

부산지역 상세 격자별 대기오염도 생성을 통한 동네예보 확대 방안에 관한 연구

I 연구목적 및 필요성

- 광화학수치모델(CMAQ)에서 필연적으로 발생하는 오차를 개선하고 예측 정확도를 향상시키기 위해 관측치를 사용한 머신러닝 기반의 비선형 회귀모델을 구축하고자 함
- 이를 위해 현재 우리원에서 운영 중인 27개 도시대기측정소 대기오염 자료를 활용하여 부산지역 1km² 상세 격자 단위의 광화학수치모델 결과를 개선하기 위한 연구를 수행함
- 1km² 상세 격자별 모델 결과를 개선하여 미세먼지 농도 예측의 공간적 해상도와 예측 정확도를 향상시켜 보다 신뢰성 있는 정보를 시민들에게 제공하고자 함

II 연구개요

- 기 간 : 2025년 1월 ~ 12월(1년)
- 대 상 : 2021년 ~ 2024년간 (초)미세먼지 예측모델링 자료 및 부산 27개 도시대기측정소 실측자료
- 방 법 : 예측자료와 실측자료로 학습시킨 머신러닝 기반 비선형 회귀모델 구축 및 2025년 상반기 1km² 상세 격자별 모델링 자료에 적용 후 개선 결과 분석

III 연구결과

- PM10과 PM2.5 시계열 농도 변동성 기준으로 군집분석 수행 결과, PM10은 3개 군집, PM2.5는 5개 군집으로 분류되었음. 동일 군집 내 측정소 실측자료와의 상관분석을 통해 격자별 짝측정소를 선정함
- 격자별 CMAQ 예측값과 짝측정소 실측값을 입력자료로 하여 4개 머신러닝 기법(RF, SVM, GB, XGB)을 비교한 결과, 성능과 운영 효율성을 고려하여 가장 적합한 기법으로 XGB 모델을 선정함
- XGB 기법으로 학습한 비선형 회귀모델을 2025년 상반기 자료에 적용한 결과, PM10은 20 $\mu\text{g}/\text{m}^3$ ~ 30 $\mu\text{g}/\text{m}^3$, PM2.5는 10 $\mu\text{g}/\text{m}^3$ ~ 20 $\mu\text{g}/\text{m}^3$ 구간에서 기존 대비 각각 약 2.7배 ~ 2.9배, 1.7배 ~ 2.4배의 정확도 개선 효과를 확인함
- 하지만 저농도 및 고농도 구간에서 학습 자료 부족과 머신러닝 자체 학습 특성으로 인해 예측 정확도 개선이 특정 농도 범위에서 제한적으로 나타났으며, 이에 따라 실질적인 적용을 위해서는 고농도 사례자료 확충, 농도 구간별·계절별 특성을 고려한 학습 방법에 대한 추가 연구가 필요할 것으로 판단됨

IV 정책연계방안

- 측정소 미설치 지역까지 반영할 수 있는 1km 격자 단위의 대기오염 정보를 개선함으로써, 공간해상도가 높고 정확한 대기질 현황 파악 가능
- 동네별 대기질 정보의 신뢰성 향상 및 대기질 관리 정책 수립에 활용

V 활용계획

- 현재 운영 중인 (초)미세먼지 동네예보의 예측 지역 확대 및 대기질 진단평가 시 참고자료로 활용
- 향후 다른 대기오염물질 예측 모델의 정확도 향상을 위한 연구 기초자료로 활용