

# 데이터 분석 보고서

원장	부장	지도교수

미래융합교육원 [계룡건설]빅데이터기반 Green Tech SW개발자양성 과정

연수기간 : 2023. 07. 12 - 2023. 12. 26

프로젝트 기간 : 2023. 11. 04 - 2023. 12. 26

팀 명		
팀 원	성 명	업 무 분 담 (역할)
	송현지(PM)	DA
	사원기(PL)	DA, FE
	김동현(TM)	DB, BE
	조하은(QA)	DB, BE
	정민형	FE
목적 및 개발 대상 (분야, 목적, 대상 등)	대전시 건물데이터 기반 행정구역별, 건물 용도별, 탄소배출량 예측 및 생활 실천 방안 제안을 통한 탄소중립 실천	
프로젝트 명	대전시 건물데이터 기반 탄소배출량 예측 및 감축 방안 시스템 구축	

## 1. 프로젝트 개요

### <시스템 개요>

- 대전시 건물데이터 기반 행정구역별, 용도별 탄소배출량을 시각화로 나타낸다.
- 해당 데이터를 기반으로 행정구역별, 용도별 탄소소비량을 예측한다.
- 개인은 '탄소 절감 계산기'를 활용하여 에너지 생활 속 에너지를 절감 방안을 제시한다.
- 지자체 등 정부기관에 탄소 절감을 위한 도시계획 방안을 제안한다.

### <시스템 기능>

- 대전시 데이터의 수집 및 표시
  - 지번별 건물에너지 데이터 수집
    - 행정구역 지번 기반 전기 및 가스에너지 데이터 저장
  - 건물용도별 데이터 수집
    - 건축물대장의 표제부 데이터 저장
    - 해당 데이터를 기반으로 1.1 데이터와 연동
  - 가로수 및 탄소배출 흡수량 데이터
    - 구별 가로수 종류별 개수 데이터
    - 해당 가로수가 흡수하는 탄소흡수량
  - 산림면적 데이터 수집
    - 산림 종류별 면적과 온실가스 흡수량 데이터 저장
  - 탄소배출량, 탄소흡수량 및 감축방안 데이터 수집
    - 에너지별 탄소배출량 공식 데이터 수집
    - 행정구역별 나무의 수와 수종별 온실가스 흡수량
    - 생활속 탄소감축 방안 데이터 수집

## 1.6. 데이터 표시

1.6.1. 위 수집한 데이터는 카카오 API 대전광역시 지도와 연동하여 표시

1.6.2. 에너지 사용 데이터는 행정구역별, 건물용도별 에너지 사용량 및 탄소배출량 현황을 계산 및 출력

1.6.3. 1,6,2데이터와 가로수 현황 및 탄소 흡수량 데이터 출력

## 2. 데이터 조회

### 2.1. 지번별 건물에너지 데이터

2.1.1. 행정구역별 에너지 사용량 및 탄소배출 현황 표시(2018.01. ~ 2023.07.)

### 2.2. 건물용도별 데이터

2.2.1. 1.1 데이터와 연동하여 건물 용도별 에너지 사용량 및 탄소배출 현황 표시

2.2.1. 1.2 지번 데이터는 지도에는 표시되지만, 용도가 적합하지 않은 경우에는 DB에서 조회 불가

### 2.3. 가로수별 나무 수량 및 온실가스 흡수량 데이터

2.3.1. 행정구역별 나무 수량 및 탄소배출 흡수량 표시

### 2.4. 탄소배출량, 탄소흡수량 및 생활속 감축방안 데이터

2.4.1. 에너지별 탄소배출량 공식을 적용하여 탄소배출량 표시

2.4.2. 탄소배출 사용량에 따른 나무별 탄소흡수량 표시(온실가스 흡수량 데이터)

2.4.3. 탄소배출량에 따른 생활속 탄소감축 방안 제시

### 2.5. 사용자는 대전시의 행정구역별, 건물용도별 탄소배출 현황과 행정구역별 산림을 통한 온실가스 흡수량 조회 가능

### 2.6. 수집한 데이터와 카카오 API 연동을 통해, 사용자 입력주소의 특정 월/년도별 에너지 사용량 조회 가능

## 3. 에너지 소비 현황 분석

3.1. 수집된 1.5 탄소배출량 및 탄소흡수량 데이터를 활용하여 사용 에너지 기반 탄소배출량 계산

3.2. 수집된 1.5 생활속 탄소 감축 데이터를 활용하여 사용 에너지 기반 감축방안 계산

3.3. Kmeans 군집화를 통한 에너지 소비량 유사도 비교 및 예측

### 3.4. 분석방법

3.4.1. 데이터 전처리, 데이터 분할, 모델 학습, 모델 평가, 모델 최적화, 예측 및 결과 해석

## 4. 에너지 소비량 예측

4.1. 시스템은 과거 데이터를 분석하여 미래(월·년별)의 행정구역별, 용도별 탄소배출량 예측

4.2. 시기(달/연도)를 선택하면 해당 시기의 전력 소비량 예측

4.3. 행정구역별 예측된 전력, 가스 소비량과 실제 전력, 가스 소비량 비교 출력

4.4. 전년도들의 동월 전력 소비량의 막대그래프 등과 같은 상세표 출력

## 5. 계산기 시스템 연동

5.1. 사용자는 2의 데이터를 기반 '탄소 감축 계산기'를 활용하여 구체적인 감량 요금 목표 설정 가능

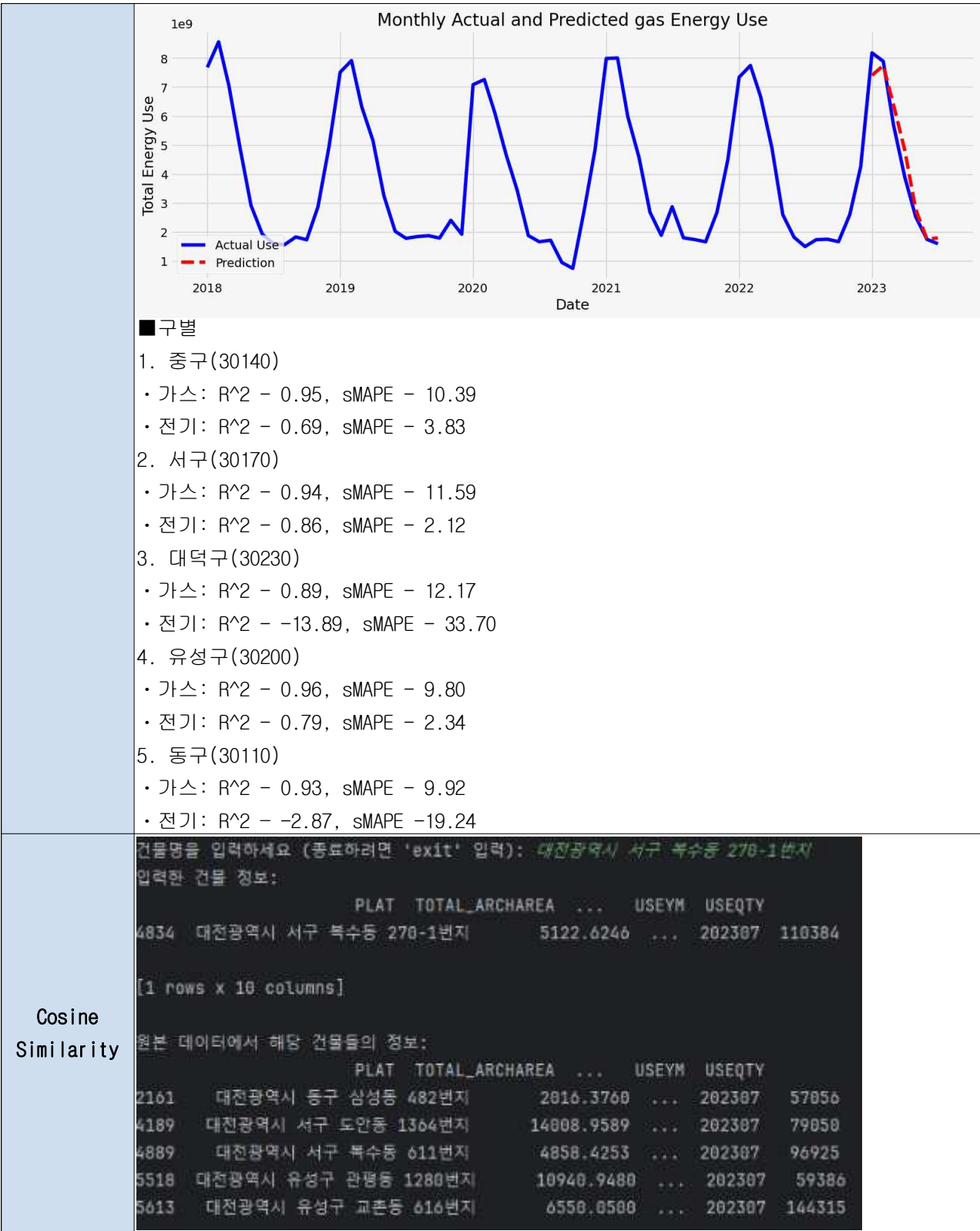
5.2. 목표에 따라 생활속 실천방안(가스 및 전기부문)을 확인 가능

## 6. 도시계획 관리 기능

6.1. 대전시 탄소배출 현황을 모니터링을 통해 행정구역별/건물 유형별 탄소배출량 현황을 지속적으로 관리 가능

6.2. 행정구역별 산림면적, 밀도를 기반으로 제안된 탄소배출 감량 방안(필요 나무 수)을 도시계획 의사결정에 활용

7. 관리자 관리	
7.1. 최고 관리자 계정은 사용자의 입력데이터 승인의 등록, 수정, 삭제가 가능	
8. 데이터 유지 및 보관	
8.1. 시스템은 매월 1일 주기적으로 에너지(가스 및 전기)사용 데이터를 백업	
8.2. 사용자가 입력한 데이터는 관리자의 승인하에 데이터를 업데이트	
2. 데이터 설명	
목표	<ul style="list-style-type: none"> <li>에너지 데이터 기반 XGBoost 모델을 구축을 통한 행정구역별, 건물 용도별, 탄소배출량 예측</li> <li>군집화를 통한 에너지 소비량 유사도 비교 및 예측</li> </ul>
핵심 개념	<ul style="list-style-type: none"> <li>XGBoost, Cosine Similarity</li> </ul>
데이터 수집	<ul style="list-style-type: none"> <li>건물별 전기(2018~2023) 사용량 데이터 : <a href="https://open.eais.go.kr/opnsvc/opnSvcInquireView.do#">https://open.eais.go.kr/opnsvc/opnSvcInquireView.do#</a></li> <li>건물별 가스(2018~2023) 사용량 데이터 : <a href="https://open.eais.go.kr/opnsvc/opnSvcInquireView.do#">https://open.eais.go.kr/opnsvc/opnSvcInquireView.do#</a></li> <li>건물 표제부 데이터 : <a href="https://open.eais.go.kr/opnsvc/opnSvcInquireView.do#">https://open.eais.go.kr/opnsvc/opnSvcInquireView.do#</a></li> <li>대전광역시 가로수 데이터, 가로수별 탄소 흡수량 : <a href="https://www.bigdata-forest.kr/frn/index">https://www.bigdata-forest.kr/frn/index</a></li> <li>산림녹지 탄소흡수량 공간정보(2021) <a href="https://www.bigdata-forest.kr/orderProduct/EC0110101">https://www.bigdata-forest.kr/orderProduct/EC0110101</a></li> </ul>
데이터 준비 및 탐색	
<ul style="list-style-type: none"> <li>파일 불러오기 : <code>pd.read_csv()</code></li> <li>정보 확인 : <code>.shape()</code> / <code>.info()</code> / <code>.head()</code> / <code>.describe()</code></li> <li>결측 데이터 제거 : <code>.fillna()</code> / <code>.notnull()</code></li> <li>레이블 확인 : <code>.value_counts()</code></li> <li>데이터 분할 : <code>.train_test_split()</code></li> <li>데이터 파악 : <code>sns.boxplot()</code> / <code>sns.kdeplot()</code> / IQR / 사분위수를 이용한 상한값 하한값</li> <li>Train set : 2018.01.~2022.12.</li> <li>Test set : 2023.01.~2023.07.</li> <li>Prediction period : 2023.08.~2023.12.</li> </ul>	
3. 분석 모델	
분석 모델 구축방법	
<ul style="list-style-type: none"> <li>상관관계분석 : <code>.corr()</code> / <code>sns.heatmap()</code> / <code>LassoCV()</code> / <code>RandomForestRegressor()</code></li> <li>XGBoost : <code>xgb.XGBRegressor()</code></li> <li>Cosine Similarity : <code>similarity_matrix</code></li> </ul>	
분석 모델 평가	
<ul style="list-style-type: none"> <li>SMAPE, <math>R^2</math></li> </ul>	
4. 분석 결과	
XGBoost	<ul style="list-style-type: none"> <li>■대전 전체</li> <li>가스: <math>R^2</math> - 0.96, SMAPE - 0.10</li> <li>전기: <math>R^2</math> - 0.64, SMAPE - 0.03</li> </ul>



위와 같이 2023년도 공공 데이터 팀 프로젝트 데이터 분석 보고서를 제출합니다.

2023 년 12 월 26 일

제출자 : 에코트리 (인)

미래융합교육원 원장 귀하