

풍향 데이터를 통한 미세먼지 예측

AI 심화 교육 프로그램

한영태 이진향 조수현 정미리



풍향 데이터를 통한 미세먼지 예측

AI 심화 교육 프로그램



풍향 데이터를 통한 미세먼지 예측

AI 심화 교육 프로그램



지구가 자전을 멈추지 않는 이상 우리나라의 서쪽,
즉 중국 쪽에서는 늘 편서풍이 불어옴

미세먼지

우리 눈에 보이지 않는 아주 작은 유해 물질로 대기 중에 오랫동안 떠다니거나 흩날려 내려오는 직경 10μm 이하의 입자상 물질

중위도 편서풍 지대

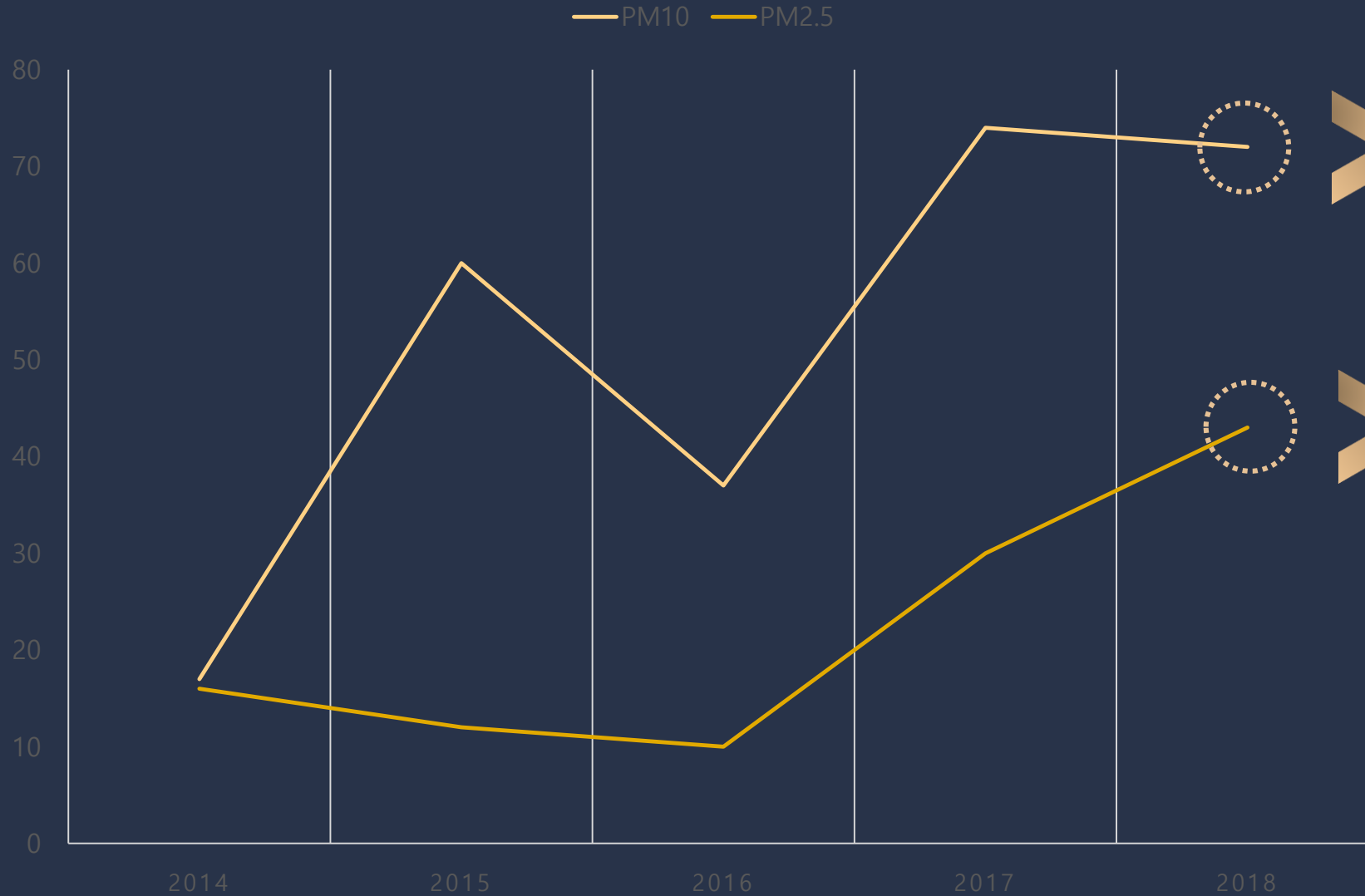
한반도는 중위도 편서풍대에 위치

미세먼지는 PM10인 미세먼지와 PM2.5인 초미세먼지로

구분

풍향 데이터를 통한 미세먼지 예측

AI 심화 교육 프로그램



미세먼지 발생률 증가

2016 년만 감소하는 추세를 보인 이후 꾸준하게 상승세

초미세먼지 위험성

인체에 더욱 해로운 영향을 주는 초미세먼지는 더욱 가파른 상승세

미세먼지에 직접적 영향을 주는 편서풍과의 연관성을 통한 예측 모델 구현을 추진

풍향 데이터를 통한 미세먼지 예측

AI 심화 교육 프로그램

데이터 수집

풍향·풍속 데이터

한국 환경 공단에서 제공 중인 임의로 선정한 20개의 시·군 지역의
일별 데이터 수집

데이터 정리

수집한 데이터 처리

학습 데이터로 사용할 수 있도록 데이터를 빈 값 처리나 인덱스 값
을 추가

LSTM 모델 구축

RNN Many To One

2017.01.01 ~ 2019.06.30 데이터를 학습
시킨 후 다음 날의 미세먼지 예측

성능 최적화

파라미터값 수정

LSTM 모델의 다양한 파라미터들을 적절히
수정하여 최적의 성능 도출

STEP 1

STEP 2

STEP 3

STEP 4

데이터 수집: 기상청 기상자료개방포털

서울특별시 서울(108) 지점목록보기 주소 입력하세요.

시간자료 일자료

- ☐ 1시간 최대강수량
- ☐ 1시간 최대 강수량 시각
- ☐ 일강수량
- ☒ 바람
 - ☐ 최대 순간 풍속
 - ☐ 최대 순간 풍속 풍향
 - ☐ 최대 순간풍속 시각
 - ☐ 최대 풍속
 - ☐ 최대 풍속 풍향
 - ☐ 최대 풍속 시각
 - ☒ 평균 풍속
 - ☐ 풍정할
 - ☒ 최대풍향
- ☐ 습도
 - ☐ 평균 이슬점온도
 - ☐ 최소 상대습도
 - ☐ 최소 상대습도 시각
 - ☐ 평균 상대습도

자료조회

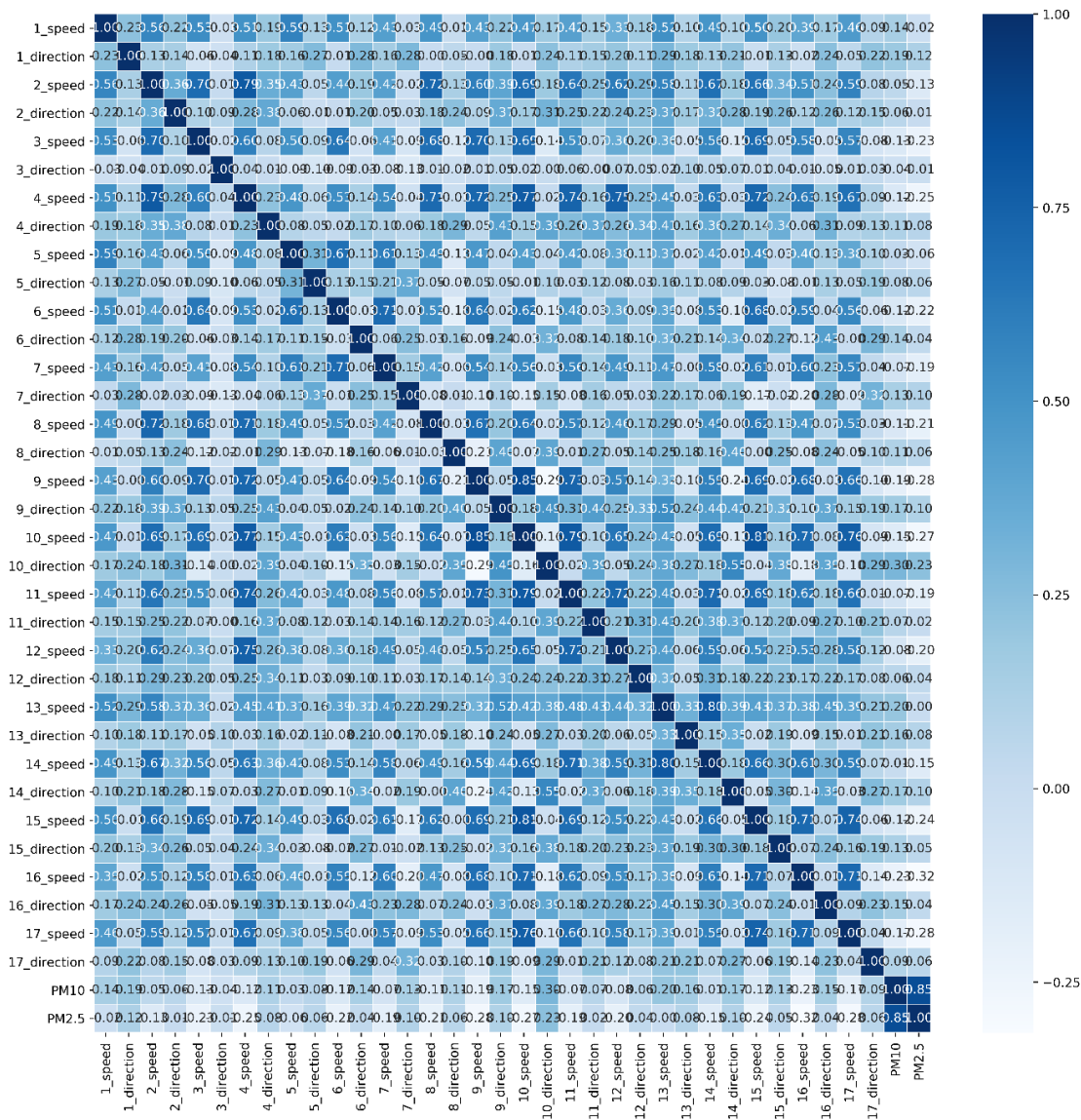
지도 확대시 방재관측 지점이 추가 표시됩니다.

데이터 셋

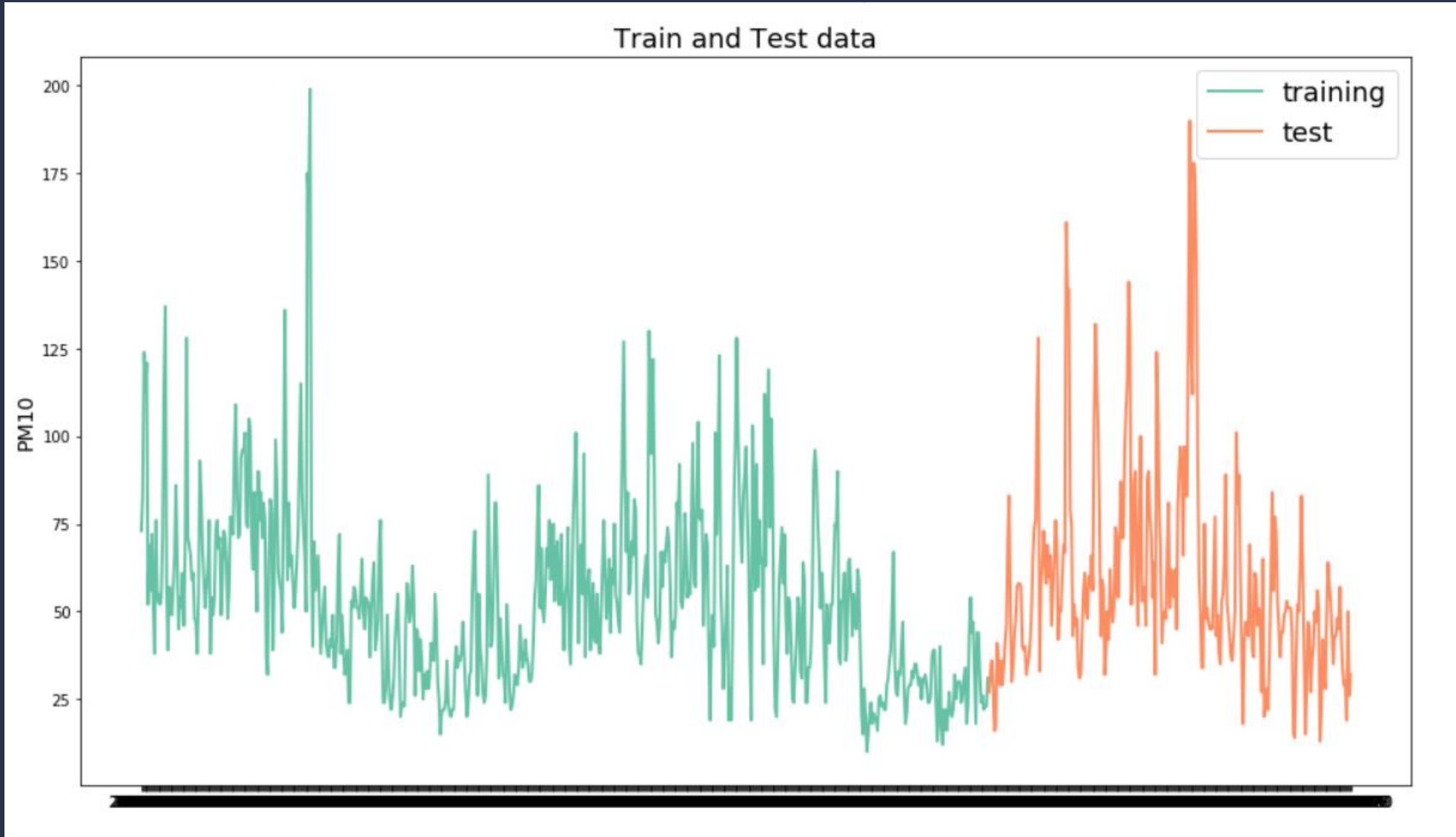
Date	속초		동두천		춘천		서산		울진		PM10	PM2.5
2017.1.1	1.7	250	0.6	140	0.5	360	0.9	200	2.7	250	73	40
2017.1.2	2.5	270	1.3	250	1	20	1.6	270	2.9	230	83	50
2017.1.3	2	290	1.2	250	1.2	200	1.3	200	2.7	230	124	83
2017.1.4	2.2	290	1	20	0.7	20	1.1	340	3	270	113	70
2017.1.5	1.5	290	1.3	20	1.8	50	1.7	70	3.3	290	121	78
2017.1.6	1.3	270	1	20	0.9	20	1.2	70	2.2	290	52	20
2017.1.7	1.5	290	1.1	20	0.8	360	1.1	50	2	250	69	35
2017.1.8	2	320	1.7	320	1	360	2.2	320	2	290	56	33
2017.1.9	1.6	290	2.8	290	2.3	230	3.2	320	2.8	250	72	41
2017.1.10	2.4	290	2.4	320	2	250	2.5	320	3.1	250	52	23
2017.1.11	2.4	200	1.7	270	1.9	180	1.4	200	3.3	250	38	12
2017.1.12	4.1	290	2.4	270	2.4	230	1.7	340	2.8	270	76	41
2017.1.13	3.2	290	2.6	290	2.2	270	2.6	320	3.1	250	53	26
2017.1.14	2.9	290	3.3	320	1.9	320	4.5	340	4.6	290	55	9
2017.1.15	1.6	290	1.3	50	0.8	20	1.6	50	2.4	250	52	14

- 2017.1~2019.6 까지 우리나라 20개 지역의 최다풍향, 평균풍속과 서울 미세먼지 농도 데이터

데이터 셋 피어슨 상관분석



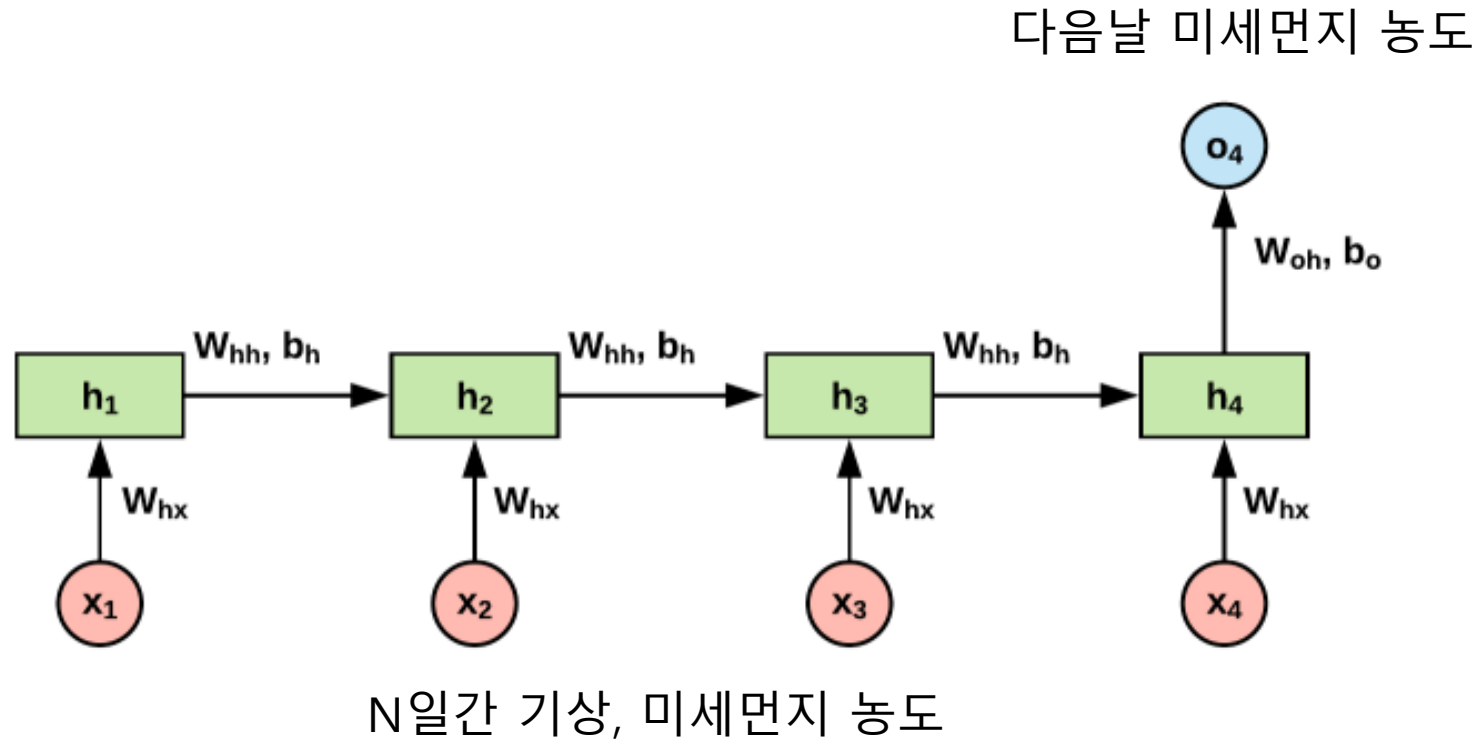
- 미세먼지와 풍속보다 미세먼지와 풍향과 상관관계가 높음
- 특히 강화지역의 풍향과 미세먼지가 상관관계가 높다



DATA

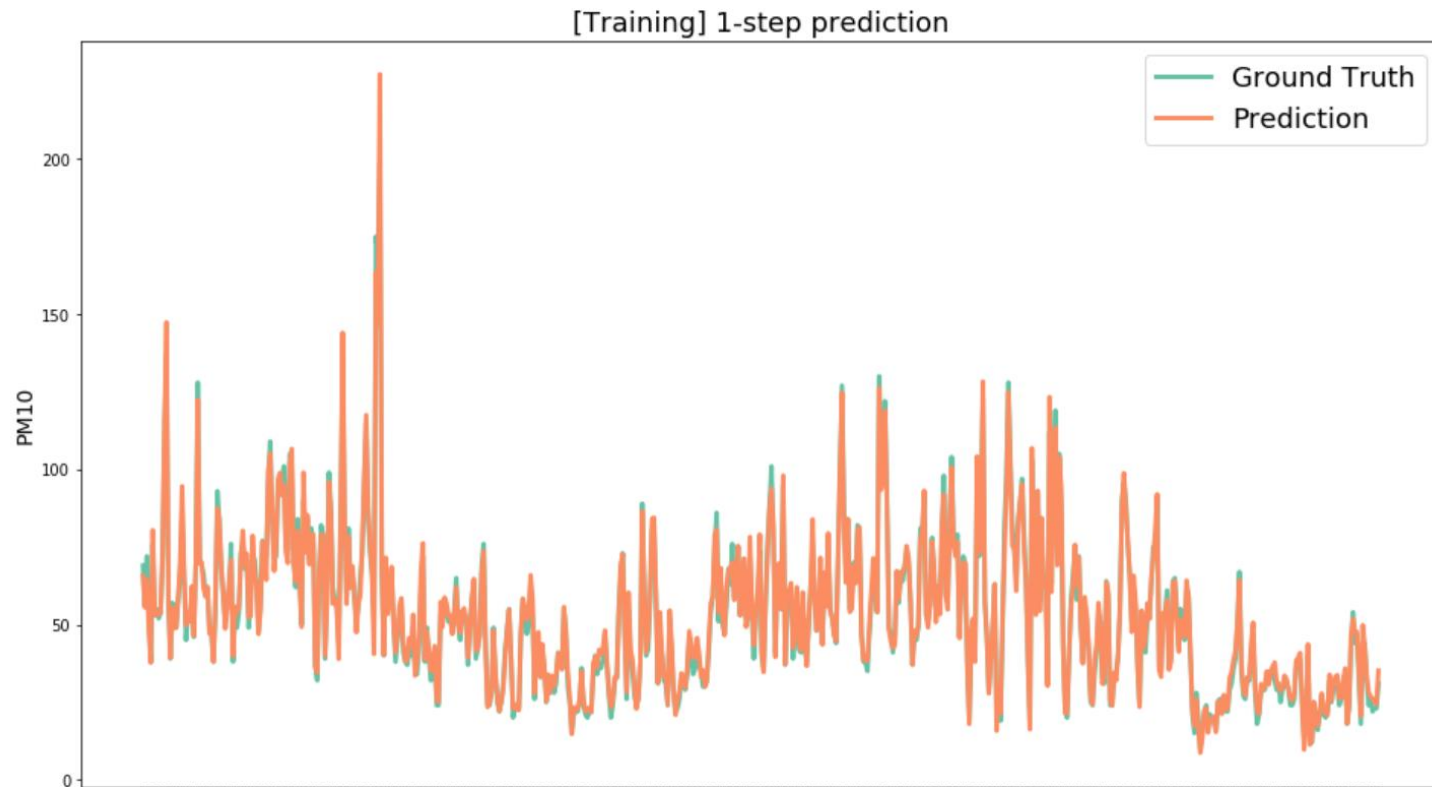
0%

사용모델: LSTM



- 손실함수 : MAE (Mean absolute error)
- 최적화 알고리즘 : Adam

학습결과 : 학습데이터에 대한 평가

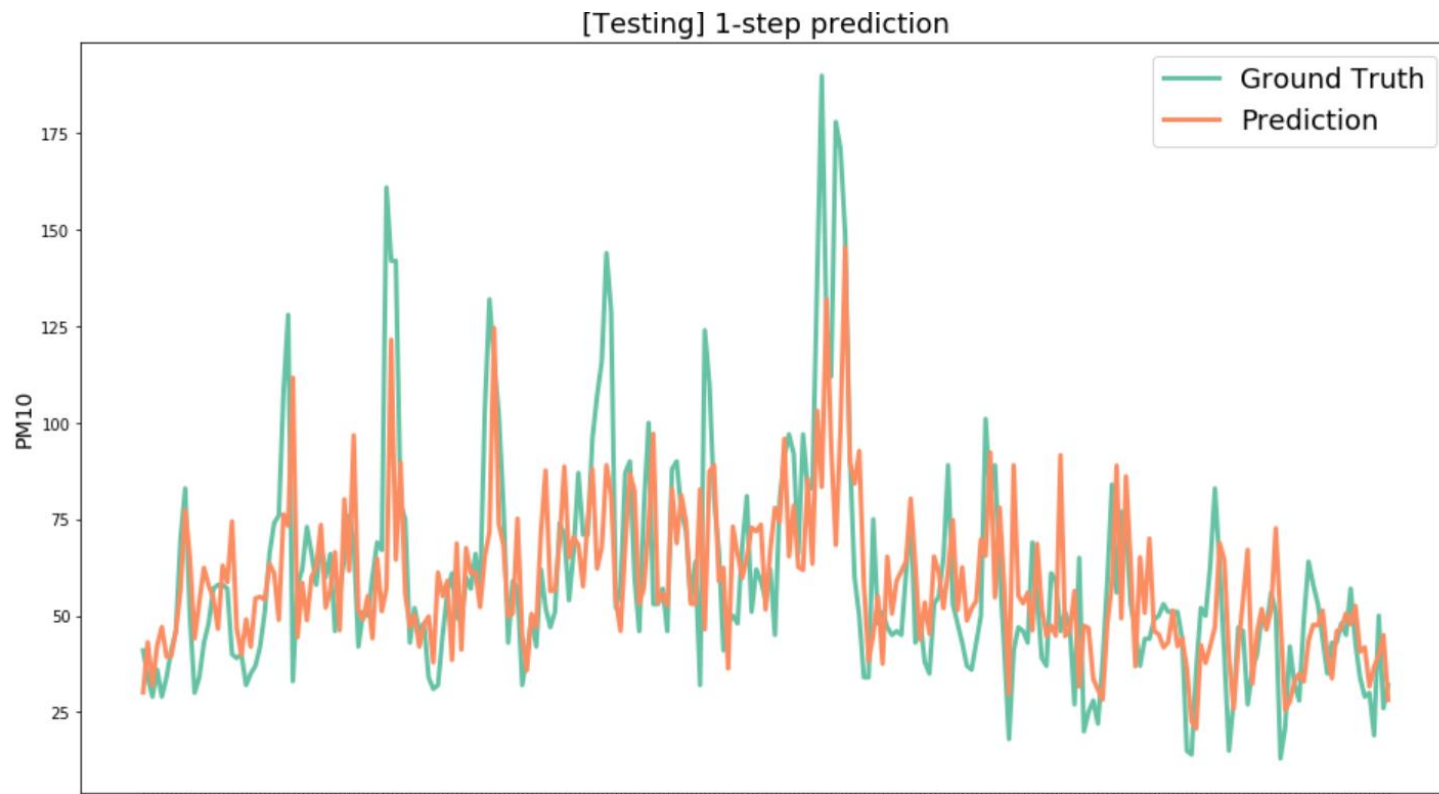


MAE = 3.2499938705299476

MSE = 16.14014997564854

RMSE = 4.017480550749255

학습결과: 테스트데이터에 대한 평가



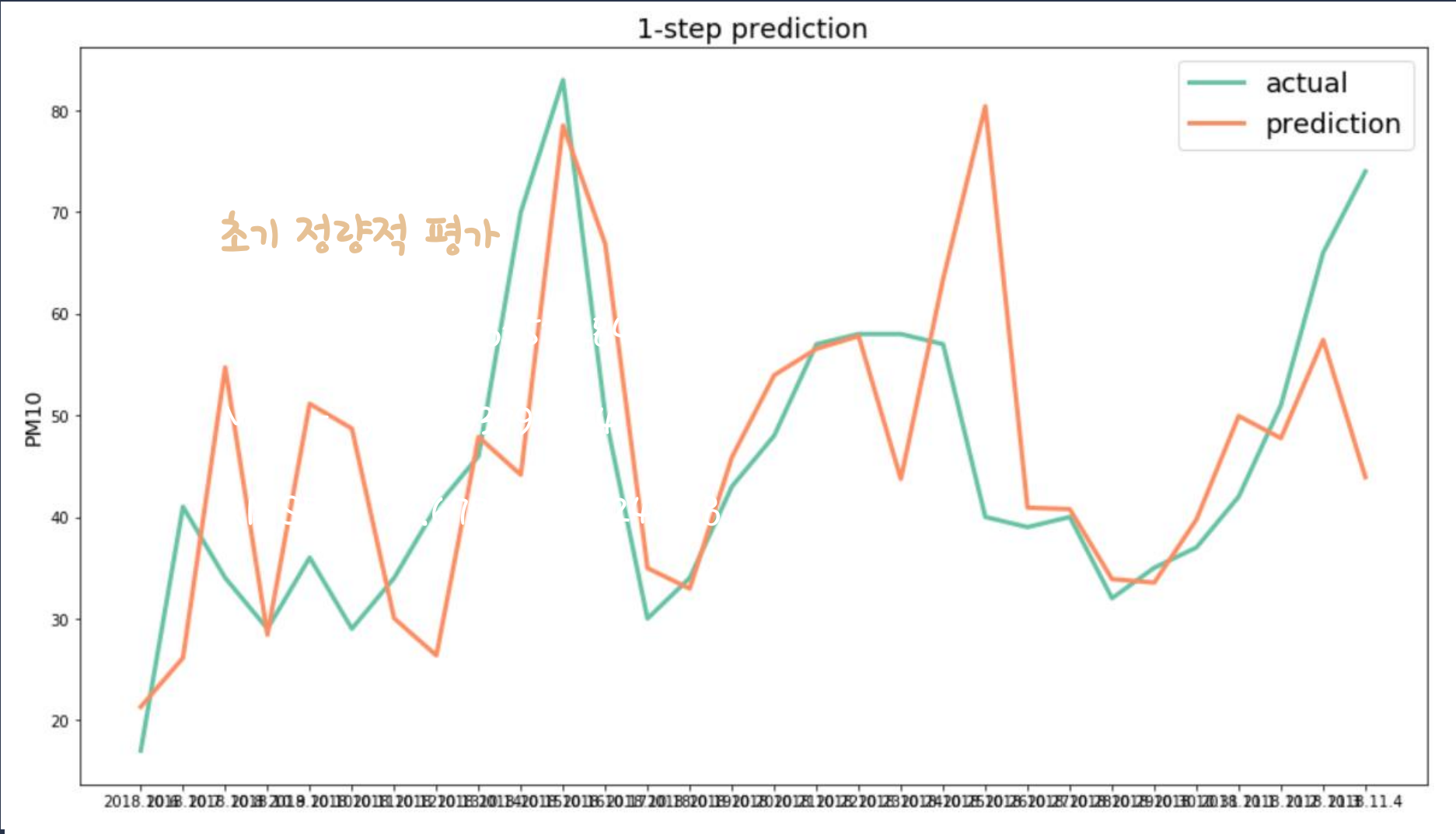
MAE = 15.158806047189548

MSE = 503.48054681551696

RMSE = 22.438372196207037

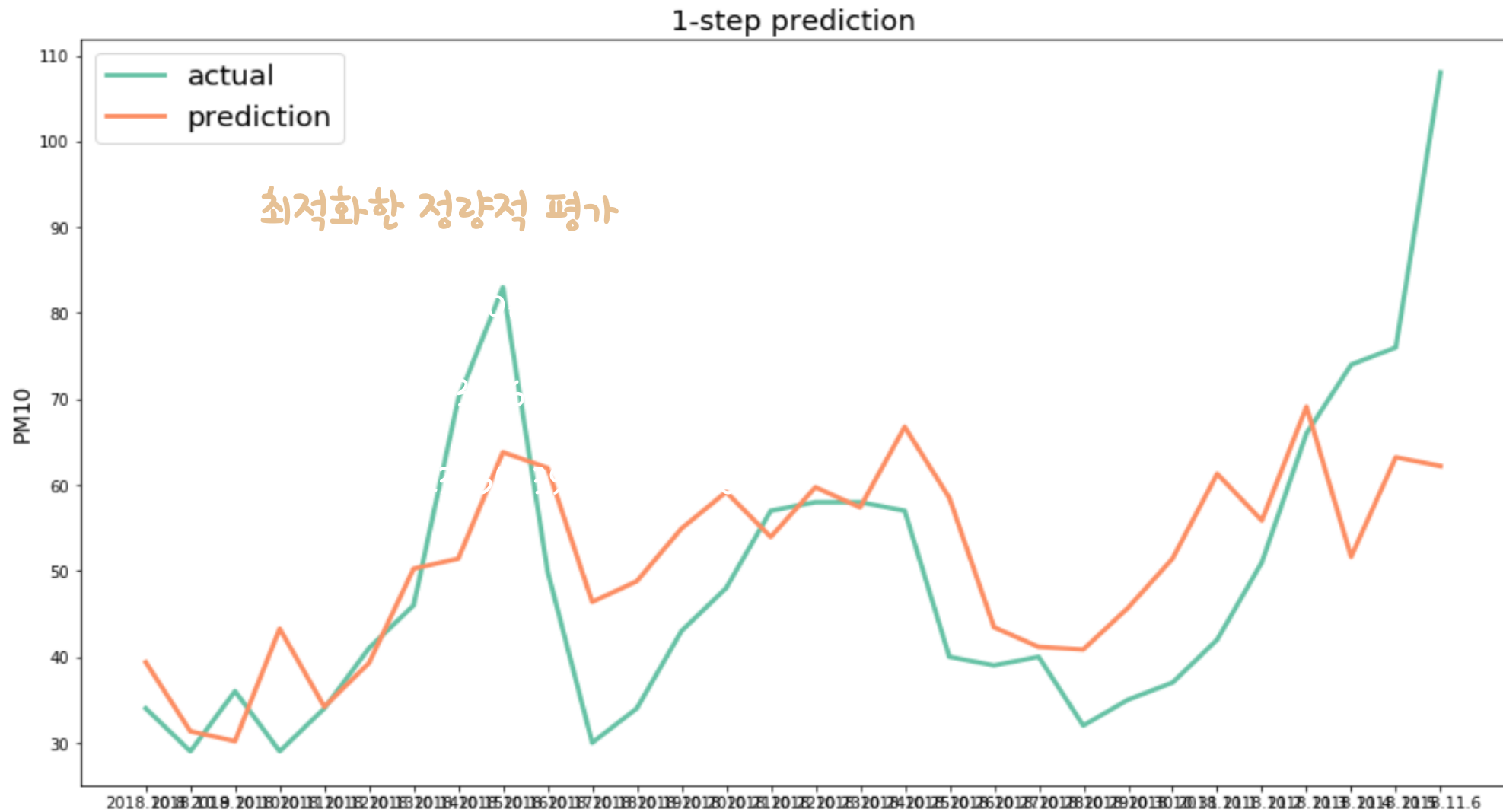
풍향 데이터를 통한 미세먼지 예측

AI 심화 교육 프로그램

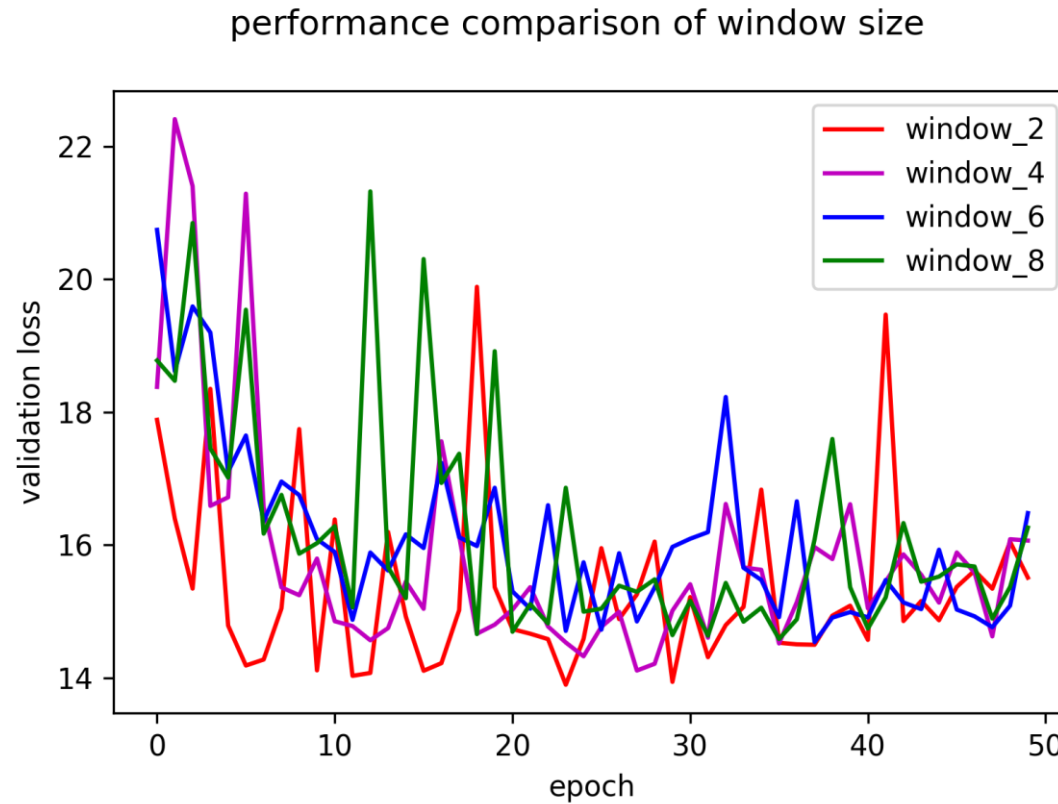


풍향 데이터를 통한 미세먼지 예측

AI 심화 교육 프로그램



Window 크기에 따른 실험결과



- 2일의 데이터로 다음날을 예측 했을 때 성능이 가장 좋음

QnA