



도시계획학 석사학위논문

다층신경망을 이용한 도시부 도로 미세먼지 농도 예측 모형 연구

Prediction of Particulate Matter Concentrations on Urban Roads Using Multi Layer Perceptrons

2023년 02월

서울대학교 환경대학원

환경계획학과 교통학전공

박성범

다층신경망을 이용한 도시부 도로 미세먼지 농도 예측 모형 연구

지도교수 장 수 은

이 논문을 도시계획학 석사학위논문으로 제출함 2022년 12월

> 서울대학교 환경대학원 환경계획학과 교통학전공

박성범

박성범의 석사 학위논문을 인준함 2023년 02월

| 위육 | 원장 | 허성호 | (인) |
|----|----|-------|-----|
| 부위 | 원장 | 한 상 진 | (인) |
| 위 | 원 | 장 수 은 | (인) |

국문초록

미세먼지 농도와 교통류 사이의 인과관계는 복잡하고 불분명한 것으로 알려져 있다. 통계모형 등 전통적인 방법론을 이용한 연구 들에서 그 관계를 명확히 규명하지 못한 이유이다. 이러한 한계를 극복하고자 최근에는 딥러닝을 이용하여 미세먼지와 교통류 사이 의 인과관계를 규명하려는 연구들이 진행되고 있다. 하지만 지역 에 따라 교통류 상황이 끊임없이 변하기 때문에 시간적, 공간적 특성을 고려한 예측 모형으로 발전하지 못하고 있다.

이에 본 연구에서는 특정 도로구간을 반복 주행하면서 실시간으 로 수집한 자료를 바탕으로 다층신경망을 구축하였다. 데이터 전 처리, 신경망 구조설정, 신경망 학습설정, 신경망 성능평가의 과정 을 거쳤다. 여러 예비 신경망 모형의 검증을 통해 가장 성능이 좋 은 신경망을 제시하고, 신경망의 성능을 향상하는 입력변수를 식 별하였다. 또한 각각의 입력변수가 미세먼지 농도에 어떻게 영향 을 미치는지 분석하였다.

연구의 주요 결과는 다음과 같다. 첫째, 교통류와 기온을 제외한 모든 기상변수를 투입한 모형의 성능이 가장 우수하였다.

둘째, 교통량, 중차량 수, 가속도의 증가는 미세먼지 농도를 증가 시키는 반면 평균속도의 증가는 미세먼지 농도를 감소시켰다. 기 상요인은 모두 증가할수록 미세먼지 농도를 감소시켰다.

셋째, 입력변수의 기여도 분석을 통해서 무거운 차량이 일반차량 에 비해, 자동차의 가속 및 감속이 평균속도에 비해 미세먼지 농 도에 더 큰 영향을 미쳤다.

연구결과를 바탕으로 미세먼지 저감을 위한 교통 관리 방안을 도

출할 수 있다. 우선 차량의 급격한 가속, 감속을 줄일 수 있는 교 통 관리 방안이 대기질 개선을 위한 현명한 전략이 될 수 있다. 다음으로 대형차에 과금하는 세제개편도 대기질을 관리할 수 있는 효과적인 방안이 될 수 있다. 장기적으로는 신소재 연구를 통해 마모에 강한 타이어와 차체를 가볍게 할 수 있는 자동차 프레임의 개발이 필요할 것이다.

주요어 : 미세먼지 농도, 연소성 미세먼지, 비연소성 미세먼지, 교통 류, 다층신경망, 대기질 학 번 : 2020-28067 목 차

| 제 1 장 서론 |
|---------------|
| 제 2 장 기존연구 고찰 |
| 제 3 장 연구방법론 |
| 제 4 장 자료 |
| 제 5 장 분석결과 |

| | 제 3 절 입력변수 기여도 측정 | 43 |
|----|-------------------|----|
| 제 | 6 장 결론 | 48 |
| 참. | 고문헌 | 50 |
| At | stract ····· | 62 |

표 목 차

| <표 4-1> | 연구 보조 장비 | 18 |
|-----------|-------------------------------|----|
| <표 5-1> | 입력 자료와 자료 유형 | 23 |
| <표 5-2> - | 출력 자료와 자료 유형 | 23 |
| <표 5-3> / | 신경망에 사용한 하이퍼 파라미터 | 27 |
| <표 5-4> . | 모형 1 성능 | 31 |
| <표 5-5> 2 | 각 모형의 입력변수와 출력변수 | 40 |
| <표 5-6> 2 | 각 모형의 성능 비교 | 41 |
| <표 5-7>] | PM10 예측에 대한 입력변수의 Shapley값 | 44 |
| <표 5-8>] | PM2.5 예측에 대한 입력변수의 Shapley값 … | 45 |

그림 목차

<그림 3-1> 다층신경망 도식화 …………………………… 10 <그림 3-2> 신경망의 일반적 분석절차와 본 연구 적용 11 <그림 4-1> 연구자료 수집에 이용한 조사 차량 18 <그림 4-2> PM10 농도와 교통류 변수 사이의 산점도 ·· 21 <그림 4-3> PM 농도의 1년 동안의 분포 21 <그림 5-1> 에폭 변화에 따른 오차 변화 25 <그림 5-2> 은닉층 개수 변화에 따른 MSE 값 변화 …… 25 <그림 5-3> 뉴런 개수 변화에 따른 MSE 값 변화 26 <그림 5-4> 모형 1의 학습데이터로 PM10 농도 예측 … 29 <그림 5-5> 모형 1의 검증데이터로 PM10 농도 예측 … 29 <그림 5-6> 모형 1의 학습데이터로 PM2.5 농도 예측 … 30 <그림 5-7> 모형 1의 검증데이터로 PM2.5 농도 예측 … 30 <그림 5-8> 모형 2의 학습데이터로 PM10 농도 예측 … 32 <그림 5-9> 모형 2의 검증데이터로 PM10 농도 예측 … 32 <그림 5-10> 모형 2의 학습데이터로 PM2.5 농도 예측 · 33 <그림 5-11> 모형 2의 검증데이터로 PM2.5 농도 예측 · 33 <그림 5-12> 모형 3의 학습데이터로 PM10 농도 예측 ·· 34 <그림 5-13> 모형 3의 검증데이터로 PM10 농도 예측 ·· 34 <그림 5-14> 모형 3의 학습데이터로 PM2.5 농도 예측 · 35 <그림 5-15> 모형 3의 검증데이터로 PM2.5 농도 예측 · 35 <그림 5-16> 모형 4의 학습데이터로 PM10 농도 예측 ·· 36 <그림 5-17> 모형 4의 검증데이터로 PM10 농도 예측 ·· 36

제 1 장 서론

제 1 절 연구 배경 및 목적

고농도 미세먼지(PM, Particulate Matter)의 빈번한 발생으로 대기 환경 에 관한 관심이 급증하고 있다. PM10은 입자의 크기가 지름 10 µm 이 하의 물질을, PM2.5는 입자의 크기가 지름 2.5 µm 이하의 물질을 각각 뜻한다.

도로 교통은 도시 지역에서 두 가지 방식으로 PM 농도를 높인다. 하나 는 차량에 의한 직접적인 배기가스 배출이다. 배기가스는 차량이 연료를 연소하는 과정에서 생기며, 연기, 분진, 그을음 등의 입자와 휘발성유기 화합물, 질소산화물 등의 기체로 구성된다. 기체, 특히 질소산화물은 공 기 중에서 수증기, 오존, 암모니아 등과 화학반응을 통해 PM으로 변한 다(Biswas et al., 2008; Bergmann et al., 2009). 다른 하나는 차량으로 인한 간접적인 PM 발생인 비배기 배출이다. 이러한 비연소 배출은 차량 이 도로를 주행할 때, 브레이크 마모, 타이어 마모, 미세먼지 재비산 등 으로 발생한다(Kumar and Goel, 2016; Timmers and Achten, 2016).

미세먼지는 대기오염을 발생시키고 대기오염은 가시성을 저하할 뿐만 아니라 호흡기 및 여러 피부 질환을 일으킨다. 이에 미세먼지를 줄이기 위해 디젤 차량에 DPF(Diesel particulate filter)를 설치한다던가(Krecl et al., 2017; Martinet et al., 2017), 전기차, 등 친환경 자동차로 전환하 는 정책(Qadir et al., 2013; Sun et al., 2014)들이 시행되고 있다.

교통류 특성이 실제 미세먼지 농도에 미치는 영향을 가늠할 수 있으면 보다 효과적인 미세먼지 저감 대책을 수립할 수 있다. 하지만 도심지역 에서는 교통류 특성이 끊임없이 변하기 때문에 미세먼지 예보 시스템보 다 시간적, 공간적으로 더 조밀한 예측 모형이 필요하다.

다층신경망(MLP, Multi Layer Perceptrons)은 방대한 데이터를 학습해

서 정확하고 빠른 예측이 필요한 분야에서 널리 사용된다. 본 연구에서 는 조밀한 시·공간적 자료를 사용하기 위해서 실도로 주행 조건에서 수 집한 교통류 데이터와 미세먼지 농도 데이터를 바탕으로 다층신경망을 이용해 미세먼지 농도변동을 예측하고 모형의 적합성을 검증하고자 한 다.

제 2 절 연구의 범위 및 방법

1. 연구의 범위

본 연구에서는 교통류 특성을 통해 미세먼지 농도를 예측할 수 있는 모 형을 개발하고자 한다. 교통류 특성을 설명하는 자료들은 실제 도로주행 환경에서 수집되었다. 도로주행의 범위는 서울시 동작대로 이수역-이수 고가차도 구간(1,190m)이다. 해당 구간은 교외 지역에서 서울 도심으로 이어지는 주요 간선도로로서, 많은 교통량과 다양한 차종으로 구성된 교 통류 특성을 보인다. 실제 도로 주행조사는 2020년 3월부터 2021년 2월 까지 수행되었으며, 오전첨두, 오후첨두, 비첨두 특성을 고려하기 위해 매주 화요일 18시부터 20시, 금요일 7시부터 9시까지 1초 간격의 GPS 데이터, 3초 간격의 교통류 이미지 데이터, 6초 간격의 PM 농도 데이터 를 수집하였다.

2. 연구의 구성

본 연구는 크게 네 부분으로 구성된다. 2장에서는 미세먼지와 교통류 특성 사이의 인과관계를 분석한 연구를 고찰해서, 미세먼지 농도에 영향 을 미치는 교통류 특성을 찾아낸다. 또한 딥러닝을 이용해서 미세먼지 농도를 예측한 연구의 한계점을 도출해 본 연구의 필요성과 차별성을 제 시한다. 3장에서는 본 연구에서 사용한 다층신경망 모형의 개요와 본 연 구 적용방안을 기술하며, 사용한 입력변수에 대한 기여도를 측정하는 방 법을 제시한다. 4장에서는 자료수집 방법과 모형화 전 필요한 전처리 과 정을 설명한다. 5장에서는 분석결과를 살펴보고 적합한 모형을 선정하며 모형 성능을 평가한다. 또한 모형에 사용된 각 변수가 미세먼지 농도에 미친 기여도를 측정한다. 6장에서는 본 연구의 의의를 정리하고 정책적 시사점을 제시한다.

제 2 장 기존연구 고찰

이 장에서는 미세먼지를 유발할 수 있는 교통류 특성을 고찰하고, 딥러 닝 기법을 이용하여 미세먼지 농도를 예측한 연구를 소개한다. 이를 바 탕으로 시사점과 본 연구의 필요성을 제시한다.

제 1 절 미세먼지 농도와 교통류에 관한 연구

첫 번째, 미세먼지를 유발하는 교통류 특성 중에 교통량이 있다. 몇몇 연구는 교통량과 PM 농도 사이에 양의 상관관계가 있음을 보여준다. 즉, 차량 대수가 증가하면 PM 농도 또한 증가한다(Giovanis, 2018; Azhari et al., 2018; Gu et al., 2022). 예를 들어 평일 PM10과 차량 대 수의 상관 계수는 0.62였다(Azhari et al., 2018). 반대로 구역 통행 제한 을 통한 교통량 감소가 미세먼지 농도의 감소로 이어지지 않는다는 연구 또한 존재했다(Mishra et al., 2019; Zhang et al., 2020; Chen et al., 2021).

두 번째, 차량의 속도가 미세먼지 농도 변화에 영향을 미친다. 선행연구 는 속도와 미세먼지 농도 사이에 약한 음의 상관관계가 있음을 보여준다 (Ježek et al., 2015; Yazdi et al., 2015; Abdull et al., 2020; Zhang et al., 2022). 다만 몇몇 연구는 도심보다 고속도로 등 속도가 높은 교통류 조 건은 미세먼지 농도와 양의 상관관계에 있다고 보고하기도 했다(Gillies et al., 2005; Belkacem et al., 2022).

세 번째, 빠르고 급격한 가속 및 감속은 더 많은 연료를 소비하여 더 많은 배기가스 배출로 이어진다(Kuppili et al., 2021; Gao et al., 2021; Munjal et al., 2022). 예를 들어, 배기가스 배출량은 -0.5~0.5m/s2의 낮 은 가속 범위 내에서 안정적이었고 이 범위를 넘어서면 증가했다 (Kuppili et al., 2021). 또한 1.3~1.5m/s2의 가속도에서의 연료 소비는 0.3~0.5m/s2 사이의 가속도에서보다 약 3배 더 많았다(Gao et al., 2021). 빈번한 감속 및 가속은 또한 더 많은 비배기 배출을 초래한다 (Ketzel et al., 2007; Karjalainen et al., 2014; Xu et al., 2020).

마지막으로 차량 중량이 무거울수록 배출되는 PM의 양이 많아진다. 버 스 같은 대형 차량은 일반 승용차보다 배기가스 EF(Emission Factor)가 더 높다(Sofowote et al., 2018; Kinnon et al., 2019). 예를 들어, 무게가 약 1,200kg인 승용차는 주행한 킬로미터당 0.8g의 PM10이, 무게가 약 18,000kg인 대형 군용 차량은 킬로미터당 48g의 PM10이 배출된다 (Gillies et al., 2005). 그리고 PM10과 PM2.5에 대해 측정된 HDV(Heavy Duty Vehicle) EF는 LDV(Light Duty Vehicle) EF보다 각각 15.2, 37.0 배 더 높았다(Raparthi et al., 2022).

차량 중량은 비연소 PM 발생과 상관관계를 갖는다(Timmers and Achten, 2016; Lin et al., 2022). 따라서 배터리 중량이 무거워진 전기 자 동차(EV, Electric Vehicle)는 내연 기관 차량(ICEV, Internal Combustion Engine Vehicle)과 동등하거나 더 많은 PM을 배출할 수 있 다(Beddows and Harrison, 2021; Liu et al., 2021).

도로교통 외에 PM 농도에 영향을 미치는 요인은 많지만 기상조건이 중요한 외부변수이다. 기존 연구를 조사한 결과 미세먼지 농도에 영향을 미치는 주요 기상변수는 습도, 풍속, 기온이다.

첫째, 습도가 높은 조건보다는 건조한 조건에서 PM 농도가 높았다(Krampah et al., 2021; Kliengchuay et al., 2021; Sarmadi et al., 2021; Nwosisi et al., 2021). 습윤한 도로는 먼지의 재비산을 감소시킬 수 있 다. 따라서 비배기 배출은 습윤한 환경보다 건조한 환경에서 증가하는 경향이 있다(Abu-Allaban et al., 2003; Amato et al., 2009).

둘째, PM 농도와 풍속 사이에는 음의 상관관계가 있다(Sritong-aon et al., 2021; Krampah et al., 2021; Sarmadi et al., 2021; Sindosi et al., 2021; Nwosisi et al., 2021). 강한 풍속은 대기를 분산시켜서 PM 농도를 감소시킨다. 반대로, 낮은 풍속은 대류를 정체하게 만들어서 국지적 대기 오염을 유발할 환경을 제공한다(Kim et al., 2017; Duan et al., 2021; Nuria Galindo et al., 2010). 셋째, 평균 기온과 PM 농도와는 음의 상관관계에 있다(Kliengchuay et al., 2021; Bodor et al., 2020). PM2.5 농도는 겨울에는 높았고 여름에는 낮았다(Cheng et al., 2021; Srimuruganandam et al., 2010; Zhang et al., 2015; Wang et al., 2021). 이는 겨울철의 기온역전 현상으로 인해 PM 확산이 어려워져 오염물질을 축적하기 때문이다(Zhao et al., 2013; Zhang and Li., 2011).

제 2 절 미세먼지 농도 예측에 관한 연구

미세먼지 예측을 위한 연구에 다양한 신경망 알고리즘이 적용되고 있다. 본 장에서는 딥러닝 알고리즘을 이용해 미세먼지를 예측한 기존연구의 과정과 결과를 분석한다.

Krishan et al. (2019)는 시간당 대기 오염물질 농도(PM2.5)를 예측하 기 위해 여러 신경망 모형을 사용하였다. 델리의 교통, 기상, 및 대기오 염 배출 데이터를 이용했다. 다층신경망 모형으로 PM 2.5를 예측했을 때 추정된 상관 계수 값은 0.47 - 0.78 범위를 나타냈다. 이 연구는 교통 류 데이터를 고정된 한 지점에서 관측한 것이고, 데이터가 차량 연료 별 (디젤, 가솔린, CNG, LPG) 교통량이라는 한계가 있다. Suleiman et al. (2019)는 머신러닝을 사용하여 도로변 PM10 및 PM2.5 감소 시나리오의 효율성을 평가하였다. 다층신경망은 PM10 농도를 예측할 때 다른 머신 러닝 알고리즘보다 좋은 성능을 보였으며 교통량이 가장 기여도가 큰 변 수였다. 이 연구 또한 교통류 자료를 이용했지만 시간 단위의 교통량, 속 도를 이용하는 데 그쳤으며 조밀한 간격의 교통류 변화를 고려하지 못했 다. Kujawska et al. (2022)는 2017년에서 2019년 사이에 기록된 고정된 관측지점에서 매시간 수집한 기상 및 대기 질 데이터를 사용했다. 기상 데이터에는 기온(T), 상대 습도(RH), 풍속(WS) 및 풍향(WD)이 포함되 었으며 다층신경망 모형은 PM10의 농도를 MSE(Mean Squared Error) 68.09 수준으로 예측했다. Bozdağ et al. (2020)은 터키 앙카라에 위치한 7개 관측소에서 2009.01.01.~2017.12.31 사이에 측정된 PM10 데이터를

사용했다. 여러 머신러닝 알고리즘에 따른 지역별 대기오염 변동을 예측 했다. 다층신경망 모형을 사용했을 때 최상의 성능(RMSE = 20.8, MAE = 14.43)을 나타냈다. Nagendra et al. (2005)는 3년 동안 8시간씩 고정된 지점에서 수집한 대기 질, 기상, 교통류 데이터를 사용했다. 기상 데이터 에는 일조 시간, 가시성, 온도, 풍속, 풍향 및 습도 등이 있었고 교통류 데이터는 이륜차, 삼륜차, 휘발유 사륜차, 디젤 사륜차로 분류했다. 기상 및 교통 특성 변수를 모두 포함하는 모형과 기상변수만을 포함한 모형이 검증데이터에서 MSE 가 5~11 수준으로 최상의 성능을 나타내었다.

제 3 절 본 연구의 차별성

제 1, 2절에서 살펴본 바와 같이, 미세먼지 농도에 영향을 미치는 교통 류 특성과 기상 특성이 다양하다. 몇몇 특성의 경우 기존 연구에서 사용 한 통계모형 등의 방법론을 통해서는 미세먼지 농도와의 인과관계가 일 관되지 않은 것으로 나타났다. 또한, 주로 지점에서 관측된 기상 데이터 를 바탕으로 미세먼지 농도를 예측하였고, 교통류 특성을 이용한 연구에 서도 지점에서 관측한 한정적인 교통류 데이터를 이용해 예측하는 알고 리즘에 치우쳐 있었다.

본 연구에서는 이를 보완하기 위해 실시간 도로주행을 통해 일정 공간 적 범위 내에서 조밀한 시간 간격으로 얻은 교통류 데이터를 가공해서 연구에 적용하려고 한다. 또한 여러 교통류 특성, 기상 특성과 미세먼지 농도 간의 복잡하고 불분명한 관계를 규명하기 위해 다층신경망 모형을 이용하고자 한다. 즉, 실시간으로 관측한 여러 교통류 데이터와 기상 데 이터를 신경망 모형에 적용해서 미세먼지 농도 예측 모형을 제안하고 이 를 평가하고자 한다.

제 3 장 연구방법론

본 연구에서는 실제 도로주행을 통해서 수집한 교통류 자료와 다층신경 망을 이용해서 실시간 미세먼지 농도를 예측하였다. 다층신경망은 연관, 분류 및 예측 분야에서 복잡한 문제를 처리하는 데 효율성이 입증된 데 이터 모델링 방법이다(Hornik et al., 1989). 다층신경망은 지도학습 중 회귀예측 모형에 적용하기 쉬우며 여러 은닉층을 만듦으로써 좀 더 정확 한 결과를 예측할 수 있다. 따라서 회귀예측 모형의 특징을 가지며, 변인 사이의 복잡하고 불분명한 관계를 나타낼 것으로 예상되는 미세먼지 농 도 예측에 적합한 방법이다.

제 1 절 다층신경망의 개요

신경망은 입력에 대한 응답으로 정보를 정교화하는 간단하고 고도로 상 호 연결된 처리 요소로 구성된 컴퓨터 시스템이다. 인공신경망은 딥러닝 의 가장 핵심적인 기술로 인간 두뇌의 신경세포를 모방한 컴퓨터 알고리 즘 네트워크 구조이며 그중 여러 계층을 가지는 신경망을 다층신경망이 라 한다.

다층신경망은 뉴런들이 다수 모여서 이루어지는 계층 구조로서 복잡하 게 연결되어 있다. 다층신경망은 입력층, 은닉층, 출력층 등 최소 세 개 의 층으로 구성되며 상호작용하는 여러 계산 요소들로 이루어져 있다. 학습된 정보로부터 가중치를 계산하고 주어진 정보의 특성을 일반화한 다.

입력층은 입력데이터가 들어가는 곳이다. 또한 여러 개의 뉴런으로 구 성된 은닉층이 있다. 각 뉴런은 입력의 가중치 합산을 수행한 다음 뉴런 함수라고도 하는 비선형 활성화(activation) 함수를 통과한다. 가중치의 합은 연산자를 통해 계산되며 이 값은 활성화 함수를 사용해 구현되는데 본 연구에서 활성화 함수는 <식 3-2>의 렐루(ReLU) 함수가 사용된다.

 $z_{j} = \sum_{i=1}^{n} w_{ij} u_{i} - b \dots < A 3-1 >$ $ReLU(x) = \max(0, x) \dots < A 3-2 >$

여기서.

u; : 입력층의 i 번째 뉴런의 값

w_{ij}: 입력층의 i 번째 뉴런에서 출력층의 j 번째 뉴런으로 연결된 가중
 치

z_i : 출력층의 j 번째 뉴런의 값

b : 오차

ReLU(*x*) : 렐루 함수

x : 은닉층 뉴런의 입력값

다층신경망의 예측 성능을 결정하는 요소 중 뉴런 간 연결 강도가 있으 며 학습 규칙에 따라 최적화된다. 인공신경망은 학습하는 동안 실제 계 산값과 예측값 사이의 오차를 합산해서 총 오차를 계산한다. 네트워크 훈련은 실제 값과 모델링 된 값 사이의 오차를 최소화하기 위해 신경망 의 매개변수를 수정하는 과정이다. 다층신경망의 훈련은 일반적으로 두 단계를 포함하는 역전파 알고리즘으로 수행된다(Rumelhart et al., 1986). 최초 순전파 단계에서 가중치 값들을 무작위로 초기화해 계층을 거쳐 계 산된다. 다음으로 오차를 평가하고 역전파 단계에서 각 가중치 별 현재 오차에 미치는 영향이 계산되고 오차를 줄이기 위해 가중치를 업데이트 하는 과정이 수행된다.





<그림 3-1> 다층신경망 도식화



<그림 3-2> 신경망의 일반적 분석절차와 본 연구 적용

<그림 3-2>는 통상적으로 사용되는 신경망 분석절차와 본 연구에 적 용한 내용을 나타낸 것이다. 자세한 사항은 제 3장 제 2절에서 서술한다.

제 2 절 다층신경망의 적용과 평가

일반적으로 신경망을 사용하여 예측을 수행할 때 먼저 학습데이터 세트 에 대해 모형을 훈련하고 이 정보를 기반으로 새로운 데이터 세트에 대 한 결과를 예측한다. 과적합(overfitting)은 신경망 모형이 학습데이터에 대해서는 정확히 예측하지만 새로 입력한 데이터에 대해서는 예측 성능 이 악화할 때 발생하는 현상이다. 과적합을 방지하는 방법으로 여러 가 지가 사용되고 있다. 본 연구에서는 드롭아웃(drop out)을 설정해서 과적 합을 방지하고자 한다. 드롭아웃은 인공신경망 학습 시 매번의 학습 과 정에서 서로 연결된 계층에서 0부터 1까지 확률로 뉴런을 버리는 것이 다.

에폭(epoch)은 데이터 세트가 신경망을 통과한 횟수를 의미한다. 너무 많은 에폭은 과적합을, 너무 적은 에폭은 과소적합(underfitting)을 일으 킨다.

배치 사이즈(batch size)는 학습데이터 세트를 여러 조각으로 나누었을 때, 하나의 조각에 속하는 데이터 수를 의미한다. 배치 사이즈가 작으면 학습은 잘 되겠지만 가중치를 업데이트하는 과정이 여러 번 늘어나면서 시간이 오래 걸리는 단점이 있다. 배치 사이즈가 클수록 한 번에 많은 데이터를 저장해야 하므로 큰 용량을 요구한다.

<그림 3-3>은 본 연구의 자료 수집 구간을 도식화한 것이다. 조사차량 은 조사구간(Survey route)을 일정 시간 동안 반복적으로 주행하는데 차 량이 조사구간을 1회 왕복하는데 대략 6~7분 정도 소요되었다. 모든 변 수의 데이터를 1분 단위로 집계했으므로 날짜마다 차이가 존재하지만 대 략 6~7분은 6~7개의 데이터 개수를 의미한다. 조사 차량은 같은 구간 을 반복주행하기 때문에 주기가 연속되는 것이므로 배치 사이즈를 6~10 사이를 변경해가며 분석했다. 신경망에 몇 개의 은닉층이 있어야 하는지 를 정의하는 규칙은 없다. 따라서 본 연구에서는 신경망의 성능을 비교 해가며 경험적으로 최적의 은닉층과 뉴런의 개수를 찾는다.



<그림 3-3> 자료 수집 지역

신경망 모형의 활성화 함수는 시그모이드(sigmoid) 함수 대신 렐루 함 수를 이용했다. 즉, 내부 은닉층에는 렐루 함수를 적용하고, 마지막 출력 층에는 시그모이드 함수를 적용하여 정확도를 높이고자 한다.¹⁾

비용함수(Loss fuction)는 실제값과 예측값의 차이를 수치화하는 함수 이다. 본 연구에서 비용함수는 MSE를 사용했으며 간단한 형태여서 신 경망의 속도를 향상할 수 있다(Xia et al., 2019).

옵티마이저(Optimizer)는 비용함수의 최솟값을 찾는 과정을 뜻한다. 본 연구의 옵티마이저는 아담(Adam)을 사용했는데, 간단한 구현과 효율적 인 연산이 가능하며 메모리 요구 사항이 거의 없다는 장점이 있다 (Kingma et al., 2021).

모형의 성능은 MAE(Mean Absolute Error)와 MAPE(Mean Absolute Percentage Error)을 통해 평가한다. MAE는 실제값과 예측값의 차이를 절대값으로 변환해 평균화한 것이며, MAPE는 MAE를 퍼센트로 변환한 것으로 각각 <식 3-3>과 <식 3-4>처럼 나타낸다. MAE와 MAPE는 모 형의 정확도를 측정하는 지표로서, 이 매개변수의 값이 작을수록 모형의 정확도가 높아진다(Wang et al., 2020; Khullar and Singh., 2021).

 $MAE = \frac{1}{m} \sum_{i=1}^{m} \left| Y_i - \hat{Y}_i \right| \qquad < \Delta 3-3 >$ $MAPE = \frac{100}{m} \sum_{i=1}^{m} \left| \frac{Y_i - \hat{Y}_i}{Y_i} \right| \qquad < \Delta 3-4 >$

여기서, m은 테스트 세트의 샘플수이고 Y_i 과 \hat{Y}_i 는 *i*번째 샘플의 실제값 과 예측값을 나타낸다.

¹⁾ https://hunkim.github.io/ml/

제 3 절 입력변수의 상대적 기여도 측정

변수의 기여도 분석은 모형의 전체 성능에 가장 큰 영향을 주는 입력변 수를 선별하는 방법이다. 각 입력변수의 상대적 중요도를 이해하고 각 입력변수의 변화에 대한 신경망의 반응을 연구하기 위해 기여도 분석을 수행한다.

본 연구는 SHAP(Shapley Additive exPlanations)을 사용해 변수의 기 여도를 측정한다. SHAP는 Shapley값을 사용하여 각 변수의 기여도를 계산하고 변수의 중요도를 측정한다(Futagami et al., 2021; Liu et al., 2022).

하나의 변수에 대한 기여도를 알기 위해 여러 변수의 조합을 구성하고 해당 변수의 유무에 따른 평균적인 변화를 통해 얻어낸 값이 바로 Shapley값이다. Shapley값은 가능한 모든 조합에 대해서 한 변수의 기여 도를 합한 값이다.

$$\varPhi_i = \sum_{S \in F|\{i\}} \frac{|S|!(|F| - |S| - 1)!}{|F|!} (f(S \cup \{i\}) - f(S)) \dots < A 3-5 >$$

𝗘_i: 가능한 모든 조합에 대한 i 변수의 평균 기여도(Shapley값)
i: 기여도를 구하고자 하는 변수
|S|: i를 제외한 변수로 이루어진 부분집합의 원소의 수
|F|: 모든 변수의 개수
f(S): S가 입력변수일 때 모형의 출력값

즉, ∅_i 는 i를 제외한 F의 모든 부분집합에 대해서, i를 추가했을 때의 출력값 변화량의 평균을 측정한 값이다. 특정 변수의 유무만 다른 모든 모형 쌍들을 찾고, 각 쌍의 결괏값의 차이의 가중합을 구한다. 모든 변수 의 조합에서 특정 변수의 유무에 따른 차이를 구하기 때문에 모든 변수 의 조합되는 상호작용도 고려한 특정 변수의 기여도를 얻을 수 있다.

$$I_j = \sum_{i=1}^n \left| \boldsymbol{\Phi}_j^{(i)} \right| \quad \cdots \quad < \geq 1 \quad 3-6 >$$

 I_j : 변수의 전역적인 기여도(global importance) $arphi_j$: j 변수의 Shapley값

각 변수의 Shapley값의 절댓값 평균을 통해서 전역적 변수 기여도를 구할 수 있다. 이것을 통해 각 변수의 출력변수에 대한 기여도를 측정한 다.

제 4 장 자료

본 연구는 2019.06~2022.02년에 걸쳐 수행된 한국연구재단의 기본연구 '교통류 특성에 기반한 도시 미세먼지 저감 계획·운영 전략 개발'과제에 서 수집된 자료를 사용하였다. 제 4장의 제 1절과 제 2절을 통해 자료 수집 방법과 데이터 전처리 과정을 간략히 소개한다.

제 1 절 자료 수집

여러 장치가 탑재된 차량을 활용해서 실제 도로를 주행하며 PM 농도 데이터, GPS(Global Positioning System) 데이터, 교통류 이미지 데이터 를 수집하였다. <그림 4-1>은 조사 차량으로서 차량 상단에 이미지 데 이터 수집 장비인 DSLR(Digital Single-Lens Reflex) 카메라를 설치하 고, 3초 간격으로 촬영할 수 있도록 보조 장비인 간격 측정기 (Intervalometer)를 함께 설치하였다. 또한 GPS를 설치하여 1초 간격으 로 차량의 좌표를 기록하며 차량 상단에 PM 농도 측정기를 두어 6초 간격으로 PM 농도를 측정하였다. 기상자료는 AWS(Automatic Weather Station)에서 관측한 자료를 사용하였다. AWS는 서울에 29개 관측지점 이 있고 그 중 현충원 지점의 관측자료를 사용하였다. 현충원 관측지점 은 서울특별시 동작구 사당동(위도:37.50036, 경도:126.97652)에 위치하며, 관측지점이 조사구간 근처이므로 조사구간의 기상을 나타내는 자료로 사 용하였다. 자료 수집을 위해 사용된 장비는 <표 4-1>과 같다.

제 1장 제 2절에서 기술한 시·공간적 범위에 따라서 조사 차량을 왕복 주행하며 연구자료들을 수집하였다. 교통류 자료의 수집은 <그림 3-2> 와 같이 서울시 동작대로 이수역-이수고가차도 구간에서 수행되었다. 조 사구간을 왕복하며 주행할 때 U턴 구간 및 진입을 위한 1차로 주행 이



<그림 4-1> 연구자료 수집에 이용한 조사 차량

| 장비 | 조사 키라 | 미세먼지 농도 | 교통류 2 | 준사 장비 |
|----------|---------------------------------------|------------|------------------------|--|
| 구분 | · · · · · · · · · · · · · · · · · · · | 측정장비 | GPS | IMAGE |
| 장비 명 | 기아 레이 | GRIMM 1109 | ASCEN 742 | NIKON D850 |
| 장비 사진 | | | Ascentaria B 3 B | Nikon * Bio * |

<표 4-1> 연구 보조 장비

제 2 절 데이터 전처리

본 연구에 필요한 데이터는 PM 농도, 교통류, 기상 데이터이다. 이 데 이터들의 관측 간격은 관측에 사용된 장비의 특성에 따라 차이가 있다. PM 농도 데이터는 6초 간격으로 측정되었고, 이미지 데이터는 3초 간격 으로 촬영되었으며, GPS 데이터는 1초 간격으로 기록되었다. 한편 기상 데이터는 1분 단위로 제공된다. 연구자료들 사이의 시간 단위 일치를 위 해 모든 자료를 1분 단위로 집계하였다.

촬영된 이미지를 영상 처리하여 실시간 교통량을 추정하였다. 이미지를 통해 교통량을 추정하는 방법에는 물체 감지 알고리즘인 YOLO(You-Only-Look-Once)를 적용했다(Chang et al., 2022).

추정된 모든 자동차 수를 더하여 교통량 변수로, 버스와 트럭만을 따로 분리해 중차량 수 변수로 각각 이용했다. 차량의 GPS에 기록된 좌표와 해당 시간을 이용해서 평균속도를 계산했으며 이것을 교통류의 평균속도 로 가정하였다. 또한 차량의 평균속도를 계산한 후 두 시점 사이의 속도 차이를 이용해서 가속도를 계산했다.

교통류 특성에 의해 발생한 PM이 PM 농도 측정기기에 반영되기까지 일정 시간이 필요하다. 이를 타임 윈도우(Time window)라 한다 (Yli-Tuomi et al., 2005; Ježek et al., 2015; Lešnik et al., 2019). 선행연 구에서는 이 반응 시간을 10초 내외로 설정했다(Yli-Tuomi et al., 2005; Ježek et al., 2015). 본 연구에서는 6초 간격으로 PM 농도를 측정하므로 타임 윈도우를 12초로 설정했다.

관측 차량이 신호 대기로 정지해 있고 풍향이 주행 방향과 같을 때 측 정된 PM 농도는 이상치로 간주해 제거했다. 이 조건에서는 전방 교통류 특성으로 인한 PM 농도 변화를 측정할 수 없기 때문이다(Bukowiecki et al., 2002).

제 3 절 데이터 특성

교통류 특성과 미세먼지 농도 사이의 일반적인 관계를 관찰하기 위해서 2021년 2월 2일에 관측된 데이터를 통해 PM10 농도와 교통류 변수 사 이의 관계를 확인하였다. <그림 4-2>는 교통류 변수인 교통량, 중차량 수, 속도, 가속도와 PM10 농도 사이의 산점도를 나타낸 것이다. 교통류 변수와 PM10 농도와의 단일 상관관계는 기존의 선행연구와 비슷한 경 향을 보이지만 각 산점도의 분산이 크다. 다층신경망에서는 여러 변수들 이 미세먼지 농도에 영향을 미쳤을 때의 결과를 예측하고, 어떤 변수가 가장 기여도가 큰지를 확인할 수 있다.

<그림 4-3>은 조사기간 동안의 PM10과 PM2.5의 농도 분포를 나타낸 것이다. PM2.5 농도는 PM10 농도보다 작은 값을 나타내지만 거의 유사 한 분포를 나타내고 있다. 또한 계절별로 구분해서 보면 주로 봄과 겨울 에 미세먼지 농도가 높고 여름에 비교적 미세먼지 농도가 낮다. 이처럼 계절별로 다르게 나타나는 미세먼지 농도는 교통류 요인뿐만 아니라 외 부요인인 기상요인과도 관련이 있을 것이다. 예를 들어 여름철의 습윤한 환경이 미세먼지 농도를 낮출 수 있고, 봄과 겨울의 짙은 황사와 미세먼 지 농도를 높일 수 있다.



<그림 4-2> PM10 농도와 교통류 변수 사이의 산점도



<그림 4-3> PM 농도의 1년 동안의 분포

제 5 장 분석결과

Python의 Keras 모듈을 이용해서 본 연구를 진행하였다. 교통류 데이 터 세트를 학습해서 적합한 파라미터를 추정한 후 본 연구에 적합한 신 경망을 구축하였다. 이 신경망을 MAE와 MAPE를 활용해 평가하고, 어 떤 변수가 가장 큰 영향을 미쳤는지 분석하였다.

무작위의 하이퍼 파라미터를 이용해 초기 신경망 모형을 생성하였다. 2020년 3월부터 2021년 2월까지의 데이터를 사용했고, 7:3의 비율로 약 3,460개의 훈련데이터와 약 1,500개의 검증데이터 세트로 나누었다. <표 5-1>과 <표 5-2>는 분석에 이용한 입력데이터와 출력데이터 형태를 제 시한 것이다.

분석을 진행할 때는 차원 차이의 영향을 제거하고 모형의 수렴 속도를 향상하기 위해 모든 데이터를 Min-Max 스케일링 방정식에 의해 변환했 다(Yan et al, 2021; Krishan et al, 2019). 이 변환에 따라 모든 변수는 [0,1] 사이의 값을 갖는다.

| 입력 자료 | 자료 유형 |
|---------|------------|
| 교통량 | float(실수형) |
| 중차량 수 | float(실수형) |
| 속도 | float(실수형) |
| 가속도의 크기 | float(실수형) |
| 습도 | float(실수형) |
| 풍속 | float(실수형) |
| 기온 | float(실수형) |

<표 5-1> 입력 자료와 자료 유형

| 출력 자료 | 자료 유형 |
|-------|------------|
| PM10 | float(실수형) |
| PM2.5 | float(실수형) |

<표 5-2> 출력 자료와 자료 유형

제 1 절 하이퍼 파라미터 최적화

다층신경망을 이용해서 데이터를 학습하기 위해서는 하이퍼 파라미터를 설정해야 한다. 일반화된 절차가 정립되어 있지 않아 연구에 맞는 파라 미터를 경험적으로 찾아야 한다. 또한 파라미터를 설정할 때 과적합이 일어나는지도 지속적으로 확인해야 한다.

에폭은 데이터의 학습횟수를 의미한다. 에폭의 크기는 오차의 변화를 확인하면서 결정한다. 다만 오차가 크게 감소하지 않고 학습이 반복되면 과적합을 초래할 수 있고 신경망의 성능이 저하되므로 주의해야 한다.

배치 사이즈는 한 번 학습에 사용되는 데이터의 수를 나타내며 본 연구 에서는 조사구간 왕복 주기를 고려해 7로 설정했다. 또한 적정 은닉층과 뉴런의 개수는, 모형의 성능을 비교하면서 경험적으로 추정하였다. 이러 한 하이퍼 파라미터의 최적화를 통해 모형의 성능을 평가해 최종 모형을 구축하였다.



<그림 5-1> 에폭 변화에 따른 오차 변화



<그림 5-2> 은닉층 개수 변화에 따른 MSE 값 변화



<그림 5-3> 뉴런 개수 변화에 따른 MSE 값 변화

<그림 5-1>은 학습이 반복됨에 따라 나타나는 오차를 나타낸 그래프 이다. 에폭이 증가함에 따라 오차가 점점 감소하는 것을 알 수 있다. 신 경망의 최대 반복횟수를 300회로 설정했지만 250회 이후 오차가 0.006으 로 수렴하는 것을 알 수 있다. 그 이상의 반복은 과적합을 초래할 수 있 으므로 본 연구는 에폭을 250으로 설정하였다.

또한 은닉층과 뉴런의 개수는 각각 임의값에서 변경해가며 추정했다. 은닉층은 2개부터 9개까지 점차 증가시키며 MSE의 변화값을 비교하였 고, 뉴런의 개수는 10개부터 10개 단위씩 증가시키며 MSE 변화값을 비 교했다. <그림 5-2>와 <그림 5-3>의 그래프를 보면 은닉층은 4개, 뉴 런의 개수는 100개일 때 예측값과 실제값의 차이가 가장 작았다. 하이퍼 파라미터의 최적값을 설정해서 신경망의 구조를 완성하였고 사용한 함수 와 파라미터를 <표 5-3>에 정리했다.
| 하이퍼 파라미터 | 신경망에 사용한 값 | |
|----------|------------|--|
| 에폭 | 250 | |
| 배치 사이즈 | 7 | |
| 활성화 함수 | ReLU | |
| 은닉층의 개수 | 4 | |
| 뉴런의 개수 | 100 | |
| 옵티마이저 | Adam | |
| 비용함수 | MSE | |
| 드롭아웃 | 0.1 | |

<표 5-3> 신경망에 사용한 하이퍼 파라미터

제 2 절 모형 성능 평가

첫 번째 신경망(모형 1)은 교통류 데이터만을 이용해 학습하였다. <표 5-4>는 개선된 신경망의 성능평가 결과이다. PM10 예측을 위한 수정된 신경망의 MAE와 MAPE는 7.93, 29.75가 나타나 초기 임의 신경망의 24.31, 69.41에 비해 오차가 크게 줄었고, PM2.5의 경우에도 19.50, 84.54 에 비해 8.98, 21.80으로 오차가 감소했다.

<그림 5-5>와 <그림 5-7>에서 파란색 선은 실제 PM 값을 의미하고 주황색 선은 예측 PM 값을 나타낸다. 관측값과 예측값 사이에 상당한 오차가 발견되며, 교통류 외 다른 요인이 미세먼지 농도에 영향을 미치 고 있음을 시사한다.



<그림 5-4> 모형 1의 학습데이터로 PM10 농도 예측



<그림 5-5> 모형 1의 검증데이터로 PM10 농도 예측



<그림 5-6> 모형 1의 학습데이터로 PM2.5 농도 예측



<그림 5-7> 모형 1의 검증데이터로 PM2.5 농도 예측

| | 최초 신경망 | | 수정한 신경망 | |
|------|--------|-------|---------|-------|
| | PM10 | PM2.5 | PM10 | PM2.5 |
| MAE | 24.31 | 19.50 | 7.93 | 8.98 |
| MAPE | 69.41 | 84.54 | 29.75 | 21.80 |

<표 5-4> 모형 1 성능

도로 미세먼지 농도에 영향을 주는 요소로 교통류 요인 외 기상요인이 있다. 제 4장 제 3절의 데이터 특성에서 계절별로 미세먼지 농도가 차이 를 보이는 것에서 기상 요소의 영향을 유추할 수 있으며, 제 2장의 선행 연구고찰에서 습도, 풍속, 기온이 미세먼지 농도에 영향을 미치는 것을 확인할 수 있었다. 따라서 기상요인을 입력데이터로 추가하여 미세먼지 농도를 예측하였다.

모형 2는 모형 1에 습도, 풍속, 기온 데이터를 추가해서 학습을 진행한 것이다. 그 결과를 모형 1과 비교했을 때, 신경망의 성능이 오히려 나빠 진 것을 확인할 수 있다. 따라서 실제 관측된 기상변수 중 미세먼지 농 도와 인과관계가 약한 변수를 찾기 위해 입력변수에 각 기상변수를 추 가, 삭제해가며 여러 가지 경우의 수를 확인하였다. 그 결과 기온변수가 신경망의 성능을 저하시킨다는 것을 확인하였다. 모형 3은 교통량, 중차 량 수, 속도, 가속도, 습도, 풍속을 이용해서 미세먼지 농도를 예측하는 신경망이며 각 모형의 설명과 성능을 <표 5-5>에 정리하였다. 모형 4는 기온변수를 변환해서 신경망을 구성했다. 날짜별 기온의 상대적인 효과 를 관찰하기 위해서 각 조사날짜의 평균기온에서 1분 평균기온이 차지하 는 비율을 변수로 신경망을 구성했다. 모형 5는 각 조사날짜의 최대 기 온에서 1분 평균기온이 차지하는 비율을 변수로 신경망을 구성했다.



<그림 5-8> 모형 2의 학습데이터로 PM10 농도 예측



<그림 5-9> 모형 2의 검증데이터로 PM10 농도 예측



<그림 5-10> 모형 2의 학습데이터로 PM2.5 농도 예측



<그림 5-11> 모형 2의 검증데이터로 PM2.5 농도 예측



<그림 5-12> 모형 3의 학습데이터로 PM10 농도 예측



<그림 5-13> 모형 3의 검증데이터로 PM10 농도 예측



<그림 5-14> 모형 3의 학습데이터로 PM2.5 농도 예측



<그림 5-15> 모형 3의 검증데이터로 PM2.5 농도 예측



<그림 5-16> 모형 4의 학습데이터로 PM10 농도 예측



<그림 5-17> 모형 4의 검증데이터로 PM10 농도 예측



<그림 5-18> 모형 4의 학습데이터로 PM2.5 농도 예측



<그림 5-19> 모형 4의 검증데이터로 PM2.5 농도 예측



<그림 5-20> 모형 5의 학습데이터로 PM10 농도 예측



<그림 5-21> 모형 5의 검증데이터로 PM10 농도 예측



<그림 5-22> 모형 5의 학습데이터로 PM2.5 농도 예측



<그림 5-23> 모형 5의 검증데이터로 PM2.5 농도 예측

| 모형 비교 | 입력변수 | 출력변수 |
|-------|---|-------------|
| | | |
| 모형 1 | 총 차량수, 중차량 수, 속도, 가속도 | PM10, PM2.5 |
| 모형 2 | 총 차량수, 중차량수, 속도, 가속도, 기온, 습도, 풍속 | PM10, PM2.5 |
| 모형 3 | 총 차량수, 중차량수, 속도, 가속도, 습도, 풍속 | PM10, PM2.5 |
| 모형 4 | 총 차량수, 중차량수, 속도, 가속도, (기온/조사날짜의 평균기온), 습도, 풍속 | PM10, PM2.5 |
| 모형 5 | 총 차량수, 중차량수, 속도, 가속도, (기온/조사날짜의 최대기온), 습도, 풍속 | PM10, PM2.5 |

<표 5-5> 각 모형의 입력변수와 출력변수

| 모형 비교 | | 신경망 성능 | | | |
|-------|-------|--------|-------|-------|-------|
| | | MAE | | MAPE | |
| | - | 학습 | 검증 | 학습 | 검증 |
| 모형 1 | PM10 | 5.13 | 7.93 | 11.87 | 29.75 |
| | PM2.5 | 7.10 | 8.98 | 15.38 | 21.80 |
| 모형 2 | PM10 | 4.41 | 13.75 | 8.48 | 34.13 |
| | PM2.5 | 6.50 | 12.60 | 12.84 | 32.86 |
| 모형 3 | PM10 | 5.80 | 6.88 | 11.97 | 21.54 |
| | PM2.5 | 5.86 | 6.58 | 12.97 | 20.04 |
| 모형 4 | PM10 | 5.02 | 13.43 | 11.36 | 36.57 |
| | PM2.5 | 4.15 | 13.90 | 9.15 | 37.52 |
| 모형 5 | PM10 | 5.12 | 13.99 | 11.72 | 34.31 |
| | PM2.5 | 4.36 | 15.62 | 10.75 | 36.86 |

<표 5-6> 각 모형의 성능 비교

<표 5-6>에서 보면 검증데이터의 MAE와 MAPE의 값이 모형 2,4,5에 서 높으므로 신경망의 성능이 좋지 않다. 모형 1과 3 중에서는 모형 3의 성능이 더 좋다고 판단할 수 있다. 모형 1과 3의 학습지표 값은 유사하 지만, 모형 3의 검증지표 값이 모형 1의 검증지표 값보다 더 낮기 때문 이다. 모형 1과 모형 2,4,5의 MAE와 MAPE의 값을 비교하면, 학습데이 터의 지표는 비슷하지만 모형 2,4,5의 검증지표 값이 커져서 신경망의 학 습이 제대로 이루어지지 않은 것을 알 수 있다. 이는 신경망이 과적합일 때 나타나는 현상이며 변수선택(Feature selection)을 통해 신경망을 수 정해야 한다. 변수선택은 성능을 저하하는 변수를 제거해서 신경망을 개 선하는 방법이다. 이를 통해서 새롭게 구축된 모형 3의 성능이 향상된 것을 알 수 있다.

모형 3의 검증데이터 지표가 모형 1 값보다 낮은 것은 교통류 변수만으 로 미세먼지 농도를 예측하는 것보다 기상변수를 추가할 때 예측력이 좋 아진다는 것을 의미한다. 실제 교통환경에서도 미세먼지가 교통 요인뿐 만 아니라 기상 등 외부요인에 의해 복합적으로 영향을 받는다. 하지만 선행연구나 연구가설과는 다르게 기온은 미세먼지 농도 예측에 영향을 끼치지 않는 것으로 나타났다.

제 3 절 입력변수 기여도 측정

제 2절에서 미세먼지 농도를 예측하기 위해 여러 입력변수를 통해 신경 망을 훈련하고 검증했다. 본 절에서는 모형의 성능에 가장 큰 영향을 주 는 입력변수를 식별하기 위해서 기여도 분석을 시행하였다. 기여도 분석 에는 SHAP를 사용했는데, Shapley값을 통해서 계산되며 최종적으로 각 변수의 개별적인 특성과 전역적인 중요도를 알 수 있다.

<그림 5-24>와 <그림 5-26>은 각 변수의 전역적 기여도를 나타낸 것 이다. Shapley값의 절대값이 클수록 출력값에 대한 기여도가 큰 변수이 다. <그림 5-25>와 <그림 5-27>은 각 변수의 개별적인 특성을 나타낸 것이며 각 입력변수가 PM 농도에 미치는 영향을 측정할 수 있다.



<그림 5-24> PM10 예측에 대한 입력변수의 전역적 기여도

| 입력변수 | Shapley값 |
|------|----------|
| 교통량 | 1.63 |
| 중차량수 | 3.09 |
| 속도 | 1.25 |
| 가속도 | 1.88 |
| 습도 | 4.12 |
| 풍속 | 1.54 |

<표 5-7> PM10 예측에 대한 입력변수의 Shapley값



<그림 5-25> 입력변수의 값이 Shapley값에 미치는 영향(PM10 예측 시)



<그림 5-26> PM2.5 예측에 대한 입력변수의 전역적 기여도

| 입력변수 | Shapley값 |
|------|----------|
| 교통량 | 1.03 |
| 중차량수 | 2.28 |
| 속도 | 1.16 |
| 가속도 | 1.58 |
| 습도 | 3.47 |
| 풍속 | 1.03 |

<표 5-8> PM2.5 예측에 대한 입력변수의 Shapley값



<그림 5-27> 입력변수의 값이 Shapley값에 미치는 영향(PM2.5 예측 시)

<그림 5-24>와 <그림 5-26>에서 PM10 농도와 PM2.5 농도에 공통적으로 가장 큰 영향을 미치는 상위 3개 변수는 습도, 중차량 수 그리고 가속도이다. <표 5-7>과 <표 5-8>에서 PM10 예측의 교통류 입력변수의 Shapley값을 보면 교통량, 중차량 수, 속도, 가속도가 각각 1.63, 3.09, 1.25, 1.88이며, PM2.5에 대해서는 각각 1.03, 2.28, 1.16, 1.58이다. 이 수치들은 각각의 변수에 대해 모든 Shapley값의 평균의 절대값을 취한 값이며 전역적인 측면에서 기여도를 의미하는 것으로 값이 클수록 상대적으로 기여도가 큰 변수임을 의미한다.

공통적으로 중차량 수와 가속도가 기여도가 높은 변수로 식별되었다. 교통량보다는 중차량 수가 미세먼지 농도에 더 많은 영향을 미치고, 속 도보다는 가속도가 더 많은 영향을 미친다는 것을 알 수 있다. 또한 기 상 지표에서는 습도가 풍속보다 높은 기여도를 보였다.

<그림 5-25>와 <그림 5-27>은 PM 농도 예측에 대한 입력변수의 효 과를 보여준다. 교통류 입력변수 중 상위 변수인 중차량 수와 가속도를 보면 중차량 수가 많아지고 가속도가 커질수록 PM10과 PM2.5의 농도가 증가하는 경향을 보였다. 교통량은 증가할수록, 속도는 감소할수록 PM 농도가 증가하는 경향이 있지만 해당 기여도는 중차량 수와 가속도에 비 해 작았다. 기상변수는 모두 값이 작을수록 PM 농도가 증가하는 경향이 있으며 습도가 풍속보다 중요한 변수로 나타났다.

중차량은 경차보다 배기 PM을 더 많이 배출할 뿐만 아니라, 재비산 먼 지와 같은 비배기 PM을 더 많이 배출한다. 따라서, 교통량을 줄이기 위 한 정책뿐만 아니라, 차체를 경량화하고 마모에 더 강한 내구성을 가진 도로 포장 및 타이어의 개발은 PM 농도를 저감하기 위한 현명한 전략 이 될 수 있다. 또한 차량의 평균속도보다는 가속과 감속의 크기가 PM 배출에 더 큰 영향을 미치는 것으로 나타났다. 따라서 교차로 신호관리 등을 통해 가감속도를 줄일 수 있는 교통 관리 전략이 필요하다.

습도와 풍속 또한 PM 농도에 영향을 미치는 것으로 나타났다. 습도가 높아지면 도로가 습윤해져서 도로의 재비산을 줄이는 경향이 있다. 풍속 이 높으면 미세먼지가 넓은 공간으로 확산하면서 PM 농도가 줄어들 수 있다. 다만 본 연구의 분석에서는 습도의 상대적 기여도가 높은 반면 풍 속의 기여도는 낮았다.

제 6 장 결론

기존 미세먼지 농도 예측 연구는 도로에서 일어나는 실시간 시·공간적 특성을 고려하지 못하는 한계가 있다. 지역에 따라 교통류 상황이 끊임 없이 변하기 때문에 시간적, 공간적 특성을 고려한 예측 모형이 필요하 다. 본 연구에서는 실도로 주행을 통해 관측한 자료와 다층신경망을 이 용해 미세먼지 농도를 예측하였다.

데이터 전처리, 신경망 구조설정, 신경망 학습설정, 신경망 성능평가의 과정을 거쳤다. 신경망에 사용한 데이터는 전처리를 통해서 학습과 검증 데이터로 분류하였다. 신경망의 구조와 학습설정에서는 신경망 내에서 동작하는 여러 함수와 하이퍼 파라미터를 설정하였다. 이후, 모형마다 입 력변수를 다르게 해서 모형의 성능을 비교했다.

연구의 주요 결과는 다음과 같다. 첫째, 교통류와 기상변수를 모두 투입 한 모형에서는 과적합이 발생했다. 교통류와 기온을 제외한 모든 기상변 수를 투입한 모형의 성능이 가장 우수하였다.

둘째, 교통량, 중차량, 가속도의 증가는 미세먼지 농도를 증가시키는 반 면 평균속도의 증가는 미세먼지 농도를 감소시키는 경향을 보였다. 기상 요인은 모두 증가할수록 미세먼지 농도를 감소시켰다.

셋째, 입력변수의 기여도 분석을 통해서 무거운 차량이 일반차량에 비 해, 자동차의 가속 및 감속이 평균속도에 비해 미세먼지 농도에 더 큰 영향을 미쳤다. 기상변수의 경우 습도가 풍속보다 더 중요한 변수였다.

다만, 연구 과정에서 한계가 존재하였다. 더 정확한 분석이 되기 위해서 는 날짜별 배경 대기의 자료가 필요하다. 날짜별로 수집된 미세먼지 농 도가 그 날의 평균적인 미세먼지 농도를 고려하지는 못한다. 조사구간 근처에서 배경 대기를 관측해 연구를 진행한다면 더 정확한 연구가 될 수 있을 것이다.

연구결과를 바탕으로 미세먼지 저감을 위한 교통 관리 방안을 도출할

수 있다. 우선 차량의 급격한 가속, 감속을 줄일 수 있는 교통 관리 방안 이 대기 질 개선을 위한 현명한 전략이 될 수 있다. 다음으로 대형차에 과금하는 세제개편도 대기 질을 관리할 수 있는 효과적인 방안이 될 수 있다. 장기적으로는 신소재 연구를 통해 마모에 강한 타이어와 차체를 가볍게 할 수 있는 자동차 프레임의 개발이 필요할 것이다.

참 고 문 헌

- Abdull, N., Yoneda, M., Shimada, Y., 2020. Traffic characteristics and pollutant emission from road transport in urban area. Air Quality, Atmosphere & Health volume 13, 731 - 738.
- Abu-Allaban, M., Gillies, J. A., Gertler, A. W., Clayton, R., Proffitt, D., 2003. Tailpipe, resuspended road dust, and brake-wear emission factors from on-road vehicles. Atmospheric Environment 37, 5283-5293.
- Amato, F., Pandolfi, M., Escrig, A., Querol, X., Alastuey, A., Peya, J., Perez, N., Hopke, P. K., 2009. Quantifying road dust resuspension in urban environment by multilinear engine: A comparison with PMF2. Atmospheric Environment 43, 2770–2780.
- Azhari, A., Latif, M. T., Mohamed, A. F., 2018. Road traffic as an air pollutant contributor within an industrial park environment. Atmospheric Pollution Research 9, 680–687.
- Beddows, D.C.S., Harrison, R.M., 2021. PM10 and PM2.5 emission factors for non-exhaust particles from road vehicles: Dependence upon vehicle mass and implications for battery electric vehicles. Atmospheric Environment 244, 117886.
- Belkacem, I., Helali, A., Khardi, S., & Slimi, K. 2022. Investigations on vehicle non-exhaust particle emissions: real-time measurements. International Journal of Environmental Science

and Technology, 19(12), 11749-11762.

- Bergmann, M., Kirchner, U., Vogt, R., Benter, T., 2009. On-road and laboratory investigation of low-level PM emissions of a modern diesel particulate filter equipped diesel passenger car. Atmospheric Environment, 43, 1908–1916.
- Biswas, S., Hu, S., Verma, V., Herner, J. D., Robertson, W. H., Ayala, A., Sioutas, C., 2008. Physical properties of particulate matter (PM) from late model heavy-duty diesel vehicles operating with advanced PM and NOx emission control technologies. Atmospheric Environment, 42, 5622–5634.
- Bodor, Z., Bodor, K., Keresztesi, Á., & Szép, R. 2020. Major air pollutants seasonal variation analysis and long-range transport of PM 10 in an urban environment with specific climate condition in Transylvania (Romania). Environmental Science and Pollution Research, 27, 38181–38199.
- Bukowiecki, N., Dommen, J., Prévôt, A.S.H., Richter, R., Weingartner, E., Baltensperger, U., 2002. A mobile pollutant measurement laboratory-measuring gas phase and aerosol ambient concentrations with high spatial and temporal resolution. Atmospheric Environment 36, 5569–5579.
- Chang, J. S., Hwang, J., & Choi, M. (2022). Vehicle Detection Approach Adjusting Road Curves to Estimate Local Traffic Density under Real Driving Conditions. Transportation Research Record, 03611981221123809.

- Chen, Z., Zhang, X., & Chen, F. 2021. Have driving restrictions reduced air pollution: evidence from prefecture-level cities of China. Environmental Science and Pollution Research, 28, 3106–3120.
- Cheng, B., Ma, Y., Feng, F., Zhang, Y., Shen, J., Wang, H., Guo, Y., Cheng, Y., 2021. Influence of weather and air pollution on concentration change of PM2.5 using a generalized additive model and gradient boosting machine. Atmospheric Environment 255, 118437.
- Ceylan, Z., & Bulkan, S. E. R. O. L. 2018. Forecasting PM10 levels using ANN and MLR: A case study for Sakarya City. Glob. Nest J, 20, 281–290.
- Futagami, K., Fukazawa, Y., Kapoor, N., & Kito, T. 2021. Pairwise acquisition prediction with SHAP value interpretation. The Journal of Finance and Data Science, 7, 22–44.
- Gao, J., Chen, H., Liu, Y., Laurikko, J., Li, Y., Li, T., Tu, R., 2021. Comparison of NOx and PN emissions between Euro 6 petrol and diesel passenger cars under real-world driving conditions. Science of The Total Environment 801, 149789.
- Gillies, J. A., Etyemezian, V., Kuhns, H., Nikolic, D., Gillette, D. A., 2005. Effect of vehicle characteristics on unpaved road dust emissions. Atmospheric Environment 39, 2341–2347.

- Giovanis, E., 2018. The relationship between teleworking, traffic and air pollution. Atmospheric Pollution Research 9, 1–14.
- Gu, J., Deffner, V., Küchenhoff, H., Pickford, R., Breitner, S., Schneider, A., ... & Cyrys, J. 2022. Low emission zones reduced PM10 but not NO2 concentrations in Berlin and Munich, Germany. Journal of Environmental Management, 302, 114048.
- Hornik, K., Stinchcombe, M., & White, H. 1989. Multilayer feedforward networks are universal approximators. Neural networks, 2(5), 359–366.
- Ježek, I., Drinovec, L., Ferrero, L., Carriero, M., Močnik, G., 2015. Determination of car on-road black carbon and particle number emission factors and comparison between mobile and stationary measurements. Atmospheric Measurement Techniques 8, 43 - 55.
- Karjalainen, P., Pirjola, L., Heikkilä, J., Lähde, T., Tzamkiozisc, T., Ntziachristosc, L., Keskinena, J., Rönkkö, T., 2014. Exhaust particles of modern gasoline vehicles: A laboratory and an on-road study. Atmospheric Environment 97, 262–270.
- Ketzel, M., Omsstedt, G., Johansson, C., Düring, I., Pohjola, M., Oettl, D., Gidhagen, L., Wåhlin, P., Lohmeyer, A., Haakana, M., Berkowicz, R., 2007. Estimation and validation of PM2.5/PM10 exhaust and non-exhaust emission factors for practical street

pollution modelling. Atmospheric Environment 41, 9370 - 9385.

- Khullar, S., & Singh, N. 2021. Water quality assessment of a river using deep learning Bi-LSTM methodology: forecasting and validation. Environmental Science and Pollution Research, 1–15.
- Kingma, D. P., & Ba, J. 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.
- Kinnon, M. M., Zhu, S., Carreras-Sospedra, M., Soukup, J. V., Dabdub, D., Samuelsen, G. S., Brouwer, J., 2019. Considering future regional air quality impacts of the transportation sector. Energy Policy 124, 63 - 80.
- Krampah, F., Amegbey, N., Ndur, S., Ziggah, Y. Y., Hopke, P. K., 2021. Fractal Analysis and Interpretation of Temporal Patterns of TSP and PM10 Mass Concentration over Tarkwa, Ghana. Earth Systems and Environment 5, 635 - 654.
- Krecl, P., Johansson, C., Targino, A. C., Ström, J., Burman, L., 2017. Trends in black carbon and size-resolved particle number concentrations and vehicle emission factors under real-world conditions. Atmospheric Environment, 165, 155–168.
- Krishan, M., Jha, S., Das, J., Singh, A., Goyal, M. K., & Sekar, C., 2019. Air quality modelling using long short-term memory (LSTM) over NCT-Delhi, India. Air Quality, Atmosphere & Health, 12(8), 899–908.

- Kliengchuay, W., Worakhunpiset, S., Limpanont, Y., Meeyai, A. C., Tantrakarnapa, K., 2021. Influence of the meteorological conditions and some pollutants on PM10 concentrations in Lamphun, Thailand. Journal of Environmental Health Science and Engineering 19, 237 - 249.
- Kumar, P., Goel, A., 2016. Concentration dynamics of coarse and fine particulate matter at and around signalised traffic intersections. Environmental Science 18, 1220–1235.
- Kujawska, J., & Pawłowska, M. 2022. The effect of amendment addition drill cuttings on heavy metals accumulation in soils and plants: Experimental study and artificial network simulation. Journal of Hazardous Materials, 425, 127920.
- Kuppili, S. K., Dheeraj Alshetty, V., Diya, M., Shiva Nagendra, S.M., Ramadurai, G., Ramesh, A., Gulia, S., Namdeo, A., Maji, K., Bell, M., Goodman, P., Hayes, E., Barnes, J., Longhurst, J., De Vito, L., 2021. Characteristics of real-world gaseous exhaust emissions from cars in heterogeneous traffic conditions. Transportation Research Part D: Transport and Environment 95, 102855.
- Lešnik, U., Mongus, D., Jesenko, D., 2019. Predictive analytics of PM10 concentration levels using detailed traffic data. Transportation Research Part D 67, 131-141.
- Lin, S., Liu, Y., Chen, H., Wu, S., Michalaki, V., Proctor, P., &

Rowley, G. 2022. Impact of change in traffic flow on vehicle non-exhaust PM2. 5 and PM10 emissions: A case study of the M25 motorway, UK. Chemosphere, 303, 135069.

- Liu, Y., Chen, H., Gao, J., Li, Y., Dave, K., Chen, J., ... & Perricone, G. 2021. Comparative analysis of non-exhaust airborne particles from electric and internal combustion engine vehicles. Journal of Hazardous Materials, 420, 126626.
- Liu, Y., Liu, Z., Luo, X., & Zhao, H. 2022. Diagnosis of Parkinson's disease based on SHAP value feature selection. Biocybernetics and Biomedical Engineering, 42(3), 856–869.
- Martinet, S., Liu, Y., Louis, C., Tassel, P., Perret, P., Chaumond, A., André, M., 2017. Euro 6 unregulated pollutant characterization and statistical analysis of after-treatment device and driving-condition impact on recent passenger-car emissions. Environmental Science & Technology 51, 5847–5855.
- Mishra, R. K., Pandey, A., Pandey, G., Kumar, A., 2019. The effect of odd-even driving scheme on PM2.5 and PM1.0 emission. Transportation Research Part D 67, 541–552.
- Munjal, A., Sharma, S., Nema, A. K., & Kota, S. H. 2022. Factors affecting particulate matter levels near highway toll plazas in India. Transportation Research Part D: Transport and Environment, 110, 103403.
- Nayeb Yazdi, M., Delavarrafiee, M., Arhami, M., 2015. Evaluating

near highway air pollutant levels and estimating emission factors: Case study of Tehran, Iran. Science of The Total Environment 538, 375–384.

- wosisi, M. C., Oguntoke, O., & Taiwo, A. M. 2021. Dispersion modeling of PM10 from selected flow stations in the Niger Delta, Nigeria: implications on soot pollution. Environmental Systems Research, 10(1), 1–16.
- Qadir, R., Abbaszade, G., Schnelle-Kreis, J., Chow, J., Zimmermanna, R., 2013. Concentrations and source contributions of particulate organic matter before and after implementation of a low emission zone in Munich, Germany. Environmental Pollution 175, 158–167.
- Raparthi, N., Debbarma, S., & Phuleria, H. C. 2022. Determination of heavy-duty vehicle emission factors from highway tunnel measurements in India: Laboratory vs. real-world study. Atmospheric Pollution Research, 13(11), 101581.
- Rumelhart, D. E., Hinton, G. E., & Williams, R. J. 1986. Learning representations by back-propagating errors. nature, 323(6088), 533-536.
- Sarmadi, M., Rahimi, S., Rezaei, M., Sanaei, D., & Dianatinasab, M. 2021. Air quality index variation before and after the onset of COVID-19 pandemic: a comprehensive study on 87 capital, industrial and polluted cities of the world. Environmental Sciences Europe, 33(1), 1–17.

- Sindosi, O. A., Hatzianastassiou, N., Markozannes, G., Rizos, E. C., Ntzani, E., & Bartzokas, A. 2021. PM 10 Concentrations in a Provincial City of Inland Greece in the Times of Austerity and Their Relationship with Meteorological and Socioeconomic Conditions. Water, Air, & Soil Pollution, 232, 1–19.
- Sofowote, U. M., Healy, R. M., Su, Y., Debosz, J., Noble, M., Munoz, A., Jeong, C. -H., Wang, J. M., Hilker, N., Evans, G. J., Hopke, P. K., 2018. Understanding the PM2.5 imbalance between a far and near-road location: Results of high temporal frequency source apportionment and parameterization of black carbon. Atmospheric Environment 173, 277–288.
- Srimuruganandam, B., & Nagendra, S. M. S. 2010. Analysis and interpretation of particulate matter - PM10, PM2. 5 and PM1 emissions from the heterogeneous traffic near an urban roadway. Atmospheric Pollution Research, 1(3), 184–194.
- Sritong-aon, C., Thomya, J., Kertpromphan, C., Phosri, A., 2021. Estimated effects of meteorological factors and fire hotspots on ambient particulate matter in the northern region of Thailand. Air Quality, Atmosphere & Health volume 14, 1857 - 1868.
- Suleiman, A., Tight, M. R., & Quinn, A. D. 2019. Applying machine learning methods in managing urban concentrations of traffic-related particulate matter (PM10 and PM2. 5). Atmospheric Pollution Research, 10(1), 134–144.

- Sun, C., Zheng, S., Wang, R., 2014. Restricting driving for better traffic and clearer skies: Did it work in Beijing? Transport Policy 32, 34 - 41.
- Timmers, V. R. J. H., Achten, P. A. J., 2016. Non-exhaust PM emissions from electric vehicles. Atmospheric Environment 134, 10–17.
- Wang, X., Yuan, J., Wang, B., 2021. Prediction and analysis of PM2.5 in Fuling District of Chongqing by artificial neural network. Neural Computing and Applications 33, 517–524.
- Wang, J. Q., Du, Y., & Wang, J. 2020. LSTM based long-term energy consumption prediction with periodicity. Energy, 197, 117197.
- Wu, S., Feng, Q., Du, Y., & Li, X. 2011. Artificial neural network models for daily PM10 air pollution index prediction in the urban area of Wuhan, China. Environmental Engineering Science, 28(5), 357–363.
- Xia, Y., Braun, S., Reddy, C. K., Dubey, H., Cutler, R., & Tashev, I. 2020, May. Weighted speech distortion losses for neural-network-based real-time speech enhancement. In **ICASSP** 2020-2020 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP) (pp. 871-875). IEEE.

- Xu, J., Saleh, M., Hatzopoulou, M., 2020. A machine learning approach capturing the effects of driving behaviour and driver characteristics on trip-level emissions. Atmospheric Environment 224, 117311.
- Yan, R., Liao, J., Yang, J., Sun, W., Nong, M., & Li, F. 2021. Multi-hour and multi-site air quality index forecasting in Beijing using CNN, LSTM, CNN-LSTM, and spatiotemporal clustering. Expert Systems with Applications, 169, 114513.
- Yli-Tuomi, T., Aarnio, P., Pirjola, L., Mäkelä, T., Hillamo, R., Jantunen, M., 2005. Emissions of fine particles, NOx, and CO from on-road vehicles in Finland. Atmospheric Environment 39, 6696–6706.
- Zhang, H., Wang, Y., Hu, J., Ying, Q., Hu, X.-M., 2015. Relationships between meteorological parameters and criteria air pollutants in three megacities in China. Environmental Research 140, 242–254
- Zhang, M., Shan, C., Wang, W., Pang, J., & Guo, S. 2020. Do driving restrictions improve air quality: Take Beijing-Tianjin for example?. Science of the Total Environment, 712, 136408.
- Zhang, Q., Yang, L., Zhang, Y., Fang, X., Wu, L., & Mao, H. 2022. PM2. 5–PAHs and PM10–PAHs at roadside environment: levels, meteorological impact, source apportionment, and health risks. Air Quality, Atmosphere & Health, 15(7), 1287–1301.

- Zhao, H., Che, H., Zhang, X., Ma, Y., Wang, Y., Wang, H., Wang, Y., 2013. Characteristics of visibility and particulate matter (PM) in an urban area of Northeast China. Atmospheric Pollution Research 4, 427–434.
- Zhou, S., Lin, R., 2019. Spatial-temporal heterogeneity of air pollution: The relationship between built environment and on-road PM2.5 at micro scale. Transportation Research Part D 76, 305–322.

Abstract

Prediction of Particulate Matter Concentrations on Urban Roads Using Multi Layer Perceptrons

Sungbeom Park

Department of Environmental Planning Graduate School of Environmental Studies Seoul National University

The causality between particulate matter (PM) concentrations and traffic characteristics is known to be complex and unclear. This is why studies using traditional methodologies such as statistical models have not clearly identified the causality. In order to overcome these limitations, studies have recently been conducted to investigate the causality between PM concentrations and traffic characteristics using deep learning. However, since the traffic situation changes constantly depending on the region, a prediction model considering the temporal and spatial characteristics is needed.

Therefore, in this study, a multi layer perceptron was built based on data collected in real time while repeatedly driving a specific road
section. The process of building a multi layer perceptron consists of data pre-processing, neural network structure setting, neural network learning setting, and neural network performance evaluation. Validation of several preliminary neural network models presents the best performance neural network and identifies input variables that improve the performance of the neural network. In addition, how each input variable affects the PM concentrations was analyzed.

The main findings of the study are as follows. First, the performance of the model with traffic variables and all weather variables except temperature was the best.

Second, the increase in total vehicles, number of heavy vehicles, and acceleration increased the PM concentrations, while the increase in average speed decreased the PM concentrations. As all weather factors increased, the PM concentrations decreased.

Third, through the analysis of the contribution of input variables, it was found that heavy vehicles affect PM concentrations more than general vehicles, and acceleration and deceleration of vehicles affect PM concentrations more than average speed.

Based on the research results, traffic management measures can be derived to reduce PM concentrations. First of all, traffic management measures that can reduce the rapid acceleration and deceleration of vehicles can be a smart strategy for improving air quality. Next, reform of taxation that imposes more penalties to heavy vehicles can also be an effective measure that can manage air quality. In the long run, more research in terms of new materials should be conducted on light but strong car frames and wear-resistant tires.

keywords : particulate matter concentrations, exhaust emissions, non-exhaust emissions, traffic, multi layer perceptrons, air quality

Student Number : 2020-28067