

계절별 데이터와 농도별 데이터의 학습에 대한 LSTM 기반의 PM_{2.5} 예측 모델 성능 평가

Performance Evaluation of LSTM-based PM_{2.5} Prediction Model for Learning Seasonal and Concentration-specific Data

정 용 진 · 오 창 현*

한국기술교육대학교 전기전자통신공학과

Yong-jin Jung · Chang-Heon Oh*

Department of Electrical, Electronics and Communication Engineering, Korea University of Technology and Education (KOREATECH), Chungcheongnam-do, 31253, Korea

[요 약]

미세먼지에 대한 연구는 실시간으로 발전하고 있으며, 예측 모델의 정확도를 향상시키기 위해 다양한 방법이 연구되고 있다. 또한 미세먼지의 정확한 원인과 영향을 파악하기 위해 이러한 다양한 요소들을 고려하는 연구들이 활발히 이루어지고 있다. 본 논문에서는 PM_{2.5}와 상관성이 있는 데이터를 계절을 기준으로 구분하여 학습하는 예측 모델과 특정 농도를 기준으로 저농도와 고농도를 구분하여 학습하는 모델을 통해 예측 성능의 비교 및 분석을 진행하였다. 기상데이터와 대기오염 물질 데이터를 사용하였으며 PM_{2.5}와 상관관계를 확인하여 학습 및 평가를 위한 데이터를 구성하였다. 계절별 예측 모델과 농도별 예측 모델은 LSTM으로 설계하였으며, 세부 파라미터는 하이퍼 파라미터 탐색을 통해 적용하였다. 예측 모델의 성능 평가는 정확도, RMSE, MAPE, 저농도와 고농도 구간에서의 정확도 그리고 AQI를 기준으로 4개의 범위에 대한 정확도로 진행하였다. 성능 평가 결과, 농도별 학습을 진행한 예측 모델이 AQI 기준 “나쁨” 구간의 정확도에서 91.02%의 정확도를 보였으며, 계절별 학습을 진행한 예측 모델보다 전반적으로 좋은 성능을 보였다.

[Abstract]

Research on particulate matter is advancing in real-time, and various methods are being studied to improve the accuracy of prediction models. Furthermore, studies that take into account various factors to understand the precise causes and impacts of particulate matter are actively being pursued. This paper trains an LSTM model using seasonal data and another LSTM model using concentration-based data. It compares and analyzes the PM_{2.5} prediction performance of the two models. To train the model, weather data and air pollutant data were collected. The collected data was then used to confirm the correlation with PM_{2.5}. Based on the results of the correlation analysis, the data was structured for training and evaluation. The seasonal prediction model and the concentration-specific prediction model were designed using the LSTM algorithm. The performance of the prediction model was evaluated using accuracy, RMSE, and MAPE. As a result of the performance evaluation, the prediction model learned by concentration had an accuracy of 91.02% in the “bad” range of AQI. And overall, it performed better than the prediction model trained by season.

Key word : Deep learning, LSTM, Correlation analysis, Particulate matter.

<https://doi.org/10.12673/jant.2024.28.1.149>



This is an Open Access article distributed under the terms of the Creative Commons Attribution Non-Commercial License (<http://creativecommons.org/licenses/by-nc/3.0/>) which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited.

Received 5 February 2024; Revised 26 February 2024

Accepted (Publication) 28 February 2024 (29 February 2024)

*Corresponding Author, Chang-Heon Oh

Tel: +82-41-560-1187

E-mail: choh@koreatech.ac.kr

I. 서론

먼지의 크기에 따라 지름이 $10\ \mu\text{m}$ 이하의 먼지는 미세먼지 (PM_{10}), 지름이 $2.5\ \mu\text{m}$ 이하의 먼지를 초미세먼지($PM_{2.5}$)라고 분류한다. 초미세먼지의 경우 매우 작은 크기로 인해 장시간 노출이 될 경우 건강에 치명적인 요인으로 작용할 수 있다. 고령자의 경우 미세먼지로 인한 환경성질환으로 사망의 위험이 13.9% 증가한다는 연구도 있으며[1], 고령자가 아니더라도 호흡기질환, 심혈관질환 등의 발생과도 연관성이 있는 위험한 물질이다[2]-[4]. 미세먼지의 영향은 단순히 건강에 대한 문제만 있는 것이 아니다. 미세먼지가 사회적 이슈가 됨에 따라 많은 매체들이 다양한 정보를 제공하고 있으며, 많은 사람들이 미세먼지에 대한 위험성을 알게 되었다. 이에 따라 미세먼지에 대한 예보를 찾아보고 마스크를 착용하거나 야외 활동을 자제하고 있다. 이로 인해 미세먼지가 경제 활동 저하 문제의 원인 중 하나로 분석되고 있다. 미세먼지에 대한 다양한 문제는 세계적으로 이슈가 되고 있다. WHO (World Health Organization)의 IARC (International Agency for Research on Cancer)에서는 고위험군의 발암물질로 지정하여 발표하였으며, 미국, 캐나다 등의 보건부에서는 미세먼지 및 주변 대기오염에 대한 정책을 시행하고 있다. 한국도 미세먼지 저감 조치에 대한 특별법을 2019년에 제정하였으며, 2024년까지 미세먼지를 저감 및 관리하기 위한 프로그램과 보건 및 양질의 예보 서비스 제공을 위해 노력하고 있다[5]-[7]. 대중들이 중요하게 생각하는 것은 정확한 미세먼지의 예보이다. 미세먼지에 대한 예보의 정확성을 높이기 위해 다양한 연구들이 진행되고 있으며, 해당하는 연구들 중에는 머신러닝, 딥러닝 등을 활용하고 있다.

김삼용의 연구에서는 기상데이터와 대기오염물질 데이터를 이용하여 딥러닝 알고리즘의 예측 성능을 비교하였다. 비교를 위한 알고리즘은 MLP (multi-layer perceptron), RNN (recurrent neural network), LSTM (long short term memory), GRU (gated recurrent unit) 그리고 CNN (convolution neural network)을 사용하였으며 단일층으로 구성한 후 하이퍼 파라미터 탐색을 통해 최적의 파라미터를 적용하였다. 예측의 성능 평가를 위해 MAE (mean absolute error)와 RMSE (root mean square error)를 사용하였다. 성능 평가 결과, GRU가 다른 알고리즘 보다 예측 성능이 우수한 것을 확인하였다[8]. 이주용의 연구에서는 $PM_{2.5}$ 의 예측을 위한 모델인 CMAQ (community multiscale air quality)의 정확도 개선을 위해 DNN (deep neural network) 모델을 사용하였다. 도시대기측정망 데이터와 CMAQ 예측데이터를 사용한 DNN-1과 DNN-1에 $PM_{2.5}$ 의 성분측정데이터를 추가한 DNN-2 모델을 개발하여 예측 및 평가를 진행하였다. 평가 결과, 기존의 CMAQ보다 개선된 것을 확인하였다[9]. 이와 더불어 예측 성능에 대한 연구 뿐만 아니라 미세먼지에 영향을 주는 요인과 환경에 대한 연구도 진행되고 있다. 허유경의 연구에서는 PM_{10} 에 해당하는 미세먼지 발생이 겨울과 초봄에 집중되어 있는 것을 주목하여 계절에 따라 고농도 미세먼지 발생에 영

향을 미치는 요인에 대해 분석을 진행하였다. 연구 결과, 봄철 평균 기온이 PM_{10} 의 나쁨 이상에 대한 농도에 유의한 영향을 미친다는 것을 확인하였다[10].

본 논문에서는 미세먼지 농도의 예측에 있어 계절에 대한 특성과 농도에 대한 특성을 구분하여 학습된 예측 모델의 성능평가를 진행한다. 기상데이터와 대기오염 물질 데이터를 사용하여 학습 및 평가를 진행하며, 계절별로 구분된 데이터와 $PM_{2.5}$ 의 $36\ \mu\text{g}/\text{m}^3$ 농도 값을 기준으로 저농도와 고농도에 대한 데이터를 이용하여 계절과 농도에 따른 두 가지 예측 모델의 학습 및 평가를 진행한다. 예측 모델의 알고리즘은 LSTM을 동일하게 사용하고 하이퍼 파라미터 탐색을 통해 예측 모델을 설계한다. 설계된 모델의 예측 값을 통해 RMSE, MAPE (mean absolute percentage error), 농도별 정확도 등을 기준으로 계절별 예측 모델과 농도별 예측모델의 성능을 비교 및 분석한다.

II. 상관 분석 및 데이터 구성

2-1 데이터 수집 및 상관 분석

예측 모델의 학습 및 평가에 사용하기 위해 대기오염물질 데이터와 기상데이터를 사용하였다. 2019년 1월부터 2023년 8월까지 시간단위로 측정된 데이터로 기상데이터의 경우 천안시 측정소에서 측정되었으며, 대기오염물질 데이터는 기상데이터 측정소와 가장 근거리에 있는 충북 오창읍의 측정소에서 측정되었다. 수집한 데이터는 온도, 풍속, 풍향, 강수량, 습도, 증기압, 현지기압, 해면기압, 이슬점 온도, 전운량, CO , NO_2 , O_3 , SO_2 , PM_{10} , $PM_{2.5}$ 이다. 수집한 데이터 중 예측 모델의 예측 대상인 $PM_{2.5}$ 와 나머지 데이터들 간의 상관성을 알아보기 위해 상관 분석을 진행하였다. 상관 분석은 전체 기간, 봄, 여름, 가을, 겨울에 해당하는 5가지의 경우로 구분하여 진행하였다. 계절을 구분함에 있어 기상청에서 적용하고 있는 자연 계절을 기준으로 하였다. 자연 계절은 일평균 기온, 일 최저 기온, 일 최고 기온으로 계절을 구분한다. 봄의 경우 일평균 기온이 $5\sim 10^\circ\text{C}$, 일 최저 기온 0°C 이상, 여름의 경우 일평균 기온 $20\sim 25^\circ\text{C}$, 일 최고 기온 25°C 이상, 가을의 경우 일 최고 기온 25°C 이하, 겨울의 경우 일평균 기온 5°C 이하, 일 최저 기온 0°C 이하를 기준으로 한다[11].

표 1은 상관 분석 결과이다. 전체 기간의 경우 기상 데이터 중 온도, 풍속, 증기압, 현지기압, 해면기압, 이슬점 온도가 상관성이 있음을 확인하였으며, 대기오염 물질 중 CO , NO_2 , SO_2 , PM_{10} 의 상관성이 있음을 확인하였다. 봄의 경우 풍속, 습도, CO , NO_2 , SO_2 , PM_{10} 과 상관성을 보였으며, 여름의 경우 풍속, 증기압, 이슬점 온도, 전운량, CO , NO_2 , O_3 , SO_2 , PM_{10} 와 상관성을 보였다. 가을의 경우 온도, 풍속, 습도, 해면기압, CO , NO_2 , SO_2 , PM_{10} 과 상관성을 보였으며, 겨울의 경우 풍속, 증기압, 이슬점 온도, CO , NO_2 , PM_{10} 과 상관성을 보였다.

표 1. 수집 데이터 및 상관분석 결과

Table 1. Collected data correlation analysis result.

| variable | | correlation analysis | | | | |
|-------------------------|--------------------|----------------------|--------|--------|-------|--------|
| | | total | spring | summer | fall | winter |
| Meteorological elements | Temperature | -0.30 | -0.12 | -0.11 | -0.21 | 0.16 |
| | Wind speed | -0.24 | -0.29 | -0.23 | -0.35 | -0.25 |
| | Wind direction | -0.09 | -0.08 | 0.05 | -0.17 | -0.19 |
| | Precipitation | -0.09 | -0.08 | -0.12 | -0.09 | -0.04 |
| | Humidity | 0.02 | 0.20 | -0.12 | 0.21 | 0.18 |
| | Vapor pressure | -0.30 | 0.07 | -0.24 | -0.11 | 0.24 |
| | Station pressure | 0.25 | 0.13 | 0.03 | 0.19 | -0.14 |
| | Sea-level pressure | 0.25 | 0.12 | 0.03 | 0.20 | -0.15 |
| | Dew-point Temp. | -0.25 | 0.11 | -0.23 | -0.06 | 0.29 |
| | Cloud cover | -0.04 | 0.02 | -0.20 | 0.01 | 0.15 |
| Air pollutants | PM_{10} | 0.80 | 0.61 | 0.89 | 0.86 | 0.89 |
| | O_3 | -0.10 | -0.10 | 0.29 | -0.14 | -0.13 |
| | CO | 0.63 | 0.56 | 0.58 | 0.63 | 0.56 |
| | NO_2 | 0.56 | 0.45 | 0.34 | 0.54 | 0.49 |
| | SO_2 | 0.27 | 0.26 | 0.28 | 0.20 | 0.19 |

상관 분석 결과에 따라 각 예측 모델의 학습에 사용할 데이터를 구분하였다. 데이터들의 경우 측정값의 표현 방법과 범위가 상이하다. 이러한 경우 예측 모델의 학습에 영향을 줄 수 있기 때문에 min-max scaling을 통해 데이터의 전처리 후 학습 및 테스트 데이터를 구성하였다.

2-2 데이터 구성

전처리 과정이 끝난 데이터를 이용하여 그림 1과 같이 모델의 학습과 평가를 위한 데이터로 구성하였다. 학습 과정을 위한 training set은 19년도 1월부터 22년도 8월까지의 데이터로 구성하였으며, 학습만을 위해 사용될 데이터인 train set은 training set 중 80%, 학습의 검증을 위한 validation set은 20%로 나누어 구성하였다. 평가를 위한 test set은 22년 9월부터 23년 8월까지 1년의 데이터로 구성하였다.

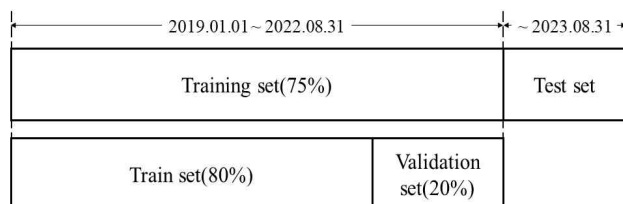


그림 1. 예측 모델의 데이터셋 구조

Fig. 1. Dataset structure of prediction models.

계절별 $PM_{2.5}$ 을 예측하기 위한 모델의 경우 계절에 대한 상관 분석 결과를 기반으로 구분된 데이터를 이용하여 동일한 data set 구조로 4개의 data set을 구성하였다. 그리고 $PM_{2.5}$ 농도 값의 $36 \mu g/m^3$ 값을 기준으로 저농도와 고농도를 구분하여 동일한 data set 구조로 농도에 따른 예측 모델을 위한 data set을 구성하였다.

III. 예측 모델 설계

예측 모델의 경우, 계절과 농도에 따른 예측 결과를 비교 및 평가하기 위해 동일하게 2layer 구조의 LSTM 알고리즘을 이용하여 기본 구조를 설계하였다. 각 layer의 활성화 함수는 relu로 설정하였으며, 최적화 함수와 손실 함수는 adam과 MSE (mean squared error)를 적용하였다. 세부 파라미터의 경우, 각 모델의 최적화를 위해 layer의 unit 수와 dropout rate, batch size는 keras tuner의 random search를 이용하여 하이퍼 파라미터 탐색을 통해 도출된 값으로 설정하였다. Unit의 하이퍼 파라미터 탐색 범위는 20부터 200까지 20의 간격으로 설정하였으며, dropout rate는 0.1부터 0.5까지 0.1의 간격으로 설정하였다. 그리고 batch size는 20부터 200까지 10의 간격으로 설정하여 탐색을 진행하였다. epoch는 200으로 설정하였으며 loss값을 기준으로 학습의 일정 수 동안 변의 폭이 적을 경우, 학습을 멈추도록 early stopping을 설정하였다.

표 2. 하이퍼 파라미터 탐색 결과(계절별)

Table 2. Hyper parameter search result(season)

| parameter | spring | summer | fall | winter |
|--------------|--------|--------|------|--------|
| units | 60 | 180 | 180 | 60 |
| dropout rate | 0.5 | 0.4 | 0.3 | 0.0 |
| batch size | 20 | 100 | 30 | 20 |

표 3. 하이퍼 파라미터 탐색 결과(농도별)

Table 3. Hyper parameter search result(concentration)

| parameter | low concentration | high concentration |
|--------------|-------------------|--------------------|
| units | 100 | 20 |
| dropout rate | 0.5 | 0.5 |
| batch size | 40 | 20 |

3-1 계절별 예측 모델 설계

계절별 예측 모델은 4계절에 대해 개별로 학습을 진행하는 모델이다. 따라서 예측 모델은 계절에 맞춰 4개로 구성하였다. 하이퍼 파라미터도 계절에 대한 4개의 예측 모델에 각각 적용하여 진행하였으며 탐색 결과는 표 2와 같다. 탐색 결과 도출된 값을 적용하여 봄, 여름, 가을, 겨울에 대한 $PM_{2.5}$ 예측 모델을 설계하였으며, 설계된 모델을 통해 미리 구성한 training set을 이용하여 모델 학습 후 test set을 이용하여 예측을 진행하였다. 계절별로 예측된 값을 모아 하나의 결과로 구성하여 평가를 진행하였다.

3-2 저농도와 고농도 예측 모델 설계

농도별 예측 모델의 경우 $36\mu g/m^3$ 의 농도 값을 기준으로 2개의 저농도 예측 모델과 고농도 예측 모델로 구성하였다. 따라서 2개의 모델에 대한 하이퍼 파라미터 탐색을 진행하였으며 탐색 결과는 표 3과 같다. 하이퍼 파라미터 탐색 결과 도출된 값을 기반으로 $PM_{2.5}$ 의 저농도와 고농도에 대한 예측 모델을 설계하였다. 모델 학습의 경우 training set의 $PM_{2.5}$ 농도 값을 기준으로 분리하여 저농도 예측 모델과 고농도 예측 모델의 학습을 진행하였으며, 학습된 2개의 모델의 평가를 위해 저농도와 고농도로 분리된 test set을 이용하여 예측을 진행하였다. 농도별로 예측된 값을 모아 하나의 결과로 구성하여 평가를 진행하였다.

IV. 성능 평가

성능 평가를 위해 각 모델들을 통해 $PM_{2.5}$ 의 예측을 진행하였다. 계절별 예측 모델의 경우 4계절에 대한 4개의 예측 모델

을 이용하여 계절로 구분된 test set을 이용하여 모델 평가를 위한 예측 값을 도출하였다. 농도별 예측 모델의 경우, 저농도와 고농도로 구분된 test set을 이용하여 예측을 진행하였다. 예측 모델의 성능 평가를 위해 정확도, RMSE, MAPE, AQI(air quality index) 별 정확도를 확인하였다.

그림 2와 표 4는 계절별 예측 모델과 농도별 예측 모델의 성능 평가이다. 계절별 예측 모델의 경우 전체 7,803개의 $PM_{2.5}$ 중 5,641개의 예측을 성공하였으며 72.29%의 정확도를 보였다. RMSE와 MAPE로부터 확인한 예측 오차의 경우 8.91, 29.82를 확인하였다. 농도별 예측 모델의 경우 전체 7,745개의 $PM_{2.5}$ 중 6,409개의 예측을 성공하였으며 82.75%의 정확도로 계절별 예측 모델보다 향상된 정확도를 보였다.

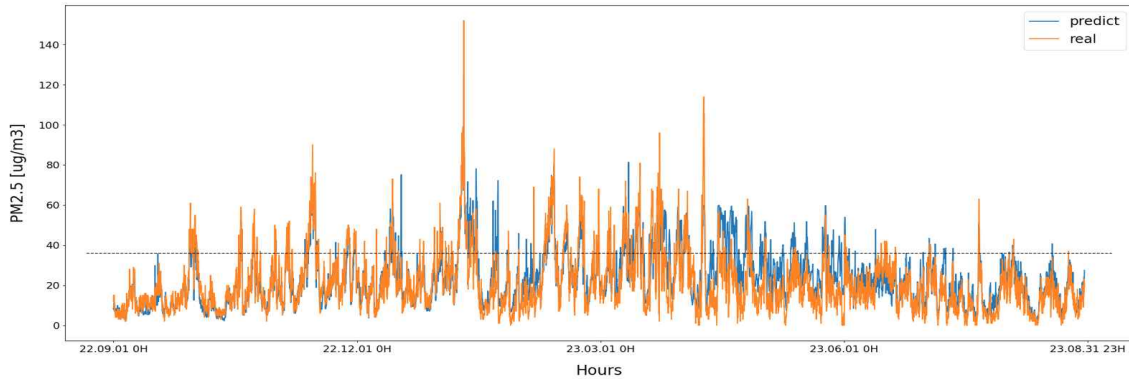
RMSE와 MAPE의 값도 7.74, 25.03으로 계절별 예측 모델보다 오차의 범위가 작은 것을 확인하였다. $PM_{2.5}$ 값의 경우 0부터 시작되며 때로는 200이상의 값도 측정된다. 매우 범위가 넓은 수치형 데이터로 실제 값과 동일한 값을 예측한 것인지 평가하기 위한 정확도 지표의 경우, 해당 모델을 평가하기 위한 지표로 올바르게 못하다.

따라서 $36\mu g/m^3$ 의 농도 값을 기준으로 저농도와 고농도 예측 정확도와 AQI 기준의 농도 값의 범위로 해당 모델들의 예측 성능을 확인하였다. 계절별 예측 모델의 경우 $36\mu g/m^3$ 미만의 농도 예측에서 71.96%의 정확도를 보였으며, $36\mu g/m^3$ 이상의 농도 예측에서 74.37%의 정확도를 보였다.

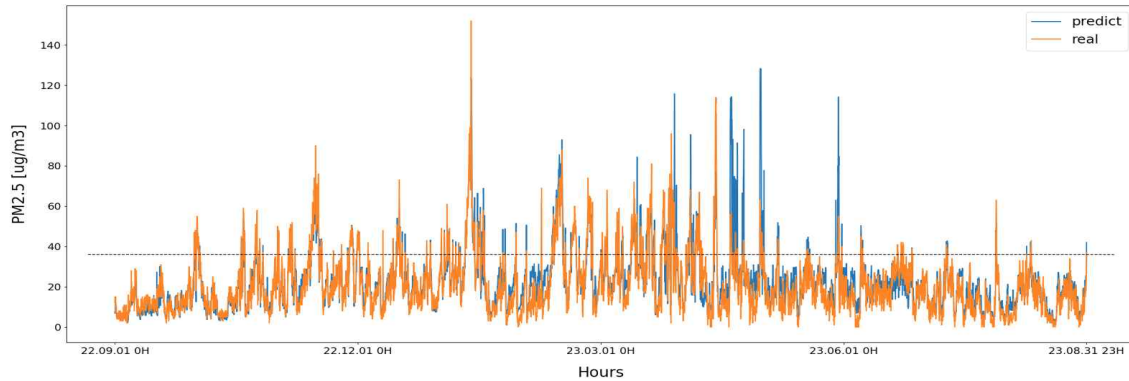
표 4. 예측 성능

Table 4 Prediction performance.

| indicator | prediction model by season | prediction model by concentration |
|---|----------------------------|-----------------------------------|
| accuracy | 72.29% (5641 / 7803) | 82.75% (6409 / 7745) |
| RMSE | 8.91 | 7.74 |
| MAPE | 29.82 | 25.03 |
| low concentration accuracy ($< 36\mu g/m^3$) | 71.96% (4846 / 6734) | 81.67% (5463 / 6689) |
| high concentration accuracy ($36\mu g/m^3 \leq$) | 74.37% (795 / 1069) | 89.58% (946 / 1056) |
| 'Good' accuracy ($< 16\mu g/m^3$) | 69.59% (2334 / 3354) | 73.59% (2455 / 3336) |
| 'Moderate' accuracy ($< 36\mu g/m^3$) | 74.32% (2512 / 3380) | 89.71% (3008 / 3353) |
| 'Bad' accuracy ($< 76\mu g/m^3$) | 75.54% (775 / 1026) | 91.02% (922 / 1013) |
| 'Very bad' accuracy ($76\mu g/m^3 \leq$) | 46.51% (20 / 43) | 55.81% (24 / 43) |



(a) prediction model by season



(b) prediction model by concentration

그림 2. 계절 계절별 예측 모델(a)과 농도별 예측 모델(b)의 PM_{2.5} 예측 결과Fig. 2. PM_{2.5} prediction results of seasonal prediction model(a) and concentration-specific prediction model(b).

농도별 예측 모델의 경우 $36\mu\text{g}/\text{m}^3$ 미만의 농도 예측에서 81.67%의 정확도를 보였으며, $36\mu\text{g}/\text{m}^3$ 이상의 농도 예측에서 89.58%의 정확도를 보였다. AQI 기준의 농도 범위에 따른 ‘ 좋음’, ‘보통’, ‘나쁨’, ‘매우 나쁨’에 대해 계절별 예측 모델은 69.59%, 74.32%, 75.54%, 46.51%의 정확도를 보였다. 농도별 예측 모델은 73.59%, 89.71%, 91.02%, 55.51%를 보였다. 농도별 예측 모델이 계절별 예측 모델보다 좋은 예측 성능을 보였다. 특히, ‘나쁨’의 경우 91.02%의 정확도로 75.54%의 계절별 예측 모델보다 좋은 성능을 보였다. 그러나 ‘매우 나쁨’의 경우, 46.51%, 55.81%의 정확도로 두 모델이 좋지 않은 예측 성능을 보였다. ‘매우 나쁨’에 해당하는 농도의 경우, 모델이 학습하기 위한 데이터의 양이 적음에 따라 발생하는 문제로 판단된다.

V. 결론

본 논문에서는 PM_{2.5}에 영향을 주는 데이터를 계절을 기준으로 구분하여 학습하는 모델과 특정 농도를 기준으로 저농도와 고농도를 구분하여 학습하는 모델을 통해 예측 성능의 비교 및 분석을 진행하였다. 이를 위해 2019년 1월부터 2023년 8월까지 측정된 기상 데이터와 대기오염 물질 데이터를 수집하여

PM_{2.5}와 상관분석을 통해 온도, 풍속, 증기압, 현지기압, 해면기압, 이슬점 온도, CO, NO₂, SO₂, PM₁₀ 데이터가 상관성이 있음을 확인하였다. 해당 데이터들을 학습과 평가에 사용하기 위해 전처리 과정을 진행하여 train set, validation set, test set으로 구성하였다.

예측 모델의 경우 LSTM을 기반으로 설계하였으며 4계절에 대한 4개의 예측 모델과 PM_{2.5} 농도 $36\mu\text{g}/\text{m}^3$ 을 기준으로 구분한 저농도와 고농도 예측 모델을 설계하였다. 각 모델들의 최적의 파라미터를 적용하기 위해 random search를 이용하여 하이퍼 파라미터 탐색을 진행하여 총 6개의 모델을 설계하였다. 설계된 모델을 이용한 평가는 정확도, RMSE, MAPE, 저농도 범위의 예측 정확도, 고농도 범위의 예측 정확도, AQI 기준 4개의 범위에 해당하는 정확도를 이용하였다. 계절별 예측 모델과 농도별 예측 모델의 성능을 비교한 결과 농도별 예측 모델의 성능이 보다 좋은 것을 확인할 수 있었다. 특히 농도별 예측 모델이 $36\mu\text{g}/\text{m}^3$ 이상의 고농도 예측 성능에 계절별 예측 모델에 비해 높은 정확도를 보이는 것을 확인하였으며, PM_{2.5}에 대해 계절에 대한 특성이 미치는 영향이 크지 않다는 것을 확인하였다. AQI 기준 ‘매우 나쁨’ 정확도의 경우, 해당 데이터의 표본 수가 매우 적음으로 학습이 잘 되지 않는 것으로 판단하였다. 향후 이러한 데이터 불균형에 따른 문제가 발생하지 않도록

추가 연구를 진행할 계획이며, XAI를 활용하여 어떤 데이터들이 예측 모델에 실질적으로 영향을 주는지 파악하고 신뢰성을 높이기 위한 연구를 진행할 계획이다.

Acknowledgments

This paper was supported by the Education and Research Promotion Program of KOREATECH in 2022.

References

[1] I. C. Hwang, "The long-term effect of fine particulate matter(PM2.5) on risk of death from environmental disease among older adults," *Journal of Environmental Policy and Administration*, Vol. 31, No. 3, pp. 65-85, Sep. 2023.

[2] S. Y. Park, D. K. Yoon, H. Jang, S. W. Yoon, and C. M. Lee, "Indoor PM2.5 concentration distribution and health risk assessment according to the implementation of a seasonal management system," *Journal of Environmental Health Sciences*, Vol. 49, No. 4, pp. 218-227, Jul. 2023.

[3] D. W. Dockery and C. A. Pope, "Acute respiratory effects of particulate air pollution," *Annual Review of Public Health*, Vol. 15, No. 1, pp. 107-132, 1994.

[4] K. Katsouyanni, G. Touloumi, C. Spix, J. Schwartz, F. Balducci, S. Medina, G. Rossi, B. Wojtyniak, J. Sunyer, L. Bacharova, J. P. Schouten, A. Ponka, and H. R. Anderson, "Short term effects of ambient sulphur dioxide and particulate matter on mortality in 12 European cities: results

from time series data from the APHEA project," *BMJ*, Vol. 314, No. 795, Jun. 1997.

[5] J. Y. Shin, "Analysis of the impact of fine dust on health and policy implications," *Korea Institute for Health and Social Affairs*, Vol. 435, pp. 1-10, Apr. 2023.

[6] W. G. Do, D. Y. Kim, H. J. Song, and G. J. Cho, "A study on the PM2.5 forecasting method in busan using deep neural network," *Journal of Environmental Science International*, Vol. 32, No. 8, pp. 595 – 611, 2023.

[7] H. J. Lee, M. K. Jee, H. D. Kim, T. H. Jun, and C. W. Kim, "Early prediction of fine dust concentration in seoul using weather and fine dust information," *Journal of Broadcast Engineering*, Vol. 28, No. 3, pp. 285-292, May. 2023.

[8] S. Y. Kim, "A study on PM10 forecasting model using machine learning," *Journal of the Korean Data & Information Science Society*, Vol. 34, No. 5, pp. 763-773. Sep. 2023.

[9] J. Y. Lee, C. Y. Lee, M. W. Jung, J. Y. Ahn, K. H. Wang, D. R. Choi, and H. Y. Yun, "XAI analysis of DNN using PM2.5 component input data and improvement of PM2.5 prediction performance," *Journal of Korean Society for Atmospheric Environment*, Vol. 39, No. 4, pp. 411-426, Aug. 2023.

[10] Y. G. Hur and M. G. Kang, "The effects of urban spatial structure and meteorological factors on the high concentration of fine dust pollution," *Journal of Korea Planning Association*, Vol. 57, No. 1, pp. 145-160, Feb. 2022.

[11] Encyclopedia of Korean Culture. Spring [Internet]. Available: <https://encykorea.aks.ac.kr/Article/E0023778/>.



정 용 진 (Yong-Jin Jung)

2018년 3월 ~ 현재 한국기술교육대학교 전기전자통신공학과 박사과정
 2016년 2월 한국기술교육대학교 전기전자통신공학과 공학석사
 2014년 2월 공주대학교 전기전자제어공학부 전자공학·나노정보공학전공 전자공학트랙 공학사
 ※관심분야 : 미세먼지 예측, 기계 학습, 인공지능경량, 심층신경망



오 창 현 (Chang-Heon Oh)

1999년 3월 ~ 현재 한국기술교육대학교 전기전자통신공학부 교수,
 2006년 8월 ~ 2007년 7월 방문교수(University of Wisconsin-Madison)
 1993년 10월 ~ 1999년 2월 삼성전자(주) CDMA 개발팀 선임연구원
 1990년 2월 ~ 1993년 8월 한진전자(주) 기술연구소 선임연구원
 1996년 2월 한국항공대학교 항공전자공학과 (공학박사)
 1990년 2월 한국항공대학교 항공통신정보공학과 (공학석사)
 1988년 2월 한국항공대학교 항공통신공학과 (공학사)
 ※관심분야 : 무선/이동통신, IoT, 기계학습 기반 통신시스템