

특집논문 (Special Paper)

방송공학회논문지 제28권 제3호, 2023년 5월 (JBE Vol.28, No.3, May 2023)

<https://doi.org/10.5909/JBE.2023.28.3.285>

ISSN 2287-9137 (Online) ISSN 1226-7953 (Print)

기상 및 미세먼지 정보를 활용한 서울시의 미세먼지 농도 조기 예측

이한주^{a)} 지민규^{b)}, 김학동^{c)}, 전태흠^{d)}, 김청원^{e)†}

Early Prediction of Fine Dust Concentration in Seoul using Weather and Fine Dust Information

HanJoo Lee^{a)}, Minkyu Jee^{b)}, Hakdong Kim^{c)}, Taeheul Jun^{d)}, and Cheongwon Kim^{e)†}

요 약

최근 미세먼지가 건강에 미치는 영향은 큰 화두가 되고 있다. 미세먼지는 코의 점막에 걸러지지 않고 인체 내부까지 침투하여 호흡기에 영향을 미치기 때문에 위험하다. 미세먼지는 산업과 직접적으로 연관되어있기 때문에 미세먼지를 제거하는 것은 사실상 불가능하다. 따라서 미세먼지 농도를 사전에 예측할 수 있다면 사전 조치를 취해 인체에 미치는 영향을 줄일 수 있다. 미세먼지는 하루 600km 이상 이동할 수 있는 특성을 가진다. 이러한 특성으로 인해 미세먼지는 인접 구뿐만 아니라 멀리 떨어져있는 구에도 직접적인 영향을 미친다. 본 논문에서는 풍향, 풍속 데이터와 시계열 예측 모델을 이용하여 서울특별시의 미세먼지 농도를 예측하고, 서울특별시의 미세먼지 농도와 지역별 미세먼지 농도의 상관관계를 확인했다. 또한, 각 지역별 미세먼지 농도와 서울특별시의 미세먼지 농도를 이용하여 예측을 진행했다. 예측 결과에서 가장 낮았던 MAE(평균 절대 오차)는 12.13으로 선행연구에서 제시된 MAE인 14.3 보다 약 15.17% 더 예측성능이 향상된 것을 확인했다.

Abstract

Recently, the impact of fine dust on health has become a major topic. Fine dust is dangerous because it can penetrate the body and affect the respiratory system, without being filtered out by the mucous membrane in the nose. Since fine dust is directly related to the industry, it is practically impossible to completely remove it. Therefore, if the concentration of fine dust can be predicted in advance, pre-emptive measures can be taken to minimize its impact on the human body. Fine dust can travel over 600km in a day, so it not only affects neighboring areas, but also distant regions. In this paper, wind direction and speed data and a time series prediction model were used to predict the concentration of fine dust in Seoul, and the correlation between the concentration of fine dust in Seoul and the concentration in each region was confirmed. In addition, predictions were made using the concentration of fine dust in each region and in Seoul. The lowest MAE (mean absolute error) in the prediction results was 12.13, which was about 15.17% better than the MAE of 14.3 presented in previous studies.

Keyword : E-health, Fine dust, Time series, Informer

I. 서론

미세먼지는 인체에 직접적인 영향을 미칠 수 있기 때문에 최근 국내외에서 중요한 이슈가 되고 있다. 미세먼지는 복잡한 성분을 가진 대기 중에서 부유하고 있는 물질로, 자동차 배기가스, 도로의 먼지 등으로 부터 발생한다. 미세먼지의 대표적인 영향은 기도의 염증 반응이다. 미세먼지는 주로 세기관지에서 염증반응을 일으켜 천식, 만성기관지염, 기도폐쇄 등을 일으키거나 악화시킨다. 미세먼지는 호흡기에 영향을 미칠 뿐만 아니라 활성산소와 산화스트레스를 증가시키는 것으로 알려져 있다^[1]. 특히 미세먼지는 코털과 기도의 섬막에서 걸러지지 않고 인체 내부인 폐까지 직접 침투하여 매우 위험하다. 미세먼지는 호흡기뿐만 아니라 심혈관, 뇌혈관 질환에 의한 사망 원인이고 그 영향력은 점점 증가하고 있다^[2]. 세계보건기구 산하의 국제 암 연구소 (IARC)에서는 미세먼지를 인간에게 발암성이 있는 것으로 확인된 1급 발암물질로 분류했다^[3]. Park et al.^[4]는 대기오염물질 농도와 미세먼지 농도가 알레르기 비염 외래이용에 미치는 영향을 분석했고 알레르기비염의 외래이용건수에 통계적으로 유의한 영향을 준다는 결과를 제시했다. 하지만 미세먼지는 산업과 직접적으로 연관돼 있어 완전히 제거하는 것은 사실상 불가능하다. 따라서 미세먼지를 줄이는 정책을 조금씩 실행해 나감과 동시에 미세먼지 농도를 사전에 정확하게 예측하여 대비해야 한다. 만약 미세먼지 농도를 사전에 정확하게 예측할 수 있다면 그에 맞게 유기적인 대응이 가능하다. 국내의 경우 산업의 급속한 성장과 인구의 증가로 인해 미세먼지가 꾸준히 증가해왔다. 2016

년 예일대학교에서 발표한 Environmental Performance Index(EPI)에 따르면 대한민국의 Air Quality 분야는 조사한 180개국 중 173위에 위치했다^[5]. 하지만 미세먼지가 국내에서 중요한 이슈가 되면서, 이것을 개선하려는 노력이 지속되고 있다. 미세먼지 감소 정책부터 연구 등 활발하게 개선을 위한 방안이 제시되고 있다. EPI에서 가장 최근 발표한 2020년도 자료에 따르면 국내 Air Quality는 180개국 중 28위까지 상승했다. 특히 아시아에서는 일본 다음으로 2위를 차지했다. 또한 2020년도의 평균 초미세먼지 농도는 $19\mu\text{g}/\text{m}^3$ 으로 낮아졌으며, 365일 중 미세먼지가 나쁨 이상인 일수는 27일로 감소하였다. 이것은 2019년도에 비해 20일이나 감소한 값이다. 감축 정책이 효과를 나타내는 것에 비해 미세먼지 예측에 관해서는 추가적인 연구가 필요하다. 국내에서는 미세먼지에 대비하기 위해 Air Korea에서 미세먼지 농도 정보를 제공하고 있다. 데이터는 각 지역별로 측정된 미세먼지와 대기 오염물질의 농도 및 주간 미세먼지 위험단계 예보 정보를 제공한다. 미세먼지와 대기오염물질 농도는 ppm과 $\mu\text{g}/\text{m}^3$ 등 각각의 단위를 가지는 값들을 제공한다. 위험단계 예보는 미래의 미세먼지 농도를 좋음, 보통, 나쁨, 매우 나쁨의 네 단계로 나누어 예측한다. 본 논문에서는 지역별 미세먼지 농도의 상관관계를 확인하고 Informer 모델을 이용한 예측 결과를 확인했다. 예측 결과 내에서 가장 낮았던 MAE는 12.13으로 선행연구에서 제시된 MAE인 14.3 보다 약 15.17% 더 예측성능이 향상된 것을 확인했다.

II. 선행연구

Jeon et al.^[6]은 미세먼지와 관계된 요소들을 분석하고, 해당 요소를 활용해 딥러닝 신경망을 구성하여 미세먼지를 예측하였다. 모델은 Fully connect layer 모델을 사용했으며, 3개의 은닉층으로 구성됐다. Soh et al.^[7]는 KNN 모델을 활용해 기상 정보 및 미세먼지 정보를 그룹화하고, 입력된 기상 및 미세먼지 정보를 활용해 미세먼지 농도를 예측하였다. Kim et al.^[8]은 Long-Short Term Memory(LSTM)을 이용하여 미세먼지 및 대기오염 물질의 농도를 예측하였다. Zamani et al.^[9]는 다양한 기상 정보를 활용해 PM2.5의 농도를 예측하는 모델을 제시하였다. Randomforest, 딥

a) 세종대학교 일반대학원 소프트웨어학과(Department of Software, Sejong University)

b) (주)셀버스(Selvers Inc.)

c) 세종대학교 일반대학원 디지털콘텐츠학과(Department of Digital Contents, Sejong University)

d) 세종대학교 일반대학원 소프트웨어융합학과(Department of Software Convergence, Sejong University)

e) 세종대학교 소프트웨어학과(Department of Software, Sejong University)

‡ Corresponding Author : 김청원(Cheongwon Kim)
E-mail: wikim@sejong.ac.kr
Tel: +82-2-3408-3795
ORCID:https://orcid.org/0000-0002-1489-8427

※ This work is financially supported by Korea Ministry of Environment (MOE) as 「Graduate School specialized in Climate Change.
· Manuscript March 28, 2023; Revised May 18, 2023; Accepted May 18, 2023.

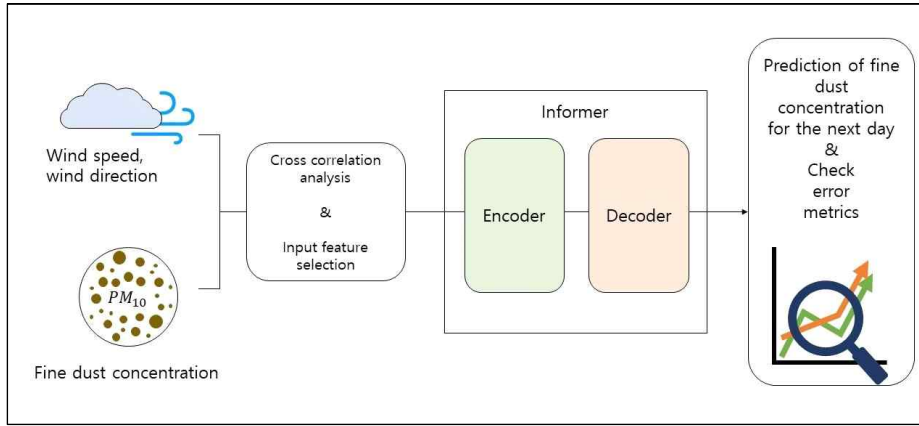


그림 1. 논문의 개요
 Fig. 1. Overview of the paper

러닝 등 다양한 모델을 활용해 학습하고 결과를 비교 분석하였다. 모델은 이전 시간들의 미세먼지 및 대기 오염물질 농도를 LSTM에 입력하고 미래의 농도를 예측하도록 학습했다. Xayasouk et al.^[10]은 LSTM 모델과 DAE(Deep Autoencoder)를 이용한 예측 모델을 제시하였다. 모델은 기상 정보를 이용해서 미세먼지 농도를 예측하고 LSTM이 더 나은 예측 성능을 보였다는 결과를 제시하였다. Lee et al.^[11]는 1-D CRNN 모델을 활용해 미세먼지 위험도 레벨을 예측하였다. 모델은 1-D CNN을 활용해 입력된 기상 및 미세먼지 정보의 Feature Map을 추출하여 RNN 및 LSTM에 활용하였다. Kim et al.^[12]은 다중회귀모델을 활용해 미세먼지 농도를 예측하였다. 모델은 획득한 데이터 중에서 분석을 통해 미세먼지 예측에 유의미한 관계를 가지는 데이터를 분석하여 활용하였다. Park^[13]는 방향성 그래피컬 모델을 이용하여 미세먼지의 이동경로를 분석하였다. 해당 논문은 미세먼지 농도는 인접 지역 간에 연관성이 높으며 계절성 풍향의 영향을 받는다는 결과를 제시한다. 또한, 미세먼지의 일별 평균 농도를 이용하였기 때문에 인접 구뿐 아니라 멀리 떨어져 있는 구끼리도 영향을 받는다는 것을 제시한다. Dairi et al.^[14]는 Variational Autoencoder(VAE) 와 Integrated Multiple Directed Attention (IMDA)를 이용하여 대기오염물질의 농도를 예측하는 모델을 제시한다. 2016년 환경부에서 발표한 국내 대기질 예보 정확도 향상 연구에서는 국내 현업에서 사용 중인 13개의 예측 모델에 대한 정확도를 분석하고 예보 정확도의 향상을 위한 중장기 로

드맵을 제안하였다. 해당 연구에 따르면, 국내에서 사용 중인 13개의 예측모델에서 미세먼지 예보의 정확도는 수도권 및 강원권에 대해 63%, 고농도일 때에는 44%의 정확도를 보인다. 또한, 미세먼지 농도의 측정값과 예측 값의 오차율은 $\pm 27\%$ 임을 제시하였다. Jo et al.^[15]는 LSTM과 풍향 풍속 데이터를 이용하여 다음날 서울시의 미세먼지 농도를 예측하였다. 또한, 풍향과 풍속이 LSTM을 학습시키는 데에 적합한 데이터이고 다음날 서울시의 미세먼지 농도를 예측한 결과 MAE는 14.3임을 제시하였다.

III. 데이터

1. 기상 데이터

미세먼지 농도 예측을 위해 기상청에서 제공하는 데이터셋을 활용하였다. 기상청은 각 관측소에서 측정된 기상 정보를 1시간 단위로 제공한다. 데이터셋은 관측소의 번호와 측정 일시를 기반으로 온도, 습도, 풍속, 풍향, 기압 등 해당 지역의 다양한 기상 데이터를 제공한다. 데이터는 2015년부터 1월 1일부터 2019년 12월 31일까지 총 5년간의 모든 관측소에서 측정된 기상 데이터를 수집하였다. 풍속은 m/s 의 단위로 풍향은 16방위를 기준으로 데이터를 제공하고 있다. Jo et al.^[14]는 풍향과 풍속과 미세먼지 농도의 상관관계를 정량적으로 분석하고 LSTM을

이용해 미세먼지 농도를 예측하였다. 해당 논문은 풍향과 풍속이 미세먼지 농도를 예측하는 데에 적합하다는 결과를 제시하였다. 본 논문에서는 기상 데이터 중 풍속과 풍향을 이용하여 서울특별시의 미세먼지 농도를 예측한 결과를 확인하였다. 풍향과 풍속은 광역지방자치단체별로 일별 평균 데이터를 이용하였다.

2. 미세먼지 데이터

미세먼지 농도 데이터는 에어코리아에서 제공하는 데이터셋을 활용하였다. 에어코리아는 각 지역별 관측소에서 측정된 대기 오염물질 농도를 제공한다. 에어코리아는 전국의 대기측정망에서 측정되는 아황산가스(SO₂), 일산화탄소(CO), 이산화질소(NO₂), 오존(O₃), 미세먼지(PM₁₀), 초미세먼지(PM_{2.5}) 등 대기정보 자료를 수집하고 관리하는 국가대기오염정보관리시스템(NAMIS)을 2004년 4월부터 구축하여 정보를 제공하고 있다. 대기오염물질 농도는 1시간 단위의 측정된 데이터를 제공한다. 본 논문에서는 2015년 1월 1일부터 2019년 12월 31일까지 총 5년간의 데이터를 이용하였다. 세종특별자치시의 경우 2018년 이후에 설치된 관측소들로 인해 결측값이 많아 데이터에서 제외하였다. 세종특별자치시를 제외한 지역에 미세먼지 농도 결측값이 있을 경우 전날의 미세먼지농도 평균값을 이용하여 결측값을 채워주었고 일별 평균 미세먼지 농도를 이용하였다.

3. 상관관계 확인

지역별 미세먼지 농도가 서울특별시의 미세먼지 농도에 미치는 영향을 확인하기 위해 상관관계를 확인하였다. 데이터는 2015년 1월 1일부터 2019년 12월 31일까지의 일별

표 1. 상관계수에 따른 상관관계 판단 기준
Table 1. Criteria for determining the correlation relationship according to the correlation coefficient

| Correlation coefficient (r) | Correlation |
|-----------------------------|---|
| More than ±0.9 | correlation relationship is very high |
| ±0.7 ~ 0.9 | correlation relationship is high |
| ±0.4 ~ 0.7 | correlation relationship exists |
| ±0.2 ~ 0.4 | correlation relationship exists, but is low |
| Less than ±0.2 | no correlation relationship |

평균 미세먼지 농도를 이용하여 상관관계를 확인하였다. 상관계수는 아래의 Table 1을 기준으로 확인하였다.

Table 2는 지역별 미세먼지 농도와 서울특별시 미세먼지 농도의 상관관계이다. 상관계수가 가장 높게 나타난 지역은 경기도와 인천광역시로 서울특별시와 인접한 지역이 상관계수가 가장 높게 나타났다. 경기도, 인천광역시, 충청남도, 충청북도의 상관계수는 양의 상관관계를 가지며 그 정도가 아주 높은 수치인 것으로 나타났다.

표 2. 지역별 미세먼지 농도와 서울특별시의 미세먼지 농도 상관관계
Table 2. Correlation between fine dust concentrations in different regions and fine dust concentration in Seoul

| Region | Correlation coefficient |
|-------------------|-------------------------|
| Gyeonggi-do | 0.9881 |
| Incheon | 0.9629 |
| Chungcheongnam-do | 0.9137 |
| Chungcheongbuk-do | 0.9042 |
| Gangwon-do | 0.8976 |
| Daejeon | 0.874 |
| Jeollabuk-do | 0.8387 |
| Gyeongangbuk-do | 0.8012 |
| Daegu | 0.7854 |
| Gwangju | 0.7696 |
| Jeollanam-do | 0.748 |
| Gyeongsangnam-do | 0.7242 |
| Ulsan | 0.6761 |
| Busan | 0.675 |
| Jeju island | 0.6175 |

과거 데이터를 활용하여 예측을 진행하기 위해 각 지역별 미세먼지 농도가 다음날 서울특별시의 미세먼지 농도에 미치는 영향을 확인하였다. Table 3은 지역별 미세먼지 농도와 다음날 서울특별시의 미세먼지 농도의 상관계수 표이다. 다음날 서울특별시의 미세먼지 농도와의 상관계수는 강원도 지역이 가장 높게 나타났다. Park^[13]는 하루에 600km 이상 이동 가능한 미세먼지의 특성으로 인해 인접 지역뿐만 아니라 멀리 떨어져 있는 지역에도 영향을 미친다는 것을 제시하였다. 이러한 미세먼지의 특성으로 인해 거리가 멀리 떨어져 있는 강원도 지역의 미세먼지 농도는 다음날의 서울특별시 미세먼지 농도와의 상관계수가 가장 높게 나타났다.

표 3. 지역별 미세먼지 농도와 다음날 서울특별시의 미세먼지 농도 상관관계
 Table 3. Correlation between fine dust concentrations in different regions and fine dust concentration in Seoul on the following day

| Region | Correlation coefficient |
|-------------------|-------------------------|
| Gangwon-do | 0.653 |
| Chungcheongbuk-do | 0.6302 |
| Gyeonggi-do | 0.6238 |
| Gyeongsangbuk-do | 0.6236 |
| Daejeon | 0.6105 |
| Seoul | 0.6104 |
| Jeollabuk-do | 0.6061 |
| Chungcheongnam-do | 0.5931 |
| Daegu | 0.5887 |
| Incheon | 0.5772 |
| Gyeongsangnam-do | 0.5657 |
| Busan | 0.5585 |
| Ulsan | 0.5535 |
| Gwangju | 0.5418 |
| Jeollanam-do | 0.5413 |
| Jeju island | 0.4593 |

IV. 실험 및 결과

1. 이용 모델

Informer 모델은 Transformer 기반의 모델로 Zhuo et al.^[16]이 제시한 모델이다. 해당 모델은 기존 Transformer 모델이 self-attention mechanism으로 인해 연산이 복잡하고 메모리 사용량이 크다는 문제점을 제시하며 이를 해결하기 위해 Sparse attention을 이용한다. 모델은 단변량(univariate)과 다변량(multivariate) 데이터를 입력 값으로 이용하여 예측이 가능하다. 본 논문에서는 서울특별시의 미세먼지 농도 예측에 Informer 모델을 이용하였다. 모든 예측 값은 100 epoch 동안 손실함수 MAE를 기준으로 학습된 모델의 예측 결과 값을 비교했다.

표 5. 상관계수가 가장 높았던 강원도 지역의 데이터를 이용하여 서울특별시 미세먼지 농도 예측 결과
 Table 5. Prediction results for fine dust concentration in Seoul using data from Gangwon Province, which had the highest correlation coefficient

| Input data | MAE ($\mu\text{g}/\text{m}^3$) | MAPE (%) | MSE ($\mu\text{g}/\text{m}^3$) | RMSE ($\mu\text{g}/\text{m}^3$) |
|---|----------------------------------|----------|----------------------------------|-----------------------------------|
| Fine dust concentration in Seoul | 12.25 | 37.23 | 309.72 | 17.60 |
| Fine dust concentration in Seoul, Fine dust concentration in Gangwon-do | 12.23 | 35.69 | 305.28 | 17.47 |

2. 서울특별시의 데이터를 이용한 예측 결과

본 절에서는 서울특별시의 데이터인 미세먼지 농도, 풍향과 풍속을 Informer 모델에 입력하였을 때 예측 결과를 확인하였다. 2015년 1월1일부터 2019년 12월 31일 까지의 데이터를 이용했고 총 1826개 데이터에 대하여 train set, validation set, test set의 비율을 7:1:2로 설정하고 100 epoch 동안 학습한 모델을 기준으로 예측을 진행하였다. Test set에 해당하는 서울특별시의 미세먼지 농도는 2019년 1월 1일부터 12월 31일까지의 데이터로 일별 평균 농도는 약 $42.66\mu\text{g}/\text{m}^3$ 이었다. 서울특별시의 미세먼지 농도를 이용한 예측 결과는 MAE가 $12.25\mu\text{g}/\text{m}^3$ 으로 나타났다. 이는 평균이 $42.66\mu\text{g}/\text{m}^3$ 인 test set에 대하여 Informer 모델의 예측 값이 평균적으로 약 $\pm 12.25\mu\text{g}/\text{m}^3$ 의 오차를 보였음을 나타낸다. 서울특별시의 풍속을 이용한 예측 결과에서 MAE는 $14.47\mu\text{g}/\text{m}^3$ 으로, 풍향을 이용한 결과에서는 $16.23\mu\text{g}/\text{m}^3$ 으로 나타났다. 서울특별시의 미세먼지 농도만을 이용한 예측 결과보다 각각 약 18% 와 32% 더 높은 어려움을 보였다. 이를 통해, 풍향과 풍속 데이터는 Informer 모델의 학습에 적합하지 않다는 것을 확인하였다.

표 4. 서울특별시 입력 데이터에 따른 예측 결과
 Table 4. Prediction results based on input data for Seoul

| Input data | MAE ($\mu\text{g}/\text{m}^3$) | MAPE (%) | MSE ($\mu\text{g}/\text{m}^3$) | RMSE ($\mu\text{g}/\text{m}^3$) |
|-------------------------|----------------------------------|----------|----------------------------------|-----------------------------------|
| Fine dust concentration | 12.25 | 37.23 | 309.72 | 17.60 |
| Wind speed | 14.47 | 45.60 | 419.99 | 20.49 |
| Wind direction | 16.23 | 55.13 | 447.49 | 21.15 |

3. 강원도 지역의 데이터를 이용한 예측 결과

Table 5는 다음날 서울특별시의 미세먼지 농도와 상관계수 0.65로 가장 높았던 강원도 지역의 데이터를 이용한 예

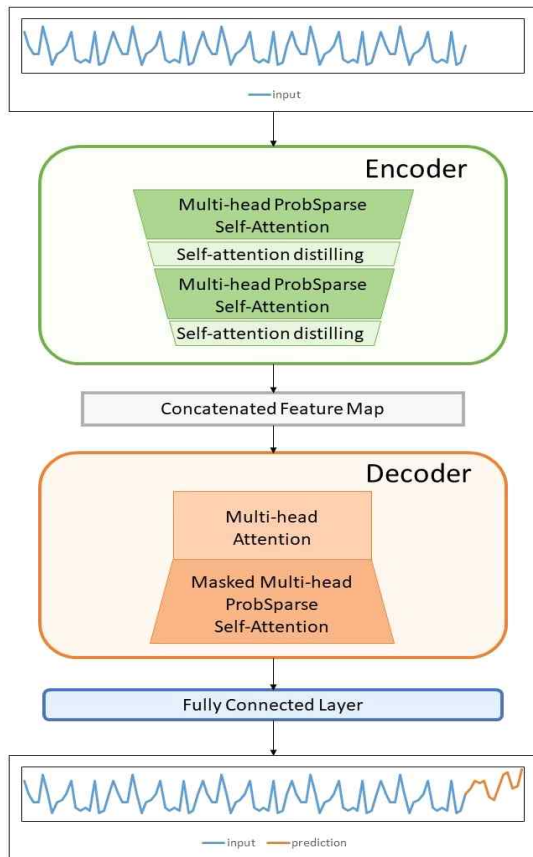


그림 2. Informer 모델 구조
Fig. 2. Informer model structure

측 결과이다. 서울특별시의 미세먼지 농도 데이터와 강원도 지역의 미세먼지 농도를 이용하여 예측하였다. 강원도 지역의 미세먼지 농도와 서울특별시의 미세먼지 농도를 이용한 예측 결과에서는 서울특별시의 미세먼지 농도만을 이용했을 경우보다 MAE를 기준으로 약 0.02 감소했고, MAPE를 기준으로 약 1.5% 더 감소한 것으로 나타났다.

4. 강원도 지역의 데이터를 이용한 예측 결과

본 절에서는 지역별 미세먼지 농도와 서울특별시의 미세먼지 농도를 이용한 예측결과를 확인하였다. Table 6은 상관계수가 높았던 지역 순으로 정렬된 예측한 결과이다. MAE는 충청남도의 미세먼지 농도를 이용한 예측 결과에서 $12.13 \mu\text{g}/\text{m}^3$ 으로 가장 낮았다. MAPE는 상관계수가 가

표 6. 지역별 미세먼지 농도와 서울특별시의 미세먼지 농도를 입력 값으로 이용한 예측 결과

Table 6. Prediction results using fine dust concentrations in different regions and fine dust concentration in Seoul as input values

| Region | MAE ($\mu\text{g}/\text{m}^3$) | MAPE (%) | MSE ($\mu\text{g}/\text{m}^3$) | RMSE ($\mu\text{g}/\text{m}^3$) |
|-------------------|----------------------------------|----------|----------------------------------|-----------------------------------|
| Gangwon-do | 12.23 | 35.69 | 305.28 | 17.47 |
| Chungcheongbuk-do | 12.44 | 37.27 | 319.91 | 17.89 |
| Gyeonggi-do | 12.2 | 36.66 | 313.52 | 17.71 |
| Gyeongsagbuk-do | 12.85 | 38.85 | 335.12 | 18.31 |
| Daejeon | 12.25 | 35.89 | 319.58 | 17.88 |
| Jeollabuk-do | 12.3 | 36.32 | 316.93 | 17.8 |
| Chungcheongnam-do | 12.13 | 36.78 | 298.8 | 17.29 |
| Daegu | 12.42 | 36.49 | 323.76 | 17.99 |
| Incheon | 12.23 | 37.82 | 307.41 | 17.53 |
| Gyeongsangnam-do | 12.72 | 37.71 | 325.28 | 18.04 |
| Busan | 12.54 | 36.49 | 321.14 | 17.92 |
| Ulsan | 12.86 | 38.05 | 332.43 | 18.23 |
| Gwangju | 12.89 | 40.74 | 326.99 | 18.08 |
| Jeollanam-do | 12.7 | 39.09 | 321.52 | 17.93 |
| Jeju island | 13.07 | 40.87 | 331.44 | 18.21 |

장 높았던 강원도 지역의 미세먼지 농도를 이용한 예측에서 35.69%로 가장 낮게 나타났다.

V. 결론

본 논문에서는 Informer 모델을 이용하여 다음날 서울특별시의 미세먼지 농도를 예측하였다. 충청남도의 미세먼지 농도와 서울특별시의 미세먼지 농도를 이용한 예측 결과에서 MAE가 $12.13 \mu\text{g}/\text{m}^3$ 으로 가장 낮게 나타났다. 이는 풍향과 풍속 데이터를 이용한 선행연구^[15]의 예측 결과보다 약 15.17% 더 낮아진 수치였다. 다음날 서울특별시 미세먼지 농도와 상관계수가 가장 높았던 강원도 지역의 미세먼지 농도를 이용한 결과에서 MAPE가 35.69%로 가장 낮게 나타났다. 미세먼지는 하루에 600km 이상 이동할 수 있으며, 강원도와 충청남도는 서울시와 유사한 거리에 위치해 있다. 이에 따라 강원도 혹은 충청남도 지역의 미세먼지 농도를 함께 이용하여 서울시의 미세먼지 농도를 예측할 때, 예측 정확도를 높일 수 있는 것으로 나타났다. 본 연구는 서울특별시의 미세먼지 농도 예측에서 충청남도 지역과 강원도 지역의 미세먼지 농도를 이용하는 것이 예측 성능을 향상시켰음을 시사한다.

참 고 문 헌 (References)

- [1] D. Shin, "Health Effects of Ambient Particulate Matter", Journal of Korean Medical Association, Vol.50, No.2, , pp.175-182, 2007.
doi: <https://doi.org/10.5124/jkma.2007.50.2.175>
- [2] G. Shin, J. Kim, and Y. Choi, "A Study on the Data Model Design of Fine Dust Related Disease", Journal of The Korea Society of Information Technology Policy & Management (ITPM), Vol.10, No.1. pp.655-659, 2018. <http://scholarworks.bwise.kr/ssu/handle/2018.sw.ssu/31956>
- [3] D. Loomis et al., "The carcinogenicity of outdoor air pollution.", Lancet Oncology, Vol.14, No.13, pp.1262 ,2013.
doi: [https://doi.org/10.1016/S1470-2045\(13\)70487-X](https://doi.org/10.1016/S1470-2045(13)70487-X)
- [4] J. Park, Y. Park, E. Lee and K. Lee, "Analysis on the Effects of Particular Matter Distribution on the Number of Outpatient Visits for Allergic Rhinitis", Health Policy and Management Vol.30, No.1, pp.50-61, 2020.
doi: <https://doi.org/10.4332/KJHPA.2020.30.1.50>
- [5] W. Jung, "South Korea's Air Pollution: Gasping for Solutions", Policy Brief, Vol.199, pp.1-4, 2017.
- [6] S. Jeon, and Y. Son, "Prediction of fine dust PM 10 using a deep neural network model", The Korean journal of applied statistics, Vol.31, No.2, pp.265-285, 2018
doi: <https://doi.org/10.5351/KJAS.2018.31.2.265>
- [7] P. Soh, J. Chang, and J. Huang, "Adaptive deep learning-based air quality prediction model using the most relevant spatial-temporal relations", Ieee Access, Vol.6, pp.38186-38199, 2018.
doi: <https://doi.org/10.1109/ACCESS.2018.2849820>
- [8] S. Kim, J. Lee, and J. Seo, "Deep-dust: Predicting concentrations of fine dust in Seoul using LSTM", arXiv preprint arXiv:1901.10106, 2019.
doi: <https://doi.org/10.48550/arXiv.1901.10106>
- [9] Z. Joharestani et al., "PM_{2.5} prediction based on random forest, XGBoost, and deep learning using multisource remote sensing data", Atmosphere, Vol.10, No.7, pp.373, 2019.
doi: <https://doi.org/10.3390/atmos10070373>
- [10] T. Xayasouk, H. Lee, and G. Lee, "Air pollution prediction using long short-term memory (LSTM) and deep autoencoder (DAE) models", Sustainability, Vol.12, No.6, pp. 2570, 2020.
doi: <https://doi.org/10.3390/su12062570>
- [11] K. Lee, W. Hwang, and M. Choi, "Design of a 1-D CRNN Model for Prediction of Fine Dust Risk Level", The Society of Digital Policy and Management, Vol.19, No.2, pp.215-220, 2021.
doi: <https://doi.org/10.14400/JDC.2021.19.2.215>
- [12] H. Kim, and T. Moon, "Machine learning-based Fine Dust Prediction Model using Meteorological data and Fine Dust data", Vol.24, No.1, pp.92-111, 2021.
doi: <https://doi.org/10.11108/kagis.2021.24.1.092>
- [13] G. Park, "Discovering a fine dust pathway via directed acyclic graphical models", Journal of the Korean Data & Information Science Society, Vol.30, No.1, pp.67-76, 2019.
doi: <https://doi.org/10.7465/jkdi.2019.30.1.67>
- [14] A. Dairi et al., "Integrated multiple directed attention-based deep learning for improved air pollution forecasting", IEEE Transactions on Instrumentation and Measurement, Vol.70, pp.1-15, 2021.
doi: <https://doi.org/10.1109/TIM.2021.3091511>
- [15] S. Jo, M. Jeong, J. Lee, I. Oh, and Y. Han, "Analysis of Correlation of Wind Direction/Speed and Particulate Matter(PM10) and Prediction of Particulate Matter Using LSTM", Proceeding of Korean Institute of Information Scientists and Engineers, pp.1649-1651, 2020.
- [16] H. Zhou et al., "Informer: Beyond efficient transformer for long sequence time-series forecasting", Proceedings of the AAAI conference on artificial intelligence, Vol.35, No.12, pp.11106-111115, 2021.
doi: <https://doi.org/10.1609/aaai.v35i12.17325>

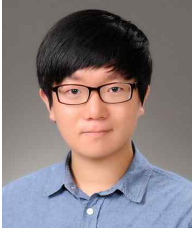
저 자 소 개

이 한 주

- 2021년 : 세종대학교 컴퓨터공학과 학사
- 2022년 ~ 현재 : 세종대학교 소프트웨어학과 석사과정
- ORCID : <https://orcid.org/0009-0005-3512-8594>
- 주관심분야 : 머신러닝, 딥러닝, 컴퓨터 비전



저 자 소 개



지 민 규

- 2018년 : 세종대학교 천문우주학과 학사
- 2018년 ~ 2022년 : 세종대학교 소프트웨어융합학과 석사
- 2022년 ~ 현재 : (주)셀버스 과장
- ORCID : <https://orcid.org/0000-0002-3089-1452>
- 주관심분야 : 머신러닝, 딥러닝, 자연어 처리



김 학 동

- 2016년 : 경성대학교 컴퓨터공학과 학사
- 2017년 ~ 현재 : 세종대학교 디지털콘텐츠학과 석, 박사 통합과정
- ORCID : <https://orcid.org/0000-0003-3816-1224>
- 주관심분야 : 머신러닝, 딥러닝, 컴퓨터 비전



전 태 흘

- 2021년 : 세종대학교 디지털콘텐츠학과 학사
- 2021년 ~ 현재 : 세종대학교 소프트웨어융합학과 석사과정
- ORCID : <https://orcid.org/0009-0006-1265-4425>
- 주관심분야 : 머신러닝, 딥러닝, 컴퓨터 비전



김 청 원

- 1981년 12월 ~ 1985년 7월 : (주)대한항공 전산실 재무 시스템 개발원
- 1982년 : 한양대학교 공과대학 금속공학 학사
- 1987년 : 미국 일리노이주 서던일리노이대학교 컴퓨터 과학 학사
- 1990년 : 미국 일리노이주 서던일리노이대학교 대학원 컴퓨터 과학 석사
- 1994년 : 미국 인디애나주 인디애나 대학교 대학원 컴퓨터 과학 전공
- 2000년 : 미국 뉴욕주 시러큐스 대학교 대학원 컴퓨터 & 정보과학 공학박사
- 2000년 1월 ~ 2001년 3월 : 미국 펜실베이니아주 외인시 소재 Bhasha, INC Technical Staff (연구원)
- 2002년 3월 ~ 2003년 8월 : 아주대학교 정보통신전문대학원 BK 교수
- 2003년 9월 ~ 2017년 2월 : 세종대학교 전자정보공학대학 교수
- 2017년 3월 ~ 현재 : 세종대학교 소프트웨어융합대학 교수
- ORCID : <https://orcid.org/0000-0002-1489-8427>
- 주관심분야 : 인공지능, 지능형 시스템, 딥러닝 등