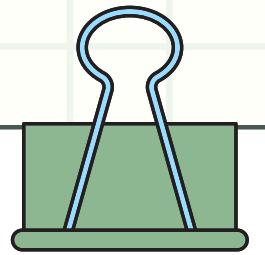


2025 전북대학교 컴퓨터인공지능학부  
산학실전캡스톤 프로젝트 예선 발표

가우리안 확산 모델을 활용한 인공지능 기반  
공장 및 발전소 배출가스의  
도시 대기 영향 분석

**TEAM GAIA**  
**(GAUSSIAN AI AIR-IMPACT ANALYSIS)**  
김세원 김수민 박은송 전찬민



# 목차

## 01 서론

1-1. 과제 배경

1-2. 과제 목적

1-3. 과제 계획

## 02 본론

2-1. 사용 데이터 및 모델

2-2. 프로젝트 구조 설계

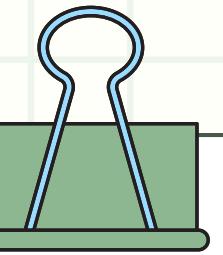
2-3. 프로젝트 과정

## 03 결론

3-1. 결과 분석

3-2. 개선 방향 및 결론

3-3. 기대 효과 및 전망



# 1-1 과제 배경 산업화와 대기오염 문제의 심각성

**“국내 석탄발전 오염물질로  
1983년부터 최대 1만3천명 조기사망”**

[에너지청정대기연구센터 한국 보고서]

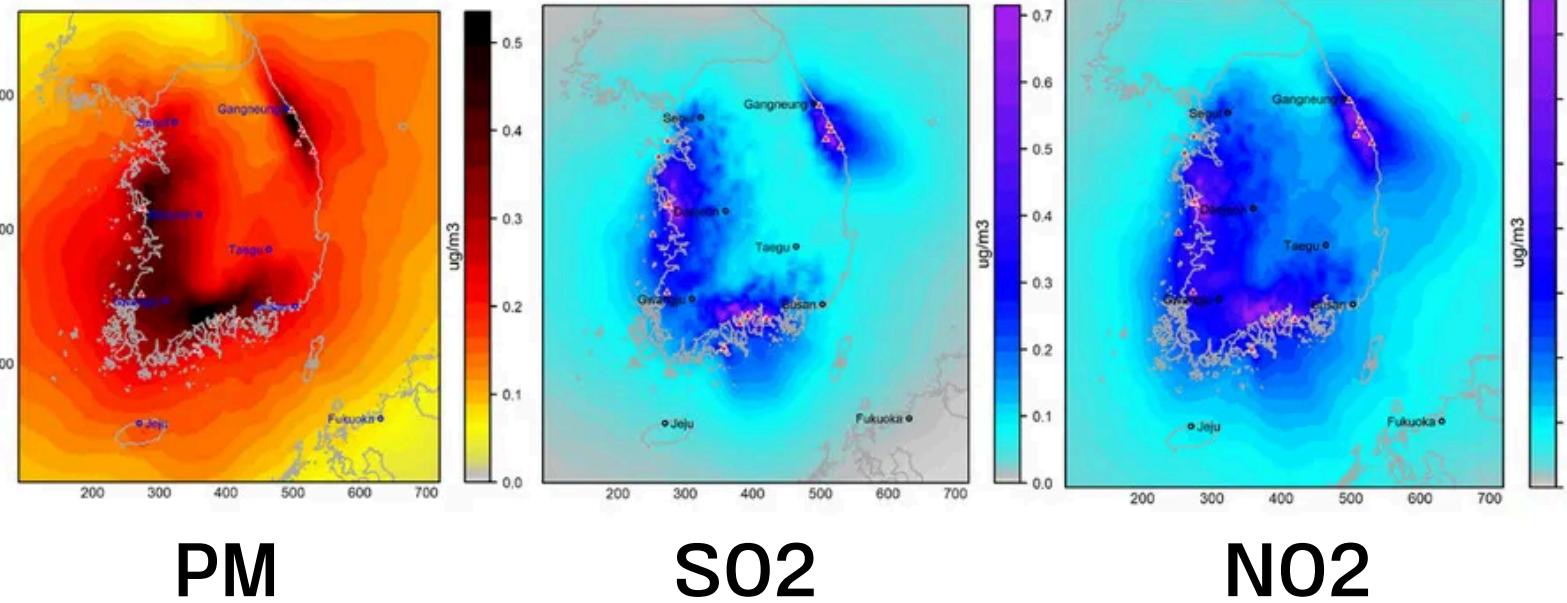
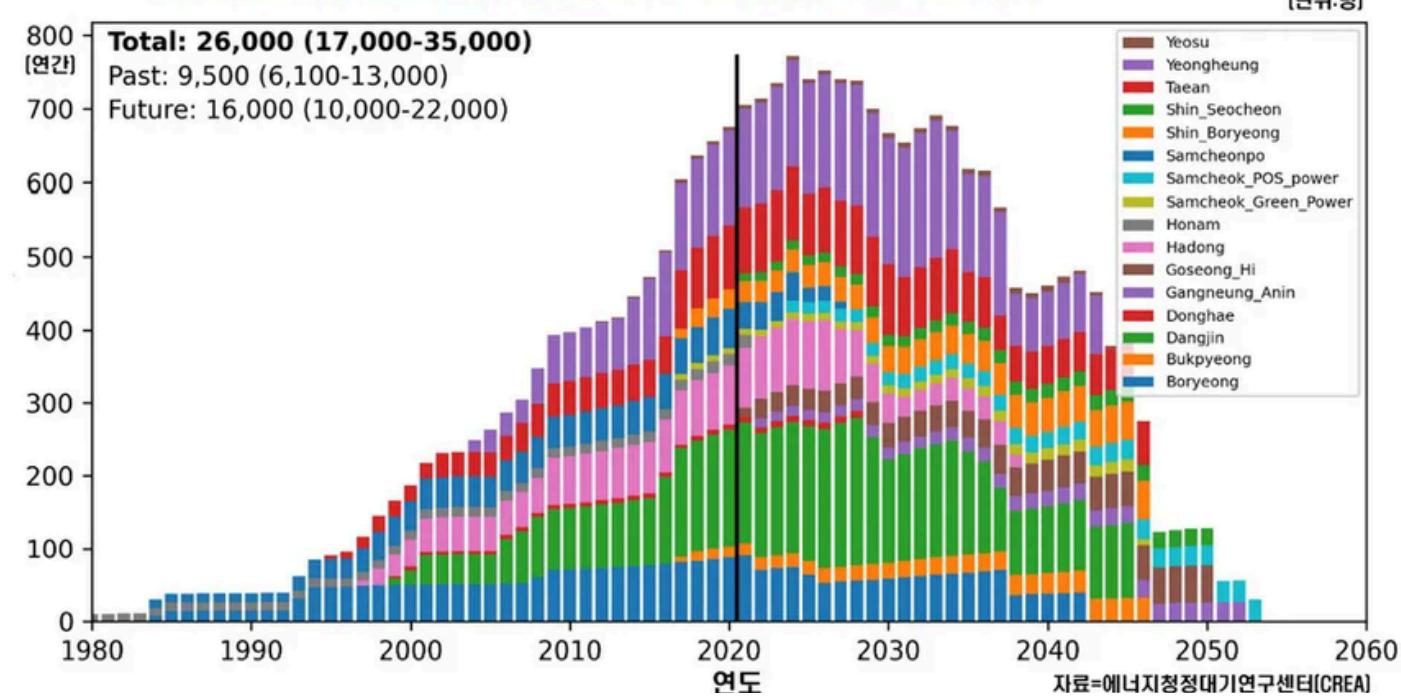
신규발전 7기 가동 멈추는 2054년까지

건강피해·노동생산성 감소 손실액만 58조원

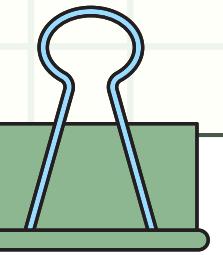
최대 3만5천명 조기 사망할 수 있어

이근영 기자

■ 석탄발전소 배출 대기오염물질에 의한 조기사망[1983~2054년]

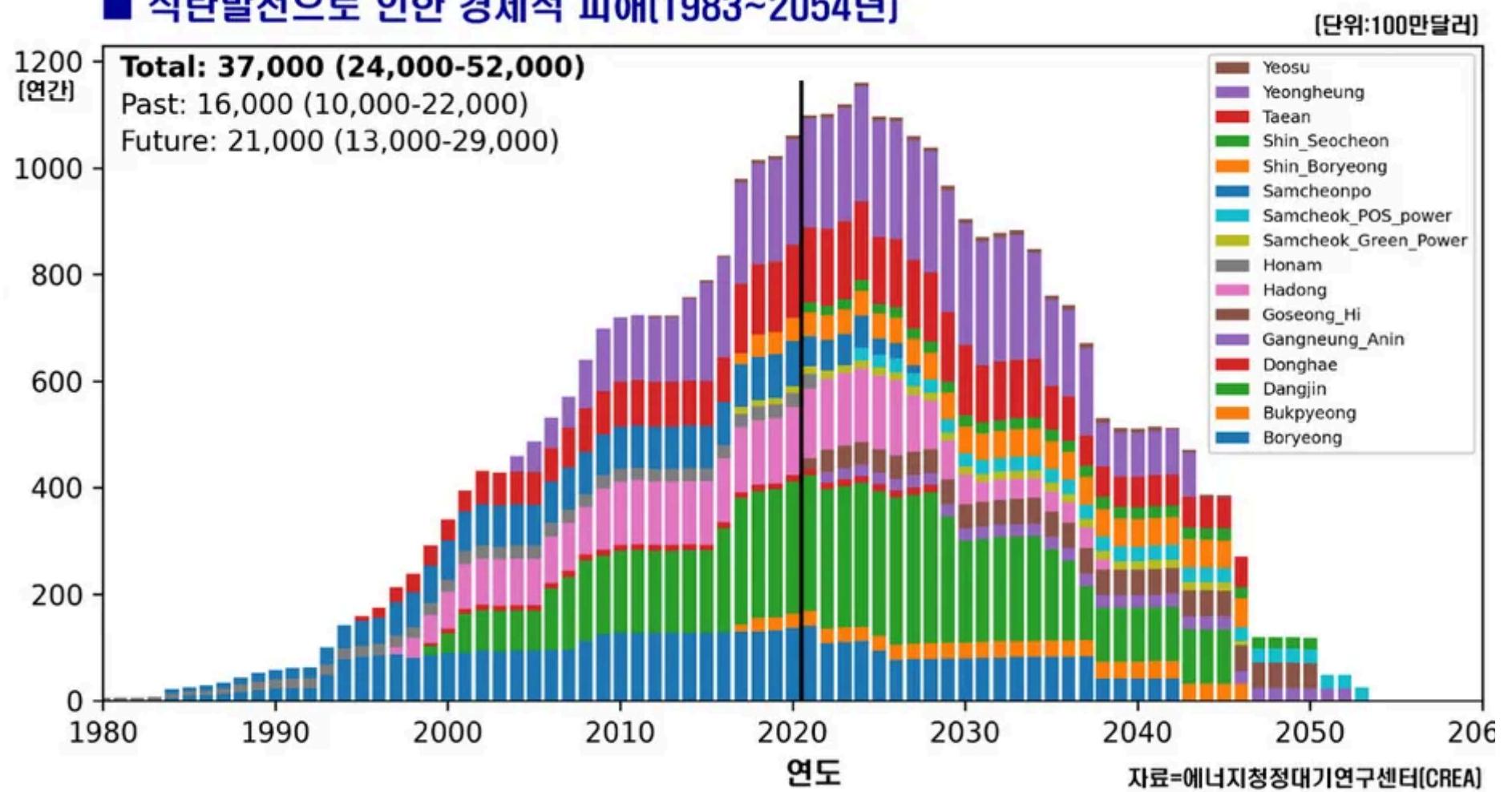


← 석탄발전소 배출 대기오염물질  
연간 평균 농도



# 1-1 과제 배경 사회적·경제적 피해 비용

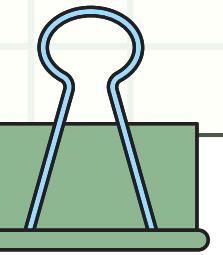
## ■ 석탄발전으로 인한 경제적 피해[1983~2054년]



1983년~2020년까지  
석탄발전으로 인한 사회적 비용:

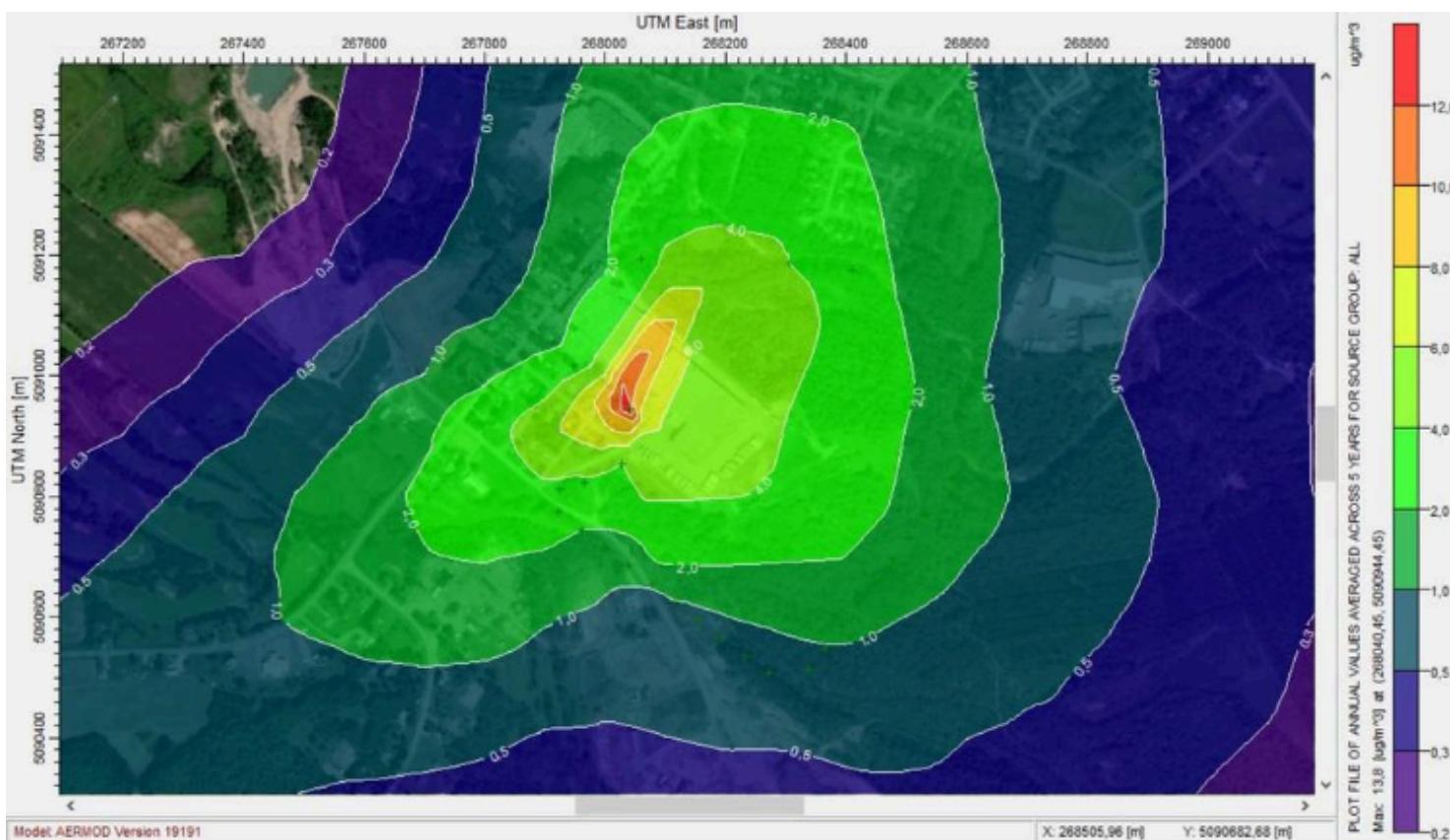
"17조9천억원"

- 간접 피해: 의료비 증가, 노동력 상실, 생산성 저하
- 단순 사후 대응 → 예측 기반 정책으로 전환 필요
- 오염 확산 경로 분석 및 실시간 대응체계 중요



# 1-1 과제 배경

## 기존 모델(AERMOD)의 한계/AI 기반 접근의 필요성

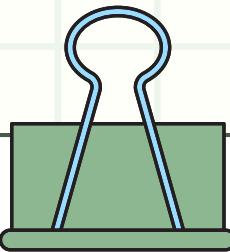


[AERMOD 실행 결과 예시]

전통적인 가우시안 확산 모델  
**AERMOD의 한계**

- 정적인 입력값 기반 → 실시간 예측 어려움
- 국지적 환경 변화(풍향, 기상 조건) 반영 미흡
- 실제 확산과 예측 결과 간 오차 발생

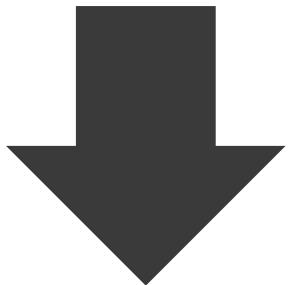
→ 더 정확하고 실시간 예측이 가능한 모델 필요



## 1-2 과제 목적

$$c(x, y, z, He) = \frac{Q}{2\pi\sigma_y\sigma_z U} \times \exp\left[-\frac{1}{2}\left(\frac{y}{\sigma_y}\right)^2\right] \times [\exp\left(-\frac{1}{2}\left(\frac{z-He}{\sigma_z}\right)^2\right) + \exp\left(-\frac{1}{2}\left(\frac{z+He}{\sigma_z}\right)^2\right)]$$

가우시안 확산 모델

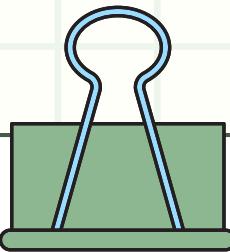


AERMOD

+ Transformer Model

### AI기반 모델의 장점

- 실시간 데이터 반영 가능
- 더 정밀한 예측 가능
- 대기질 개선 정책 수립을 위한  
데이터 기반 의사결정에 기여 가능



# 1-3 과제 계획

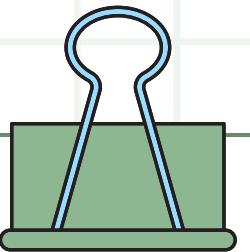
공장 및 발전소 배출 오염물질의 도시 대기 영향 정량 분석

실시간 대응 가능한 AI 기반 확산 예측 시스템 구축

정책 수립 및 환경 규제 대응을 위한 시각화 기반 플랫폼 개발

특정 지역(보령·신보령·신서천) 실측 데이터 기반 예측 모델 검증

지도 기반 웹 대시보드를 통한 오염농도 시각화 및 사용자 인터페이스 구현



# 2-1 사용 데이터 및 모델 사용 데이터

## 굴뚝자동측정기기 (TMS)

- 수집 항목:** TSP, SO<sub>2</sub>, NOx, CO, HCl, HF, NH<sub>3</sub> 등 총 7개 주요 오염물질
- 활용 목적:** 배출원별 오염물질 배출 특성 분석 모델 입력값으로 활용

## 대기질 측정소

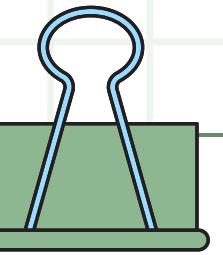
- 수집 항목:** PM10, PM2.5, O<sub>3</sub>, NO<sub>2</sub>, CO, SO<sub>2</sub> 등
- 활용 목적:** 확산 결과에 따른 도심 대기질 변화 모니터링 및 정답 데이터로 사용

## 기상 데이터

- 주요 항목:** 풍속, 풍향, 온도
- 활용 목적:** 오염물질 확산 경로에 영향을 주는 핵심 환경 변수로 입력

## 시계열

- 수집 기간:** 2018년 ~ 2024년 (7년간)
- 활용 목적:** 시계열 기반 Transformer 예측 모델 학습에 사용, 시간대별 패턴 분석



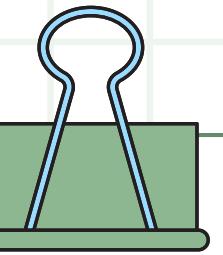
# 2-1 사용 데이터 및 모델 사용 데이터

1	api_data
2	tms_보령_f1
3	tms_보령_nox
4	tms_보령_o <sub>2</sub>
5	tms_보령_sox
6	tms_보령_tmp
7	tms_보령_tsp
8	tms_신보령_f1
9	tms_신보령_nox
10	tms_신보령_o <sub>2</sub>
11	tms_신보령_sox
12	tms_신보령_tmp
13	tms_신보령_tsp
14	tms_신서천_f1
15	tms_신서천_nox
16	tms_신서천_o <sub>2</sub>
17	tms_신서천_sox
18	tms_신서천_tmp
19	tms_신서천_tsp
20	upload_log

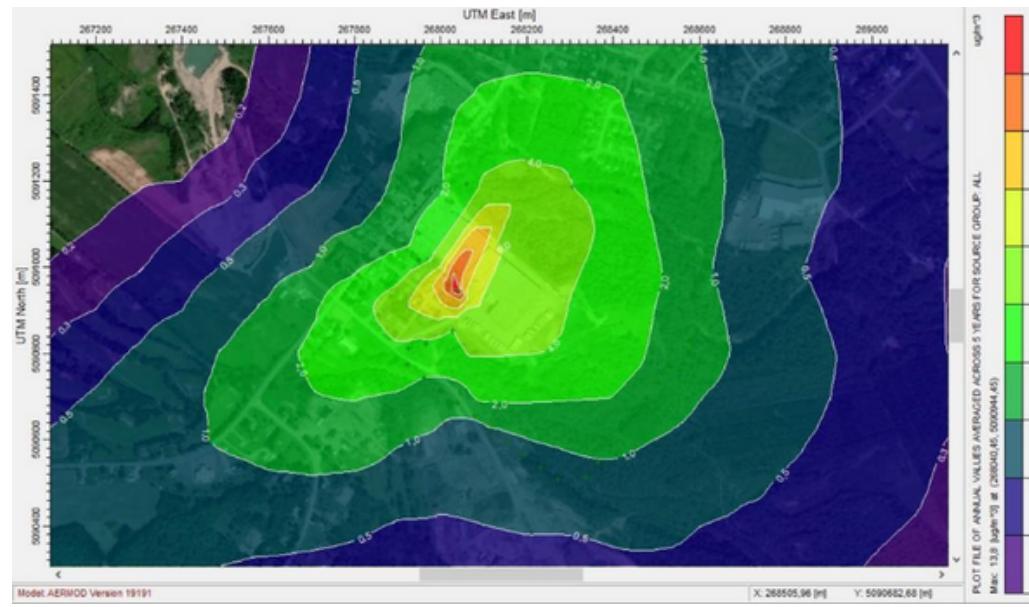
1	신보령	2022	F1	2022_F1.csv
2	신보령	2022	O <sub>2</sub>	2022_O <sub>2</sub> .csv
3	신보령	2022	TMP	2022_TMP.csv
4	신보령	2022	TSP	2022_TSP.csv
5	신보령	2022	S0x	2022_S0x.csv
6	신보령	2022	NOx	2022_NOx.csv
7	신보령	2025	TMP	2025_TMP.csv
8	신보령	2025	TSP	2025_TSP.csv
9	신보령	2025	S0x	2025_S0x.csv
10	신보령	2025	NOx	2025_NOx.csv
11	신보령	2025	F1	2025_F1.csv

region	measure_date	speed	direction
보령	2020-01-01 01:00:00	0.4	0
보령	2020-01-01 02:00:00	0.4	0
보령	2020-01-01 03:00:00	0.2	0
보령	2020-01-01 04:00:00	0.1	0
보령	2020-01-01 05:00:00	1.4	200
보령	2020-01-01 06:00:00	0.1	0
보령	2020-01-01 07:00:00	0	0
보령	2020-01-01 08:00:00	0.1	0
보령	2020-01-01 09:00:00	1.3	50
보령	2020-01-01 10:00:00	0	0
보령	2020-01-01 11:00:00	0	0
보령	2020-01-01 12:00:00	0.7	180

Tables_in_weatherCenter
bout_wind2018
bout_wind2019
bout_wind2020
bout_wind2021
bout_wind2022
bout_wind2023
bout_wind2024
processed_weather_2018
processed_weather_2019
processed_weather_2020
processed_weather_2021
processed_weather_2022
processed_weather_2023
processed_weather_2024

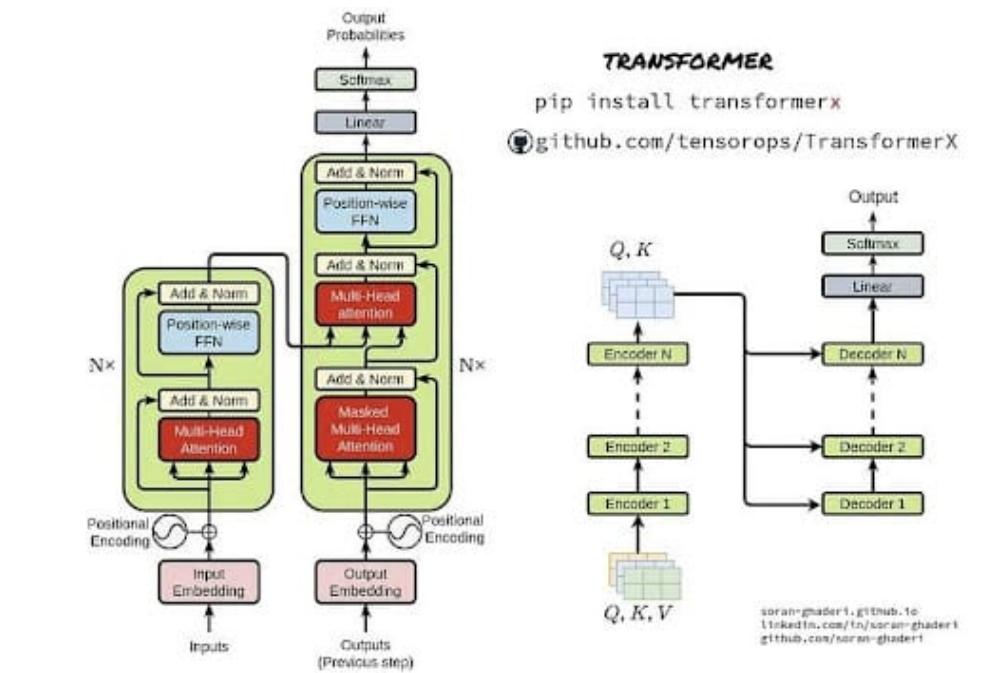


# 2-1 사용 데이터 및 모델 사용 모델



## AERMOD (가우시안 확산 모델)

- 미국 EPA에서 개발한 대표적인 대기 확산 모델
- 굴뚝(배출원)에서 배출된 오염물질이 바람과 기상 조건에 따라 확산 예측

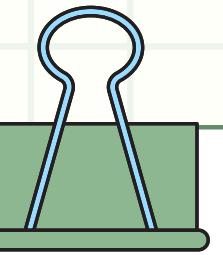


## Transformer 모델

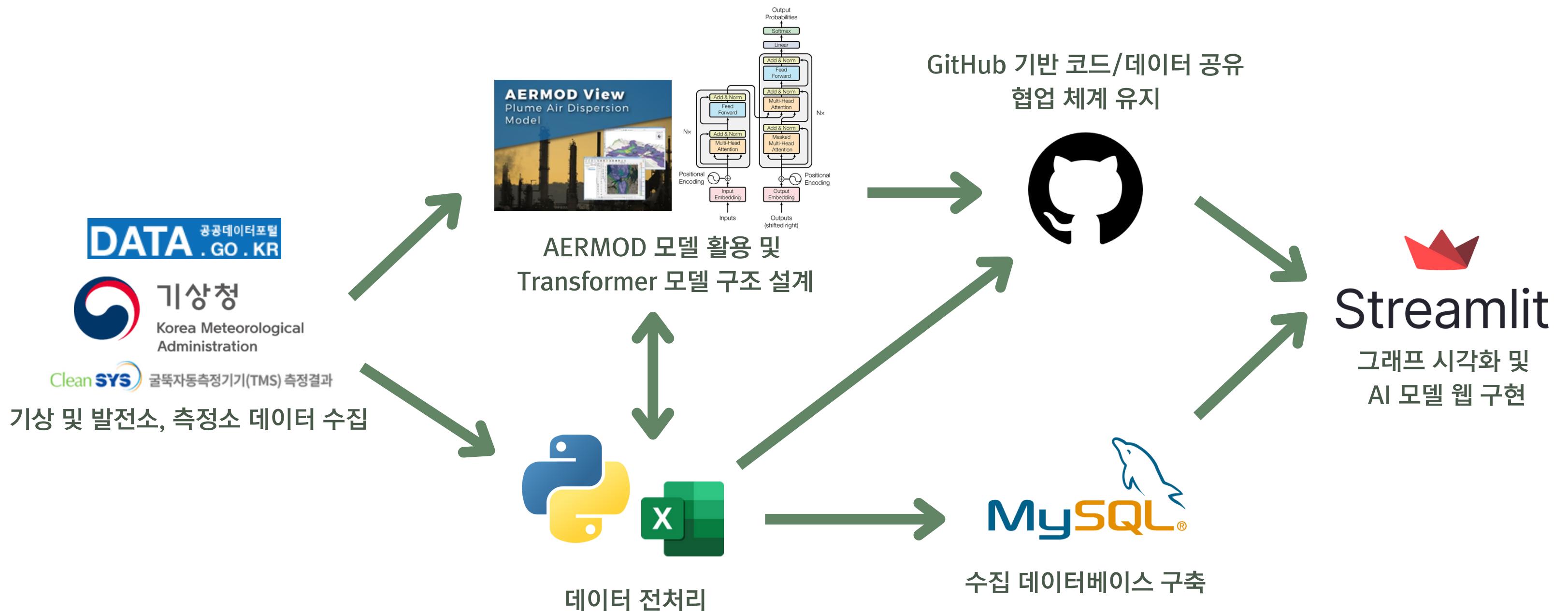
- 시계열 데이터 예측에 적합한 딥러닝 모델 구조
- 기존 확산 모델의 한계(고정 계수, 단일 시점 추정 등)를 보완
- AI 기반으로 확산 경로의 패턴과 상관관계까지 반영 가능

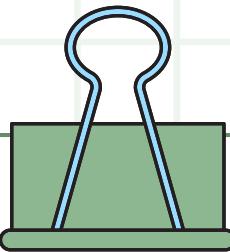
기존의 실시간 기반 AERMOD 모델 + 트랜스포머 모델(AI 학습)

→ 전통 확산 모델의 정밀도와 AI 모델의 시계열 학습 능력을 융합하여 실시간 대기질 예측 성능을 향상



## 2-2 프로젝트 구조 설계





# 2-3 프로젝트 과정

## 전처리 및 확산 모델 적용

```
def preprocess_sheet_long_format(df):
    """
    [단일 시트 전처리] - Long Format
    1. 불필요한 열 제거
    2. 대체값이 존재하면 측정값 보완
    3. 측정일시, 측정항목 기준으로 pivot (Wide-format 변환)
    4. 시간단위 평균으로 변환
    5. 결측값 보간
    """

    # 열 이름 통일
    df.columns = df.columns.str.strip().str.replace('\n', '').str.replace(' ', '')

    # 유효 열만 선택
    df = df[['측정일시', '측정항목', '측정값', '대체값']]

    # datetime 변환
    df['측정일시'] = pd.to_datetime(df['측정일시'], errors='coerce')
    df = df.dropna(subset=['측정일시', '측정항목'])

    # 대체값 우선 반영
    df['값'] = df['대체값'].combine_first(df['측정값'])

    # 시간 내림 (30분 → 시간 단위)
    df['측정일시'] = df['측정일시'].dt.floor('H')

    # 피벗 (항목별 열 생성)
    pivoted = df.pivot_table(index='측정일시', columns='측정항목', values='값', aggfunc='mean')

    # 결측값 선형 보간 + 앞뒤 채움
    pivoted = pivoted.interpolate(method='linear').fillna(method='bfill').fillna(method='ffill')

    pivoted.reset_index(inplace=True)
    return pivoted
```

전처리 과정

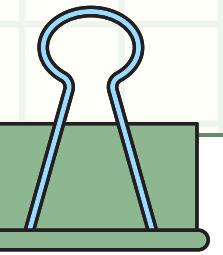
```
class GaussianPlumeModel:
    def __init__(self, Q, u, H, sigma_y, sigma_z):
        self.Q = Q
        self.u = u
        self.H = H
        self.sigma_y = sigma_y
        self.sigma_z = sigma_z

    def concentration(self, x, y, z=0):
        Q = self.Q
        u = self.u
        H = self.H
        sigma_y = self.sigma_y
        sigma_z = self.sigma_z

        part1 = Q / (2 * np.pi * u * sigma_y * sigma_z)
        part2 = np.exp(-y**2 / (2 * sigma_y**2))
        part3 = np.exp(-(z - H)**2 / (2 * sigma_z**2))
        part4 = np.exp(-(z + H)**2 / (2 * sigma_z**2))
        C = part1 * part2 * (part3 + part4)
        return C

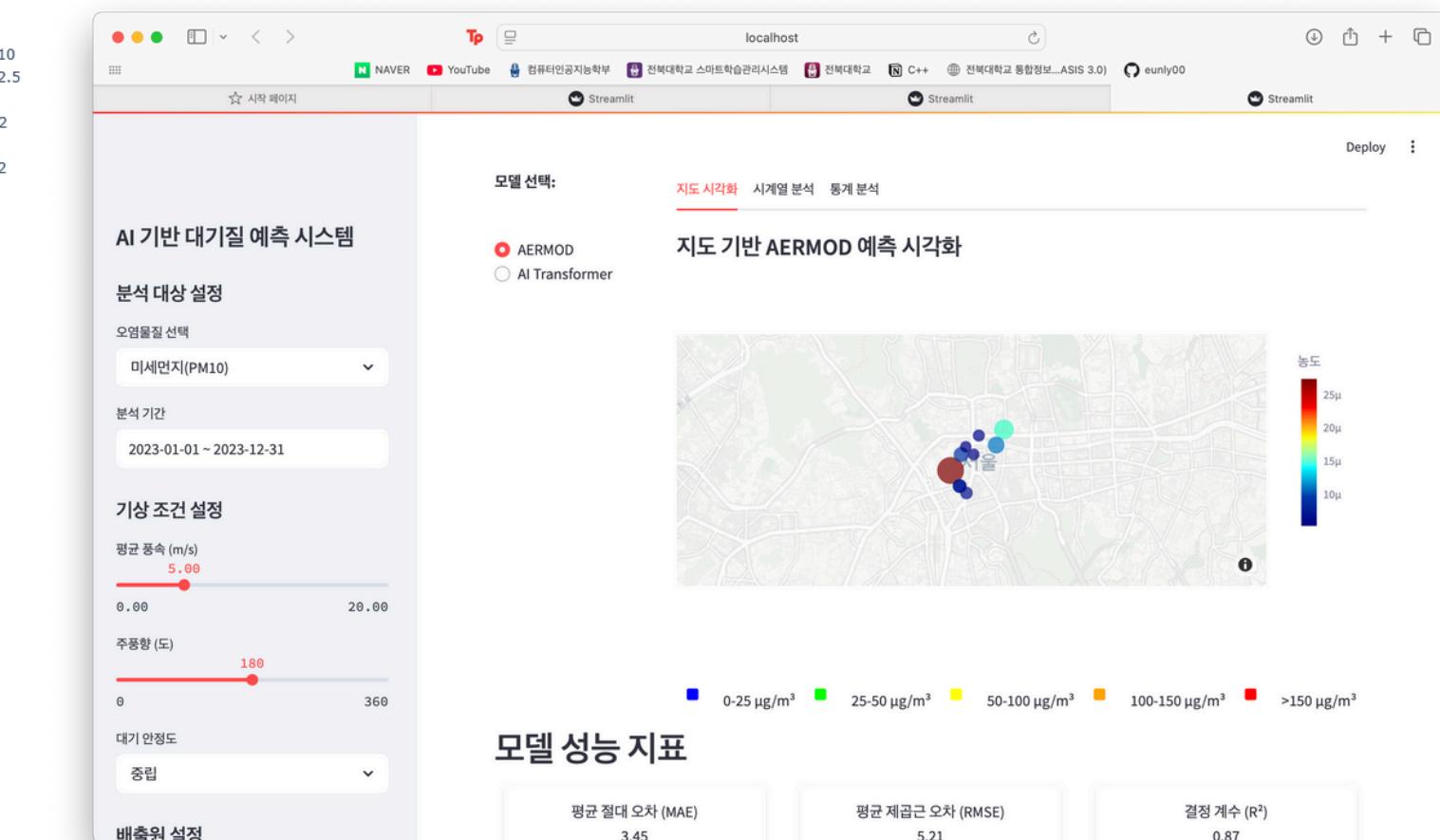
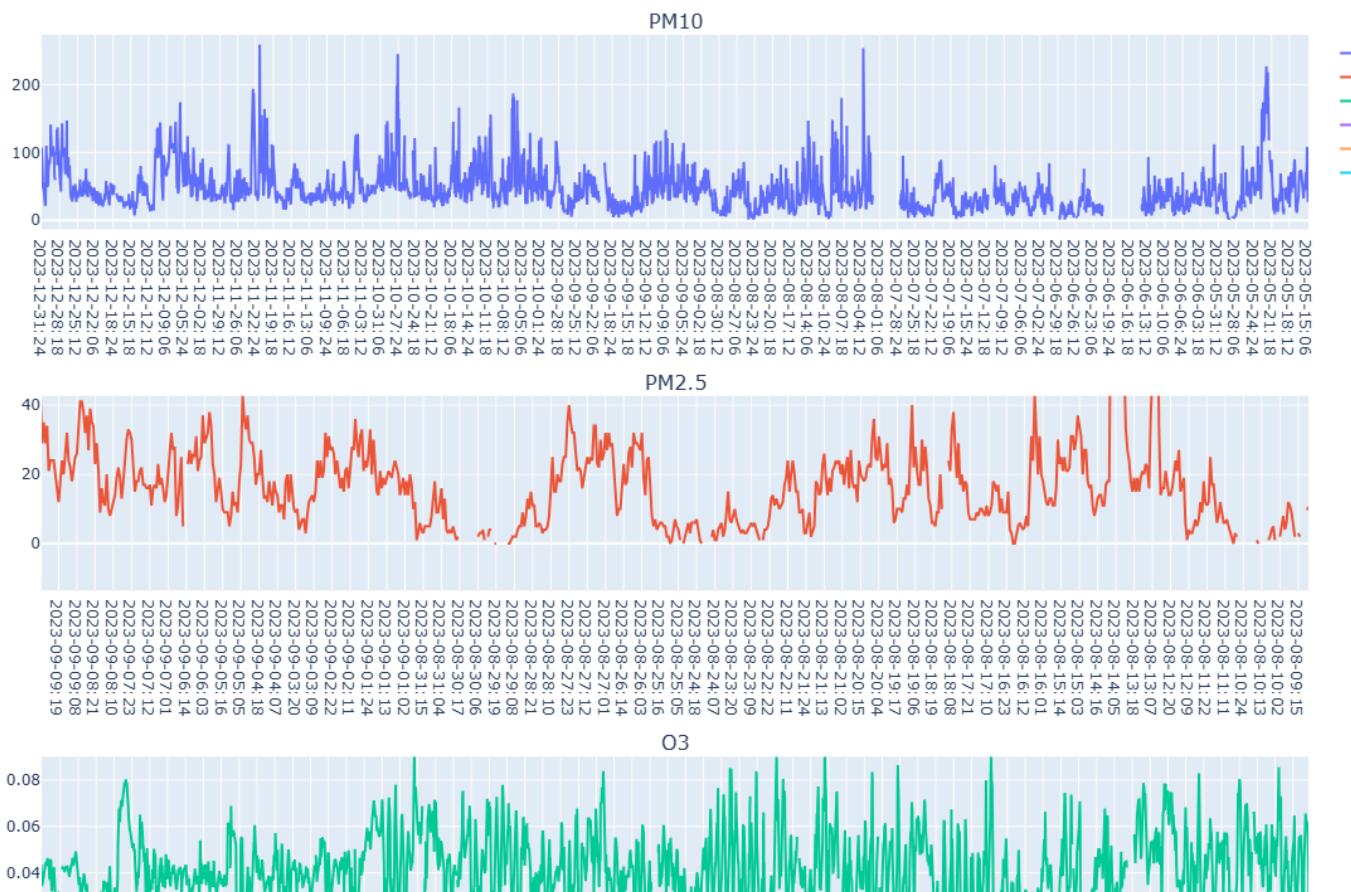
    def batch_concentration(self, points):
        results = []
        for pt in points:
            x, y, z = pt['x'], pt['y'], pt.get('z', 0)
            c = self.concentration(x, y, z)
            results.append({'x': x, 'y': y, 'z': z, 'concentration': c})
        return pd.DataFrame(results)
```

확산 모델 계산(AERMOD)

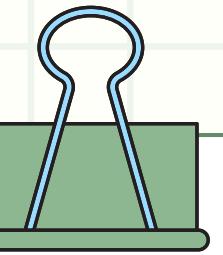


# 2-3 프로젝트 과정 웹 대시보드 시각화

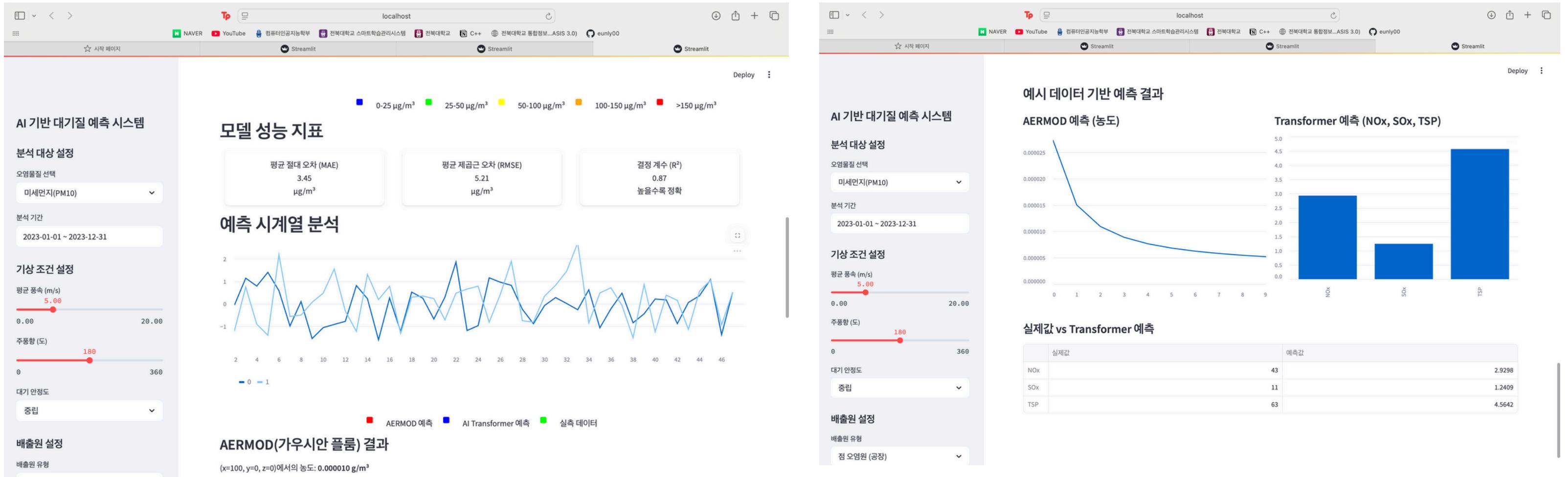
BoryeongPort2023



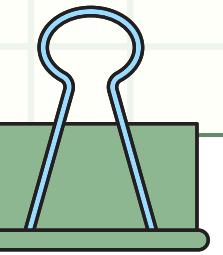
사용자 상호작용 기반, 기존 오염물질 배출 정도 그래프화  
→ 발전소 별 오염물질 배출 정도 확인



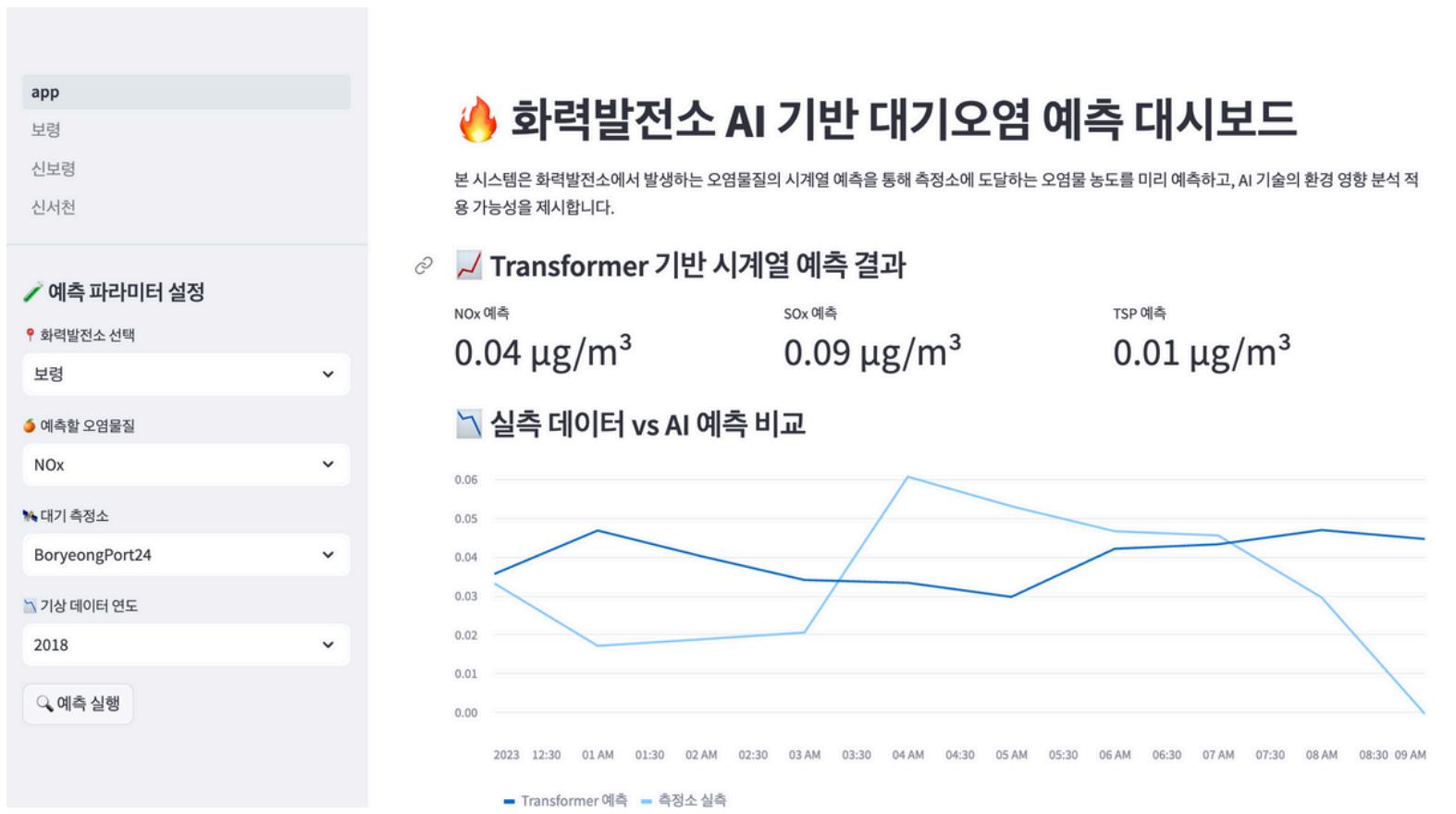
# 2-3 프로젝트 과정 웹 대시보드 시각화



사용자 상호작용 기반, 기존 오염물질 배출 정도 그래프화  
→ 발전소 별 오염물질 배출 정도 확인



# 3-1 구현 예측 대시보드



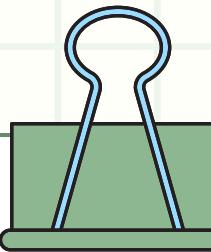
예측 성능 평가 (평가지표)

	오염물질	MAE ( $\mu\text{g}/\text{m}^3$ )	RMSE ( $\mu\text{g}/\text{m}^3$ )
0	NOx	3.2	4.1
1	SOx	2.8	3.6
2	TSP	1.5	2.1

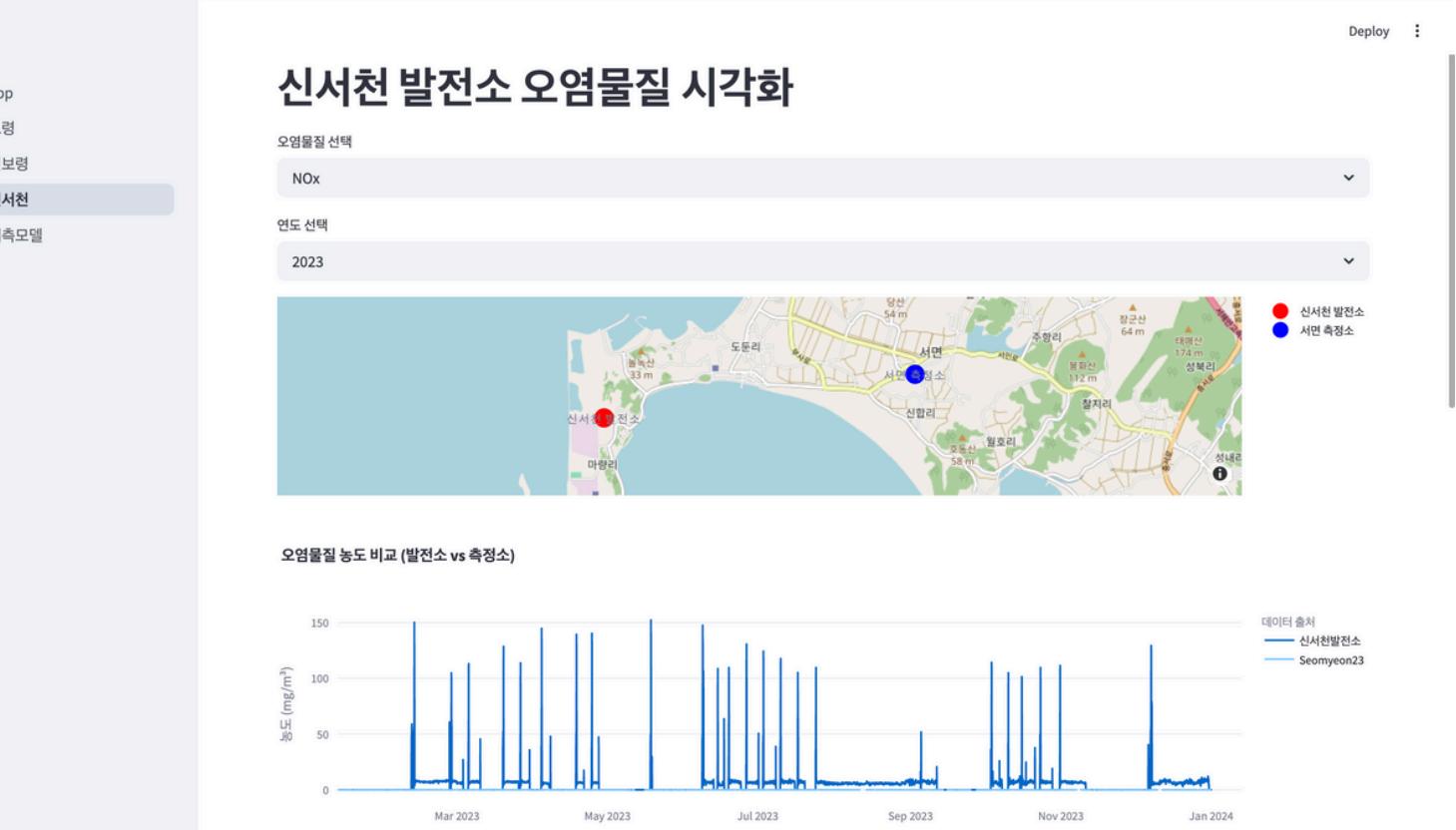
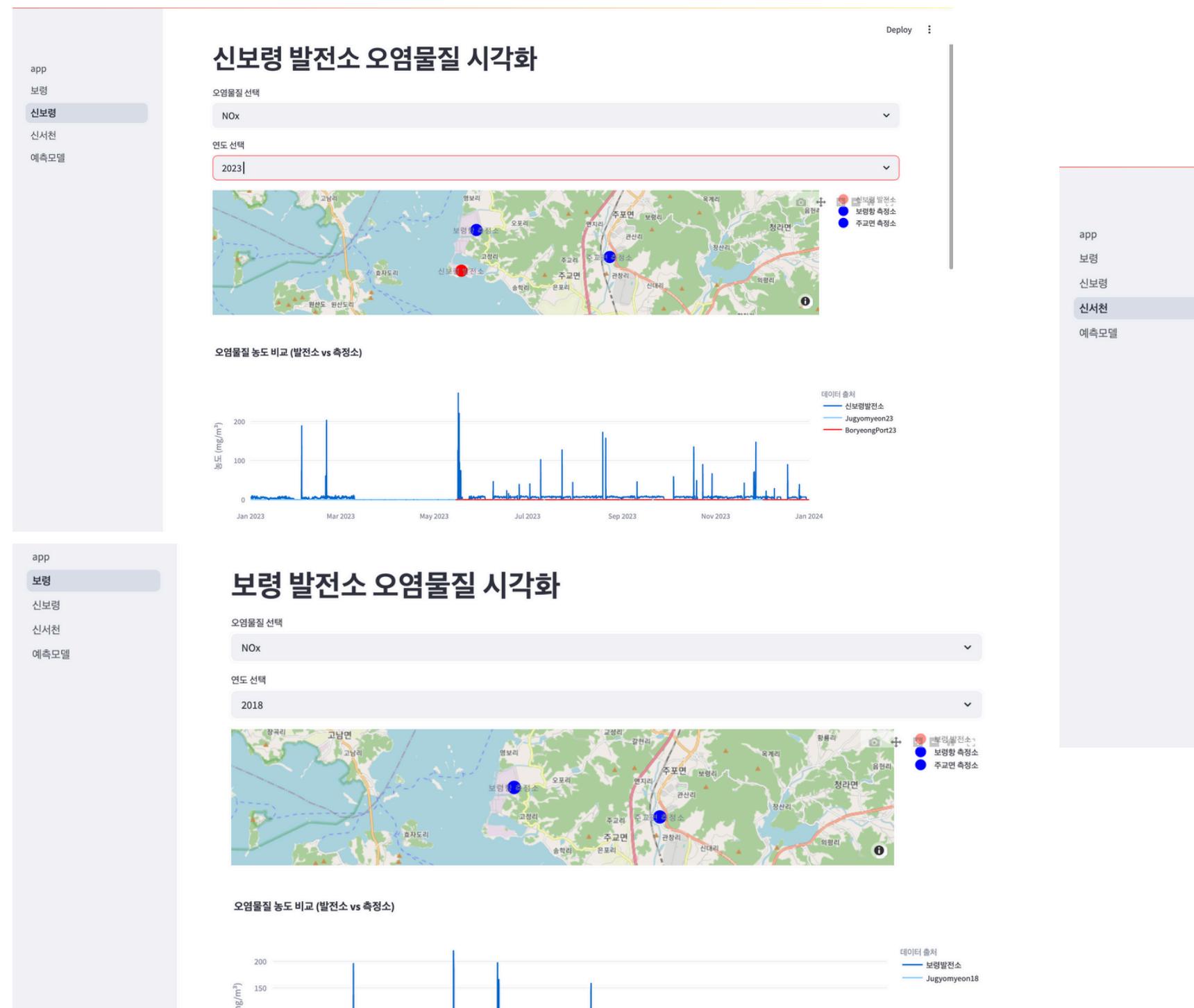
→ 과거 데이터 기반

기존 AERMOD vs 트랜스포머 모델 학습 결과 vs 실측 데이터 비교 결과 확인

- 과거 데이터를 기반으로 AERMOD 구조에 맞게 **트랜스포머 모델 학습**
- 학습 결과 기반으로 다음과 같은 파라미터 조정하여 예상 예측 결과 분석 가능
  - 화력발전소
  - 대기중 오염물질
  - 대기 측정소
  - 기상 데이터 연도

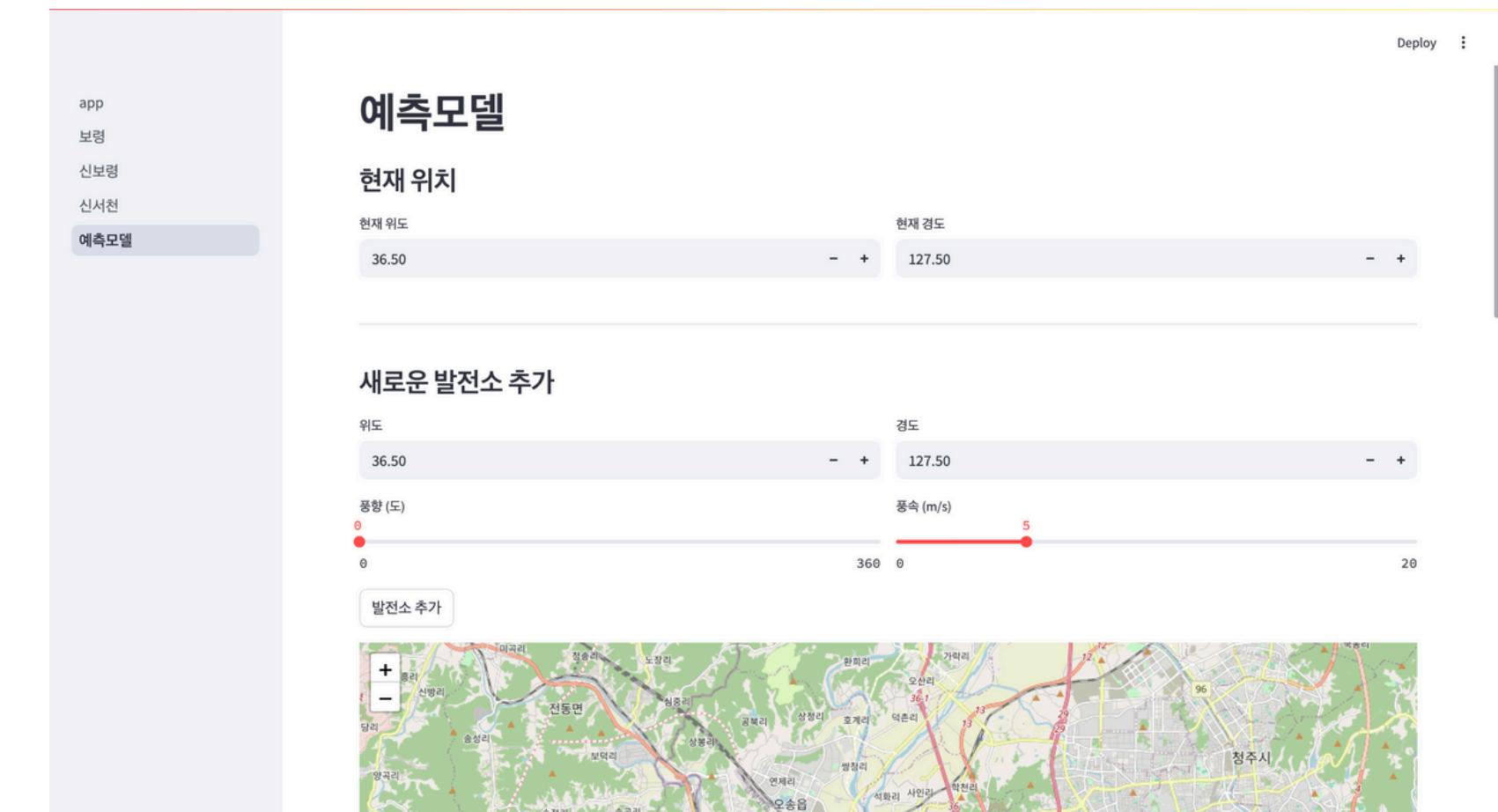
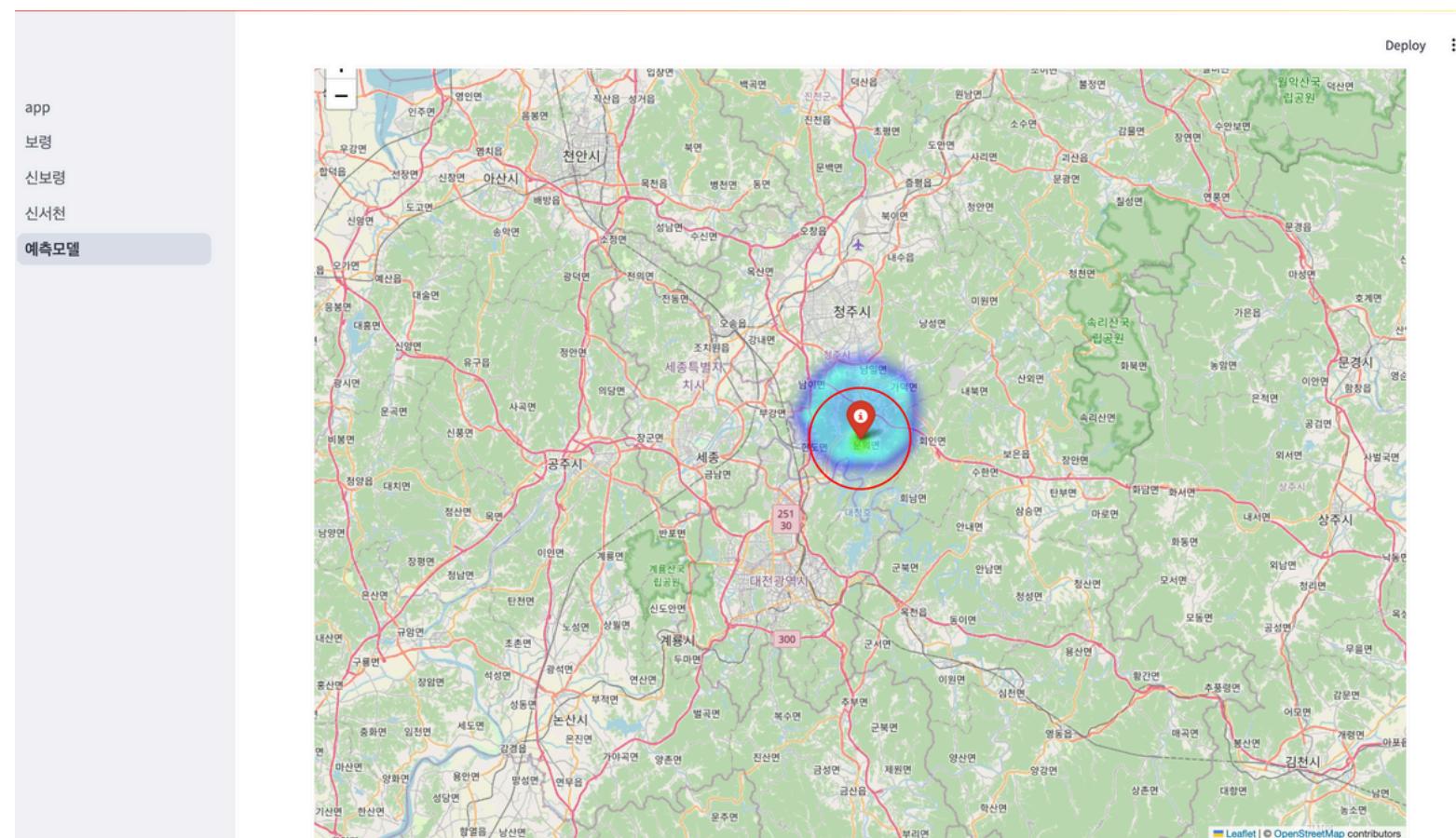


# 3-1 구현 측정 데이터



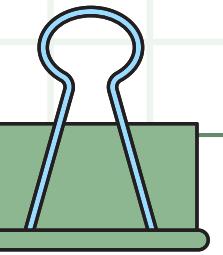
가우시안 확산 모델을 활용한 인공지능 기반 공장 및 발전소 배출가스의 도시 대기 영향 분석

# 3-1 구현 예측모델



- 학습 결과 기반으로, 추후 국내 화력발전소가 추가될 경우  
오염물질 확산 정도를 예측할 수 있는 시스템 구축

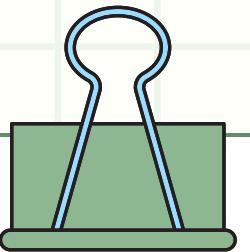
- 추가될 발전소의 위치 정보, 풍향, 풍속을 고려  
→ 대기 확산 정도 예측 가능



# 3-1 구현 시연영상



가우시안 확산 모델을 활용한 인공지능 기반 지역 공장 및 발전소 배출가스의 도시 대기 영향 분석



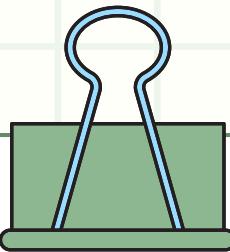
## 3-2 개선 방향 및 결론

### 1. 주요 성과 요약

- AI 기반 가우시안 확산 예측 모델 개발
- 실시간 기상 및 대기질 데이터를 반영한 정밀 예측 가능
- 오염물질의 도시 확산 영향 정량 분석

### 2. 기술적 기여

- AERMOD + 트랜스포머 모델 융합을 통한 예측 정확도 향상
- 지형·기상·오염 데이터를 통합한 분석 파이프라인 구축
- 실시간 시각화 및 해석 가능한 예측 모델 구현

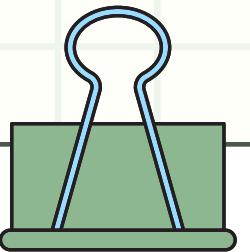


## 3-3 7|대 효과 및 전망

환경정책 및 규제 평가를  
위한 과학적 분석 도구 제공

산업 지역과 도시 간  
대기오염 확산 문제 해결 지원

탄소중립 및 기후대응  
정책 기반 마련



# 팀원 소개



김세원

팀 리더

전체 프로젝트 관리 및  
인공지능 모델 연동 구축



박은송

웹 구현 및 그래프화

웹 페이지 그래프 구현  
및 백엔드 관리



전한민

데이터 수집 및  
모델 설계

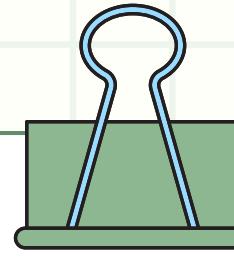
웹 크롤링 구현 및  
수집 데이터 관리



김수민

데이터 전처리 및  
모델 설계

데이터 기반 인공지능  
모델 구조 설계



# 감사합니다

# THANK YOU!