

# 업데이터 DACON 분석 스터디

[ 2022 업데이터 여름방학 프로젝트 ]

**스터디장** 이제운 (데이터과학과 21)

**전화번호** 010 2210 9133

## Description

이 스터디에서는 다양한 머신러닝 방법론(분류 및 회귀)과 이와 관련된 DACON 코드 분석 및 Python 라이브러리 사용법을 다룹니다. DACON 및 kaggle은 데이터 분석 경진 대회 사이트로써 다양한 데이터와 양질의 코드가 공유되어 있어 머신러닝을 이용한 데이터 분석을 다루기에 적절합니다. 따라서 개인 코딩 과제나 팀 코드 분석에서 DACON과 Kaggle을 활용할 예정입니다.

이 스터디는 이론 강의를 바탕으로 코드를 공부하고 팀 발표까지 진행합니다. 방학 기간 6주 동안 주 1회 2시간씩 비대면 줌으로 진행되며, 이전 주차 내용에 대한 복습 과제와 개인 코딩 과제 및 팀 코드 분석 과제가 부여될 예정입니다. 마지막 6주차에 팀별 분석 결과 발표를 마지막으로 스터디가 마무리될 예정입니다.

## Expectations and Goals

1. 데이터 분석 과정(pipeline)을 알고 python으로 구현할 수 있다.
2. 다양한 분류 및 회귀 머신러닝 모델을 이해한다.
3. python을 이용해 머신러닝 모델들의 성능을 비교할 수 있다.

## Reference

부족한 설명이나 궁금증을 해결하는 데 도움을 줄 수 있는 참고 자료들입니다.

- (optional) Hands on Machine Learning
- (optional) 파이썬을 이용한 통계적 머신러닝

스터디는 사전에 제공되는 pdf 자료와 google colab 코드와 함께 진행됩니다.

## Project Schedule

임시 계획입니다. 수업은 이론을 위주로 진행되고 실습은 주로 과제를 통해 진행됩니다.

Week	Topic	Keyword	Assignment
[Week 1] (7/11~7/15)	OT: 통계적 머신러닝 이란?	MLE, MSE	복습 및 예습 과제
[Week 2] (7/18~7/22)	데이터 분석 과정	Pipeline, 전처리, 모델링	개인 코딩 과제
[Week 3] (7/25~7/29)	Classification 모델	Logistic regression, Naive Bayes Classifier, KNN, SVM, Decision tree	개인 코딩 과제
[Week 4] (8/1~8/5)	Regression 모델	Linear Regression, Ridge, Lasso, ElasticNet	팀 코드 분석 과제
[Week 5] (8/8~8/12)	Ensemble learning	Voiting, Boosting, Bagging, XGBoost, LightGBM, RandomForest	팀 코드 분석 과제
[Week 6] (8/13~8/19)	팀 코드 분석 결과 발표		

저도 배우는 입장이어서 가르치는 것에는 많이 부족하겠지만, 제가 아는 것을 공유할 좋은 기회  
라고 생각해서 DCON 분석 스터디장을 맡게 되었습니다! 많은 지원 부탁드립니다, 7월 둘째 주에  
ZOOM에서 OT 및 수업으로 뵙겠습니다. 감사합니다. 😊