

<div> <div>프로젝트 과정</div> <div>&lt; 파일럿 프로젝트 기획서 &gt;</div> </div>			
주 제	호주 강우 예측		
분 야	기후변화 팀	일 자	2022년 09월 21일
팀 장	김태찬 <ohapjjin135@gmail.com>	팀 원	김태찬

## 1. 과제 개요

호주의 날씨 데이터를 활용해 다음날 비가 올 것인지를 예측하는 과제이다.

## 2. 데이터 수집 및 분석

- 데이터 셋 이름  
weatherAUS.csv(14.09 MB)
- 데이터 셋 출처  
<https://www.kaggle.com/datasets/jsphyg/weather-dataset-rattle-package>

## 3. 프로젝트 계획(일정, 수행과제, 목표 등)

- 목표  
주어진 데이터를 활용하여 호주의 다음날 비가 올 확률을 구하는 모델을 생성하고, 데이터 처리 및 모델 개선을 통해 오차를 최소화한다.
- 데이터셋  
23개의 columns  
**Train**  
Date : 관측 일자  
Location : 지역 이름  
MinTemp : 최저 온도(섭씨)  
MaxTemp : 최고 온도(섭씨)  
Rainfall : 일일 강우량(mm)  
Evaporation : 증발량(mm)  
Sunshine : 하룻동안 햇빛이 비추는 시간  
WindGustDir : 하룻동안 가장 강한 바람이 분 방향  
WindGustSpeed : 하룻동안 가장 강한 풍속(Km/h)  
Etc...  
**Target**  
RainTomorrow : 다음날 비가 왔는가?(Yes, No)

- 순서

1. EDA(탐색적 데이터 분석)
2. Data Featurig / 전처리
3. 데이터 시각화
4. 모델링
5. 평가
6. 결론

- 역할 분담 및 일정 계획

- 결과 도출 방법 및 가설

본문제는 다음날의 비가 온다 또는 안 온다. 즉 Target Variable이 범주형 이므로 지도학습 분류 알고리즘들을 적용해 볼 수 있다.

KNN, 의사결정트리, 앙상블\_랜덤포레스트 등의 알고리즘들을 적용해 결과를 도출 해 볼 수 있을 것을 생각된다.