

서울시 어린이 교통사고 감소를 위한 빅데이터 및 AI 기반 新 어린이 보호구역 제안



팀명 : FOR THE CHILDREN



어린이
보호구역



01/ 프로젝트 소개

- 어린이 교통사고 상황
- 現 어린이 보호구역의 지정 현황
- 어린이 보호구역 지정 효과
- 新 어린이 보호구역 제안

02/ 데이터 분석

- 과정 및 개요
- 데이터 설명 및 전처리
- 주요 전처리 소개

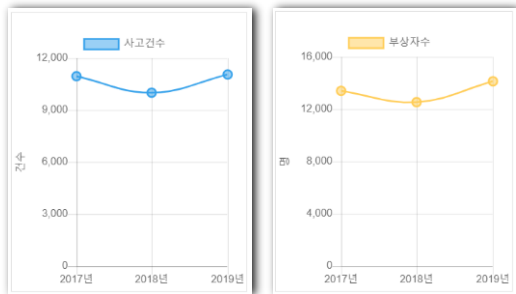
03/ 新 어린이 보호구역 선정 과정 및 결과

- 머신러닝 모델링 과정
- 현재 위험 사각 지대 파악
- 잠재적 위험 사각 지대 추정

04/ 결론 및 제언

- 활용 결과
- 프로젝트 기대 효과
- 참고 문헌 및 사용 도구

어린이 교통사고 건수 매년 상승



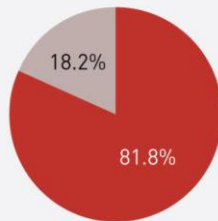
출처: TAAS 어린이 교통사고 분석 시스템

어린이 보호구역 미지정 구역 20%

스쿨존, 얼마나 지정됐나?

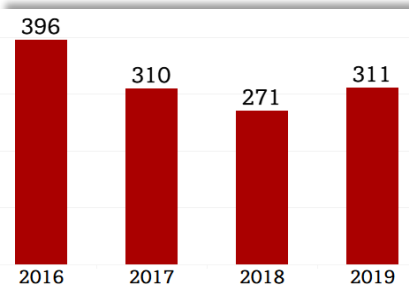
출처: 경찰청, 2019년 12월 기준

지정	16912
미지정	3771



출처: SBS

스쿨 존 제외한 사고 매년 250건 이상!



출처: TASS

어린이 교통 사고 건수는 매년 1만 건 이상 발생

2018년 대비 2019년 어린이 부상 사고 12% 상승

데이터에 근거한, 현재 어린이 보호구역 기준에 대한 심층 진단 필요!!
기존 어린이 보호 구역외 사각 지역에서 발생하는 사고에 대한 해결책은?

현재 어린이 보호구역 지정 현황은?

어린이보호구역 지정대상은 유치원, 초등학교, 특수학교, 어린이집, 학원, 외국인학교 또는 대안학교, 국제학교 등의 만 13세 미만 어린이시설 주변도로 일정구간을 보호구역으로 지정을 할 수 있다.

도로교통법 제 12조



어린이 공원 및 어린이가 자주 가는 장소라도 교육 기관이 아닌 이유로 어린이 보호구역에 포함되지 않음

지방경찰청장 또는 경찰서장이 제1항의 규정에 의한 보호구역지정신청을 받은 때에는 당해 국민학교등의 주 출입문을 중심으로 반경 300미터 이내의 도로중 일정구간을 보호구역으로 지정한다.

어린이보호구역의 지정 및 관리에 관한 규칙



어린이 교통사고 다발 지역인 경우라도 300미터에서 조금이라도 벗어난 경우 어린이 보호구역으로 포함되지 않음

어린이 보호구역 지정 효과

<참고 논문>

기본-RR-16-07

도로교통 안전사업의 효과 분석 및 제도적 개선방안

Effects and Improvements on Highway Safety Projects in Korea

임재경 · 설재훈 · 박정욱 · 성낙문 · 최재식 · 현진우

출처 : 도로교통공단



어린이 보호구역 지정으로 어린이 교통사고 13% 감소 효과

<어린이 보호구역 지정의 사고 감소 효과 및 경제성 분석>

구 분			사고감소 효과(%)	B/C	
직접적 효과	보호구역내 어린이 교통사고	교통사고 발생건수	13.23	분석기간 3년	0.30
		EPDO/개소	4.30	분석기간 5년	0.48
				분석기간 10년	0.85
직·간접적 효과	보호구역내 모든 교통사고	교통사고 발생건수	4.56	분석기간 3년	3.36
		EPDO/개소	1.23	분석기간 5년	5.31
				분석기간 10년	9.38

어린이 보호구역 지정 이후

보호구역 내 전체 교통사고 감소에 의한

B/C는 3.36~9.38로 어린이 보호구역 지정에 따른

매우 높은 경제적 타당성을 나타냄

新 어린이 보호구역 제안

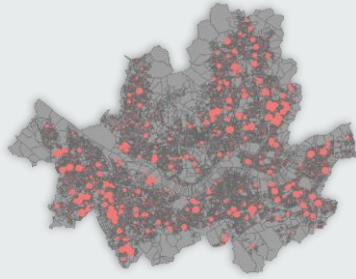
01.프로젝트 소개

- 빅데이터 및 AI에 기반한, 新 어린이 보호 구역이 필요한 이유

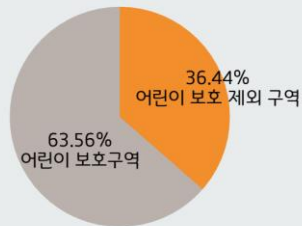
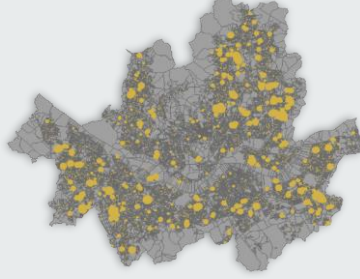
10대 유동인구 밀집 지역의
어린이 교통 사고 현황



어린이 교통 사고 주변의
초등학교 개수



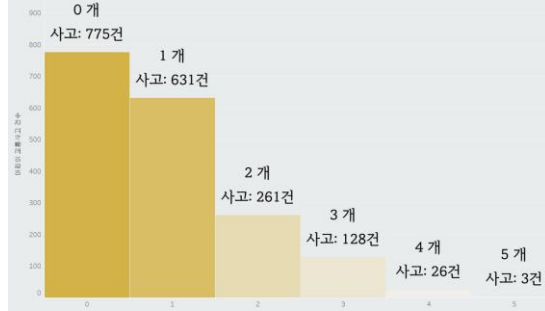
어린이 교통 사고 주변의
유치원 개수



10대 유동인구 밀집지역의 36%는
어린이 보호구역으로 지정되지 않음



근처 초등학교가 없는 곳에서도
어린이 교통사고 1,066 건 발생



근처 유치원이 없는 곳에서도
어린이 교통사고 755 건 발생

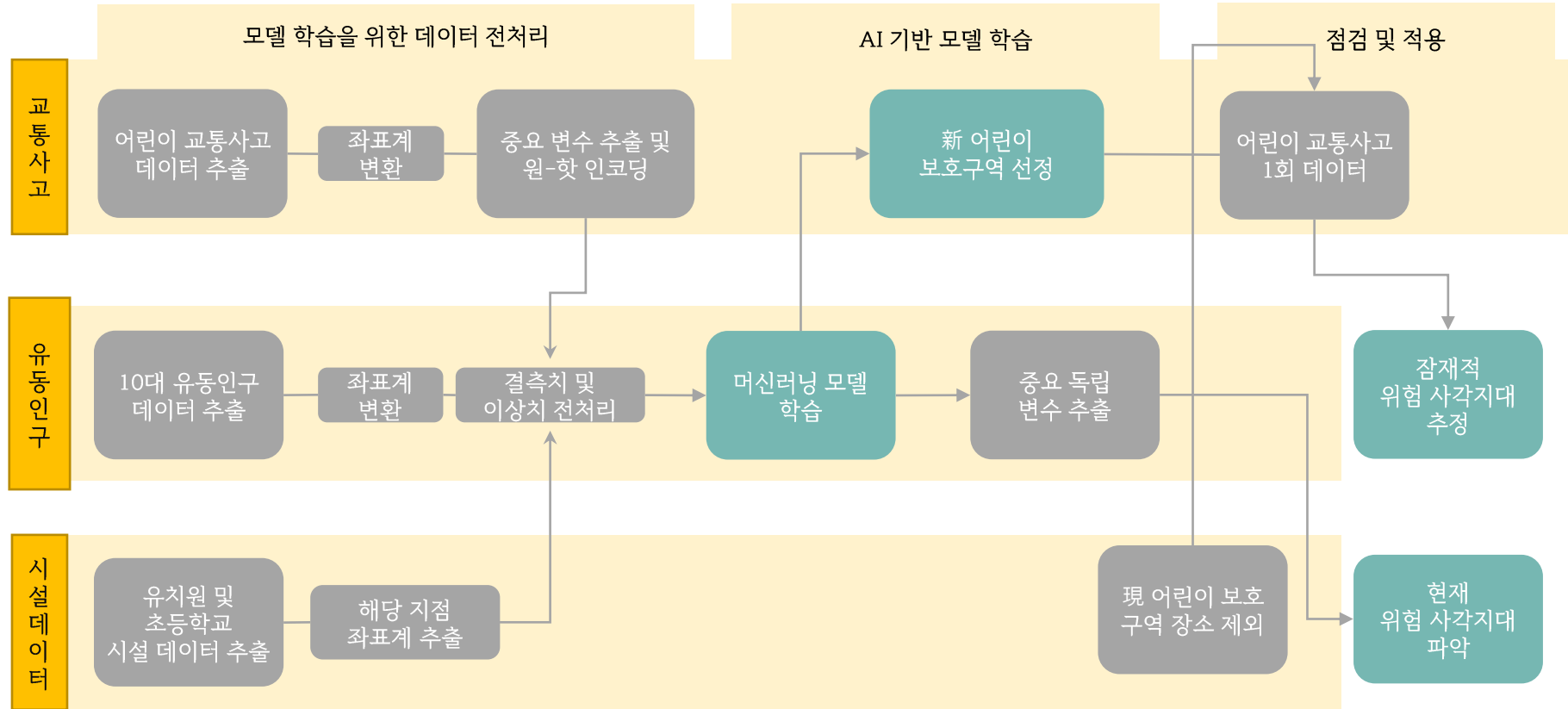


AI 기반 新 어린이 보호구역의 장점

어린이 교통사고에 주요 영향을 미치는 요인 자동 탐색

현재 위험에 노출 되었지만 어린이 보호구역으로 지정되지 않은 장소 찾기 가능 - 현황 분석

학습한 데이터를 바탕으로 미래의 잠재적 사고 위험 지역 찾기 가능 - 예측 분석



데이터 수집	데이터 용도	데이터 추출	데이터 추출 값 기준	데이터 전처리 주요 방법
서울시 TASS 교통 사고 (2014~2019)	어린이 교통사고 데이터를 통해 사고 장소 및 사고 변수들에 대해 분석할 수 있음	부상 정도	경상, 부상, 중상, 사망	원-핫 인코딩
		사고 발생 시간	00시간 ~ 24시간	결측치 처리
		날씨	맑음, 흐림, 눈 등	원-핫 인코딩
		피해자 연령	12세 이하	결측치 처리
		사고 장소 (X,Y 좌표)	해당 장소 반경 사고 횟수	경도, 위도 변환 및 반경 버퍼 (300M, 100M) 거리 계산
		도로 형태	교차로, 단일로 등	원-핫 인코딩
		차량 유형	승용, 화물, 자전거 등	원-핫 인코딩
서울시 어린이 보호구역	어린이 비보호 구역인 곳을 필터링하여 파악할 수 있음	X, Y 좌표	장소: 서울시	경도, 위도 변환
서울시 초등학교 현황	사고 지역 주변의 8세 이상 어린이의 주요 이동 장소를 파악할 수 있음	X, Y 좌표	장소: 서울시 공립 및 사립 모두 포함	경도, 위도 변환
서울시 유치원 현황	사고 지역 주변의 7세 이하 어린이의 주요 이동 장소를 파악할 수 있음	X, Y 좌표	장소: 유치원 공립 및 사립 모두 포함	경도, 위도 변환
SKT 유동 인구	10대 유동인구의 주요 밀집 지역을 분석할 수 있음	X, Y 좌표	서울시 전체 10대 유동 인구	경도, 위도 변환 유동 인구 수 EDA 및 결측치/ 이상치 처리
		유동 인구 수		

주요 전처리 소개

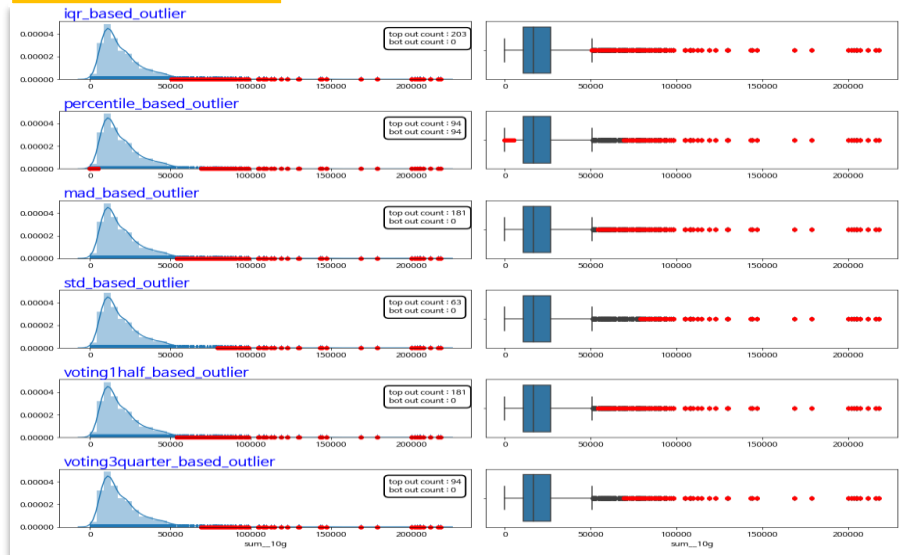
데이터 별 outlier 처리

1. Iqr(Inter Quartile Range) based
2. Percentile based
3. Mad(Median Absolute Deviation) based
4. Std (Standard Deviation) based
5. Voting based
 - I. 4가지 방식 중 2개 이상 선택된 데이터
 - II. 4가지 방식 중 3개 이상 선택된 데이터

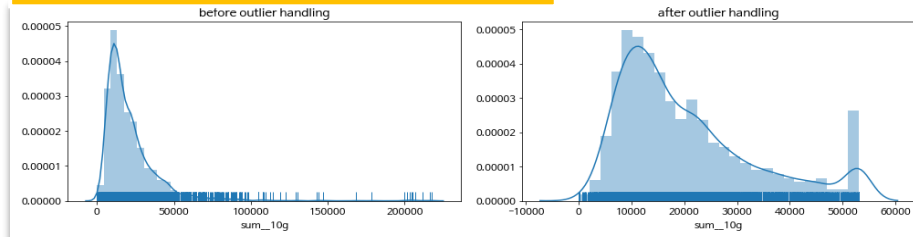


데이터 분포에 따라 평균값 혹은 중앙값으로 outlier 대체

결측치 처리 프로세스



10대 유동인구 데이터 결측치 처리 결과 전, 후



1. 머신러닝 회귀 모델?

2. 종속변수 선정

3. 독립 변수 선정

4. 최적 모델 선정

5. 모델 의사 결정 과정

1. 머신러닝 회귀 모델?

- 정의: 분류와 달리 연속된 값을 예측하는 것을 회귀(regression)라고 하고, 하나 혹은 그 이상의 독립 변수(x)들이 종속변수(y)에 미치는 영향을 추정하는 알고리즘으로 이를 이용하여 추론 및 예측 가능

사용한 머신러닝 회귀 모델 종류 (8가지)

Linear Lasso
Ridge
Elastic Net
Linear SVR

Stepwise linear
Random Forest
Ada boost
Gradient Boosting XGB



모델 선택 기준 - MAE(Mean Absolute Error)

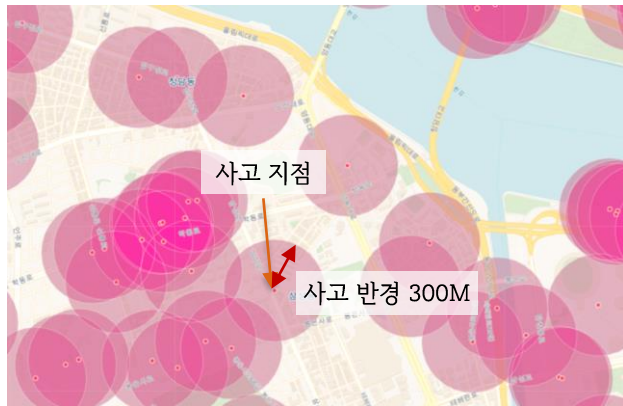
y_i = 실제 교통 사고 횟수

$h(x_i)$ = regressor를 통해 예측한 교통 사고 횟수

$$MAE = \frac{1}{n} \sum_{i=1}^N |y_i - h(x_i)|$$

2. 종속 변수 선정

- 정의: 모델을 통해 학습 및 예측하고자 하는 어린이 교통사고 수
- 기준: 어린이 교통 사고가 1번 이상 일어난 장소 (경도, 위도)
- 범위: 해당 장소 및 반경 300M이내에서 6년간 일어난 어린이 사고 횟수



머신러닝 모델링 과정

03. 新 어린이 보호구역 선정 과정 및 결과

- 1. 머신러닝 모델링

1. 머신러닝 회귀 모델?

2. 종속변수 선정

3. 독립 변수 선정

4. 최적 모델 선정

5. 모델 의사 결정 과정

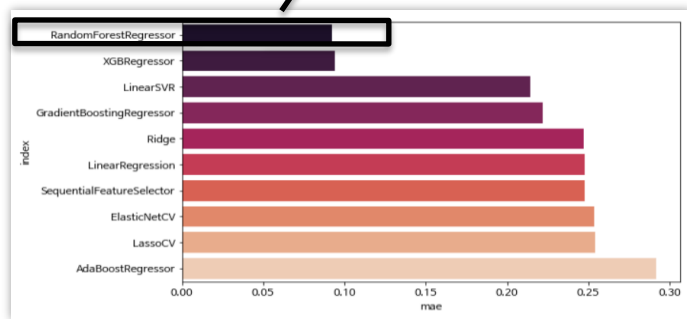
3. 독립 변수 선정

독립 변수	변수 값 설명
초등학교 수	교통 사고 반경 400미터 이내 수
유치원 수	교통 사고 반경 400미터 이내 수
유동인구 수	교통 사고 반경 100미터 이내 수
...	...

전체 독립 변수 = 56개

4. 최적 모델 선정

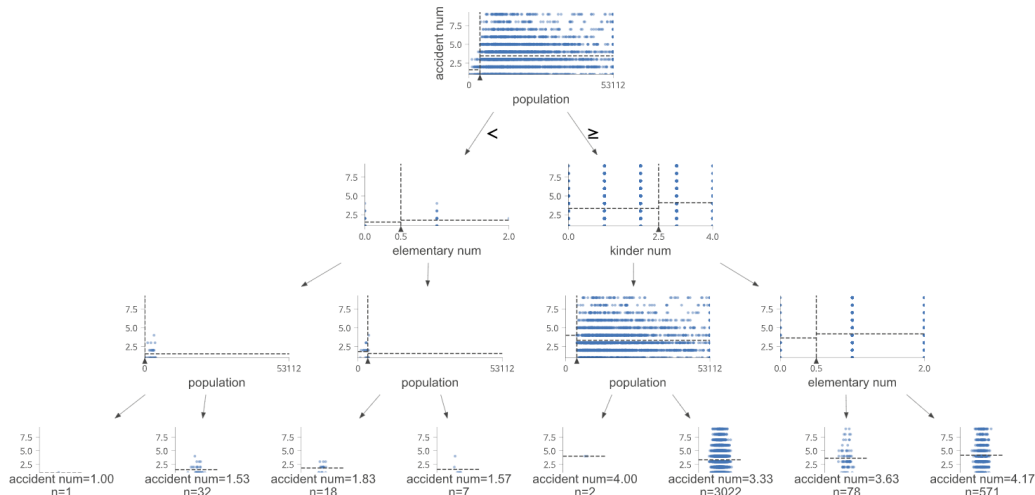
모델별 MAE(Mean Absolute Error)를 산출, 가장 작은 값을 가진 모델인 RandomForestRegressor를 최적 모델로 선정



5. 모델 의사 결정 과정의 이해

머신러닝 모델을 이용해, 어린이 보호구역 선정 기준을 새롭게 정의할 수 있음

머신러닝 모델의 의사 결정 과정은 아래와 같이 시각화 가능



RandomForestRegressor는 여러 decision tree의 앙상블 모델로, 위 그림은 하나의 decision tree를 시각화한 것임.

현재 위험 사각 지대 파악

03. 新 어린이 보호구역 선정 과정 및 결과 - 2. 현황 분석

1. 머신러닝 모델링

2. 현황 분석

3. 예측 분석

STEP 1

머신러닝 모델 결과 중요 변수 추출



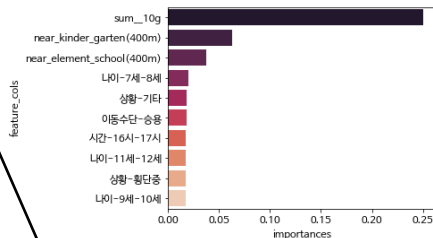
필터링에
사용될 중요 변수

STEP 2

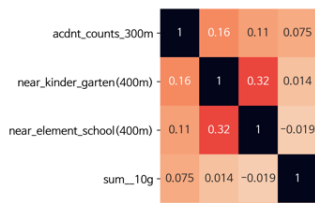
중요 변수 기반 필터링

현재 어린이 보호구역 사각지대 파악

모델의 특성 분석 및 종속 변수와의 상관 관계 분석으로 중요 변수 파악



Feature Importances



Pearson Correlation

10대 유동 인구

주변 유치원 수

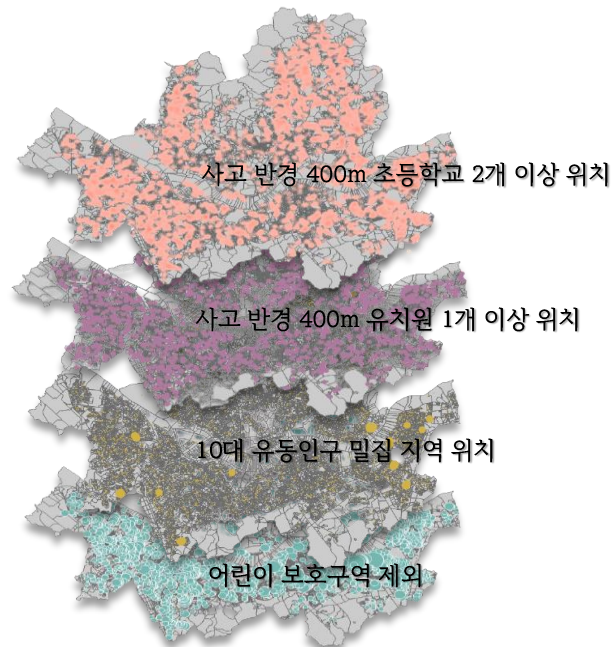
주변 초등학교 수

머신 러닝 예측 기법을 도입하여 새로운 어린이 보호구역 지정 기준 방안 제시

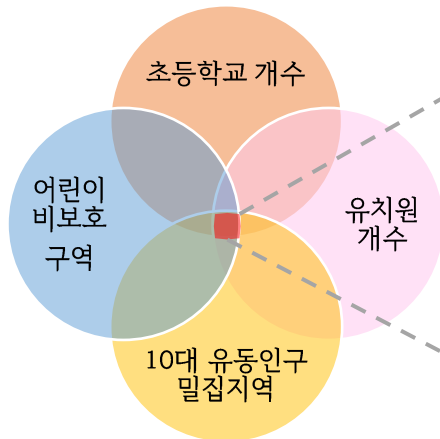
어린이 보호구역 지정을 통한 어린이 교통 사고 감소 효과 기대!

현재 위험 사각 지대 파악

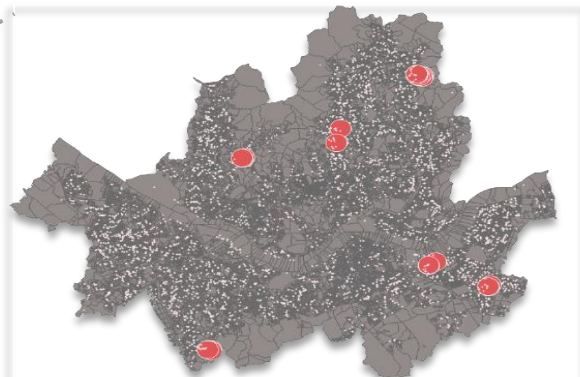
- 중요 변수 기반 필터링



AI 중요 변수 기반 필터링



어린이 보호구역 사각지대 확인



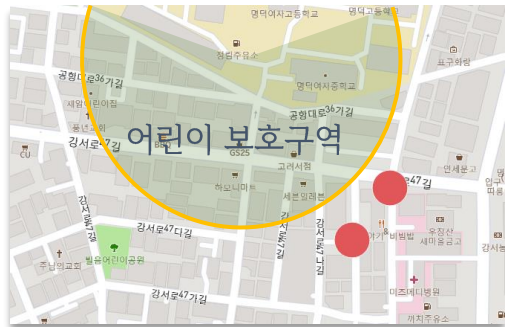
중요도	변수 이름	기준값*
1	10대 유동인구	41,366.97 (명)
2	주변 유치원 수	1 (개)
3	주변 초등학교 수	2 (개)

어린이 교통사고 다발 지역 중요 특징들을 머신 러닝을 통해 도출하여
현재 어린이 보호구역 기준으로 제외되었지만 반드시 지정되어야 하는 장소 제시

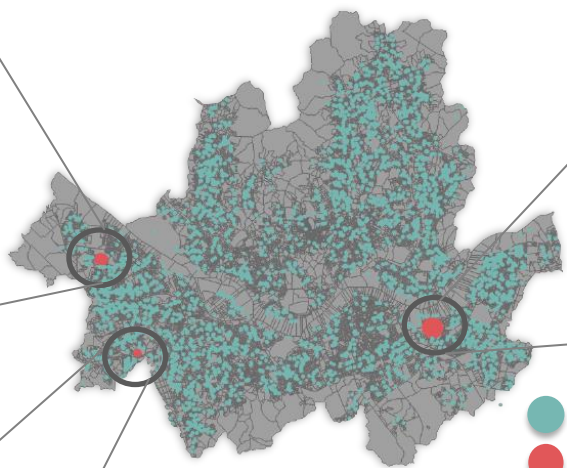
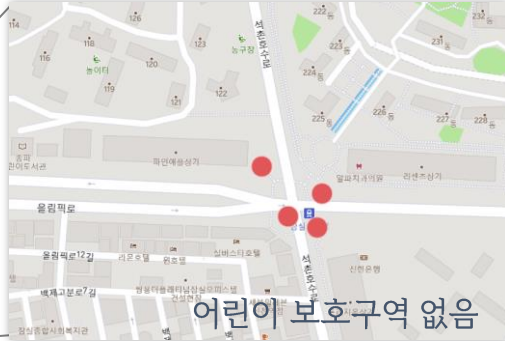
현재 위험 사각 지대 파악

03. 新 어린이 보호구역 선정 과정 및 결과 - 2. 현황 분석

2위 - 내발산동 강서로 47길

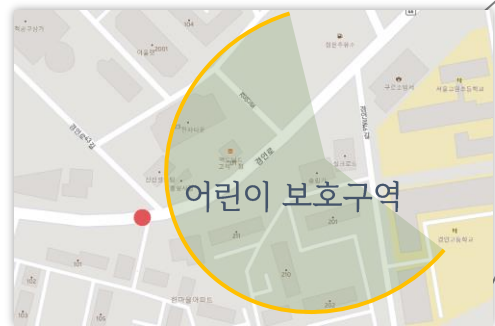


1위 - 석촌 호수로 잠실새내역 사거리



- 기준값 미만인 사고 지점
- 기준값 이상인 사고 지점

3위 - 구로구 개봉동 179-46



중요 변수 필터링 후 어린이 보호구역 지정이 시급한 TOP3 지역!
잠실새내역 사거리, 강서로 47길, 개봉동 176-46

현재 위험 사각 지대 파악

03. 新 어린이 보호구역 선정 과정 및 결과 - 2. 현황 분석

2위 - 내발산동 강서로 47길



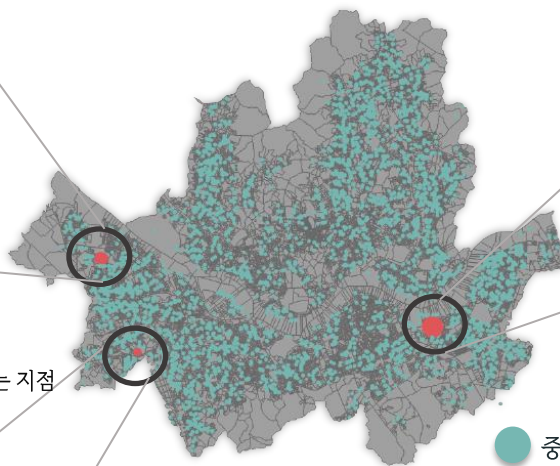
선정 후보군 8곳 중 3곳 이상 모여 있는 지점이며
어린이 교통사고 5건 이상 발생!!

어린이 보호 구역과 근접하지만 보호를 받지 못하고 있는 지점

3위 - 구로구 개봉동 179-46



10대 유동 인구 상위 10%에 속하는 밀집 지역이며
반경 400m 이내 유치원 및 초등학교 수 2 곳 이상 위치
어린이 보호 구역과 근접하지만 보호를 받지 못하고 있는 지점



● 중요 변수 기준치 미만인 사고 지점
● 중요 변수 기준치 이상인 사고 지점

1위 - 석촌 호수로 잠실새내역 사거리



선정 후보군 8곳 중 4곳 이상 모여 있는 지점이며
어린이 교통사고 5건 이상 발생

반경 400m 이내 유치원 및 초등학교 수 2 곳 이상 위치

잠재적 위험 사각 지대 추정

03. 新 어린이 보호구역 선정 과정 및 결과 - 3. 예측 분석

1. 머신러닝 모델링

2. 현황 분석

3. 예측 분석

STEP 1

Test Set - 어린이 교통사고 1건 일어난 장소 데이터 추출



STEP 2

머신 러닝 회귀 모델을 통해 해당 지점 미래 사고 건수 예측

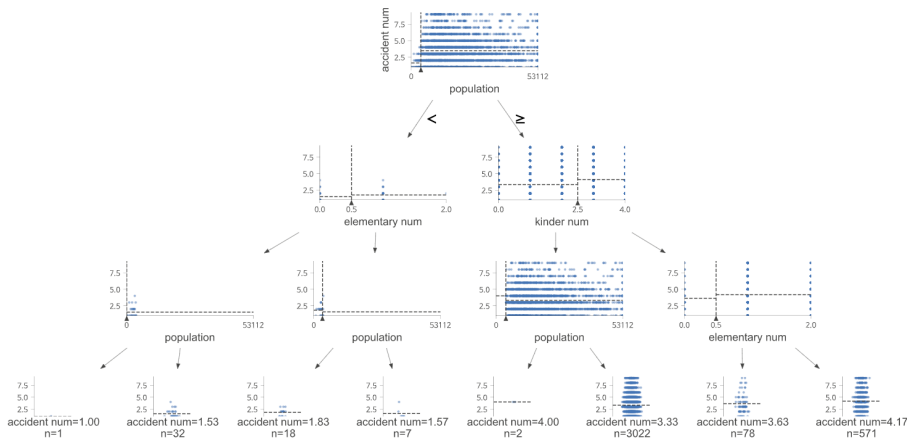


STEP 3

미래 사고 추정 건수 1건 이상 장소!



잠재적 어린이 보호구역 사각지대 추정



머신 러닝 예측 기법을 도입하여 새로운 어린이 보호구역 지정 기준 방안 제시

어린이 보호구역 지정을 통한 어린이 교통 사고 감소 효과 기대!

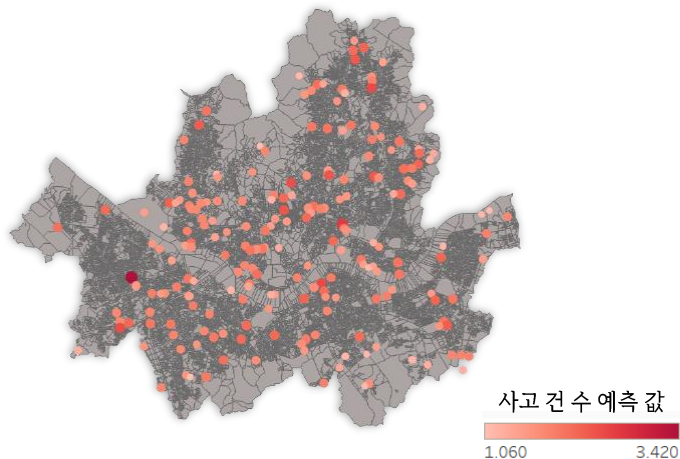
잠재적 위험 사각 지대 추정

머신러닝 모델 기반, 미래 어린이 사고 추정 건 수 파악

사고 건수 예측 값 확인

id	pred	datetime	legaldong_name	Ldong_NM	x_w84	y_w84	sum_10g
1	3.42	2015-06-01 00:00:00	서울특별시	목동	126.8644943	37.5262836	53111.75
2	2.7	2018-11-03 00:00:00	서울특별시	행당동	127.0377623	37.561151	53111.75
3	2.51	2016-05-04 00:00:00	서울특별시	압구정동	127.0200064	37.5226195	35370.19
4	2.49	2019-06-30 00:00:00	서울특별시	상계동	127.0615162	37.6500961	25156.02
5	2.47	2016-05-19 00:00:00	서울특별시	개봉동	126.8553058	37.4932715	50581.44
6	2.44	2014-06-18 00:00:00	서울특별시	명륜3가	126.9955948	37.5883072	9257.77
7	2.36	2015-05-24 00:00:00	서울특별시	안암동5가	127.0265979	37.593393	9815.79
8	2.36	2015-07-29 00:00:00	서울특별시	종로3가	126.9896119	37.5704168	53111.75
9	2.34	2015-04-25 00:00:00	서울특별시	방학동	127.0483325	37.6687671	14769.03

사고 건수 예측 값



03. 新 어린이 보호구역 선정 과정 및 결과

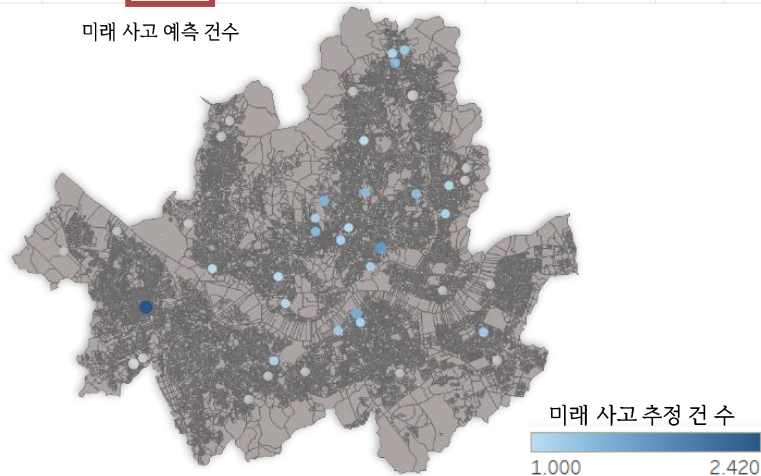
- 3. 예측 분석

유의미한 사고 건 수 기준 = 1 건 이상

미래 사고 추정 건 수 = 사고 예측 값 - 과거 교통사고 건 수 (1건)

id	pred	pred-1	datetime	legaldong_name	Ldong_NM	x_w84	y_w84	sum_10g
1	3.42	2.42	2015-06-01 00:00:00	서울특별시	목동	126.8644943	37.526284	53111.75
2	2.7	1.7	2018-11-03 00:00:00	서울특별시	행당동	127.0377623	37.561151	53111.75
3	2.51	1.51	2016-05-04 00:00:00	서울특별시	압구정동	127.0200064	37.522619	35370.19
4	2.49	1.49	2019-06-30 00:00:00	서울특별시	상계동	127.0615162	37.650096	25156.02
5	2.47	1.47	2016-05-19 00:00:00	서울특별시	개봉동	126.8553058	37.493271	50581.44
6	2.44	1.44	2014-06-18 00:00:00	서울특별시	명륜3가	126.9955948	37.588307	9257.77
7	2.36	1.36	2015-05-24 00:00:00	서울특별시	안암동5가	127.0265979	37.593393	9815.79
8	2.36	1.36	2015-07-29 00:00:00	서울특별시	종로3가	126.9896119	37.570417	53111.75

미래 사고 예측 건수



잠재적 위험 사각 지대 추정

03. 新 어린이 보호구역 선정 과정 및 결과

- 3. 예측 분석

미래 어린이 사고 추정 건 수 기준, top3 지역 추출

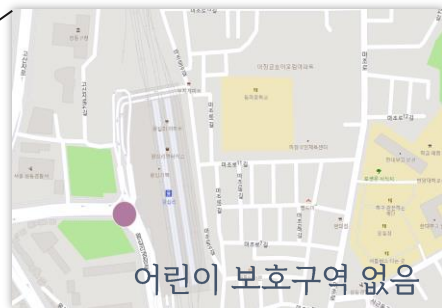
1위 - 양천구 목동 926-3



미래 사고 추정 건 수는? **2.42건**

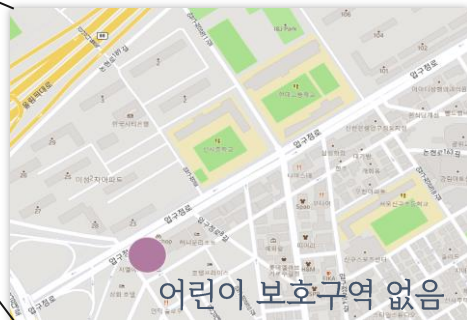
● 테스트 데이터
● 예측 값 TOP 3 지역

2위 - 성동구 행당1동 246



미래 사고 추정 건 수는? **1.7건**

3위 - 강남구 신사동 667-35



미래 사고 추정 건 수는? **1.51건**

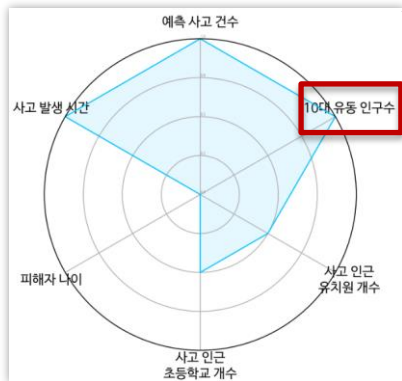
머신 러닝 모델이 예측한 미래 사고 건수 중 가장 높은 TOP3 지역!
양천구 목동 926 - 3, 성동구 행당1동 246, 강남구 신사동 667-35

잠재적 위험 사각 지대 추정

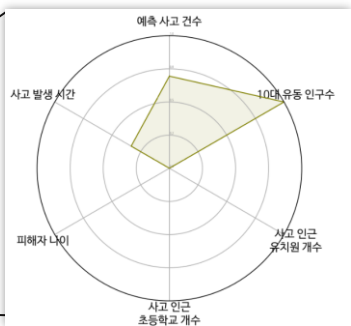
03. 新 어린이 보호구역 선정 과정 및 결과 - 3. 예측 분석

잠재 위험 지역별 중요 변수 분석

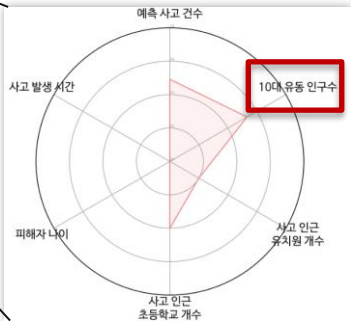
1위 - 양천구 목동 926-3



2위 - 성동구 행당1동 246



3위 - 강남구 신사동 667-35



● 테스트 데이터
● 예측 값 TOP 3 지점

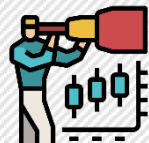
어린이 보호 추천 구역 결정의 최대 영향 변수는
10대 유동인구, 사고 발생시간, 사고 인근 초등학교 개수

“新 어린이 보호구역 실제로 어떻게 활용될까?”

현재 어린이 보호구역 지정이 시급한 장소 찾기
8곳 (사고 횟수 5번)



잠재적 어린이 보호구역 지정이 시급한 장소 찾기
216곳 중 43곳은 추후 사고 발생 확률 ▲



어린이 보호구역 반경 확대 시
50m당 어린이 비보호구역 평균 34% 감소



新 어린이 보호구역 활용에 따른 효과

- 1 기존 어린이 보호구역 기준이 가진 단점을 보완하여 서울시 어린이 교통사고 예방에 큰 도움을 주면서 어린이 보호구역 대표 모델이 될 수 있음
- 2 막연한 어린이 보호구역 지정이 아닌 데이터에 근거한 지정으로, 정책 시행시 타당성 확보와 효과 제고
- 3 AI 기반 모델의 특성상 지속 가능한 활용으로 추후 발생 가능한 사각지대도 추정 가능하며 데이터 축적에 따른 AI 모델의 보완 및 업데이트 가능
- 4 서울시 어린이 보호구역 개선 사업에서 cctv를 적재적소에 설치하는 데 이용 가능

참고 문헌

- 2011. 어린이보호구역 설치,운영의 문제점 및 개선방안, 한국공안행정학회
- 2016. 도로교통 안전사업의 효과 분석 및 제도적 개선방안, 한국교통연구원
- 2020. 어린이 보호구역내 교통사고 예방을 위한 제도 개선방안에 관한 연구. 경희법학연구소
- 2020. 어린이 교통 사망사고에 대한 현황 분석, 한국민간경비학회
- 2017. 어린이보호구역내 어린이 교통사고 발생에 미치는 영향요인 분석. 한국도로학회
- 2018. DEA를 활용한 어린이보호구역개선사업의 효율성 평가에 관한 연구. 대한토목학회
- 2015. ‘어린이 교통사고 예방을 위한 교통안전시설 연구’, 대한교통학회
- 1999. ‘초등학교 저학년 아동들의 안전사고 발생 실태 및 관련요인 분석’, 대한간호학회지
- 2008. 계층분석법(AHP)을 이용한 어린이보호구역 지정기준에 관한 연구, 대한교통학회
- 2018. 어린이보호구역 내 교통사고 다발지점의 교통안전시설에 관한 비교 연구, 한국융합과학회

사용 도구



사용 데이터

	데이터셋 목록	활용 목적	출처
1	서울시 TASS 교통 사고 데이터	최근 6년간 서울시에 발생한 어린이 교통 사고 데이터 분석	TASS 교통사고분석 시스템
2	전국 어린이 보호구역 데이터	사고 발생 지역의 보호구역 유무 파악	공공데이터 포털
3	서울시 초등학교 현황	사고 발생 지역 인근 초등학교 개수 파악	서울 열린데이터 광장
4	서울시 유치원 현황	사고 발생 지역 인근 유치원 개수 파악	서울 열린데이터 광장
5	서울시 50m간격 월별 SKT 유동인구	사고 발생 지역 인근 10대 유동 인구 파악	서울특별시 빅데이터 캠퍼스

THANK YOU
FOR LISTENING



어린이
보호구역

