

## 『 최종 프로젝트 수행 계획서 』

제 출 일	2021.10.15	담당 강사	임동조
팀 명	멋쟁이 승우처럼	팀 인원	4명
기술 스택	머신러닝 알고리즘 (데이터 탐색, 모델 구축 및 평가) - 지도학습 : Regression, Decision Tree, Ensemble (RandomForest, lightgbm, xgboost 등) - 비지도학습 : K-means, PCA 등 - 시각화 라이브러리 : matplotlib, seaborn		
프로젝트명	DACON – 가스공급량 수요예측 모델개발		

### ■ 프로젝트 목표 및 개요

#### [대회명]

DACON – 가스공급량 수요예측 모델개발

#### [대회 링크]

<https://dacon.io/competitions/official/235830/overview/description>

#### [주제]

한국가스공사의 시간단위 공급량 내부데이터와 기상정보 및 가스 외 발전량 등 외부데이터를 포함한 데이터셋을 구축하여 90일 한도 일간 공급량 예측하는 인공지능 모델 개발

#### [주최 및 주관]

한국가스공사

#### [대회 개요]

한국가스공사가 보유한 데이터 및 노하우와 빅데이터, 인공지능 분야 4차 산업 기술 융합으로 새로운 가치 창출 및 다년간 시간 단위 공급량 데이터를 기반으로 미래 공급량 예측 모델 생성

#### [외부 데이터]

<https://data.kma.go.kr/cmmn/main.do>

#### [프로젝트 목표]

- 지금까지 배운 내용을 토대로 실전 분석 대회에 참여하여 팀원 전체의 실전 역량 향상
- 데이터를 시각화를 통해 데이터의 특징점을 찾아 데이터를 한눈에 확인할 수 있도록 분석
- 데이터가 의미하는 바를 쉽게 확인할 수 있도록 시각화하여 데이터 활용도 향상
- 가스 공급량을 예측하는 머신러닝(또는 딥러닝) 모델을 구축 및 평가
- 다양한 모델의 비교를 통해 최적의 모델을 구축 및 적용

## ■ 역할 분담

성 명	분담 내용	역 할
문승우	데이터 전처리, 추가 변수 생성, 모델 성능 개선 - merging dataset and processing missing value - feature engineering (using statistics, model) - evaluate model (using Cross Validation, Grid Search)	팀장
정진우	모델 성능 향상, 발표 - compare Machine Learning Model - evaluation metrics – by MSE, RMSE, finally evaluated by NMAE - Cross Validation and Parameter Tuning	부팀장
오소영	데이터 시각화, 발표 자료 준비, PPT 자료 정리 - Matplotlib, seaborn 등을 활용한 데이터 탐색 및 시각화	팀원
강수정	데이터 탐색 및 수집	팀원

## ■ 예상 결과물

- 데이터 시각화 결과물 (PPT 정리)
- 모델별 비교 평가 결과 (PPT 정리, EXCEL 파일 등)
- 모델 파라미터 튜닝 결과 (PPT 정리)
- 최종 정리 결과물 (PPT 정리)
- [Github 소스 코드 및 자료](#)