

2025 스마트농업 AI 경진대회

- 사전테스트 문제설명서 -

1. 출제 배경

2025 년 스마트농업 AI 경진대회 예선 진출팀 선정을 위해 프로그래밍 역량 및 농업 도메인에 대한 이해도 평가를 목적으로 문제 출제

- 농업 분야에서의 탄소 중립 달성은 전 세계적인 과제로, 온실 내 탄소배출량 정밀 예측 기술을 필요로 함
- 온실에서 수집되는 다양한 데이터를 이용해 하루 간의 탄소배출량 예측 모델 개발을 목적으로 함

2. 사전테스트 문제 정의

딸기 재배 온실에서 수집된 환경/제어/생육 데이터를 이용해 1 일 단위의 온실 전체 탄소배출량의 정밀 예측 모델 개발

- x_test.csv 의 1 일 단위 구분값(Sample_Number)별 CO₂ 배출량(kg)을 예측

3. 데이터 설명

가. x_train.csv

- 1 일 단위 구분값(Sample_Number)별 10 분 단위의 환경/제어/생육 데이터

나. y_train.csv

- x_train.csv 의 1 일 단위 구분값(Sample_Number)별 온실 CO₂ 배출량(kg) 데이터

다. x_test.csv

- x_train.csv 와 동일

라. 데이터셋 상세

● 환경 데이터

번호	항목	변수명	주기	비고
1	내부온도	in_temp	10 분	
2	내부습도	in_hum	10 분	
3	내부 CO ₂	in_co2	10 분	
4	외부온도	out_temp	10 분	
5	외부습도	out_hum	10 분	
6	일사량	solar_rad	10 분	
7	풍속	wind_speed	10 분	
8	풍향	wind_direction	10 분	
9	강우감지	rain_sensor	10 분	감지 1, 미감지 0

● 생육 데이터

번호	항목	변수명	주기	비고
1	초장	plant_height	1 회/주	관부에서 가장 긴 잎의 선단까지의 길이
2	관부직경	crown_diameter	1 회/주	원줄기 관부에서 가장 굵은 부분의 길이
3	엽병장	petiole_length	1 회/주	관부에서 초장을 측정한 잎의 하단까지의 길이(초장-엽장)
4	엽수	leaf_count	1 회/주	한 개체에서 완전히 전개된 잎의 수
5	엽장	leaf_length	1 회/주	완전히 전개된 최근 3 번째 잎을 대상으로 하며, 엽육이 시작되는 지점부터 잎 끝부분까지 길이
6	엽폭	leaf_width	1 회/주	완전히 전개된 최근 3 번째 잎을 대상으로 하며, 가장 폭이 넓은 부분 측정(엽장과 수직 방향)
7	착과수	fruit_count	1 회/주	과실의 수
8	화방꽃수	flower_count	1 회/주	화방에 달린 꽃의 수

● 제어 데이터

번호	항목	변수명	주기	비고
1	유동팬	fan	10 분	작동 1, 중지 0
2	CO ₂ 발생기	co2	10 분	작동 1, 중지 0
3	난방기	heater	10 분	작동 1, 중지 0
4	천창 1	window1	10 분	구동기 현재 위치(0-100%)
5	천창 2	window2	10 분	구동기 현재 위치(0-100%)
6	커튼 1	curtain1	10 분	차광, 구동기 현재 위치(0-100%)
7	커튼 2	curtain2	10 분	차광, 구동기 현재 위치(0-100%)
8	커튼 3	curtain3	10 분	보온, 구동기 현재 위치(0-100%)
9	측커튼	side_curtain	10 분	구동기 현재 위치(0-100%)

4. 성능 평가

가. 평가 지표

- 제출된 predict 값과 true 값을 비교하는 아래 2 가지 지표를 이용해 성능 평가 수행
 - RMSE (Root Mean Squared Error, 평균 제곱근 오차)

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i)^2}$$

- R^2 (R-squared, 결정 계수)

$$R^2 = 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (y_i - \bar{y})^2}$$

나. 평가 지표 코드

- 평가 지표 코드 구현 방식에 따른 계산 결과 오차를 줄이기 위한 기준 코드 정의

- RMSE

```
import numpy as np
def rmse(y_true, y_pred):
    y_true = np.array(y_true)
    y_pred = np.array(y_pred)
    return np.sqrt(np.mean((y_true - y_pred) ** 2))
```

- R^2

```
import numpy as np
def r2_score(y_true, y_pred):
    y_true = np.array(y_true)
    y_pred = np.array(y_pred)
    ss_res = np.sum((y_true - y_pred) ** 2)
    ss_tot = np.sum((y_true - np.mean(y_true)) ** 2)
    return 1 - (ss_res / ss_tot)
```

- 제출 답안에 예측 값이 아닌 NaN, None, inf 등이 포함되는 경우, 오류로 간주하여 일정 패티를 부과
 - $\text{패널티} = (\text{해당 데이터셋에서 발생 가능한 최대 오차값}) \times (\text{오류값 개수})$

다. 평가 점수 계산식

- 2 가지 지표 평가 결과는 8:2 비율의 등수에 따른 절대값(%)으로 환산 후 점수 계산

■ 점수 환산식 예시(40 팀 참가 시)

RMSE (80%)			
등수	값	등수	값
1	80	21	40
2	78	22	38
3	76	23	36
4	74	24	34
5	72	25	32
6	70	26	30
7	68	27	28
8	66	28	26
9	64	29	24
10	62	30	22
11	60	31	20
12	58	32	18
13	56	33	16
14	54	34	14
15	52	35	12
16	50	36	10
17	48	37	8
18	46	38	6
19	44	39	4
20	42	40	2

R ² (20%)			
등수	값	등수	값
1	20	21	10
2	19.5	22	9.5
3	19	23	9
4	18.5	24	8.5
5	18	25	8
6	17.5	26	7.5
7	17	27	7
8	16.5	28	6.5
9	16	29	6
10	15.5	30	5.5
11	15	31	5
12	14.5	32	4.5
13	14	33	4
14	13.5	34	3.5
15	13	35	3
16	12.5	36	2.5
17	12	37	2
18	11.5	38	1.5
19	11	39	1
20	10.5	40	0.5

■ 점수 계산 예시

- ◆ A 팀: RMSE 10 등, R² 5 등 → 62+18=80 → A 팀 점수 80 점
- ◆ B 팀: RMSE 5 등, R² 10 등 → 72+15.5=87.5 → B 팀 점수 87.5 점
- ◆ C 팀: RMSE 4 등, R² 2 등 → 74+19.5=93.5 → C 팀 점수 93.5 점

5. 제출 방법

가. 데이터셋 공개

- 데이터셋: 대회 공식 홈페이지 공개(2025 년 7 월 30 일 수요일 0 시 이후) 및 다운로드
- 스마트농업 AI 경진대회 홈페이지(<https://agricchallenge.ai/>)
→ 대회운영 → 사전테스트

나. 제출 방법

- 제출 기간: 문제 공개후 ~ 2025 년 7 월 31 일 목요일 23:59:59
- E-mail 제출: agricchallenge25@gmail.com
- '다'에 기술된 제출 양식에서 요구되는 전체 파일을 zip 형식으로 압축 후 제출
 - 제출 파일명: {팀명}.zip (예: 스마트팜최고.zip)
- 기한 내 수신된 최종 제출물을 기준으로 평가

다. 제출 양식

(1) 학습 및 추론 코드 파일

a. train.py (또는 ipynb)

- Train 데이터를 로드한 후 모델 학습, 모델 가중치 파일을 저장하는 구조의 코드 파일 제출
- Input: x_train.csv, y_train.csv
- Output: model.확장자

b. inference.py (또는 ipynb)

- Test 데이터와 모델 가중치 파일을 로드해 예측을 수행하고, 결과를 csv 파일로 저장하는 코드 파일 제출
- Input: x_test.csv, model.확장자

- Output: submission.csv

(2) test 예측 파일

a. submission.csv

- x_test 의 각 'Sample_Number'에 대응하는 'CO2_final' 값을 예측 후 제출
- y_train.csv 파일과 동일한 구조

(3) 모델 가중치 파일

a. model 파일 (h5, pkl, pt, 또는 기타 프레임워크별 표준 모델 파일)

- train.py 파일에서 저장, inference.py 파일에서 로드

(4) 모델 및 전처리 설명 자료

a. 모델개발전략.pdf

- 개발한 탄소배출량 예측 모델의 구조, 특징, 주요 전처리 방식 등을 1 페이지 이내로 요약한 자료를 pdf 로 변환 후 제출
- 권장 작성 항목: 데이터 전처리/특성 선택 방법, 모델 아키텍처 및 주요 하이퍼파라미터, 성능 개선 및 과적합 방지 전략, 기타 본인 팀만의 주요 전략/설명 등

[제출 파일 구조 예시]

스마트팜최고.zip

├─ train.py

├─ inference.py

├─ submission.csv

├─ model.pkl

└─ 모델개발전략.pdf

- 제출 시 '다'에 기술된 파일명 엄수((1),(2),(4)는 확장자 변경 허용)

6. 문의 안내

문의처

- 카카오톡 채널 문의: http://pf.kakao.com/_MKzxin

문의 답변 가능 시간

- 2025 년 7 월 30 일 수요일 9:00~18:00
- 2025 년 7 월 31 일 목요일 9:00~18:00