

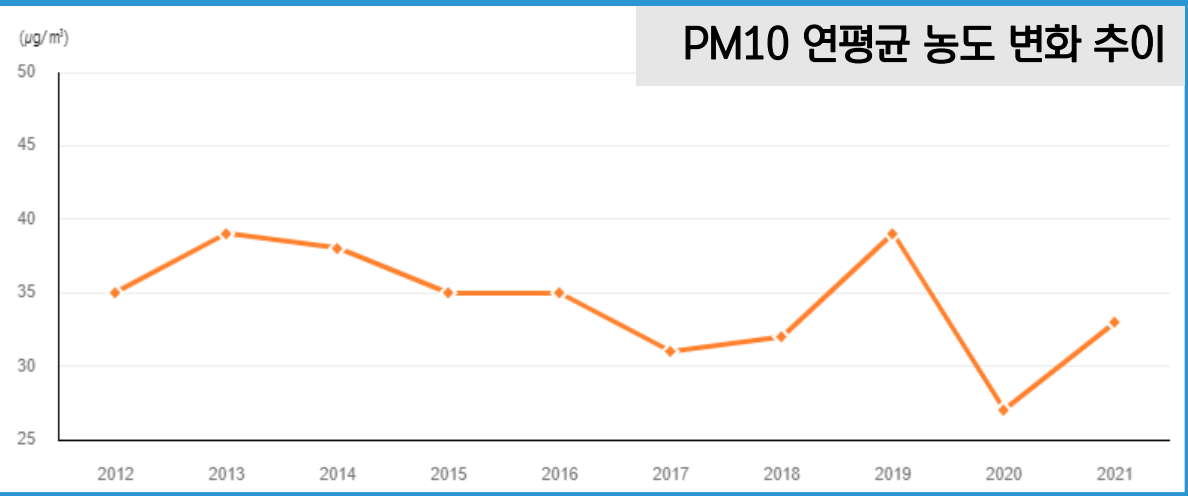
미세먼지 유발 영향인자 분석을 통한 대응방안 수립

미세먼지를 발생시키는 영향인자 선정 및 발생량 예측 모델 생성
C4 고다영

‘침묵의 살인자’라 불리는 미세먼지는 건강, 환경, 경제 등 악영향을 끼치며 다양한 문제를 일으키고 있다. 국내에서는 이에 맞서 각 지자체에서 여러 대응방안을 제시하고 있다. 대표적으로 환경부에서 운영하는 ‘미세먼지 계절관리제’는 배출가스 5등급 차량의 수도권 운행 제한, 공공기관 차량 2부제 등의 조치를 취하고 있다.

그러나 우리나라는 10년간 전 지역에서 세계보건기구(WHO)가 정한 연평균 미세먼지 수치($20\mu\text{g}/\text{m}^3$)를 초과하며 뚜렷한 효과를 거두지 못하고 있는 실정이다. **이에 따라 빅데이터 분석을 통해 미세먼지 발생/증가에 영향을 미치는 인자를 도출하고 그 영향도를 분석한다. 이와 더불어 선정된 영향인자를 활용하여 미세먼지 발생량 예측 모델을 만들고 적절한 대응방안을 수립하고자 한다.**

PM10 : 지름 $10\mu\text{m}$ 이하의 먼지 = 미세먼지



http://www.index.go.kr/potal/main/EachDtlPageDetail.do?idx_cd=1398

통계청에서 운영하는 e-나라지표 홈페이지에서는 **에어로졸(PM10)** 연평균 농도 변화 추이와 함께 아황산가스(SO_2) 오존(O_3), 이산화질소(NO_2), 일산화탄소(CO) 등 대기오염물질의 오염도에 관한 정보를 함께 제공하고 있다.

[미세먼지 계절 관리제]

[아주 정확한 팩트체크] 미세먼지 계절 관리제, 왜 겨울에만?

고기압 영향...봄·겨울 미세먼지 많아

노경조 기자 (felizkj@ajunews.com) | 입력 : 2021-11-30 12:00 | 수정 : 2021-11-29 07:44



미세먼지 [사진=연합뉴스]

정부는 29일 겨울철 고농도 미세먼지 대응을 위한 '제3차 미세먼지 계절관리제 시행 계획'을 논의했다.

고농도 미세먼지는 미세먼지가 나쁨 수준(PM_{10} $81\mu\text{g}/\text{m}^3$ 이상)이거나 초미세먼지가 나쁨 수준($\text{PM}_{2.5}$ $36\mu\text{g}/\text{m}^3$)인 상태가 1시간 이상 지속되는 것을 의미한다. 주로 대기 정체, 난방 사용, 바람에 의해 봄·겨울에 많이 발생한다.

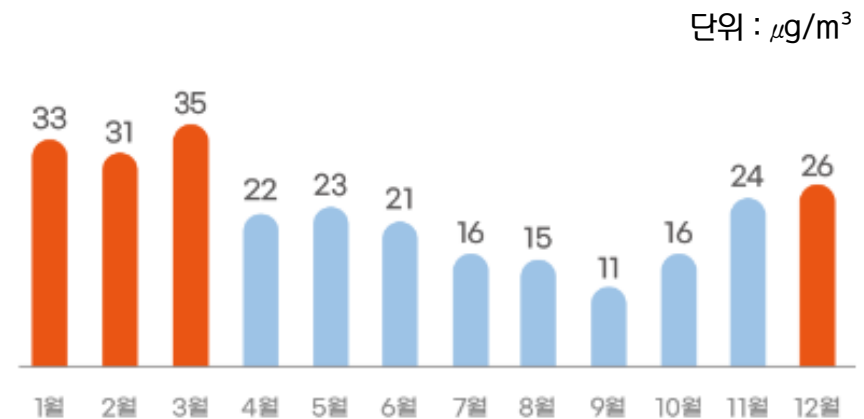
다만, 올해 우리나라 초미세먼지 농도는 정부 대책 누적 효과, 중국의 미세먼지 농도 개선, 기상 영향 등이 복합적으로 작용해 양호하다고 정부는 설명했다.

미세먼지 1차 발생원으로 매연, 자동차 배기가스, 건설현장 등이 지목 받고 있다. 그렇다면 **대기오염물질**과 미세먼지 발생량 사이에 연관성이 있지 않을까?

매년 12월부터 3월까지 미세먼지 집중관리제가 운영되고 있다. 미세먼지가 계절에 따라 변하는 이유는 **기온, 기압, 바람 등 기상정보**와 관련 있는 걸까?



[미세먼지 1차 발생원]



출처 : 서울특별시

[월별 초미세먼지 농도('18~'20)]

[분석 목표] 미세먼지 발생/증가에 영향을 미치는 인자 도출 및 대응방안 수립

데이터 수집

기상 정보

- 기온($^{\circ}\text{C}$)
- 강수량(mm)
- 풍속(m/s), 풍향
- 습도(%)
- 기압(hPa)
- 적설량, 전운량

대기 오염

- 아황산가스(SO_2) 농도
- 오존(O_3) 농도
- 이산화질소(NO_2) 농도
- 일산화탄소(CO) 농도

데이터 정제

- **데이터 품질** 확인 – 결측치, 이상치 처리
- **EDA** – 데이터 분포 확인 및 파생변수 생성

데이터 분석

- 회귀분석 및 기계학습 기법을 통한 **데이터 별 중요도 파악**
- **미세먼지 발생량 예측 모델 생성** 및 모델 평가

결론 도출

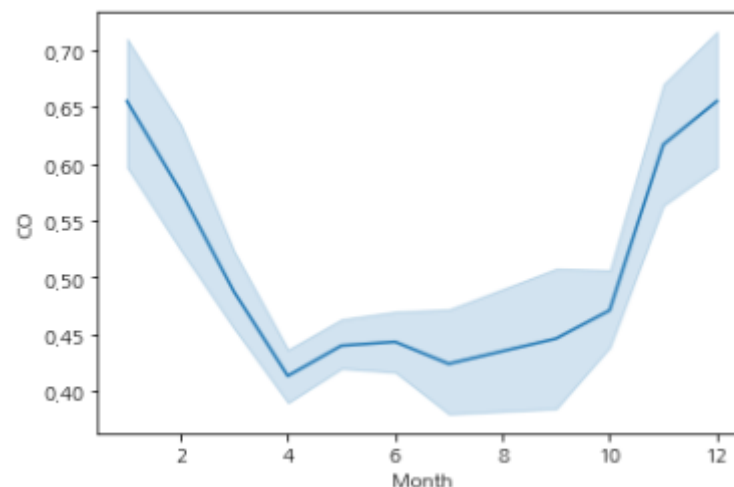
- 미세먼지 발생/증가에 영향을 미치는 **인자 도출**
- 영향인자 분석을 통한 **대응방안 수립**

[데이터 수집 결과]

MeasDate	측정일자	설명변수	연속형
PM10	미세먼지 $10\mu\text{g}/\text{m}^3$	목표변수	연속형
O3	오존 농도	설명변수	연속형
NO2	이산화질소 농도	설명변수	연속형
CO	일산화탄소 농도	설명변수	연속형
SO2	이황산가스 농도	설명변수	연속형
TEMP	기온($^{\circ}\text{C}$)	설명변수	연속형
RAIN	강수량(mm)	설명변수	연속형
WIND	풍속(m/s)	설명변수	연속형
WIND_DIR	풍향(16방위)	설명변수	연속형
HUMIDITY	습도(%)	설명변수	연속형
ATM_PRESS	현지기압(hPa)	설명변수	연속형
SNOW	적설(cm)	설명변수	연속형
CLOUD	전운량(10분위)	설명변수	연속형

① 결측치 처리

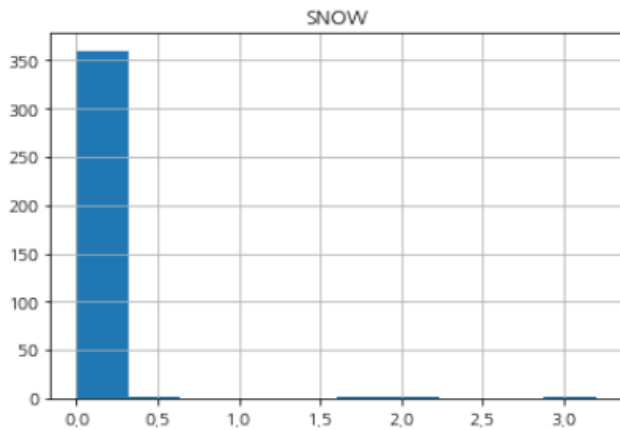
- PM10 데이터가 누락된 행은 삭제
- 2019년 7월부터 2019년 9월 CO 데이터 누락



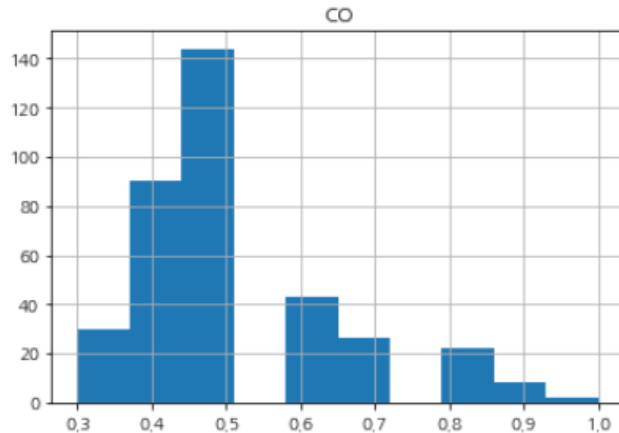
→ 일산화탄소는 도시가스가 연소되면서 발생하는 대표적인 성분으로 일반적으로 여름철보다 겨울철에 도시가스 사용량 증가로 함께 증가하는 추세를 보인다. 따라서 결측치는 3월부터 10월까지 일산화탄소 배출량 평균으로 대체한다.

데이터 탐색

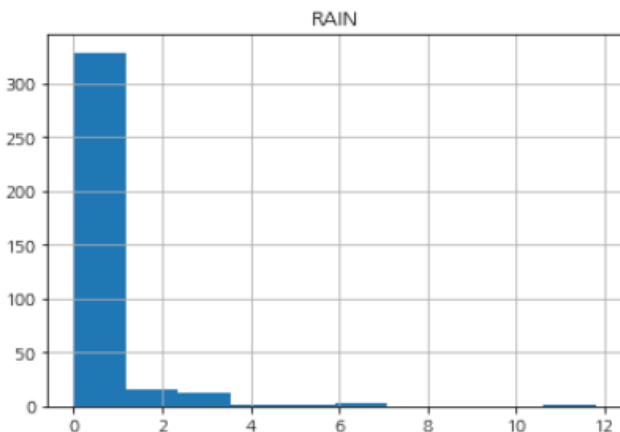
② 이상치 확인



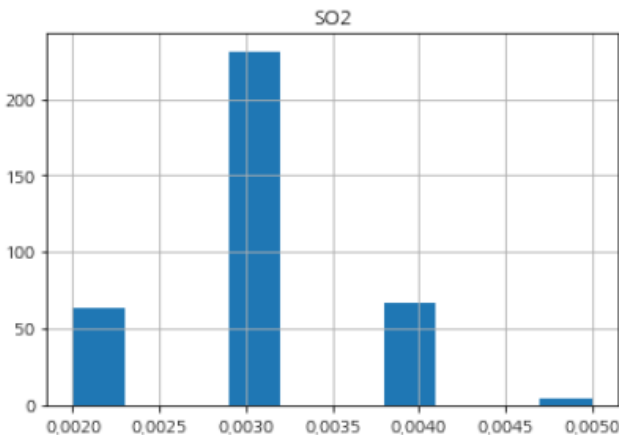
[적설량 분포]



[일산화탄소 분포]



[강수량 분포]



[아황산가스 분포]

- 대부분의 분포가 연속적으로 나타나지만 적설량, 강수량, 일산화탄소, 아황산가스에 대해서 불연속적인 분포를 확인

- 적설량, 강수량
→ 우리나라의 기후 특성 상 계절별로 편차가 심하게 나타나기 때문에 불연속적인 분포가 나타남.

- 일산화탄소, 아황산가스
→ 일산화탄소는 계절에 따라 배출량에 차이를 보이며 아황산가스는 x축 범위가 작기 때문에 불연속적인 분포가 나타난 것으로 파악됨

③ 상관관계 확인

①

	PM10	O3	NO2	CO	SO2
PM10	1.000	-0.052	0.396	0.593	0.429
O3	-0.052	1.000	-0.592	-0.520	-0.234
NO2	0.396	-0.592	1.000	0.797	0.563
CO	0.593	-0.520	0.797	1.000	0.586
SO2	0.429	-0.234	0.563	0.586	1.000

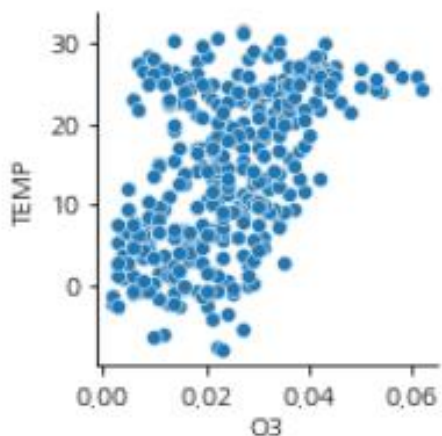
① { PM10과 대기오염물질 간의 상관관계가 상당히 높기 나타난다
대기오염물질 간의 상관관계 역시 높게 나타난다.

② 오존과 기온 사이에 양의 상관관계가 나타난다.

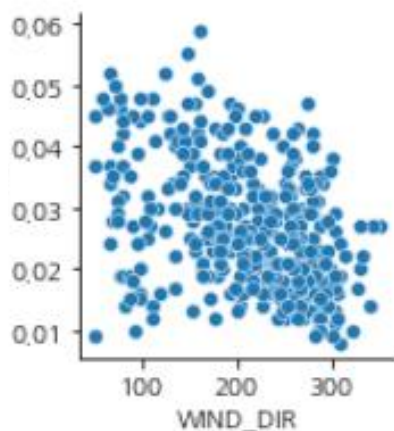
→ 여름에는 겨울보다 햇빛과 자외선이 강해져 오존 농도가 상승하는 현상과 연관

③ 풍향과 이산화질소 사이 음의 상관관계가 나타난다.

→ 서풍이 불수록 이산화질소의 농도가 감소한다는 것은 바람이 대기 중 오염물질을 이동시킨다고 유추할 수 있다.



②



③

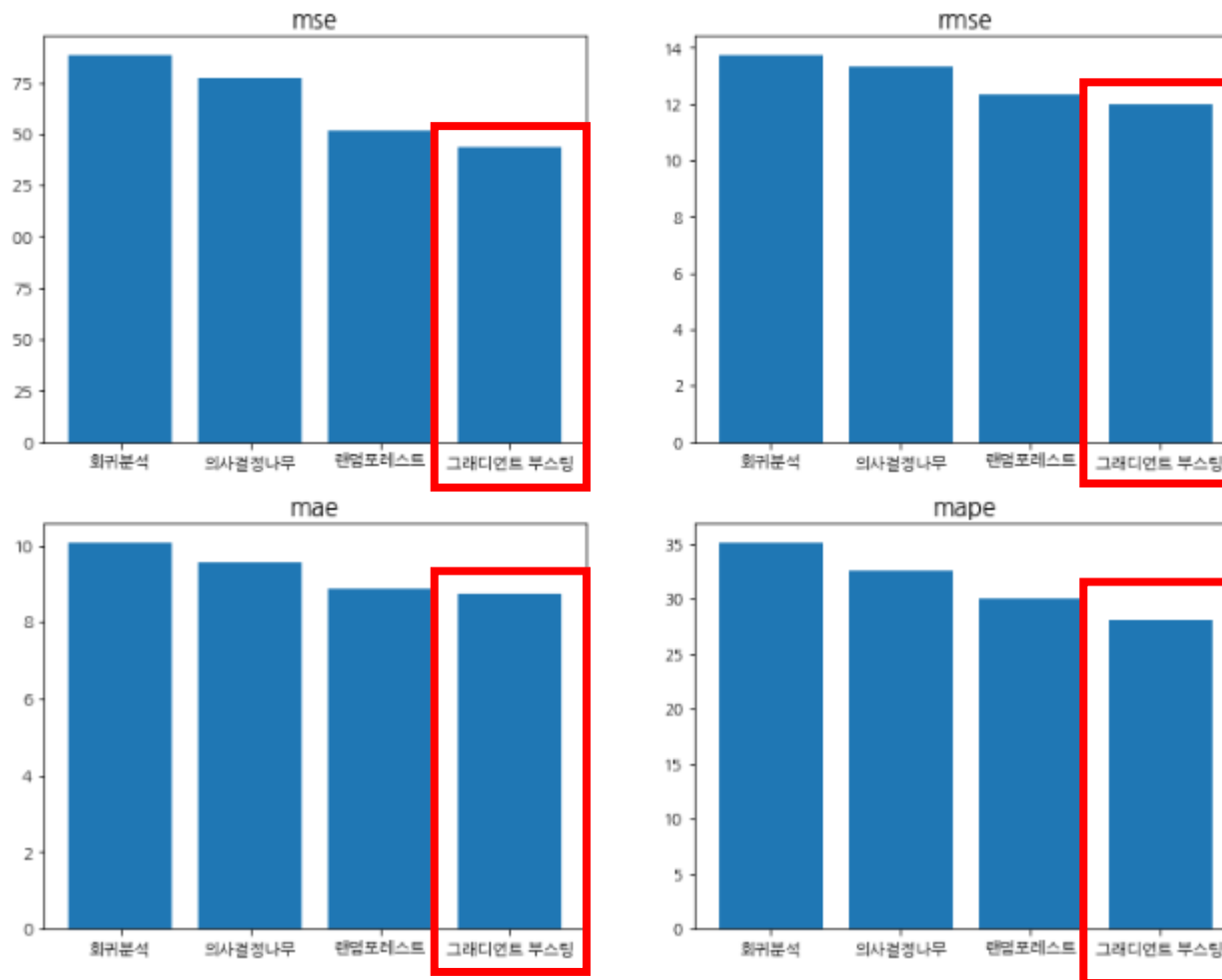
다중선형회귀분석, 의사결정나무(DT), 랜덤포레스트(RF), 그래디언트 부스팅(GB)를 통해 변수 간 중요도 순위를 파악한 결과, **최종 영향인자로 오존, 이산화질소, 일산화탄소, 기온, 풍속, 풍향, 기압을 선정**

[각 변수 별 중요도 순위]

	모델링 기법				최종 영향인자 선정
	회귀분석	DT	RF	GB	
O3	2	3	2	2	0
NO2	3	6	6	9	0
CO	1	1	1	1	0
SO2		8	11		X
TEMP	4	4	5	3	0
RAIN		10	10	10	X
WIND	7		3	5	0
WIND_DIR	6	2	4	4	0
HUMIDITY		7	7	6	X
ATM_PRESS	5	5	9	8	0
SNOW	9				X
CLOUD	8	9	8	7	X

빈 칸은 순위에서 벗어났음을 의미.

[평가지표 별 모델 성능 평가]



- 최종 선정된 영향인자를 통해 미세먼지 발생량 예측 모델을 생성한 결과, 최종 성능이 가장 우수했던 **그라디언트 부스팅 모델**을 채택한다.

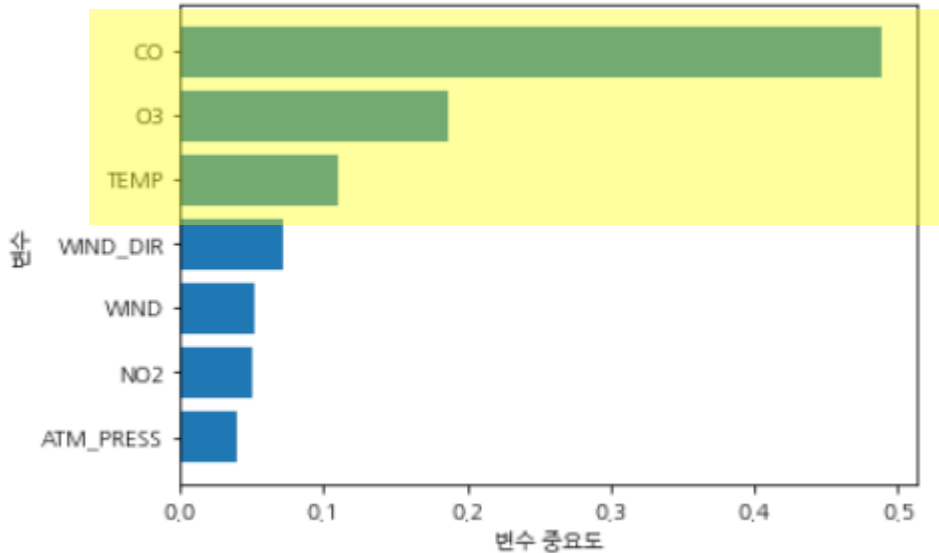
- 최종 성능**
 → Train data : **78.5 %**
 → Test Data : **45.4%**

- 차후 데이터를 더 확보하고 조정하여 모델 개선이 필요해 보인다. 풍향과 풍속 사이의 연관성을 간과하고 모델을 개발해 정확도에서 아쉬운 점을 보인다.

분석 결과 해석

최종모델에서의 영향인자 간 중요도 파악 결과, 일산화탄소가 미세먼지 발생/증가에 가장 많은 영향을 미치고, 다음으로 오존과 기온이 영향을 미친다는 것을 확인하였다. 세 영향인자를 구체적으로 분석하여 미세먼지 대응방안을 수립하고자 한다.

[최종모델에서의 영향인자 간 중요도]



■ 일산화탄소

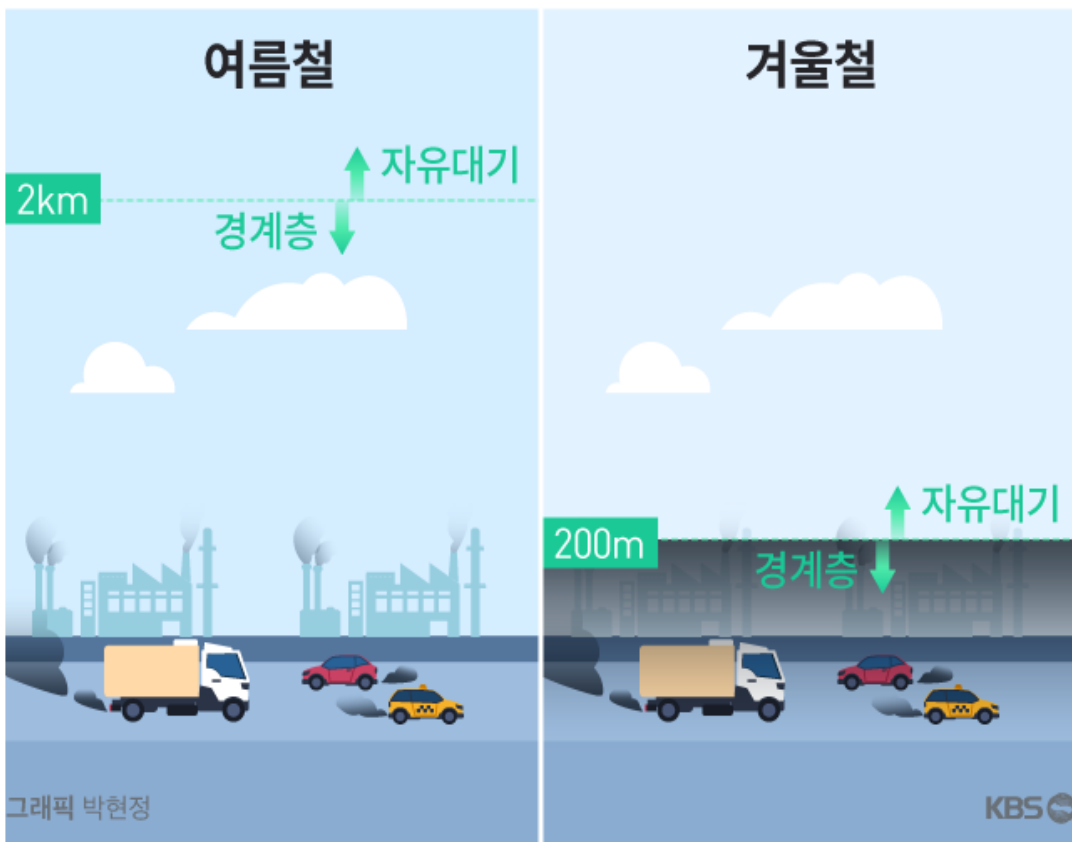
→ 주로 자동차 배기가스나 산업공정, 도시가스 등에서 불완전 연소로 발생하는 대기오염물질
→ 겨울에 도시가스 사용량이 증가하면서 겨울철에 특히 증가하는 추세를 보인다.

■ 오존

→ 자동차 등을 통해 대기 중에 배출된 질소산화물과 휘발성유기화합물(VOCs) 등이 자외선과 광화학 반응을 일으켜 생성된 2차 대기오염물질
→ 여름에 자외선이 강해져 대기 중 화학반응이 강하게 일어나 오존 농도가 증가하는 추세를 보인다.

분석 결과 해석

최종모델에서의 영향인자 간 중요도 파악 결과, 일산화탄소가 미세먼지 발생/증가에 가장 많은 영향을 미치고, 다음으로 오존과 기온이 영향을 미친다는 것을 확인하였다. 세 영향인자를 구체적으로 분석하여 미세먼지 대응방안을 수립하고자 한다.



- 겨울이 되면 밤과 낮에 기온 차이가 커지면서 기온역전현상이 발생하고 공기가 상하로 순환되지 않아 이른바 ‘대기 정체’ 현상이 발생한다.
- 더불어 교통 수단 이용률과 도시가스 사용량이 증가하면서 대기중에 일산화탄소와 질소산화물, 휘발성유기화합물(VOCs) 농도가 증가한다.
- 이때 일산화탄소와 태양열에 의해 분해되어 생성된 오존이 대기 중에 순환되지 못해 정체되며 미세먼지 농도를 증가시킨다.

친환경 자동차 통행료 감면

최근 여러 회사에서 친환경 자동차에 미세먼지 감축 장치를 부착하기 위한 연구가 진행 중이다. 이에 따라 친환경 자동차 통행료 감면 정책을 도입하여 질소산화물 및 휘발성유기화합물(VOCs)의 배출을 최소화할 수 있도록 노력해야 한다.



공용 공간 내 스마트 공기청정시스템 구축

공용 공간 내 일산화탄소 배출량 센싱 시스템을 도입하여 일정 농도 이상이 감지될 시 공기청정기를 자동으로 작동할 수 있는 시스템을 구축해야 한다. 이를 통해 순환되기 힘든 실내 공기질을 제어하여 실내 미세먼지 문제를 해결할 수 있도록 한다.



감사합니다.