

광진구 빅데이터 분석 공모전 -데이터 기반으로 '사용자 중심' 전기차 충전 인프라 입지 선정

알파브레인 2023. 05.

0. 목차



01 분석배경 02

데이터 수집 및 전처리 03

모델링을 활용한 데이터 분석 04

실증적 자료를 활용한 데이터 분석

05

모델링과 실증적 자료의 결합 06

결과해석

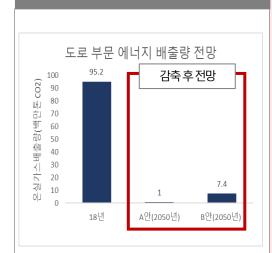
07

시사점



기후변화로 인한 정부의 2050 탄소중립 시나리오로 전기차에 대한 수요가 빠르게 증가하고 있으며, 광진구 또한 전기차의 보급률이 빠르게 증가하고 있음

정부 2050 탄소중립 시나리오



- 현재 95.2백만톤에 해당하는 도로 부문 에너지 배출량을 2050년 탄소 중립 시나리오로 7.4백만톤 이하로 감축할 것을 계획하고 있음
- 무탄소 교통강화, 내연기관차 이용 억제 등의 정책을 활용해 정부 2050 탄소중립 시나리오를 이행해 나갈 계획

자동차 회사들의 탄소 중립비전

현대차 탄소중립 로드맵 2023년 수소전기차 RV라인업확대 2030년 제네시스 전동화 비중 100% 2035년 유럽지역 전동화 100% 2040년 한국 및 주요시장 전동화 100% 2045년 탄소중립 달성

- •현대차에서는 2040년까지 한국 및 주요시장에서 생산되는 자동차를 전면 전동화 자동차로 생산해 나갈 계획임
- 앞으로 도로 부문의 시장에서 내연기관 자동차의 축소와 전동화 자동차의 확대의 흐름은 더 가속화 될 것으로 전망됨

전기차 수요의 폭발적인 성장



 ARIMA 분석기법을 적용해 자체 분석한 결과 광진구 전기차 등록현황을 예측한 결과 2022년에서 2030년까지 큰 폭으로 전기차 등록이 예상됨



하지만 광진구의 전기차 인프라 구축은, 서울시 다른 지역에 비해 더디게 구축되고 있으며, 빠르게 증가하고 있는 전기차 수요에 맞춰 더 많은 인프라 구축이 요구됨

전기차 충전기 인프라 개요



서울시 구별 전기차대수 및 충전기 개수와 비율 현황

시군구별	연료별	전기차대수	충전기	전기차 1대당 충전기 비율	순위	
성북구	전기	1482	2125	143.39%	1	
금천구	전기	1033	1232	119.26%	2	
중랑구	전기	1457	1658	113.80%	3	
동작구	전기	1203	1343	111.64%	4	
노원구	전기	1561	1728	110.70%	5	
마포구	전기	1776	1963	110.53%	6	
은평구	전기	1533	1685	109.92%	7	
성동구	전기	2239	2323	103.75%	8	
동대문구	전기	1112	1105	99.37%	9	
용산구	전기	1201	1147	95.50%	10	
종로구	전기	907	854	94.16%	11	
서대문구	전기	1221	1117	91.48%	12	
송파구	전기	3273	2799	85.52%	13	
도봉구	전기	1093	893	81.70%	14	
강북구	전기	890	704	79.10%	15	
강서구	전기	2867	2201	76.77%	16	
광진구	전기	1069	751	70.25%	17	
영등포구	전기	2182	1525	69.89%	18	비율
양천구	전기	1763	1229	69.71%	19	비슷
강동구	전기	2271	1544	67.99%	20	
관악구	전기	1355	830	61.25%	21	
서초구	전기	4757	2695	56.65%	22	상업
중구	전기	2152	1190	55.30%	23	시설
구로구	전기	5161	1833	35.52%	24	발전
강남구	전기	14026	3000	21.39%	25	지역

"현재 급속충전기와 완속충전기 모두 적정설치대수를 만족하고 있지만, 전기차 수요 증가를 고려할 때 다른 지역에 비해 전기차 인프라 구축이 늦은 상태"



이에 2025년까지 광진구에 추가로 설치될 2000여개의 전기차 충전기를 '사용자 중심 충전인 프라'를 구축한다면 더 효과적인 정책의 시행이 가능할 것임

전기차협회 "전기차 충전, 제조사보다 사용자 중심 개편해야"



광진구가 전기차 이용자의 지속적인 증가에 따른 '전기자동차 충전인프라 확충 계획'을 수립하고, 충전기 확대에 발 벗고 나섰다.





'사용자 중심 충전 인프라 가 구축된다면 정책의 효과를 극대화 할 수 있음



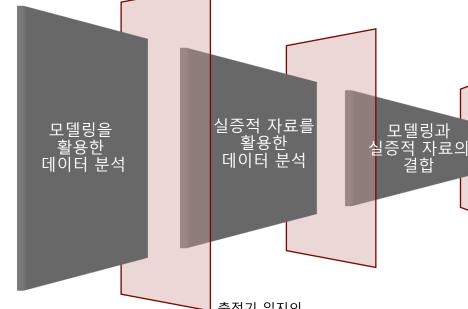
'사용자 중심의 전기차 충전기 입지'를 분석하기 위해서 본 팀은 데이터 분석과 실증적 분석을 결합해 최적의 입지를 도출하였으며 분석 로드맵은 다음과 같음

Description

여러 모델의 성능을 비교한 후, 최적의 성능을 보이는 LightGBM 모델을 활용해 전기차충전소가 입지하지 않았지만 입지가 필요한 지역 도출

급속 충전기의 입지와 완속 충전기의 입지에 영향을 미치는 유의한 변수를 활용해 충전기별 입지가 필요한 장소 도출 LightGBM 모델의 예측값, 예측한 자료의 실질적인 전기차 충전기의 수요는 검증 많지만 공급이 부족한 지역을 산정해 최적의 입지 도출

사용자 중심의 광진구 전기차 충전기 입지



최적의 입지도출 및 향후 전기차 충전 시설의 기회 및 리스크 파악

Output Result

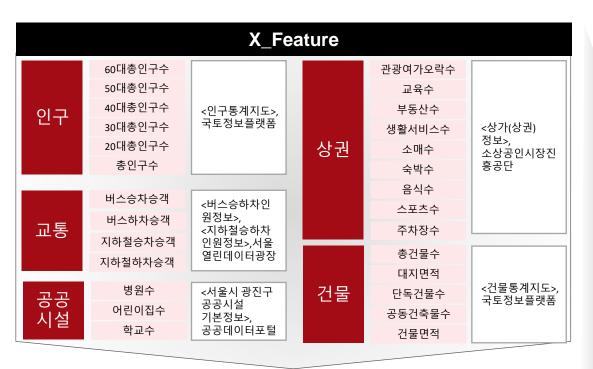
모델의 예측값을 활용해 현재 전기차 충전기가 없지만 충전기 입지가 필요한 지역 398개 도출 충전기 입지의 시급성이 떨어지는 지역을 제외하고 완속 충전기의 입지가 필요한 지역 199곳 급속 충전기의 입지가 필요한 지역 113곳 도출

2*2 Matrix를 활용해 급속충전기의 입지가 최우선적으로 필요한 지역 33곳 완속충전기의 입지가 최우선적으로 필요한 지역 55곳 도출

2. 데이터 수집 및 전처리



여러 데이터를 가공해서 Qgis와 파이썬을 활용해 광진구를 100m * 100m의 격자로 쪼개고 해당 격자 안에 관련 DATA 수치를 매핑하는 식으로 접근하였음







최적의 입지를 정밀하게 도출하기 위해서 광진구를 100m*100m의 격자로 쪼개서 총 1933개를 도출하고, 수집한 데이터를 각각의 격자에 할당

2. 데이터 수집 및 전처리



특히, 다음의 데이터 전처리 과정을 통해 모델에 학습되는 파라미터 조정을 통해 최종_광진구_데이터셋 .CSV 라는 데이터 프레임 도출

100m*100m <u>그리드</u> 속 매핑

Description

예시

- 정확한 입지 선정을 위해
 좌표데이터와 QGIS를 활용해 1800개의
 그리드 생성
- 주소데이터를 geocoding을 활용해 좌표 데이터로 변환
- EPSG 4326 -> EPSG 5181 변환해 Qgis에 매핑

X_Feature값 파생변수생성

- 데이터 불균형 문제를 해소하고, 필요한 정보를 반영하기 위한 Feature 값 파생변수 생성
- 총인구 남성 비율 / 여성 비율 생성을 통한 데이터 불균형 문제 해소
- 승하차승객수, 그리드당 전기차 기댓값

전처리 Y_Target값 파생변수생성

데이터

- Classification의 문제로 접근해 모델링을 진행하기 위해 종류별 전기차 충전기 유무의 파생변수 생성
- 전기차충전기 유무, 급속충전기 유무,
 완속충전기 유무등의 판생변수 생성

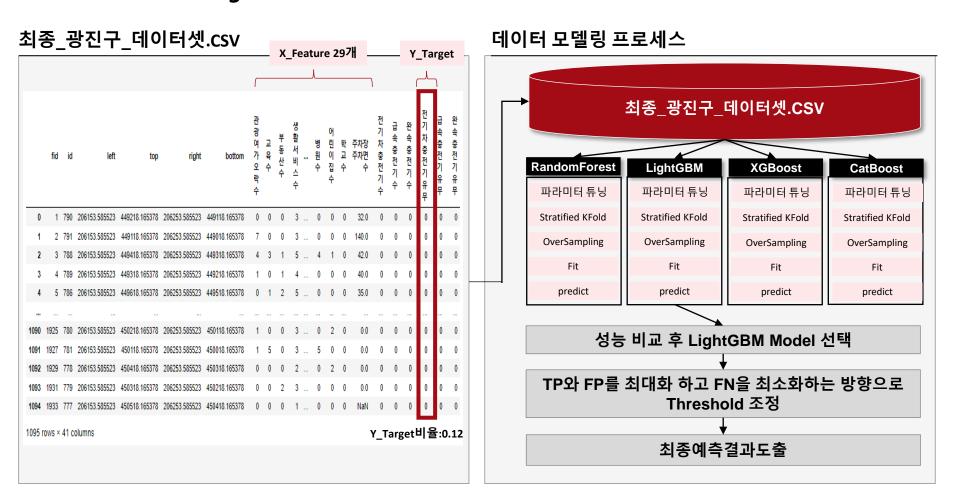
전체 데이터 전처리

- 전기차 충전기의 설치가 불가하거나 전기차 충전기가 불필요한 지역을 제거
- 상권데이터를 활용해 그리드 안 상가의 개수가 0인 지역들을 제거 전체 1933개의 그리드에서 1095개의 그리드로 축소

2. 데이터 수집 및 전처리



최종_광진구_데이터셋.CSV를 활용해 다음과 같은 프로세스로 데이터 모델링을 진행했고, 성능이 가장 좋은 LightGBM Model로 최종 예측결과 도출



3. 모델링을 활용한 데이터 분석



모델 학습의 성능을 개선하고, 데이터셋에 존재하는 한계를 해결하기 위해 다음과 같은 3가지 방법론을 도입하여 모델링을 진행함

Purpose

Method

Over Sampling

- Positive 클래스의 비율이 12.05%로 데이터셋의 클래스 레이블 불균형 문제가 존재
- 이로 인해 편향된 학습 가능성이 있음
- 클래스 불균형 문제를 해소하기 위해 Oversampling 기법인 SMOTE 사용
- Positive 클래스의 수를 132개에서 963개로 확대하여 50%비율로 맞춤

데이터 모델링 Stratified K-Fold

- 단일 데이터셋에 대해 모델 학습과 예측을 동시에 진행하기 위한 목적
- 클래스의 비율 차이를 고려한 K-fold 방식 필요
- Stratified K-Fold(K=3)를 통해 클래스 비율에 맞추어 train, test 분리
- 2개의 fold로 구성된 train set으로 학습한 후 1개의 fold로 된 test set에 대해 확률 예측 진행

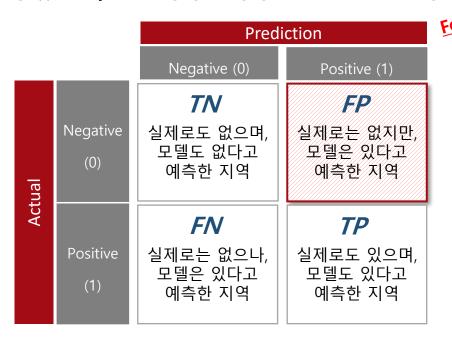
Hyper parameter Tuning

- 모델의 성능을 최적화하고 개선하여 정확도(accuracy)를 최대로 높이기 위한 작업
- 많은 모델을 학습에 사용하다 보니 튜닝에 많은 시간이 소요되어 이를 보완하는 튜닝 방법 사용
- Optuna 라이브러리를 통해 하이퍼 파라미터 튜닝 과정의 시간 단축
- 파라미터의 범위를 대략적으로 지정해주면 자동적으로 탐색 과정을 거쳐 최적의 파라미터 도출

3. 모델링을 활용한 데이터 분석



분류 모델 학습 후 도출된 Confusion Matrix에서 FP에 해당하는 영역을 1차 후보지로 선정할수 있으며, 이를 위해 아래와 같은 모델 학습 목표를 설정함



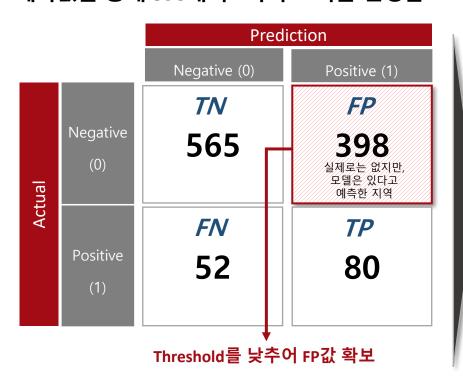


- 모델링을 활용한 입지 분석에서 초점을 맞춘 부분은 FP(False Positive) 영역으로서, FP 영역에 분류된 그리드는 실제로는 전기차 충전소가 위치하고 있지 않지만, 모델은 있을 것이라고 예측한 지역임
- 이는 모델이 분류 task를 수행하는 과정에서 전기차 충전소가 위치한 그리드에 대해 학습한 정보를
 바탕으로 출력한 유력한 후보지라고 해석이 가능함
- Threshold 값을 조정하여 사전에 정한 1차 입지 후보 수 만큼의 FP값을 확보할 수 있음
- 이 과정에서 정확도(accuracy)에 해당하는 TN + TP를 높이면서, 동시에 FN을 낮추고 FP를 높이는 방향으로 학습하는 것을 목표로 함

3. 모델링을 활용한 데이터 분석



모델링 분석 결과 가장 성능이 우수한 LightGBMClassifier 모델을 선택하였으며, 모델의 확률 예측값을 통해 398개의 1차 후보지를 선정함





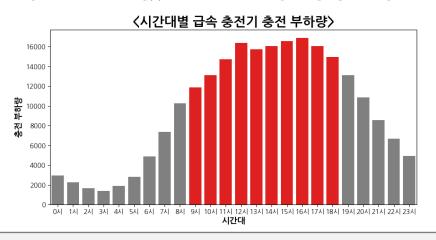
최종적으로 도출된 1차 후보지 398개

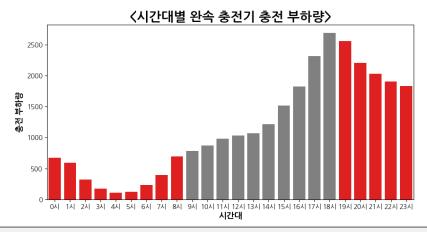
- RandomForest, XGBoost, LightGBM, CatBoost 등 총 4개의 모델을 학습한 후 각 모델의 성능을 서로 비교한 결과, accuracy와 f1_score 점수 모두 어느 정도 높은 성능을 보여준 모델은 LightGBM으로 나타남
- LightGBM이 출력한 예측 확률값 P(y=1)에 대해 threshold를 조정하여 accuracy 손실을 최소한으로 하면서 FP값을 적정 수준으로 확보함
- 모델 분석 결과 총 1095개의 그리드 중에서 398개의 그리드를 1차 후보지로 도출할 수 있었음

4. 실증적 자료를 활용한 데이터 분석



또한, 전기차 충전 시설의 실증적 자료 분석을 한 결과, 급속 충전기와 완속 충전기의 상반된특징이 발견되었고 이를 통해 1차 후보지를 급속 · 완속 충전기 후보지로 구분 가능함





급속 충전기 - 충전기 사용 패턴이 주간 시간대에 집중

완속 충전기 - 충전기 사용 패턴이 야간 시간대에 집중



4. 실증적 자료를 활용한 데이터 분석



특히 급속 충전기는 상권 위주에 입지하는 특징이 있었으며, 완속 충전기는 거주 지역 인근에 입지하는 특징이 존재해 관련 변수로 분석의 기준을 설정함

상권 구	관련 fea	ture	e –							
	관광여가오락수	교육수	부동산수	생활서비스수	소매수	숙박수	스포츠수	음식수	급속충전기유무	완속충전기유무
관광여가오락수	1.000	0.112	0.272	0.313	0.127	0.052	0.236	0.584	-0.007	-0.024
교육수	0.112	1.000	0.350	0.286	0.077	-0.001	0.291	0.197	0.019	-0.016
부동산수	0.272	0.350	1.000	0.518	0.233	0.030	0.268	0.372	0.028	0.085
생활서비스수	0.313	0.286	0.518	1.000	0.440	0.040	0.219	0.526	0.025	0.057
소매수	0.127	0.077	0.233	0.440	1.000	0.014	0.133	0.393	0.114	0.048
숙박수	0.052	-0.001	0.030	0.040	0.014	1.000	0.087	0.113	0.005	0.019
스포츠수	0.236	0.291	0.268	0.219	0.133	0.087	1.000	0.270	-0.026	0.034
음식수	0.584	0.197	0.372	0.526	0.393	0.113	0.270	1.000	0.014	0.041
급속충전기유무	-0.007	0.019	0.028	0.025	0.114	0.005	-0.026	0.014	1.000	0.262
완속충전기유무	-0.024	-0.016	0.085	0.057	0.048	0.019	0.034	0.041	0.262	1.000

인구 관련 feature												
	60 대총인 구수	50 대총인 구수	40 대총인 구수	30 대총인 구수	20대총인 구수	총인구대비 남성비율	총인구대비 여성비율	생산가능인구 남성비율	생산가능인구 여성비율	인구당전 기차수	급속충전 기유무	완속충전 기유무
							:					
생산가능인 구남성비율	0.439	0.407	0.349	0.392	0.324	0.949	0.595	1.000	0.625	0.430	-0.122	-0.014
생산가능인 구여성비율	0.394	0.383	0.360	0.432	0.367	0.732	0.814	0.625	1.000	0.439	-0.120	-0.025
인구당전기 차수	0.878	0.940	0.927	0.839	0.738	0.419	0.408	0.430	0.439	1.000	-0.052	0.097
급속충전기 유무	-0.050	-0.036	-0.022	-0.080	-0.068	-0.129	-0.100	-0.122	-0.120	-0.052	1.000	0.262
완속충전기 유무	0.082	0.108	0.108	0.053	0.042	-0.019	0.004	-0.014	-0.025	0.097	0.262	1.000

급속충전기유무와 완속충전기 유무를 극명하게 구분할 수 있는 '소매수', '인구당전기차수' feature를 실증적 분석의 기준으로 설정

4. 실증적 자료를 활용한 데이터 분석



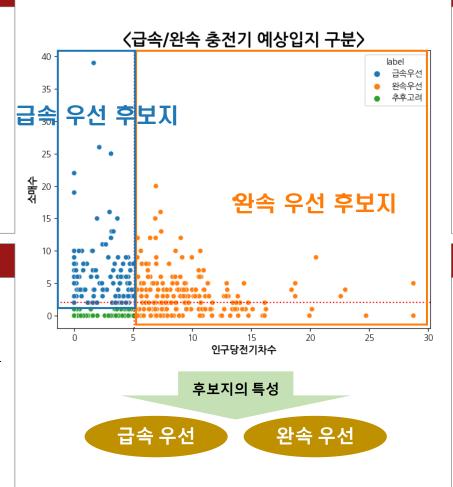
소매수와 인구당 전기차 수를 기준으로 후보지를 재배치하고 두 기준의 평균값을 기준으로 입지의 특성을 분석한 결과 급속 충전기 중심 후보지와 완속 충전기 중심 후보지로 구분이 가능함

2사분면 해석

- 소매업 수는 높고, 인구당 전기차 수는 낮은 지역
- 주거시설에 비해
 상업시설이 많은 지역으로
 해석 가능
- 주간인구가 야간인구보다 많은 지역으로 **급속 충전기를 중심으로 설치** 가능

3사분면 해석

- 소매업 수와 인구당 전기차 수 모두 적은 지역
- 상대적으로 다른 지역에 비해 주간 인구와 야간 인구 모두 적을 것으로 예상
- 다른 사분면에 비해
 중요도가 떨어지는
 지역으로 후보지에서 제거



1사분면 해석

- 소매업 수와 인구당 전기차 수 모두 높은 지역
- 주야간 인구 모두 많은 지역으로서 모든 사분면 중 가장 중요도가 높은 지역으로 지정
- 다만, 주간인구가 많을 것으로 예상되기 때문에 **완속 우선지역**으로 구분

4사분면 해석

- 소매업 수는 낮고, 인구당 전기차 수는 높은 지역
- 상업 시설에 비해 주거
 시설이 많은 지역으로 해석
 가능
- 주간 인구에 비해 야간 인구가 많기 때문에, 완속 충전기를 중심으로 설치 가능

5. 모델링과 실증적 자료의 결합



3차 분석에서는 X축에 모델이 도출한 예측확률, Y축에 충전기별 중요변수를 각각 설정해 예측확률이 높고 주요변수의 값도 높아 수요가 높은 지역을 최적의 후보지로 도출함

데이터에 기반한 1차 분석

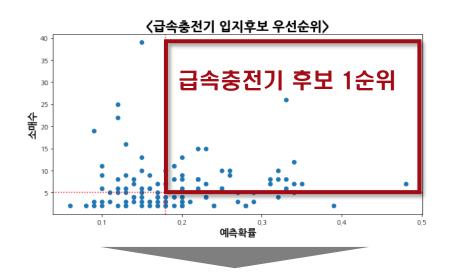
• 모델이 추론한 결과를 통해 **타당성** 확보

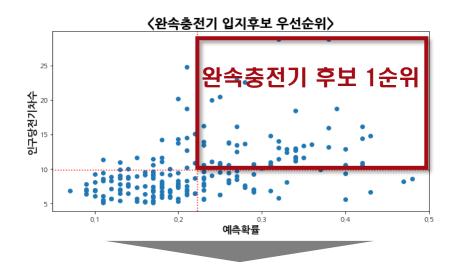
실증적 자료를 통한 2차 분석

• 급속과 완속 입지를 구분함으로서 현실성 확보



2번의 분석을 통해 얻은 타당성과 현실성에 우선순위를 부여하여 시급성 반영





최종 급속충전기 후보 33곳

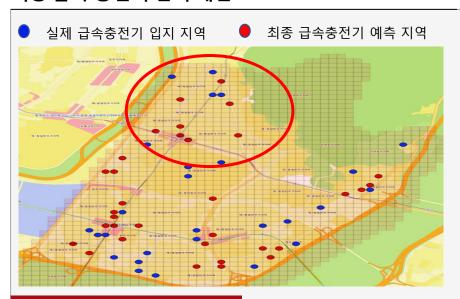
최종 완속충전기 후보 55곳

6. 결과해석



각 충전기에 도출된 지역을 같으며, 급속 충전기와 완속 충전기는 다음과 같은 특징을 지니고 있음

최종 급속 충전기 입지 제안



Key_Findings

- 상업지역이 많은 지하철역 인근에 많은 군집이 형성
- 군자역-아차산역 인근의 인프라가 다른 지역에 비해 상대적으로 부족하며, 모델과 자체적인 실증적인 분석에서도 이 지역의 입지를 제안
- 건대입구역을 근처로도 많은 수요예측이 되고 있음, 수요가 공급에 비해 많은 지역의 설치를 고려해야함

최종 완속 충전기 입지 제안



Key_Findings

- 주거 단지가 많은 지역에 군집 되어 있음
- 강변역 광나루역 사이 주거 단지에 현재 전기차 충전기 시설이 존재하지만, 더 많은 수요가 있는 지역이기에 추가적인 설치 필요
- 대규모 아파트 단지 중 (ex. 잠실대교 북단 아파트 단지, 아차산역 인근 아파트 단지) 완속 충전기가 단지 내에 입지 하지 않은 지역의 설치 필요

6. 결과해석



특히 충급속 충전기는 입지의 예시로 군자역 인근의 상권, 완속충전기는 잠실대교 북단의 아파트 단지를 검토해 보았고, 실제 충전기 공급이 부족한 지역임을 확인

급속충전기 입지-_{군자역 상권}



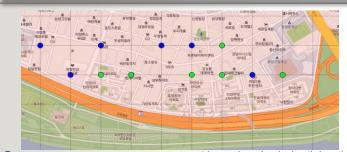
Challenge

군자역 인근 상업 단지에 완속 충전기 시설은 입지해 있으나 비개방 전기차 시설이 많으며, 인근에 상권이 발달한 정도를 고려했을 때, 급속 전기차 충전시설이 부족

Solution

주민센터 주차장을 활용해 급속 충전기를 설치한다면, 주간 시간에 더 많은 차량의 충전을 효과적으로 개선시켜 줄 것이며, 일부 충전시설을 야간에도 개방한다면, 인근 지역 주민의 전기차 충전에 대한 문제 해소를 가능하게 해줄 것임 실제적 입지

완속충전기 입지-_{잠실대교 북단}



●실제 완속충전기 입지 지역 ● 최종 완속충전기 예측 지역

Challenge

잠실대교 북단 아파트 단지에 많은 인구가 거주 중이지만, 단지 내 완속충전시설이 설치되지 않은 경우가 많아서 주민들이 인근 충전소에서 충전을 하는 불편을 겪고 있음



거주인구가 많은 아파트 단지에 우선적으로 완속충전기를 설치해 주민들의 불편을 개선해야 하며, 이 외에도, 해당 문제를 겪고 있는 아파트 단지에 완속 충전기의 도입이 검토되어야 함

7. 시사점



지금까지의 분석과정과 결과는 다음과 같은 기대효과를 지니며, 전기차 입지에서 기회를 극대화 하고 리스크를 최소화 한다면 건전한 전기차 충전 시설의 확대가 가능할 것임

분석 결과

기대효과

기회 리스크

빅데이터를 통한 최적 입지 도출

- 머신러닝을 활용해서 입지를 예측할 수 있었으며, 현실성을 고려한 실증적 분석과 함께 최적의 입지를 도출해 비교적 현실적인 입지 예측이 가능함

정책의 효과성 극대화

- 2025년까지 설치될 약 2000여개의 전기차 입지에 있어서 다양한 요소들을 고려할 수 있으며, 조금 더 과학적인 의사결정으로, 전기차 입지의 우선순위에 대해 파악해 전기차 인프라 취약지역에 대한 지원이 가능

전기차 충전을 기다리는 소비자들을 위한 소구포인트

Q. 전기차충전소에서충전 시 무엇을 많이 하시나요?

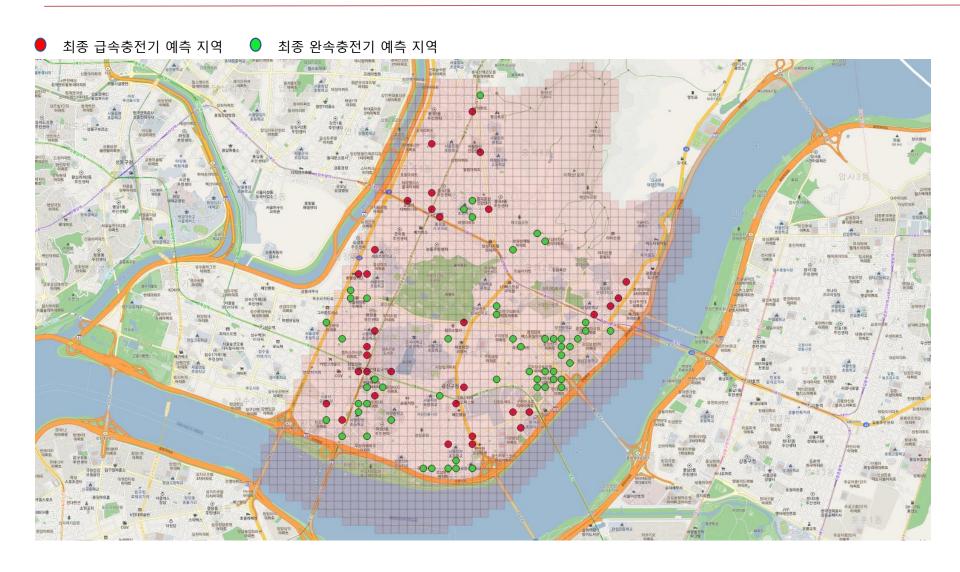


- 문화시설을 활용해 전기차 충전을 기다리는 고객을 사로잡을 수 있는 편의시설은 기회가 될 수 있음

주차장 및 전기차와 관련된 리스크

- 전기차 충전시설은 주차장에 입지하는 경우가 많아서, 주차요금이나 주차장 운영시간과 같은 요소에 영향을 받는 경우가 많음 이러한 리스크를 해소해 더 원활한 이용이 될 수 있도록 노력해야 함
- 충전기의 종류에 따라서 일부 차종은 호환이 안되는 경우가 있음 이러한 점을 고려해 전기차 인프라를 구축해야 함

광진구 최종 전기차 충전기 입지 도출





End of Document