

『 최종 프로젝트 수행 계획서 』

제 출 일	2021.10.15	담 당 강 사	임동조
팀 명	멋쟁이 수정처럼	팀 인 원	4명(문승우, 정진우, 오소영, 강수정)
기 술 스 택	- 머신러닝 알고리즘(데이터 탐색, 모델 구축 및 평가) - [지도학습] - 회귀, 의사결정트리, 앙상블(랜덤포레스트, lightgbm, xgboost) - [비지도학습] - K-means, PCA 등 - 시각화 라이브러리(matplotlib, seaborn)		
프로젝트명	데이콘 - 가스공급량 수요예측 모델개발 대회에 참여		

■ 프로젝트 목표 및 개요

[대회명] : 가스공급량 수요예측 모델개발 대회

[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 공급량을 예측하는 인공지능 모델을 개발

[주최 및 주관] 한국가스공사

[대회 개요] 한국 가스 공사가 보유한 다년간 시간 단위 공급량 데이터를 기반으로 미래 공급량을 예측하는 모델을 만든다.

[외부 데이터] <https://data.kma.go.kr/cmmn/main.do>

[프로젝트 목표]

01 지금까지 배운 내용을 토대로 실전 분석 대회에 참여하여 팀원 전체가 실전 역량을 업그레이드한다.

02 가스 공급량을 예측하는 머신러닝(또는 딥러닝) 모델을 구축하고 평가, 그리고 다양한 모델의 비교

■ 역할 분담

성 명	분담 내용	역 할
문승우	데이터 전처리, 추가 변수 생성, 모델 성능 개선 - 데이터 셋 병합 및 결측치 처리 - 피처 엔지니어링(통계량을 이용한 특성 선택, 모델을 이용한 특성 선택) - 모델 평가(교차 검증, 그리드 리서치 함수 등을 활용)	팀장
오소영	데이터 시각화, 발표 자료 준비 - matplotlib, seaborn 등을 활용한 데이터 탐색 및 시각화 - ppt 자료 정리	부팀장
정진우	모델 성능 향상, 발표 - 머신 러닝 모델 비교 평가 지표 - MSE, RMSE를 활용, 최종적으로 NMAE로 평가 - 교차 검증 및 파라미터 튜닝	팀원
강수정	데이터 수집 - 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 자료