, pandas, sklearn, scipy
개발 일정	주자 1주차 2주차 2주차 11월 14일 15일 16일 17일 18일 19일 20일 21일 22일 23일 24일 25일 25일 23일 24일 25일 21일 22일 23일 24일 25일 25일 25일 25일 25일 25일 25일 25일 25일 25