

미세먼지 예측

Deep Learning

김건효

010 5440 2281

rlarjsgy69i@naver.com

미세먼지 정의



미세먼지는 지름 10 μm 이하의 **매우 작은 크기를 가진 입자**를 뜻한다. 그 크기에 따라 10 μm 이하는 미세먼지, 2.5 μm 이하는 초미세먼지로 나뉜다.

미세먼지 원인



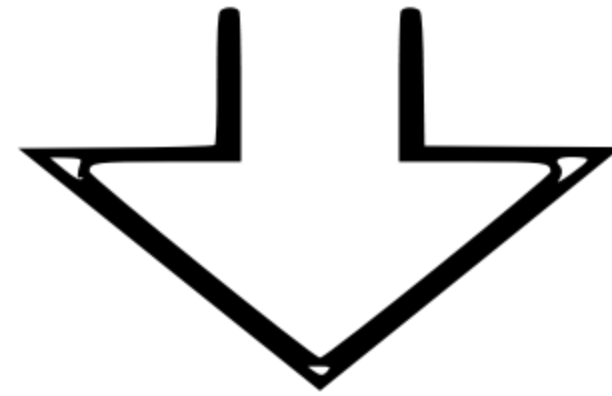
공장, 발전소



대기 반응

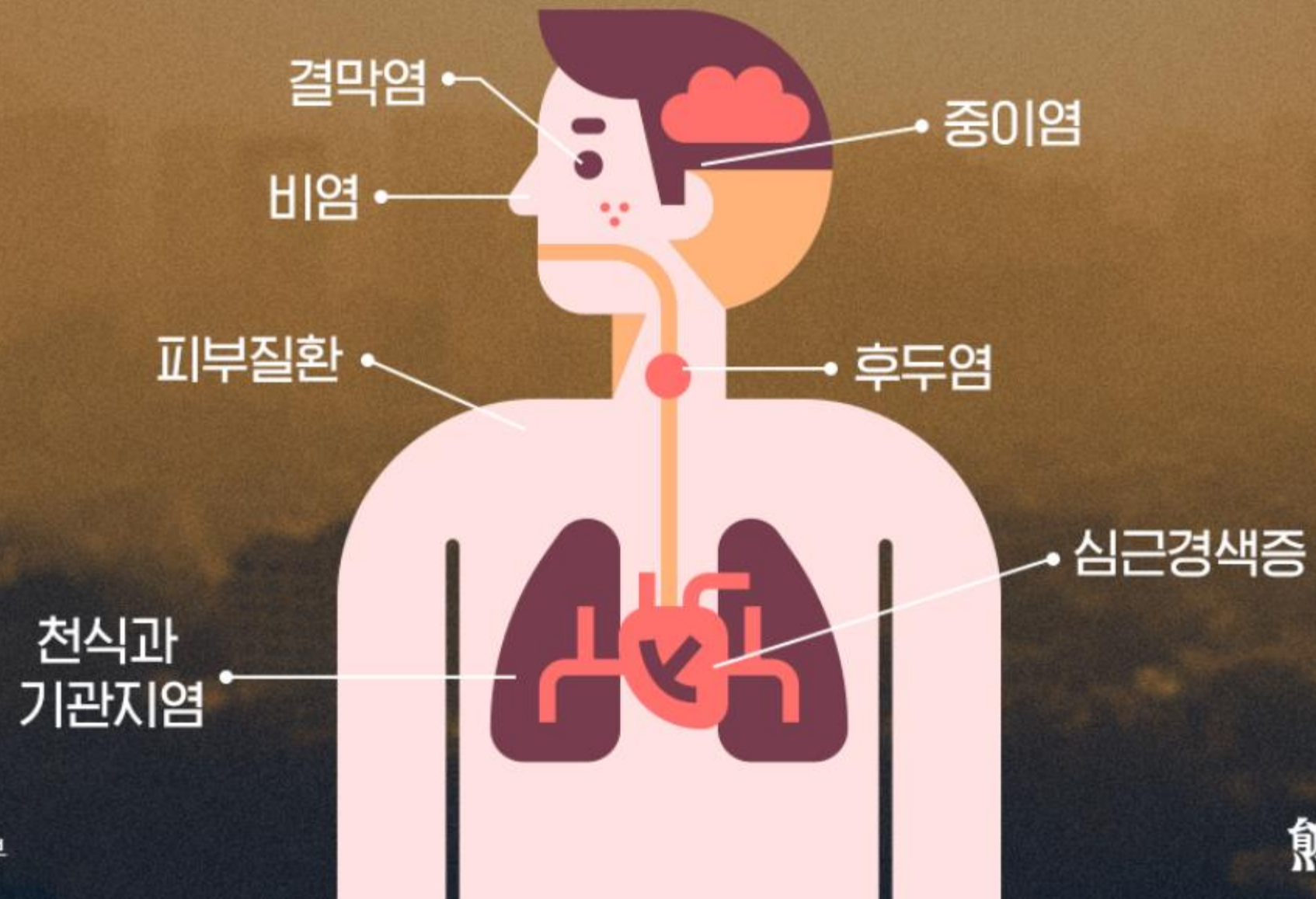


자동차 매연



중금속과 같은 화학물질 다량 포함

미세먼지가 우리 몸에 미치는 영향



미세먼지

예측의 한계

국지적 특성

- 지리·지형적 특성에도 영향을 받아 측정소마다 큰 차이보임

예측 불가 발생 원인

- 도심에서의 화재나 산불 등 예측불가한 발생원 존재

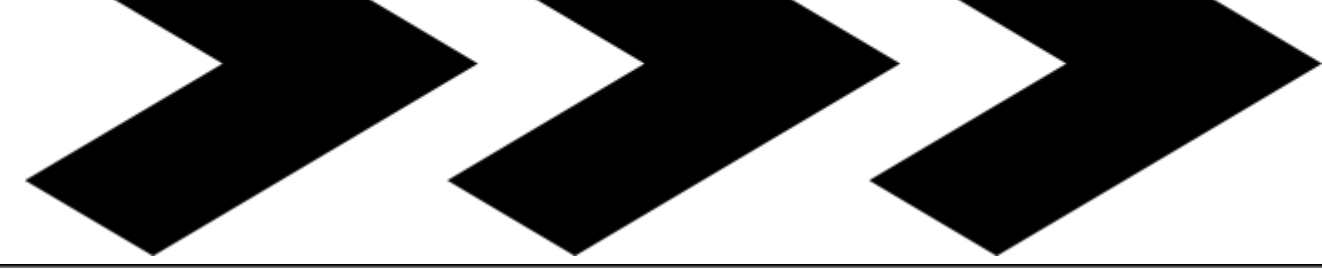
다양한 발생 원인

- 일상생활, 산업현장, 국외유입 등 너무 많은 발생원으로 인해 정확한 예측 어려움

낮은 정확도

- 시간적, 공간적 범위가 넓어질수록 정확도가 떨어짐

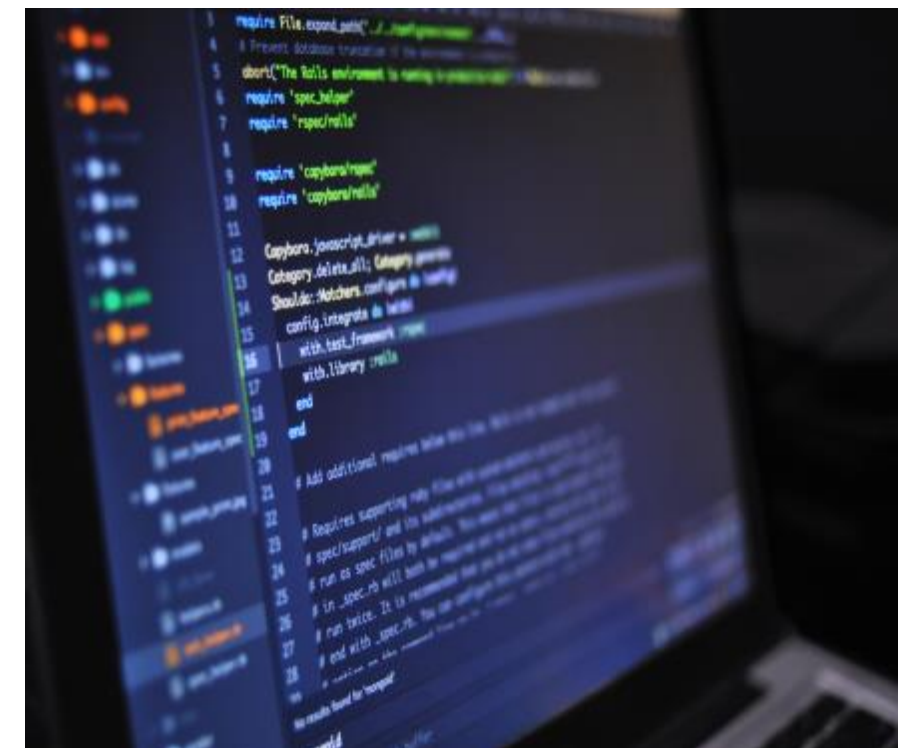
미세먼지 예측 설정 범위 및 예측 방법



기상청 미세먼지 관측소
데이터 활용



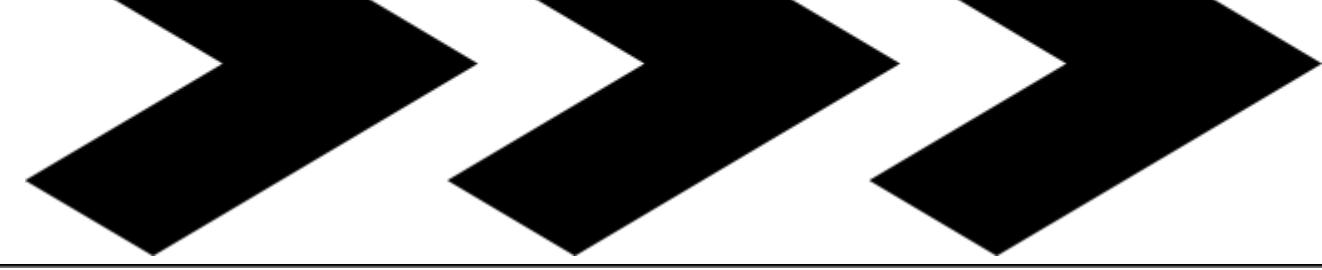
현재 미세먼지 농도를
기반으로 한 시간 후 농도 예측



파이썬을 활용한
DNN 모델 사용

미세먼지 예측을 위한 수집 데이터





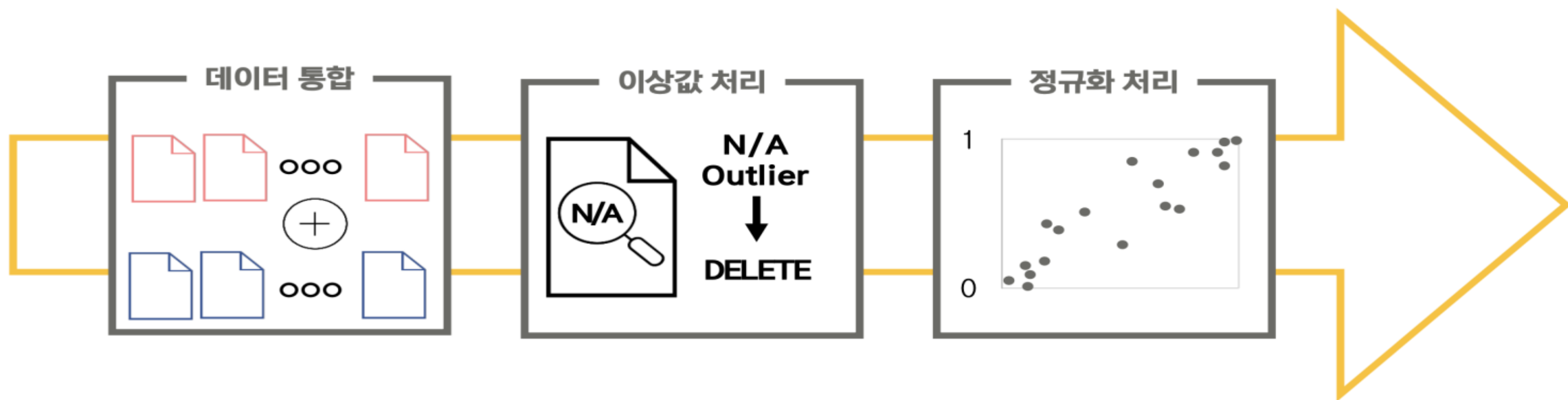
CO SO₂ NO₂ O₃ PM₁₀ PM_{2.5}
매 시간 데이터 수집



기상청

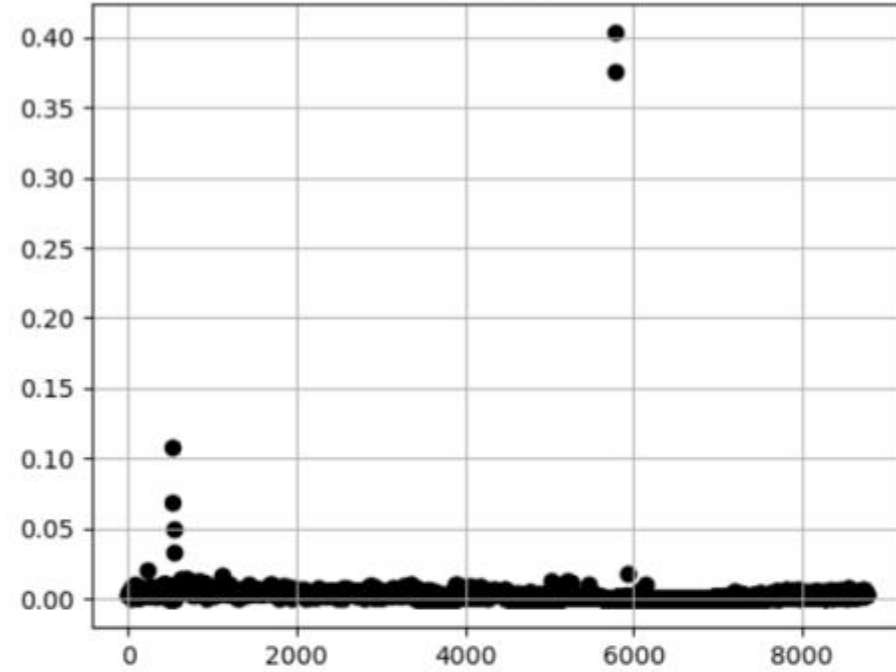
Korea Meteorological
Administration

기온 풍속 습도 강수량 매 시간 데이터 수집

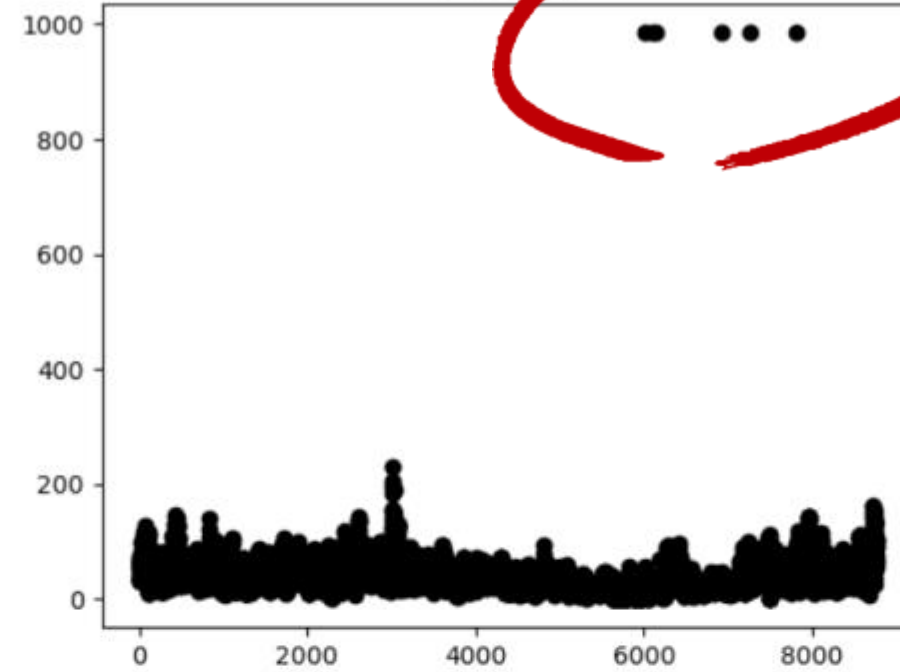


이상치 처리

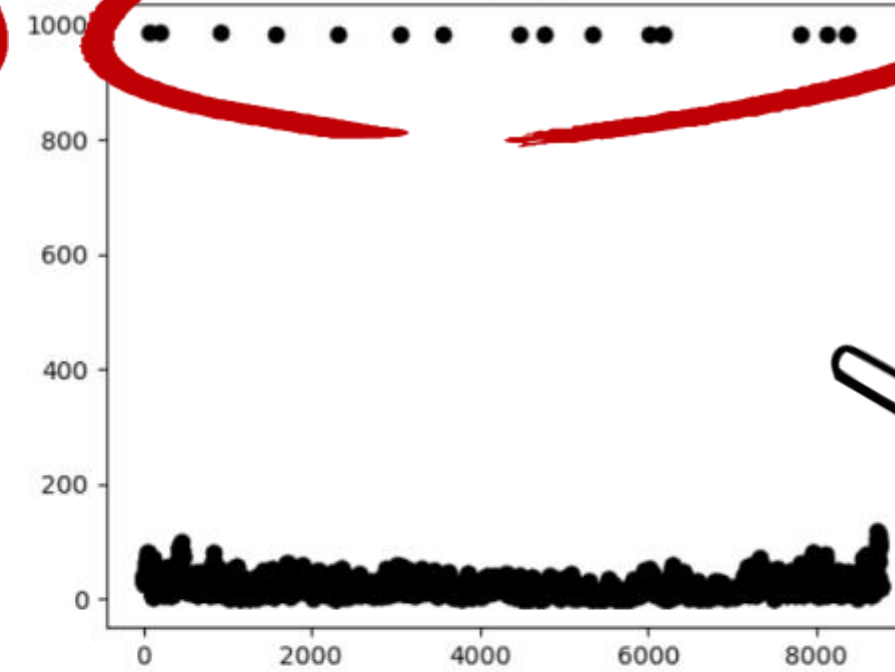
SO2



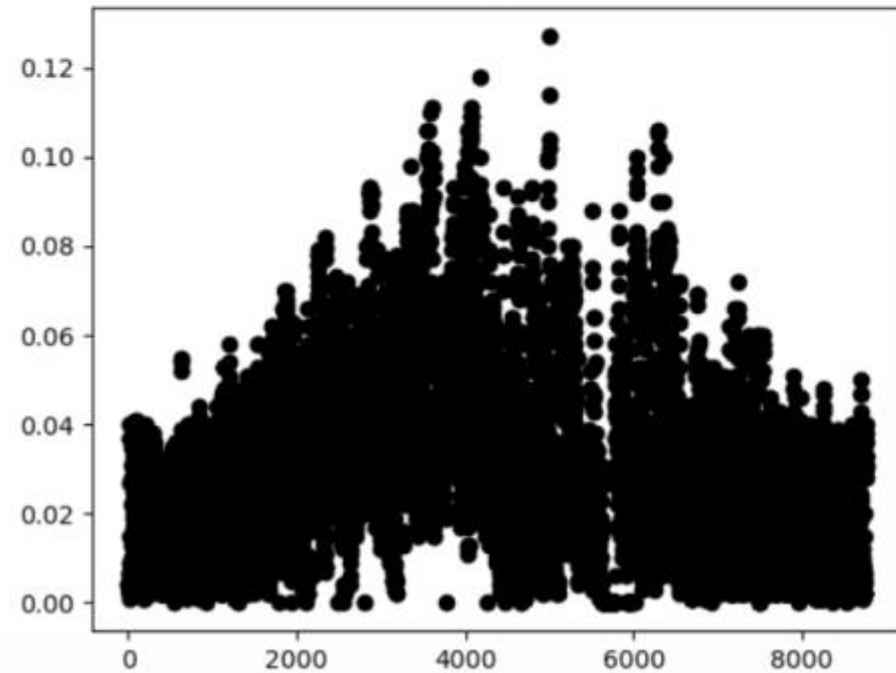
PM10



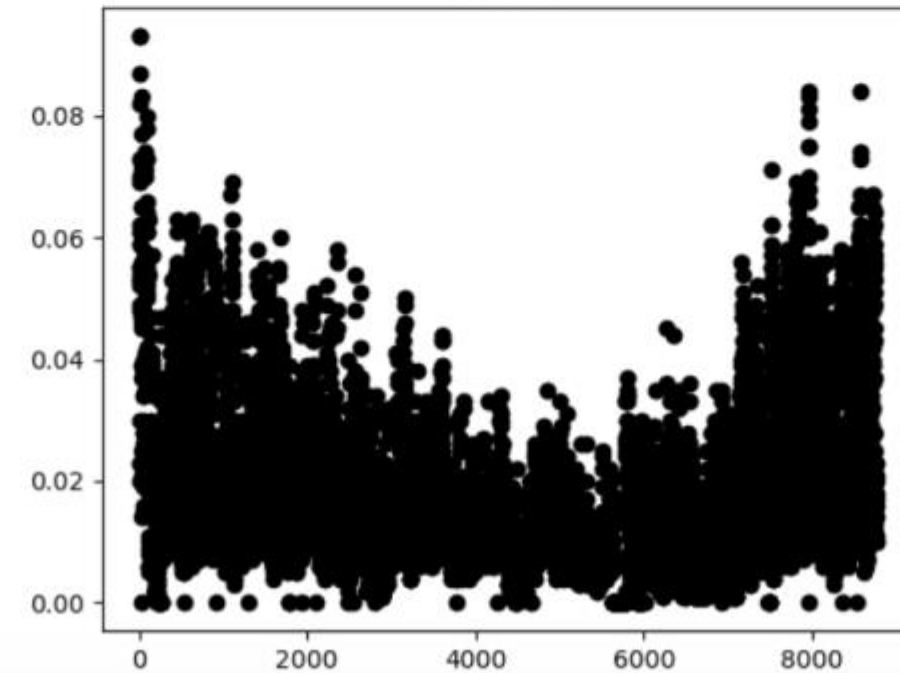
PM25



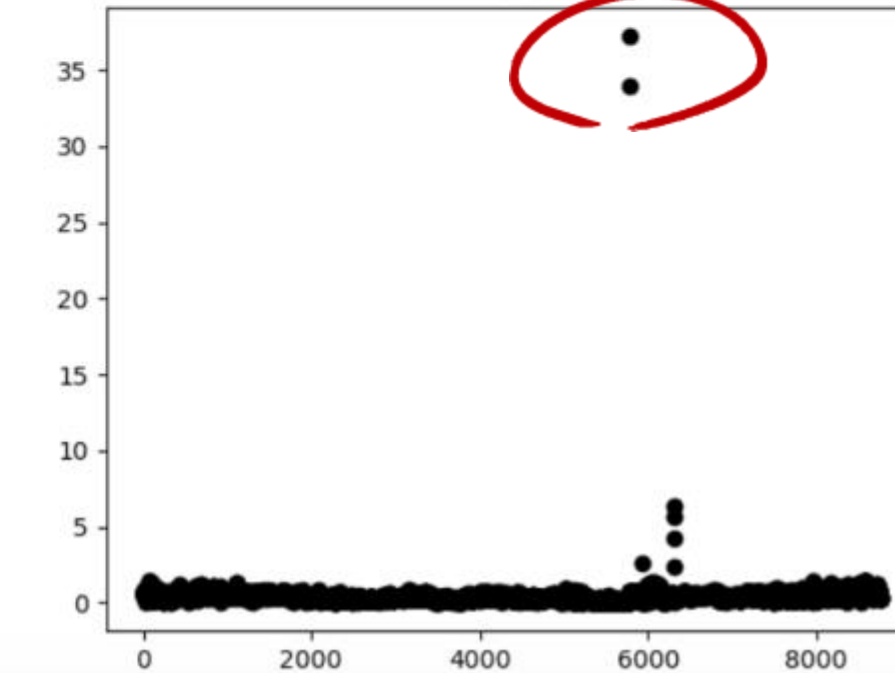
O3



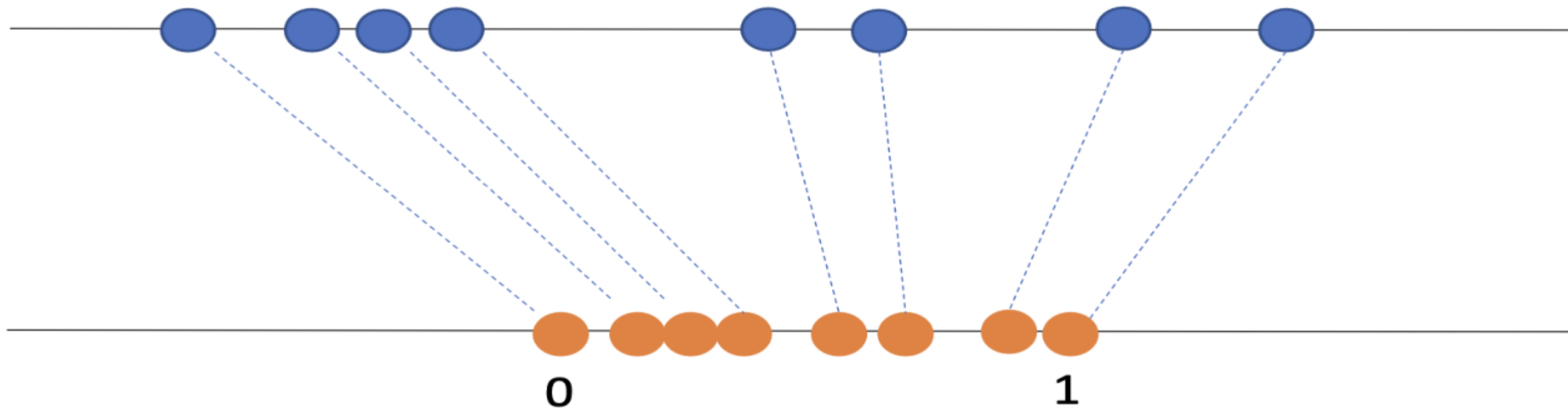
NO2



CO



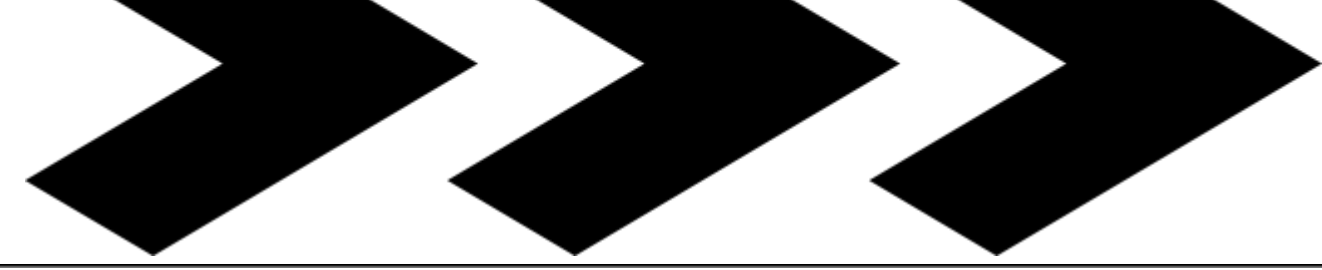
기기 오작동



본적인 전처리로 각 샘플을 평균이 0 분산이 1이 되도록 변환하는 표준 정규화 방법이 아닌,
데이터를 0에서 1사이로 압축하거나 늘리는 방법으로,
데이터에서 최소값을 빼고 최대값과 최소값의 차이로 나누어주는
Min - Max Normalization 방법 사용

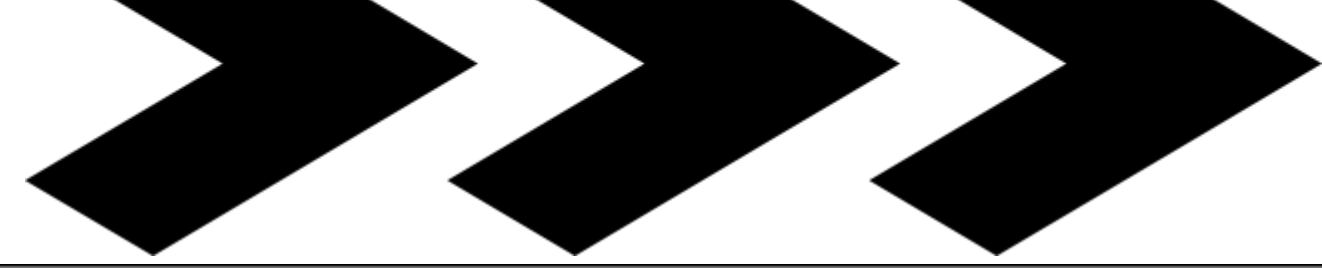
$$x_{scaled} = \frac{x - x_{min}}{x_{max} - x_{min}}$$

하이퍼 파라미터 설정

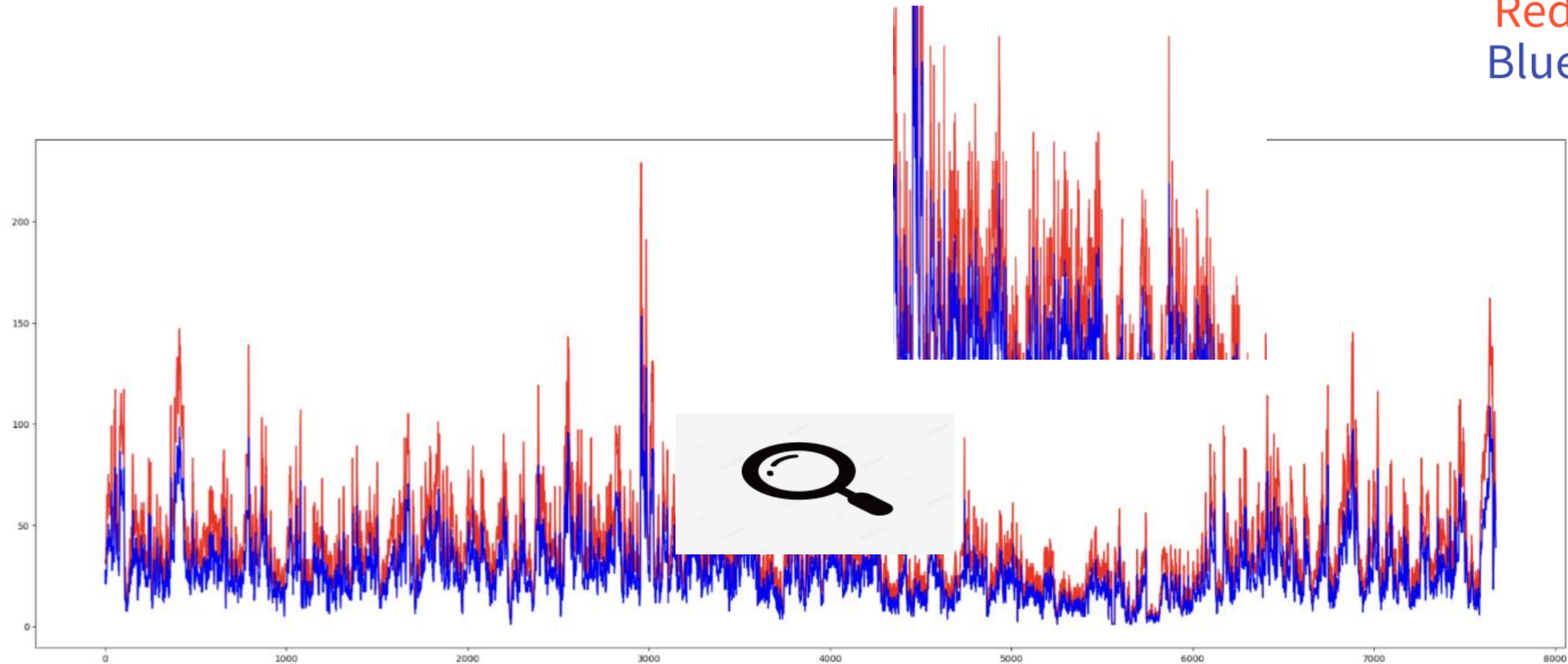


하이퍼 파라미터들은 임의의 값을 설정한 뒤, 결과치가 가장 좋은 하이퍼 파라미터로 결정

PM10 정보 활용 DNN 모델

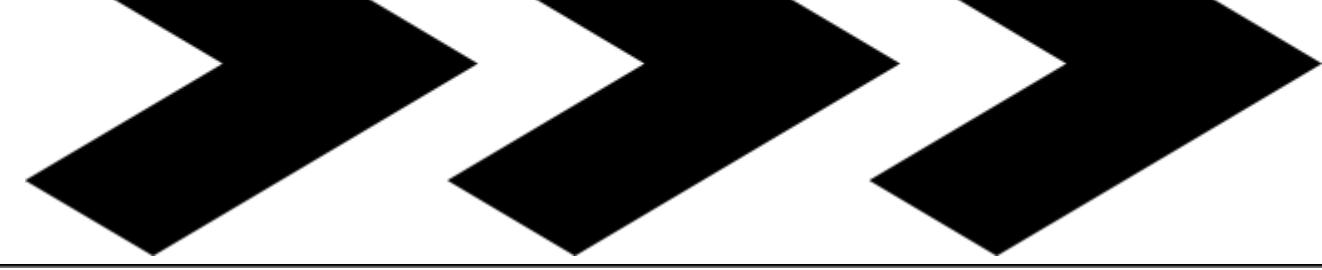


Red Line : 실제 농도
Blue Line : 예측 농도

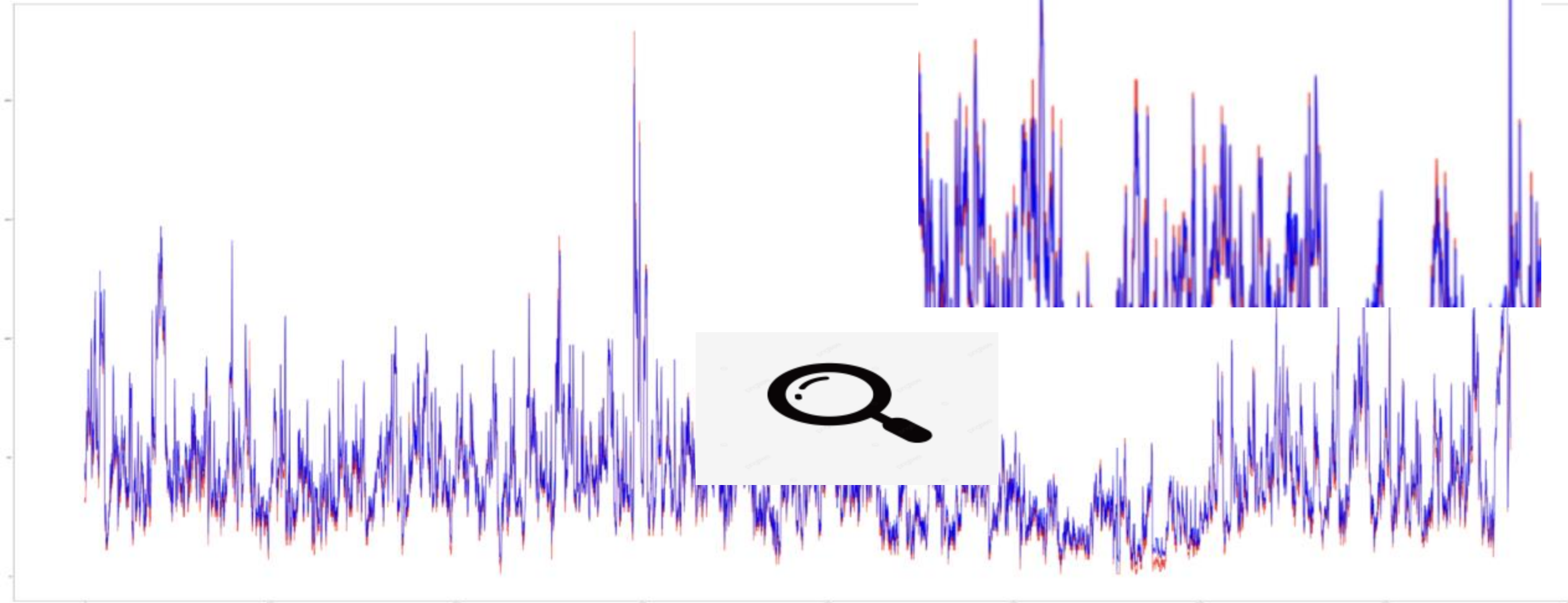


DNN 모델을 테스트 셋에 적용시킨 결과
약 75%

PM10 + 대기 + 기상 정보 활용 DNN 모델

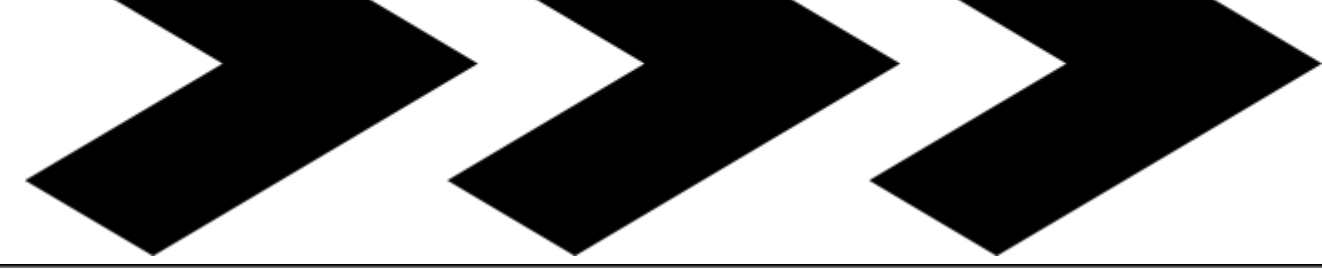


Red Line : 실제 농도
Blue Line : 예측 농도

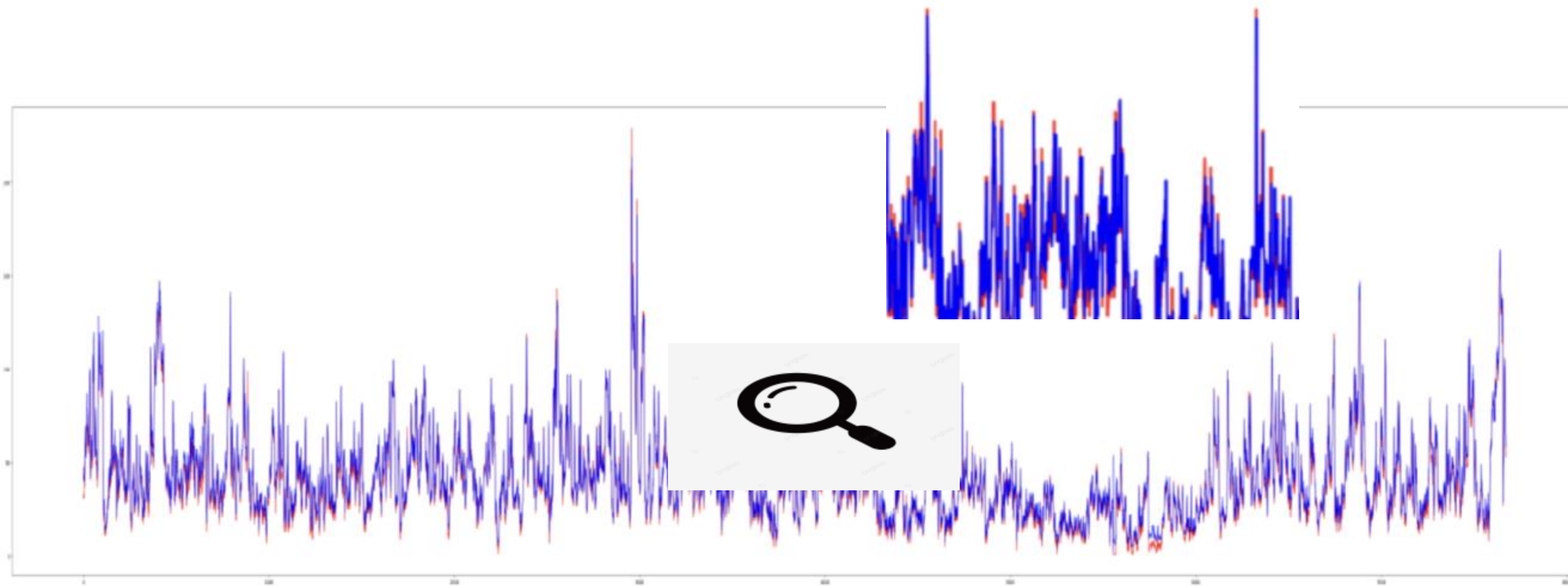


DNN 모델을 테스트 셋에 적용시킨 결과
약 85%

PM2.5 + 대기 + 기상 정보 활용 DNN 모델



Red Line : 실제 농도
Blue Line : 예측 농도



DNN 모델을 테스트 셋에 적용시킨 결과
약 90%

결론

○ 미세먼지 농도 예측 모델은 **미세먼지 정보와 대기, 기상 정보**를 함께 features로 주었을 경우, **가장 정확도가 높게** 나타났음.

○ 미세먼지 농도가 저농도에서 고농도로, 고농도에서 저농도로 **농도가 급변하는 경우 오차 값이 크게** 나타났음. 이는 앞서 미세먼지 예측 한계점에서 말했던 발생원들에 의한 농도 변화로 **수집 데이터 부족**으로 추측됨.

○ **PM2.5가 PM10**보다 더 높은 정확도를 보여주었음. 이는 위에서 말한 농도의 급변화 정도가 적기 때문에 나타난 결과로 추측됨.



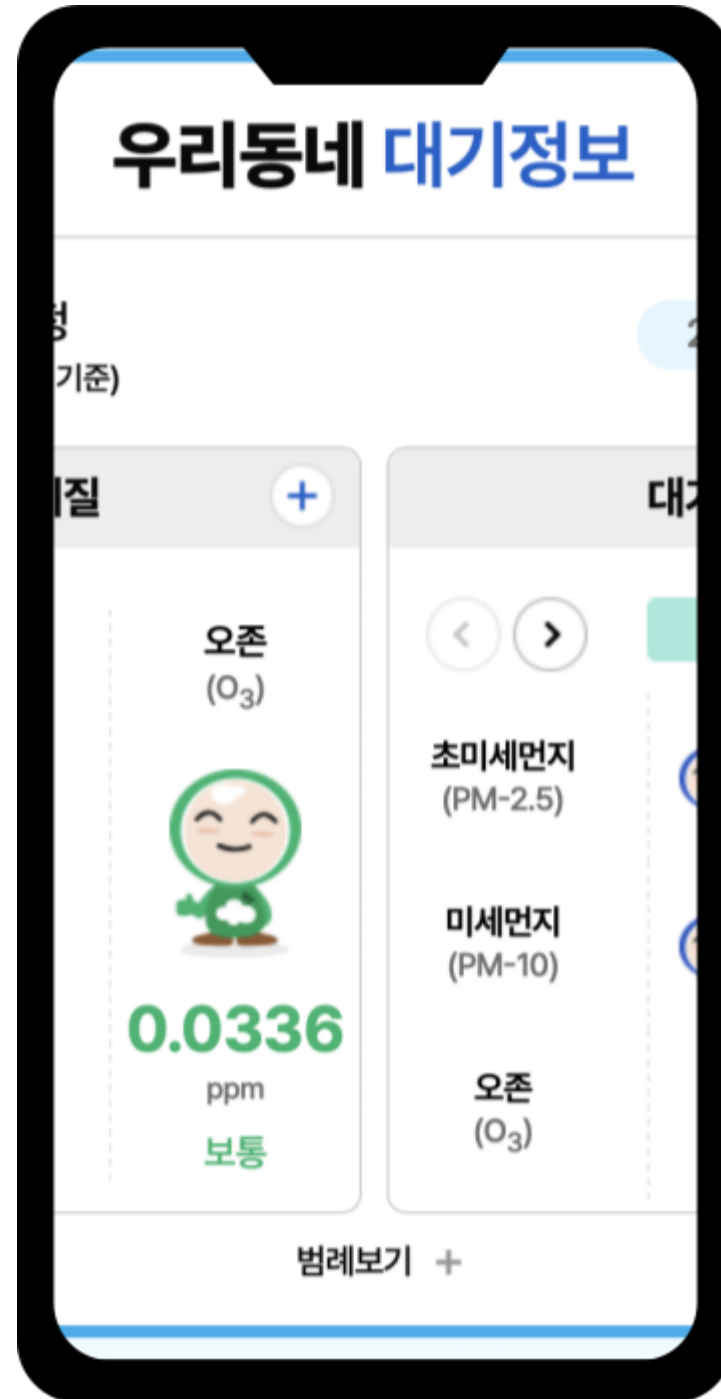
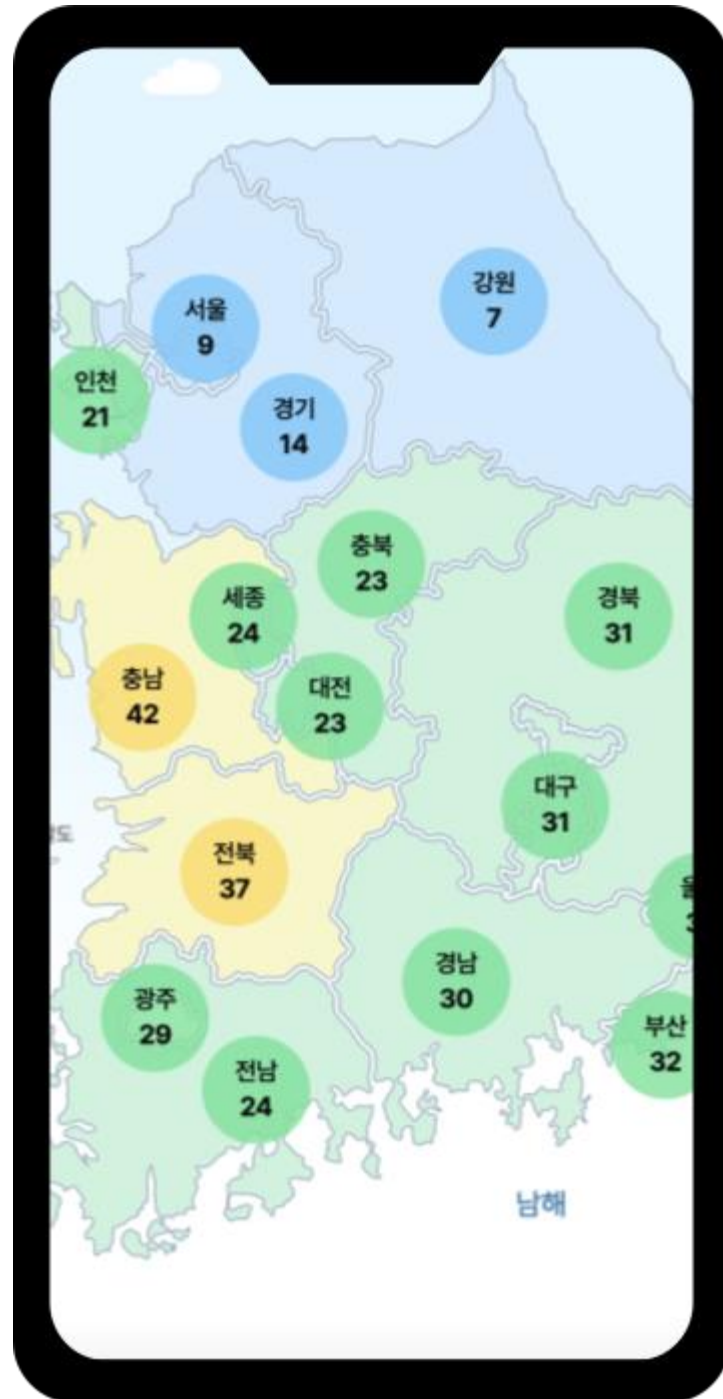
시간적 범위 확장

기존의 모델에 교통량, 발전소 주변대기정보 및 풍향 등의 features를 추가로 추가하고 1시간보다 넓은 범위의 시간에서 높은 정확도를 가진 미세먼지 예측 모델을 개발

공간적 범위 확장

기존의 지산동 측정소뿐만 아니라 전국에 위치하는 각각의 측정소별로 모델을 개발하여 한 지역에서 전국으로 공간적 범위를 확장

응용 프로그램



웹·앱 어플리케이션 개발

정확도가 높은 미세먼지 예측 모델을 개발, OpenAPI 활용하여 실시간 대기정보와 함께 실시간 예측정보를 함께 제공하는 웹·앱 어플리케이션 개발 제공

감사합니다.

김건효

010-5440-2281

rlarjsgy69@naver.com

@hyoxiki_

