



# 어린이집 폐원 예측



MIRICANVAS.COM

# 목차

**01** 사회적 배경

**02** 선행 연구

**03** 예측 설계

**04** 예측 진행

**05** 예측 분석

**06** 향후 정책적 제언

## 1. 사회적 배경

# 사회적 배경

## #1 폐원하는 어린이집, 반대하는 학부모

서초구가 국공립 어린이집 두 곳을 폐원하려고 함. 학부모의 반발로 계획 접음.

### 구청 측

출산율 저하  
한정된 예산 내 보육 내실화 다지려 함

### 학부모 측

어린이집에 대한 수요 충분  
무한정 대기로 이어짐  
행정편의주의적 사고  
가장 큰 피해자는 영유아

**어린이집 폐원 VS. 학부모**  
**끊이지 않는 갈등 양상**

## 1. 사회적 배경

# 사회적 배경

## #2 갑작스러운 폐원 통보

반포동 내 국공립 어린이집 2개의 갑작스러운 폐원 통보  
타 어린이집 역시 대기 인원 100명 이상  
1년 넘게 기다려야 하는 상황

〈영유아보육법 제43조〉

폐지하거나 일정 기간 운영 중지 시 해당 사실을 2개월 전까지  
교직원과 보호자에게 알려야 함.

원아들이 다른 기관으로 갈 수 있도록 권익보호

**해당 법 조항은 국공립 어린이집은 X**  
**이 조항마저 지켜지지 않는 경우 多**

## 1. 사회적 배경

# 사회적 배경

저출생으로 인한 불가피한 폐원  
BUT. 갑작스러운 폐원 통보로 인해  
학부모와 영유아의 권리가 보장되지 않는 현실  
→ 즉, 수요자의 알 권리를 충족하는  
**보육 맞춤형 환경 조성 필요**

## 2. 선행 연구

# 선행 연구와의 차이

## 어린이집 정원충족률에 따른 폐쇄위험과 수급조절 방안

Cox 이산형 비례 위험 모형으로  
정원충족률에 따른 폐쇄위험률을 예측

## 어린이집/유치원 단기 수요 측정

인공신경망 모형을 활용하여  
인구 관련 데이터를 바탕으로 수요 예측

어린이집 운영 정보와 인구 관련 정보를 통해  
폐원에 관한 **다각도적인 예측** 시도



### 3. 예측 설계

# 예측 설계

## 기간: 2008~2023년

- 2008년: 보육정책 확대기 → 보육정책의 질적 수준 향상시키는 수요자 중심의 보육정책으로 전환.  
EX) 아이사랑플랜
- 보육에 대한 관심도가 정부/학부모 모두 높아지기 시작한 시기

## 단위: 서울특별시/광역시별 구군

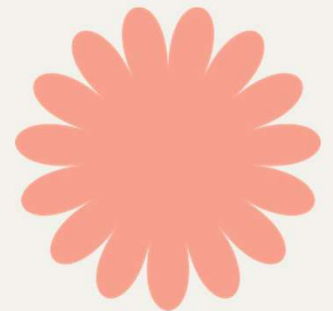
- 상대적으로 인구가 많이 밀집된 지역인 수도와 광역시를 타겟팅
- 각 지역의 최하위 자치단체인 '구'와 '군'을 기반으로 단위 선정

### 3. 예측 설계

# 예측 설계

## 데이터 설명

- 지역별 어린이집 기본정보조회
  - 어린이집 정보공개포털에서 제공하는 OPEN API
  - 매월말일 정기적으로 수집
  - 어린이집 운영과 관련한 자세한 데이터를 담고 있음
- 지역별 0~4세 인구
  - "행정동별 연령별 인구 현황" 사이트에서 제공하는 데이터
  - 보육이 필요한 연령을 지역별로 파악 가능
- 지역별 인구밀도
  - 국가통계포털에서 제공하는 구군별 인구밀도
  - 구 단위로 인구 밀도 비교 가능





## 3. 예측 설계

# 예측 설계

## ○ 변수 설정

- 목표 변수: **운영현황**
- 예측 변수

보육실 수	보육실 면적	놀이터 수	CCTV 수
정원 수	통학차량운영여부	운영연수	어린이집유형구분
학령인구감소율	폐지 당시 인구밀도	지역 내 어린이집 밀집도	보육실 면적비율



# 전처리 과정



분석 데이터: 부산 어린이집 기본정보, 부산 0~4세인구, 부산 인구밀도  
사용 모델: 랜덤포레스트, XGBoost

#### 4. 예측 진행

# 모델 향상

F1 Score	어린이집 기본정보	학령인구감소율	폐지연도 인구밀도	어린이집 밀집도	정원수에 따른 보육실 면적비
랜덤포레스트	0.6987	0.7080	0.7133	0.7798	0.7898
XGBoost	0.6904	0.7619	0.7878	0.8263	0.7852

부산 데이터를 통해 파생 변수를 추가할수록 모델이 향상되었음을 확인





## 4. 예측 진행

서울

# 지역별 모델 비교

모델	정확도	정밀도	재현율	F1 score
로지스틱 회귀	0.7905	0.7651	0.6781	0.7189
랜덤포레스트	0.9092	0.9476	0.8154	0.8765
XGBoost	0.9279	0.9568	0.8562	0.9037



## 부산

모델	정확도	정밀도	재현율	F1 score
로지스틱 회귀	0.7655	0.7555	0.3908	0.5151
랜덤포레스트	0.8717	0.8823	0.6896	0.7741
XGBoost	0.8864	0.8589	0.7701	0.8121

4. 예측 진행

# 지역별 모델 비교



#### 4. 예측 진행

## 인천

# 지역별 모델 비교

모델	정확도	정밀도	재현율	F1 score
로지스틱 회귀	0.7511	0.6821	0.6130	0.6457
랜덤포레스트	0.9008	0.9424	0.7797	0.8534
XGBoost	0.9339	0.9423	0.8750	0.9074





## 대구

모델	정확도	정밀도	재현율	F1 score
로지스틱 회귀	0.8333	0.7571	0.6794	0.7162
랜덤포레스트	0.9206	0.9677	0.7692	0.8571
XGBoost	0.9007	0.8630	0.8076	0.8344

4. 예측 진행

지역별  
모델  
비교



## 4. 예측 진행

## 대전

# 지역별 모델 비교

모델	정확도	정밀도	재현율	F1 score
로지스틱 회귀	0.7346	0.7232	0.7718	0.7467
랜덤포레스트	0.9489	0.9589	0.9395	0.9491
XGBoost	0.9217	0.8987	0.9530	0.9250



## 울산

모델	정확도	정밀도	재현율	F1 score
로지스틱 회귀	0.7459	0.7567	0.6588	0.7044
랜덤포레스트	0.8864	0.9444	0.8000	0.8662
XGBoost	0.8972	0.9459	0.8235	0.8805

4. 예측 진행

# 지역별 모델 비교



## 4. 예측 진행

## 광주

# 지역별 모델 비교

모델	정확도	정밀도	재현율	F1 score
로지스틱 회귀	0.6923	0.6000	0.4426	0.5094
랜덤포레스트	0.9112	0.9107	0.8360	0.8717
XGBoost	0.9289	0.9298	0.8688	0.8983



#### 4. 예측 진행

# 모델 평가

## 평가

- 대전, 대구 지역을 제외한 지역 - XGBoost > 랜덤포레스트 > 로지스틱 회귀 순 성능 ↑
- Q. 로지스틱 회귀보다 트리 기반 계열의 모델이 성능이 더 좋은 이유?
- A. **선형적 모델 vs. 비선형적 모델**. 변수 중 비선형적 관계가 존재하기 때문

## 한계

- 최근 설립된 어린이집을 설명하기에는 역부족

## 5. 예측 분석

# 주요 변수

각 지역별로 가장 성능이 높은 모델 내  
주요 변수를 분석  
랜덤포레스트 → 상관계수  
XGBoost → SHAP



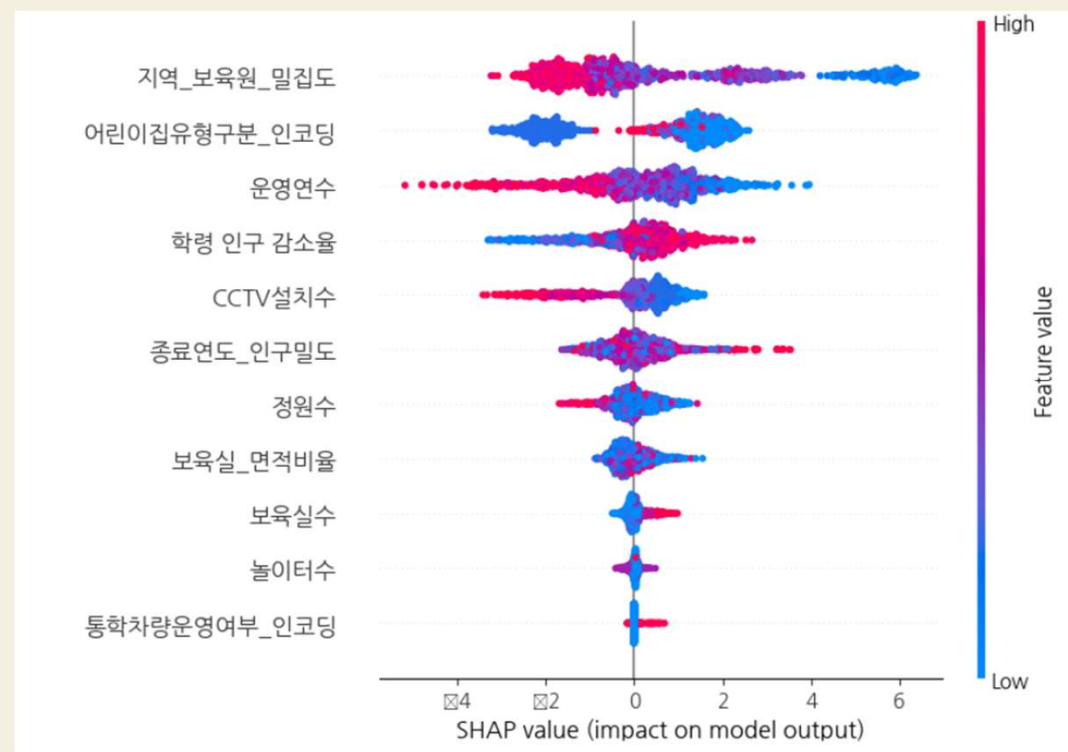
## 5. 예측 확장

## 서울

XGBoost 모델의 SHAP 값 비교

## 주요 변수

1. 지역 어린이집 밀집도
2. 어린이집 유형
3. 운영연수



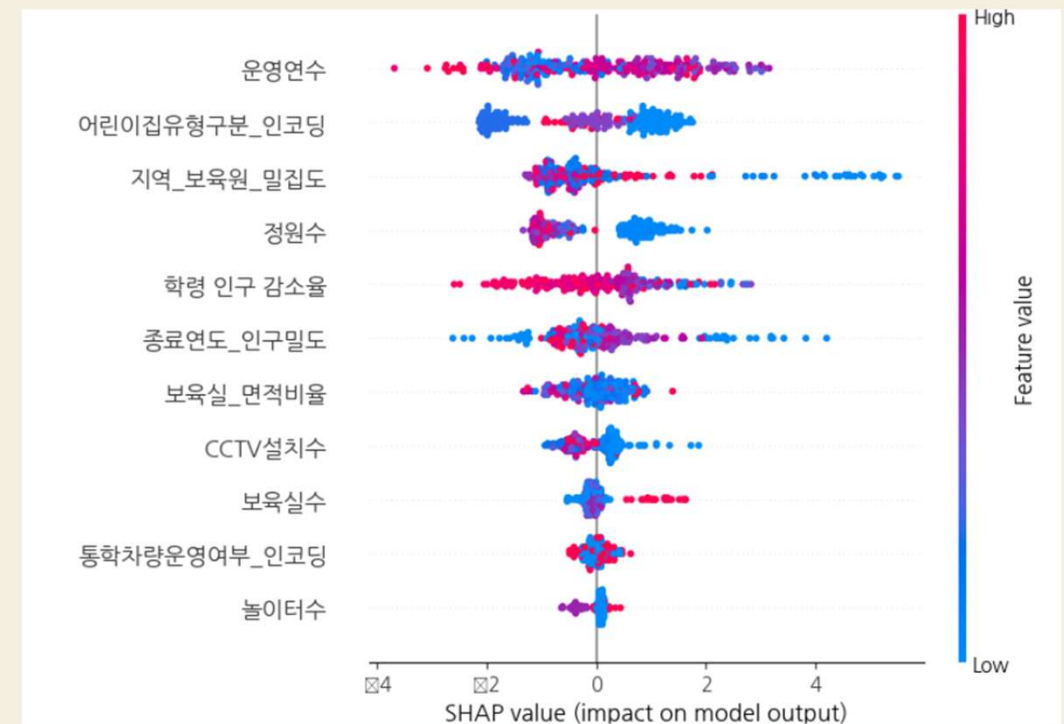
## 5. 예측 확장

## 부산

XGBoost 모델의 SHAP 값 비교

## 주요 변수

1. 운영연수
2. 어린이집 유형
3. 지역 어린이집 밀집도



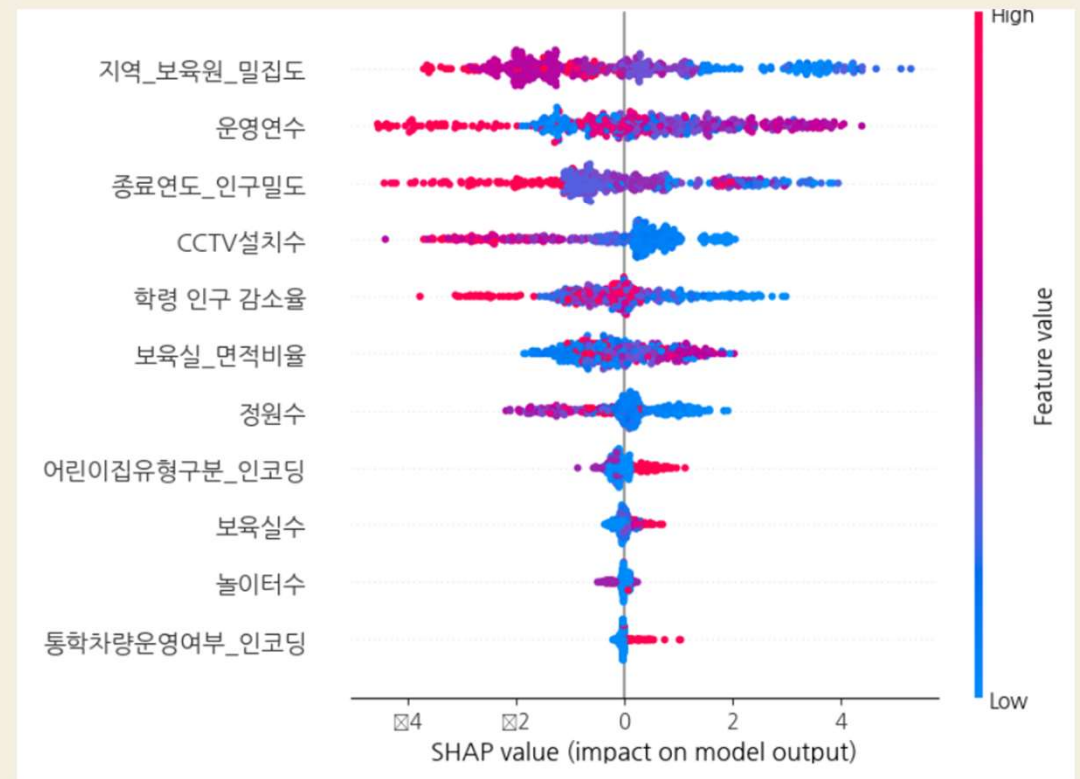
## 5. 예측 확장

## 인천

XGBoost 모델의 SHAP 값 비교

## 주요 변수

1. 지역 어린이집 밀집도
2. 운영연수
3. 폐지연도 인구밀도



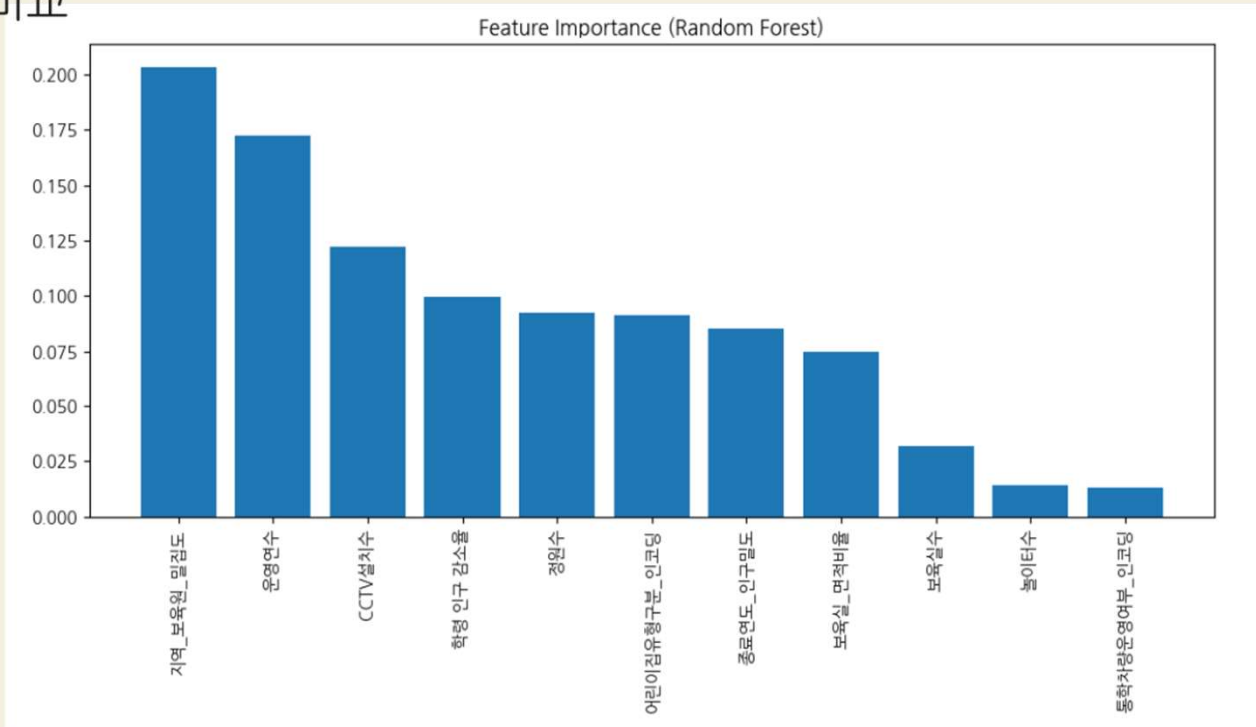
## 5. 예측 확장

# 대구

랜덤포레스트 모델 상관계수 비교

### 주요 변수

1. 지역 어린이집 밀집도
2. 운영연수
3. CCTV 설치 수



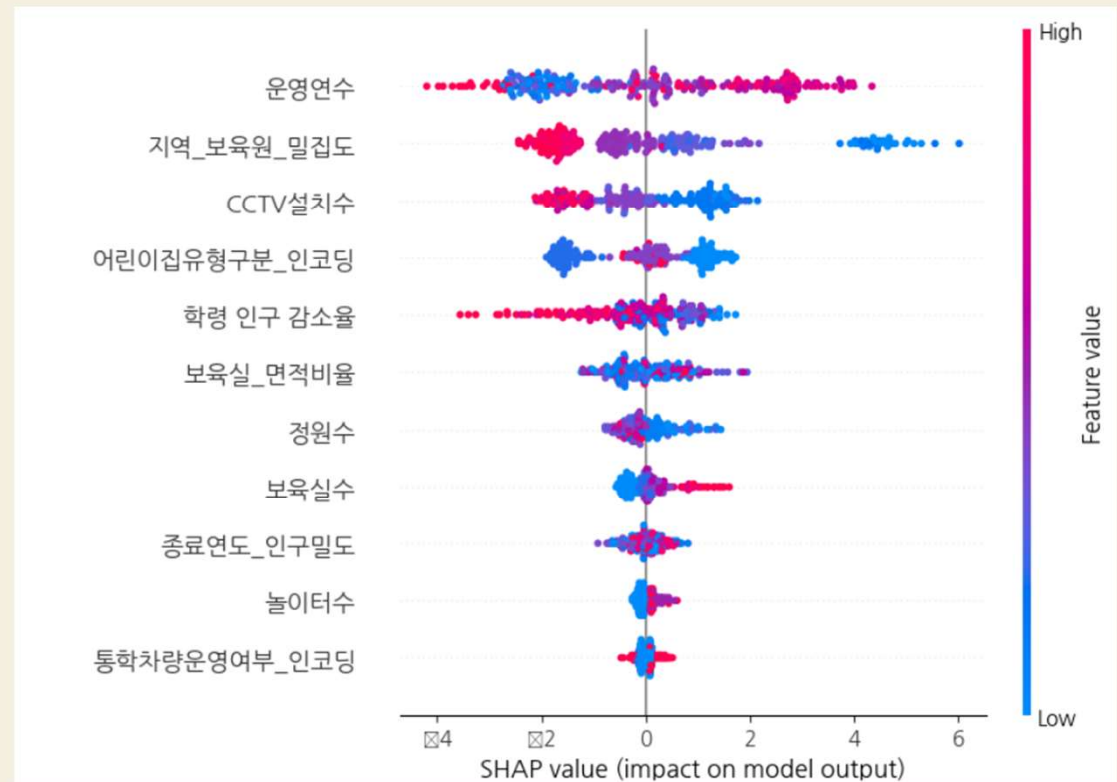
## 5. 예측 확장

## 대구

XGBoost 모델의 SHAP 값 비교

## 주요 변수

1. 운영연수
2. 지역 어린이집 밀집도
3. CCTV 설치 수



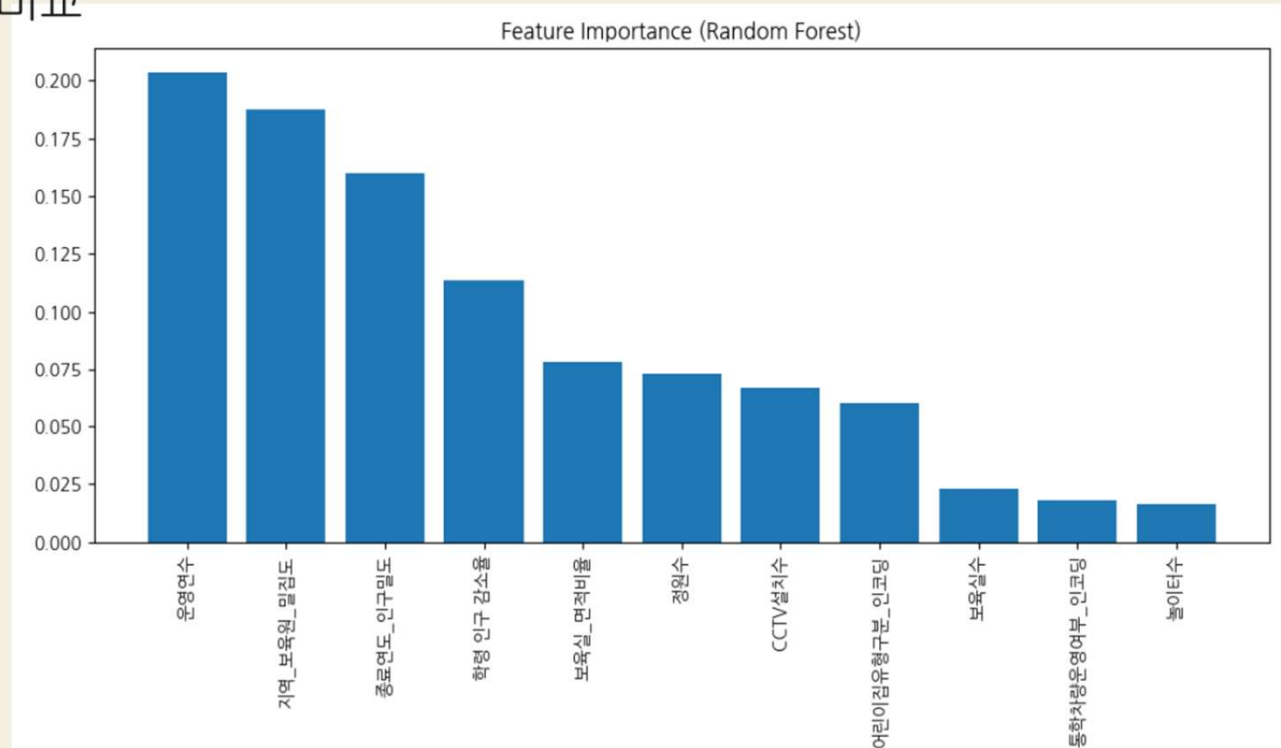
## 5. 예측 확장

# 대전

랜덤포레스트 모델 상관계수 비교

### 주요 변수

1. 운영연수
2. 지역 어린이집 밀집도
3. 폐지연도 인구밀도





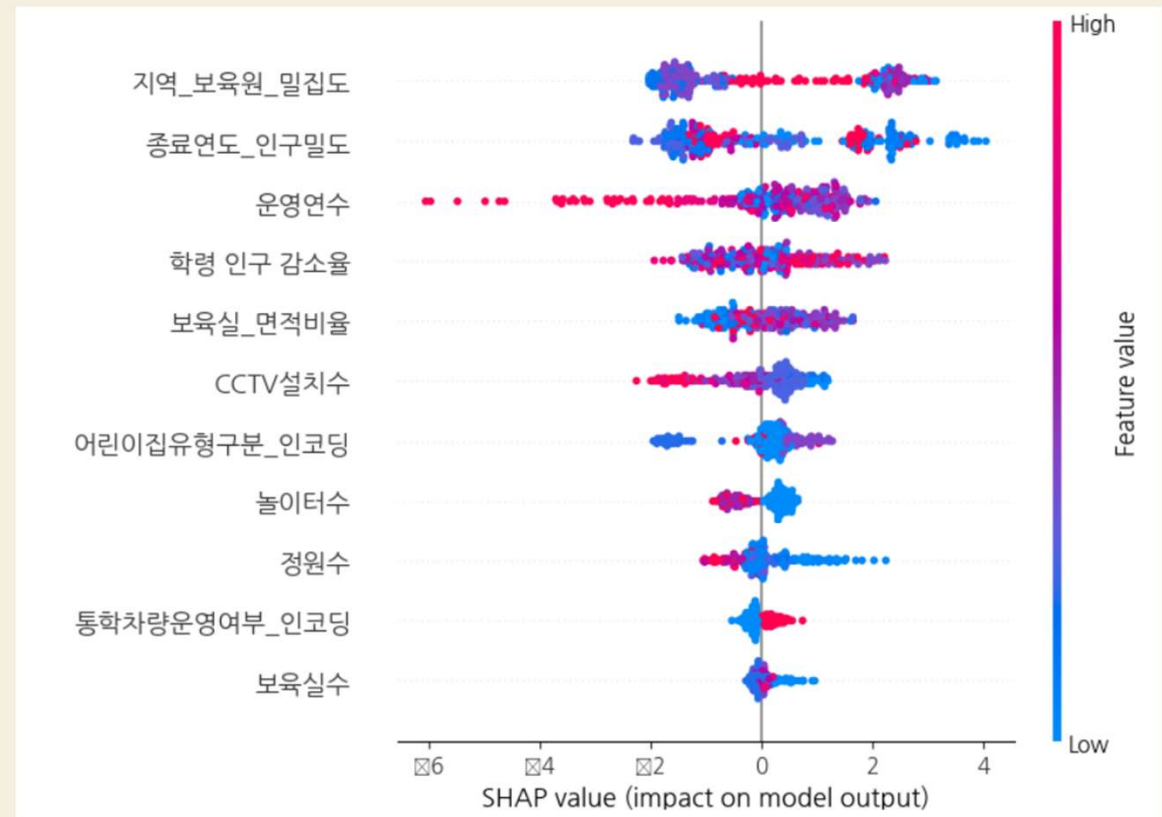
## 5. 예측 확장

## 대전

XGBoost 모델의 SHAP 값 비교

## 주요 변수

1. 지역 어린이집 밀집도
2. 폐지연도 인구밀도
3. 운영연수



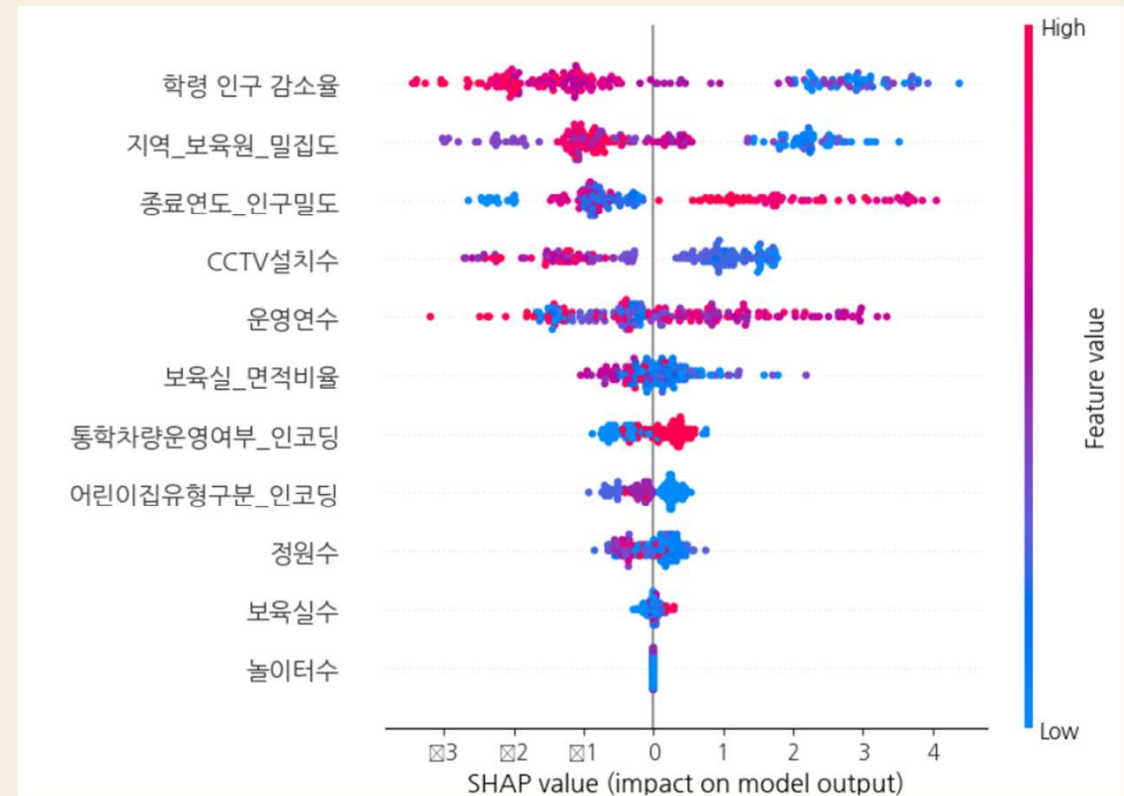
## 5. 예측 확장

## 울산

XGBoost 모델의 SHAP 값 비교

## 주요 변수

1. 학령 인구 감소율
2. 지역 어린이집 밀집도
3. 폐지연도 인구밀도



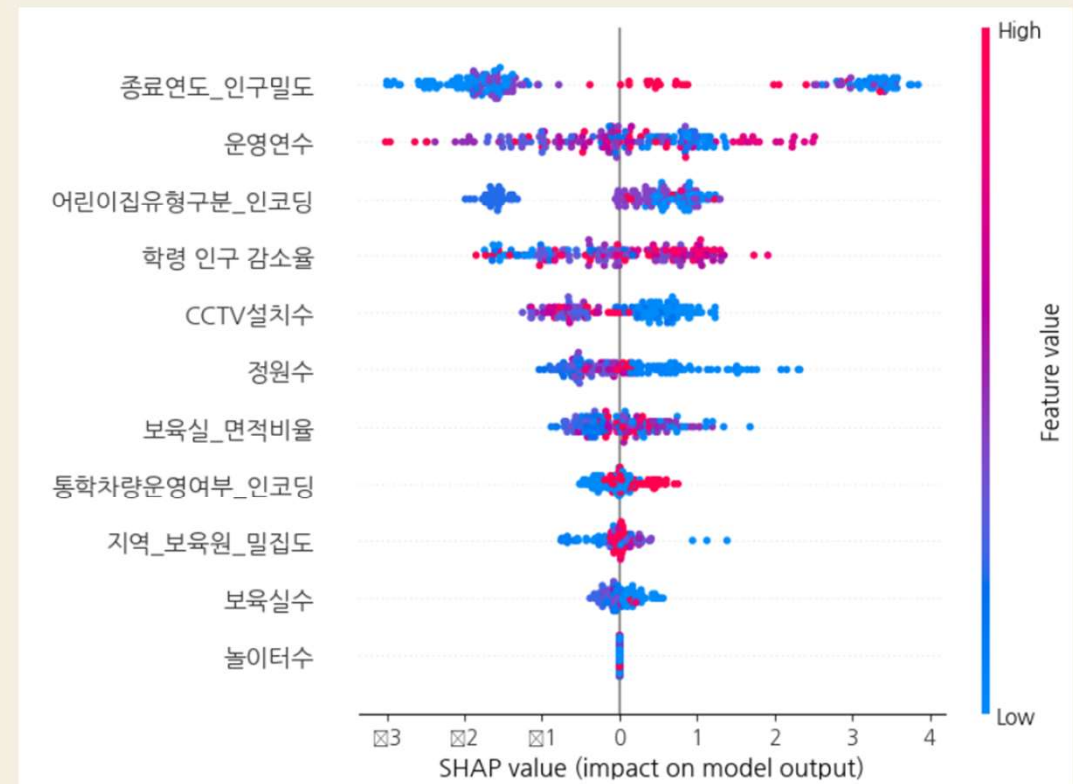
## 5. 예측 확장

## 광주

XGBoost 모델의 SHAP 값 비교

## 주요 변수

1. 폐지연도 인구밀도
2. 운영연수
3. 어린이집 유형



## 5. 예측 분석

# 분석 평가

### 평가

- 운영연수, 지역 어린이집 밀집도, 폐지연도 인구밀도 등과 같이 인구 + 어린이집 운영 기간 관련 변수가 많이 도출
- 각 변수간의 관계 파악을 중요시

### 한계

- 예상치 못한 부분이 존재 → 추가적인 사회적 조사 및 배경 연구가 필요로 됨.

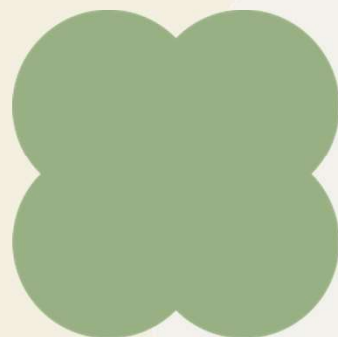
## 6. 향후 정책적 제언

# 정책적 제언



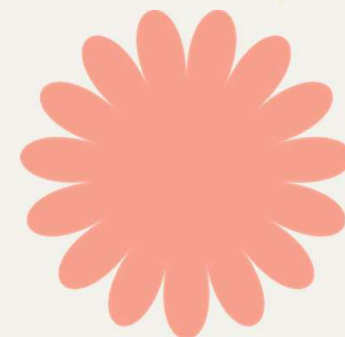
### 폐원 가능성 지도

지도를 통해 원하는  
어린이집에 대한 폐원 가능성  
폐원 관련 정보 제공



### 어린이집 내 운영 개선

폐원 예측을 통해  
미비한 운영 요인 예방  
사전 개선 가능성



### 다른 정책적 방향성

유보통합에서 벗어난 시각  
거시적으로 바라보기보다  
미시적인 문제부터 해결



THANK YOU

감사합니다