

기상위성 자료를 활용한 여름철 자외선 산출 기술 개발

참 가 번 호	220081	팀 명 ※ 반드시 참가신청 시 작성한 팀명	선크림
---------	--------	----------------------------	-----

1. 분석 데이터 설명

1-1 사용 데이터 기간

1) 오존전량 계절별 차이

- 3~5월까지 봄, 6~8까지 여름, 9~11월까지 가을, 12~2월까지 겨울로 데이터를 분류하여 계절별 오존전량의 차이를 분석한 결과, 오존전량은 한 계절 사이에 차이가 존재하며, 여름과 겨울 사이에는 차이가 없음

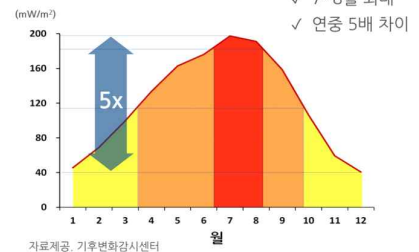
2) 오존전량과 자외선과의 관계

- 오존전량은 상공에 존재하는 오존분자의 총량을 의미
- 오존전량의 90%는 성층권에 존재하기 때문에 성층권 오존을 나타내는 값으로 사용
- 성층권오존은 오존층이라고도 하며, 지구대기로 입사하는 유해한 자외선을 흡수해 지상의 생태계를 보호

3) 자외선 B 계절별 차이

- 자외선 B는 연중변화가 크고 한 계절 사이에 차이가 존재
- 자외선 지수는 태양고도가 최대일 때 지표에 도달하는 자외선B의 복사량을 지수식으로 환산한 것
- 자외선과 오존전량을 봤을 때, 월별 차이가 존재하므로 여름철 자외선 지수를 예측하기 위해 예측데이터(6월)과 근접한 3월~8월까지의 데이터를 학습데이터로 사용

자외선B의 연중변화
(안면도, 2008-2012)



1-2 기존변수(uv와 연관성), 추가변수(풍속, 강수, 강수형태, 풍향, 기온, 습도)

1) 사용한 변수 리스트 (변수명, 출처, 간단한 설명, 기간)

- 사용 변수 기간은 2020년 03월~08월, 2021년 03월~08월, 2022년 06월로 동일

구분	사용 데이터	구분	사용 데이터
제공 데이터 (출처: 기상청)	uv.yyyymmdd	추가 데이터 (출처: 기상자료개방포털)	uv.esr
	uv.hhnn		fom (강수형태)
	uv.uv		precipitation (강수)
	uv.landtype		temperature (기온)
	uv.stn		humidity (습도)
	uv.band1~16		dir (풍향)
	uv.solarza		velocity (풍속)
	uv.sateza		

2) 변수와 자외선(uv) 간 연관성

2-1) 오존이 발생하기 쉬운 기상 조건

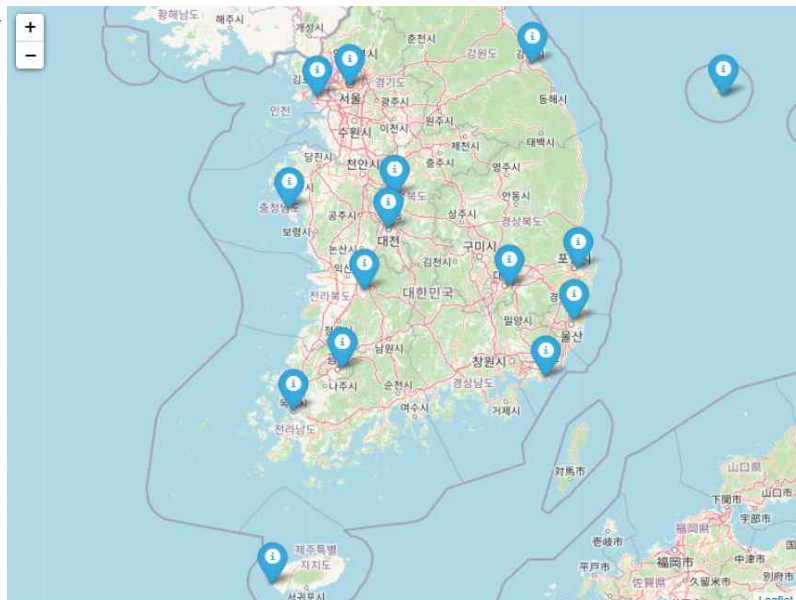
- 풍속의 경우, 지상의 평균풍속이 3.0 m/sec 미만으로 바람이 약할 때
→ velocity, dir 변수 추가
- 기온의 경우, 기온이 평년보다 높은 경우, 최고기온이 25 °C 이상으로 기온이 높을 때
→ temperature 변수 추가
- 날씨의 경우, 쾌청한 날씨가 계속될 경우
→ precipitation, form, humidity 변수 추가

2-2) 자외선에 영향을 주는 요인

- 대기 중의 구름, 수증기, 에어로졸
- 대기 중 기체들에 의한 복사량의 흡수와 산란과정
- 태양의 고도각, 날씨, 오존 전량, 대기오염도
- 오존층 두께가 1% 감소하면 지상 자외선 양은 약 1.5% 증가

2-3) 추가변수

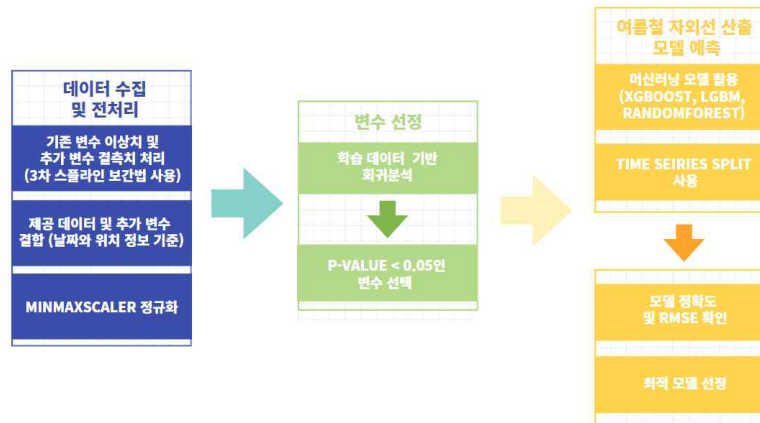
	stn	names	lon	lat
0	13	한경면	126.16000	33.30000
1	105	중앙동	128.89098	37.75147
2	108	사직동	126.96500	37.57100
3	112	동인천동	126.62490	37.47772
4	115	울릉읍	130.89864	37.48129
5	131	북대1동	127.44066	36.63924
6	132	안면읍	126.33000	36.54000
7	133	온천2동	127.37212	36.37200
8	138	송도동	129.37962	36.03259
9	143	효목1동	128.65295	35.87797
10	146	덕진동	127.11900	35.84080
11	152	병영2동	129.33472	35.58250
12	156	운암2동	126.89158	35.17294
13	159	대청동	129.03203	35.10468
14	165	연산동	126.38121	34.81689



- 제공된 uv 데이터 내 위도, 경도에 부합하는 행정구역에 대한 '풍속', '강수', '강수형태', '풍향', '기온', '습도' 데이터 수집
- 기간은 2020년 3~8월, 2021년 3~8월, 2022년 6월로 지정하여 수집

2. 분석 프로세스

2-1. 분석 프로세스



2-2. 분석 과정

1) 데이터 수집 및 전처리

1-1) 이상치 및 결측치 처리

- -999 이상값 및 시간 기준 추가 변수 내 결측값 처리
- 3차 spline 메소드 사용

1-2) 추가 데이터 수집 및 결합

- 제공 데이터와 추가 변수 간 outer join을 활용한 데이터 결합
- 기존 데이터와의 시간단위 차이로 결측치 3차 spline 메소드로 처리

1-3) 중복 변수 제거

- 의미가 동일한 변수를 중복으로 사용할 경우, 의도치 않은 가중치를 부여할 수 있음
- 따라서 stn(지역)별로 값이 동일한 lon(경도), lat(위도), height(관측높이) 변수 제거

1-4) 정규화

- 모든 데이터 포인트를 동일한 정도의 스케일(중요도)로 반영
- Min-Max Scaler (최소-최대 정규화) 사용

2) 데이터 분리

- 학습 데이터: 2020년 03월~2020년 08월, 2021년 03월 ~ 2021년 08월 (2021년 06월, 2021년 07월 제외)
- 검증 데이터: 2021년 06월, 2021년 07월
- 테스트 데이터: 2022년 06월

3) 회귀분석

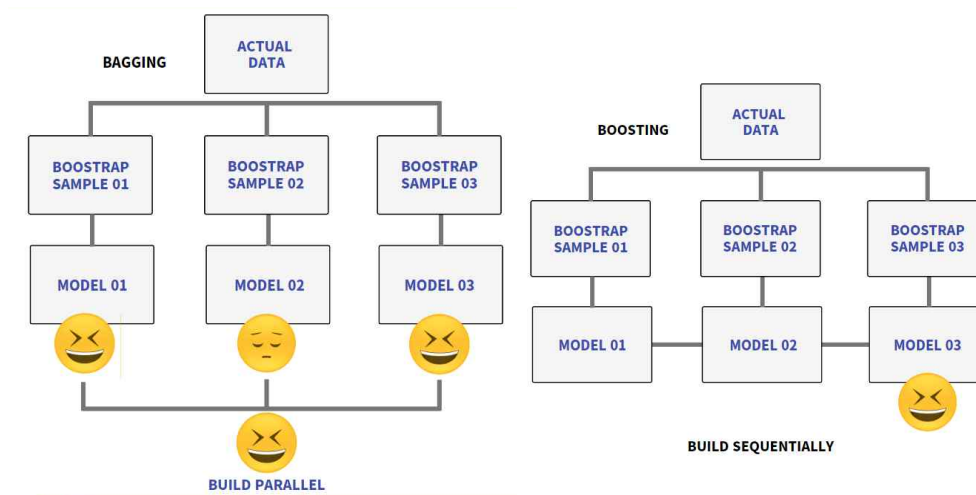
- 'yyymmdd' - 0.756, 'band15' - 0.362, 'precipitation' - 0.398, 'humidity' - 0.005, 그 외 변수 - 0.000
- 유의수준이 0.05일때 p-value가 유의수준보다 큰 변수는 회귀분석에서 유의하지 않은 변수이므로 'yyymmdd', 'band15', 'precipitation' 제거

[최종 학습 변수]

구분	사용 데이터	구분	사용 데이터
제공 데이터 (출처: 기상청)	uv.hhnn	추가 데이터 (출처: 기상자료개방포털)	uv.sateza
	uv.uv		uv.esr
	uv.landtype		fom (강수형태)
	uv.stn		temperature (기온)
	uv.band1~14		humidity (습도)
	uv.band16		dir (풍향)
	uv.solarza		velocity (풍속)

4) 활용 모델

양상블	부스팅	XGB Regression
		LightGBM Regression
	배깅	Random Forest Regression



- 봄과 여름에 해당하는 데이터(3~8월)만 사용하므로, 시계열 모형 대신 양상블 모형 사용
- 'uv.hhnn' 변수를 독립변수에 추가
- 리프 노드에 속한 데이터 값의 평균값을 구해 회귀 예측값을 계산

5) Time Series Split

- 시계열 분할 교차검증은 그 이름 대로 시계열 데이터를 사용한 분할법
- 미래와 과거의 데이터를 혼동한 훈련 세트로 학습을 실시하여, 모델이 각 시점에 어떤 일이 일어났는지 제대로 학습할 수 없다는 문제점 존재
- 고정된 시간 간격의 시계열 데이터 교차검증에 활용

3. 분석 결과

3-1. 모델 성능 평가결과 및 검증 결과

활용모델 및 검증값	모델 성능 평가결과 (RMSE)	202206 검증값 (RMSE)
XGBoost	0.0281	0.728071
Random Forest	0.0325	2.180233
LightGBM	0.0260	0.615103

3-2. 최종 선정 모델(rmse 가장 잘 나온 모델)

1) 파라미터

최종 모델(LightGBM) 사용 파라미터			
bagging_fraction	0.72	scale_pos_weight	1.5
num_iterations	20000	lambda_l1	0.1
max_depth	7	lambda_l2	0.35
min_data_in_leaf	8	early_stopping_rounds	300
learning_rate	0.03	n_splits	4
colsample_bytree	0.72		

2) RMSE

- 최종 선정 모델은 LightGBM 회귀 모델이며, 최종 검증값은 0.615103

4. 활용 방안

4-1. 자외선이 미치는 영향과 대안






1) 사람에 미치는 영향과 대안

1-1) 영향

- UV-A 노출 : 피부암 발생, 피부 멜라닌 침착
 - UV-B 노출 : 광각막염, 백내장, 익상편, 기후성각막증, 홍반 반응, 일광화상 반응
 - UV-C 노출 : 오존층에 흡수되어 위험성 없음
- '여름철 자외선 산출 모델'을 기반으로, 단계별 자외선지수 도출을 통한 대응 방안 수립

1-2) 대안

- 자외선 단계별 대응방안

단계	지수범위	대응요령
 위험	11 이상	<ul style="list-style-type: none"> ▶ 햇볕에 노출 시 수십 분 이내에도 피부 화상을 입을 수 있어 가장 위험함 ▶ 가능한 실내에 머물러야 함 ▶ 외출 시 긴 소매 옷, 모자, 선글라스 이용 ▶ 자외선 차단제를 정기적으로 발라야 함
 매우높음	8 이상 11 미만	<ul style="list-style-type: none"> ▶ 햇볕에 노출 시 수십 분 이내에도 피부 화상을 입을 수 있어 매우 위험함 ▶ 오전 10시부터 오후 3시까지 외출을 피하고 실내나 그늘에 머물러야 함 ▶ 외출 시 긴 소매 옷, 모자, 선글라스 이용 ▶ 자외선 차단제를 정기적으로 발라야 함
 높음	6 이상 8 미만	<ul style="list-style-type: none"> ▶ 햇볕에 노출 시 1~2시간 내에도 피부 화상을 입을 수 있어 위험함 ▶ 한낮에는 그늘에 머물러야 함 ▶ 외출 시 긴 소매 옷, 모자, 선글라스 이용 ▶ 자외선 차단제를 정기적으로 발라야 함
 보통	3 이상 6 미만	<ul style="list-style-type: none"> ▶ 2~3시간 내에도 햇볕에 노출 시에 피부 화상을 입을 수 있음 ▶ 모자, 선글라스 이용 ▶ 자외선 차단제를 발라야 함
 낮음	3 미만	<ul style="list-style-type: none"> ▶ 햇볕 노출에 대한 보호조치가 필요하지 않음 ▶ 그러나 햇볕에 민감한 피부를 가진 분은 자외선 차단제를 발라야 함

※ 의학자문 : 서울대학교병원운영 서울특별시 보라매병원 피부과 박현선 서울의대 교수

2) 식물에 미치는 영향과 대안

2-1) 영향

- 식물의 경우, 동물보다 공해에 민감하게 반응
- 농작물이나 유실수의 경우도 동일
- 산화작용에 따라, 엽록소의 파괴나 효소작용의 저하와 같은 문제가 발생할 수 있음

→ 식물의 성장 및 농작물로부터의 수확량 확보를 위한 대책 수립 가능

2-2) 대안

- 자외선(UV-B)의 차단제의 축적을 제어하는 단백질, UVR8와 COP1을 연구
- 식물의 세포와 여러분자(HY5, SPA 및 RUP)와 상호작용하여 자외선(UV-B)에 반응하여 자외선 차단제의 축적을 제거하는 신호를 보낼 수 있도록 함

[참고 문헌]

기상청 - 오존전량과 자외선 관계,

http://www.climate.go.kr/home/10_wiki/index.php/%EC%84%B1%EC%B8%B5%EA%B6%8C_%EC%98%A4%EC%A1%B4

보건환경연구원, <https://www.incheon.go.kr/ecopia/EC040303>

생활기상정보 - 자외서에 따른 사람 대안,

https://www.weather.go.kr/plus/life/jisudaymap_A07_1.jsp

웨더아이 - 자외선 지수, <https://www.weatheri.co.kr/forecast/forecast09.php>

오존전량, 자외선지수 예측모델의 평가 및 개선 - 이윤곤, 김준, 조희구, 최병철, 김지영, 박일수

인너셀프 - 자외선에 따른 식물 대안

<https://ko.innerself.com/%EC%83%9D%ED%99%9C/%EA%B0%80%EC%A0%95%EA%B3%BC-%EC%A0%95%EC%9B%90/%EC%9B%90%EC%98%88/12967-%EC%8B%9D%EB%AC%BC%EC%9D%B4-%ED%96%87%EB%B3%95%EC%97%90-%ED%83%80%EC%A7%80-%EC%95%8A%EB%8A%94-%EC%9D%B4%EC%9C%A0.html>

UV 지수 계산 방법, <https://ko.wukihow.com/wiki/Compute-the-UV-Index>