

# 부동산 전세가격 예측·전세가율 분석

: 적정 전세가율을 활용한 전세사기 예방 웹사이트 구축

Statistical Analysis & Machine Learning

[3조: 건물주]

김주성, 이길연, 최동원, 최선재, 최연우, 황신엽

multicampus

## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)  
주제 선정 배경  
프로젝트 수행 방향  
프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집 & 전처리  
EDA  
통계 분석  
변수 설정  
분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가  
Streamlit 구축 & 배포

### 04. 최종 결론

기대 효과  
한계점

### 05. Document

부록  
참고문헌  
출처

# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 조직 (구성원 및 역할)



김주성 (팀장)

데이터 전처리, 통계 분석, 머신 러닝

블로그: <https://jay-s1.tistory.com/>

깃허브: <https://github.com/jas21aad>



이길연

데이터 전처리, EDA, 통계 분석

블로그: <https://ls-all.tistory.com/>

깃허브: <https://github.com/Gil-Yeon>



최동원

데이터 전처리, 머신 러닝

블로그: <https://dongwonc.github.io/>

깃허브: <https://github.com/DongWonC>



최선재

데이터 전처리, 시각화, Web 연동

블로그: <https://seonjaechoi.tistory.com/>

깃허브: <https://github.com/HSYhvae>



최연우

데이터 전처리, EDA, 시각화

블로그: <https://heimmersedn.tistory.com/>

깃허브: <https://github.com/heimmersedn>



황신엽

데이터 전처리, EDA

블로그: <https://hyhvae.tistory.com/>

깃허브: <https://github.com/HSYhvae>

# 부동산 전세가격 예측·전세가율 분석

주제 선정 배경

국내

## "전세사기 막을 보완대책 시급"

'전세사기 특별법 보완 및 문제 해결을 위한 정책 포럼' 주제로 국회토론회 개최

23.07.18 17:38 | 최종 업데이트 23.07.19 09:06 | 윤종은(tomayoun) ▼

전세사기 특별법 제정 과정에서 피해자들의 의견이 반영되지 못했다는 응답은 85.4%로 대부분 소외감을 나타냈으며 특별법에 따른 피해구제책에 대해 91%가 충분하지 않다고 응답했다. 또 구제책을 아예 모르고 있거나 정책 내용을 이해하지 못하고 있으며 정책이 도움이 안 된다고 응답한 경우가 많았다.

연합뉴스 구독

## 무자본 갭투자로 126채 산 뒤 전세보증금 250억 편취한 일당

입력 2023.08.02. 오전 10:06 | 수정 2023.08.02. 오전 10:10 | 기사원문

경기북부경찰청은 부동산 업체 대표 A씨 등 11명을 검거, 8명을 구속했다. 이들은 무자력자를 내세워 주택 126채를 허위 매수하고 전셋값과 매매가를 유사하게 설정하여 전세보증금 약 253억원을 편취한 혐의를 받는다. 조직적으로 역할을 분담하여 세입자들에게 신뢰를 줬으며, 일부는 가짜 전세계약서로 청년전세대출도 받았다.

▶ 갭 투자를 비롯하여 여러 전세사기 문제가 발생

→ 부동산 시장에 불신 증가 & 불안정한 전세가격 형성에 기여

# 부동산 전세가격 예측·전세가율 분석

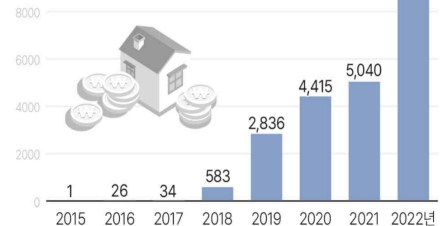
주제 선정 배경

국내

## HUG 전세보증금 대위변제액 추이

HUG가 집주인을 대신해 세입자에게 갚아준 전세보증금

단위: 억원



연암뉴스

자료: 주택도시보증공사(HUG)

사회

## 신혼부부 꿈 망친 전세사기... "소송 다 이겼지만 소용없어"

2023년 04월 26일 05시 18분 댓글

## "경매 막겠다는 말 믿었는데" 울산 오피스텔 임차인들, 전세사기 고소

남구 D오피스텔 58세대 강제경매...임대인·회사 대표 고소 예정

(울산=뉴스1) 임수정 기자 | 2023-08-29 15:59 송고

이어 6월 배씨는 "권씨는 상환 능력이 없고 실소유주인 본인이 직접 이자를 갚아 경매를 취하하겠으니 보증금 만기를 6개월 연장해달라"는 취지로 임차인들을 설득했다고 한다. 배씨의 말을 믿고 기다리다 강제경매 소식을 들은 임차인들은 허탈감을 감추지 못했다.

임차인 김모씨는 "임차인들 대부분이 20대, 30대 초반으로 부동산이나 법률에 대해 잘 알지 못하고 전세보증금반환보증에도 가입하지 않았다"며 "배씨가 도의적 책임을 지고 사태를 해결하겠다는 말을 믿으면서 아무 대처도 하지 않고 흘러보낸 시간이 후회된다"고 말했다.

H 부동산임대회사 대표 배씨는 수 차례 연락을 시도했으나 연락이 닿지 않았다.

(1) 전세사기 피해 + 전세보증금 대위변제액 급증

(2) 전세사기 피해자가 승소하더라도 임대인이 '배 패라는 식'의 행태로 대응하면 마땅한 해결책이 없는 상황

→ 우선적으로 전세사기 예방이 가장 중요하다

# 부동산 전세가격 예측·전세가율 분석

주제 선정 배경

해외

## 파월은 물론 연준 간부들 잇달아 추가 금리인상 시사

뉴스1 원문 | 기사전송 2023-08-28 06:47 최종수정 2023-08-28 08:48

파월 “물가 여전히 높아” : 파월 의장은 지난 25일 잭슨홀 회의 개막 연설에서 “소비자 지출이 견조하고, 주택 부문이 반등할 가능성이 있는 가운데 경제가 예상만큼 냉각되지 않을 수 있다는 징후에 주의를 기울이고 있다”고 말했다. 이에 따라 “적절하다고 판단되면 금리를 더 올릴 준비가 돼 있으며, 물가 상승률이 목표치를 향해 내려가고 있다는 확신이 들 때까지 긴축 정책을 계속 유지할 생각”이라고 덧붙였다.

MTO머니투데이 + 구독

PICK 1

## 中 부동산위기 진앙지 헝다, 상반기만 6조 손실..."주식 거래재개"

입력 2023.08.28. 오전 9:30 · 수정 2023.08.28. 오전 9:31 · 기사원문

중국의 대형 부동산 개발업체인 비구이위안(碧桂園·컨트리가든)이 디폴트 위기를 맞으면서 중국 부동산 업계에 ‘도미노 디폴트’ 우려가 번지고 있다. 다른 중국 주요 부동산 업체의 재정 상황도 나쁜 상황이다. 중국 부동산 시장은 중국 국내총생산(GDP)의 25%를 차지한다. 부동산 시장이 흔들리면 중국 경제도 위기에 빠질 수 있다는 의미다.

미국 금리인상 지속 + 중국 부동산 리스크 등

→ 세계 경제 정세의 불확실성 또한 국내 부동산 시장의 불안정에 영향을 끼침

## 주제 선정 배경

### 주식회사설립 등기신청서

1. 상 호 : 건물주 주식회사
2. 본 점 : ○○ ○○구 ○○동 ○○○번지
3. 등기의 목적 : 주식회사 설립등기
4. 등기의 사유 : 정관을 작성하고 공증인의 인증을 받아 발기인이 회사 설립시에 발행하는 주식의 총수를 인수받고, 2023년 9월 26일 상법 제299조의 절차를 종료하였으므로 다음 사항의 등기를 구함.

상 호 : 건물주 주식회사

본 점 : ○○ ○○구 ○○동 ○○번지

목 적 : 1) 아파트 전세 가격 예측

2) 정부의 부동산 안정화 정책의 기여

3) 개인별 맞춤 투자전략 구축

4) 위 각호에 관련된 부대사업

회사가 발행할 주식의 총수 : 40,000주

1주의 금액 : 금 5,000원

회사의 설립시에 발행하는 주식의 총수 : 10,000주

발행주식의 총수, 그 종류와 각종 주식의 내용과 수 : 10,000주 보통주식

자본의 총액 : 금 50,000,000원

광고방법 : 서울특별시 내에서 발행하는 일간 ○○일보에 게재한다.

회사의 존립기간 또는 해산사유 : 회사성립일로부터 만 30년

이사의 성명 이사 이 길 연

이사 최 동 원

이사 최 선 재

이사 최 연 우

이사 황 신 업

회사를 대표할 이사의 성명과 주소

대표이사 김 주 성

위와 같이 등기신청합니다

2023년 9월 26일

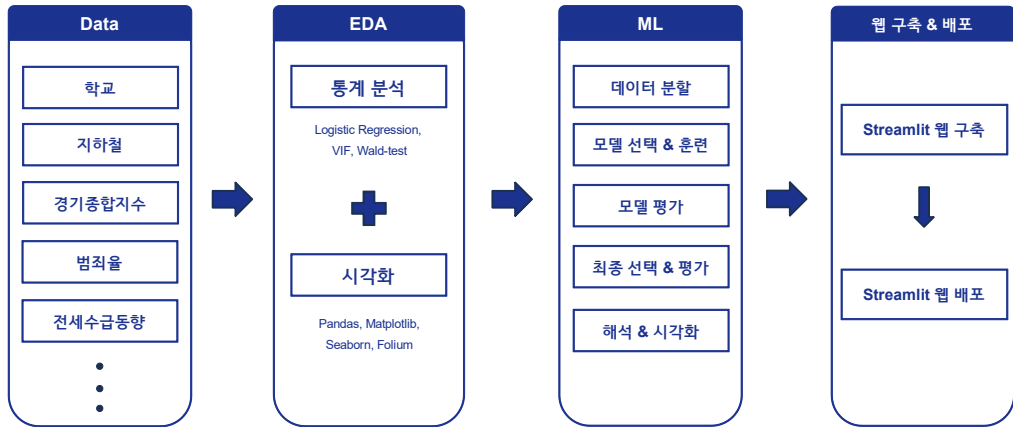
신청인 건물주 주식회사

김 주 성 (인)

(이하 이사 전원이 기명날인한다)

# 부동산 전세가격 예측·전세가율 분석

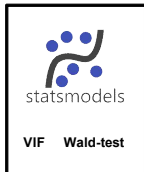
## 프로젝트 수행 방향





### Frameworks

#### Statistics

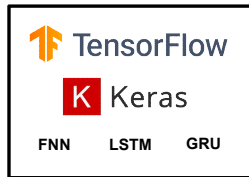


#### Machine Learning



XGBoost    LightGBM

#### Deep Learning



# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 수행 방향

### Statistical Analysis

Logistic Regression, VIF, Wald Test, Prophet  
(Multicollinearity, P-value, Prediction)

### ML

KNN, Random Forest, LightGBM, XGBoost  
(Supervised Learning)

### DL

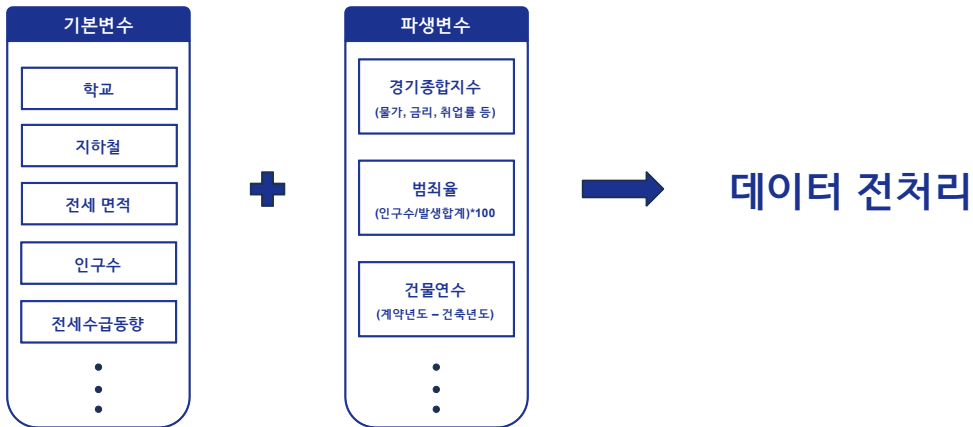
FNN, LSTM, GRU  
(ANN, RNN)



부동산 전세가격  
예측 · 분석

# 부동산 전세가격 예측·전세가율 분석

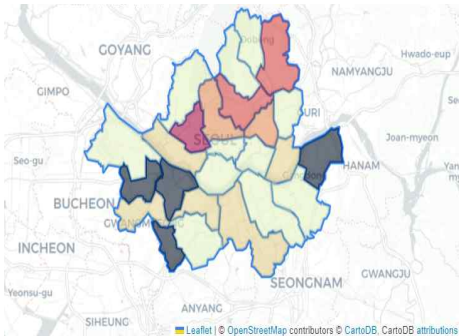
## 프로젝트 수행 방향



# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 수행 방향

지도 시각화



공간정보 분석

데이터 시각화

자치구별 위치 데이터

학교 위치 데이터

지하철역 위치 데이터

구역 설정

서울 자치구별로 분할 후 데이터 분석

기대 효과

1. 서울 자치구별로 적정 전세가격 확인
2. 전세사기 위험 매물 추정 가능

# 부동산 전세가격 예측·전세가율 분석

## 프로젝트 추진 일정

구분	1주차(9/6 ~ 9/10)			2주차(9/11 ~ 9/17)			3주차(9/18 ~ 9/24)			4주차(9/25 ~ 9/26)	
분석 기획											
데이터 수집											
논문 자료 수집											
데이터 전처리											
탐색적 데이터 분석											
데이터 시각화											
통계 분석											
모델 생성 및 평가											
Streamlit 배포											
웹 연동 및 배포											
프로젝트 발표											

## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)  
주제 선정 배경  
프로젝트 수행 방향  
프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집 & 전처리  
EDA  
통계 분석  
변수 설정  
분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가  
Streamlit 구축 & 배포

### 04. 최종 결론

기대 효과  
한계점

### 05. Document

부록  
참고문헌  
출처

# 부동산 전세가격 예측·전세가율 분석

## 데이터 수집



### 지역주택가격 변동의 장단기 결정요인에 관한 실증분석

구분	변수명	변수정의 (단위)
종속 변수	주택가격	지역별 주택매매가격종합지수 (2015.12=100)
	금리	한국은행 기준금리 (%)
설명 변수 (전국)	주식가격	KOSPI 종합지수 (월평균, 1980.1.4=100)
	주택대출	예금은행주택대출 (말잔, 십억원)
	경기상황	경기동행지수 (2010=100)
설명 변수 (지역)	물가	지역별 소비자물가지수 (2010=100)
	주택보증	지역별 주택금융신용보증 공급 (말잔, 억원)
	전세가격	지역별 주택전세가격종합지수 (2015.12=100)
	취업자수	지역별 총취업자수 (천명)
	가구수	지역별 주민등록세대수 (세대)
	고령인구 비중	지역별 65세 이상 인구 비중 (%)

### 주택시장 경기변동과 주거특성들의 아파트가격에 대한 영향력 변화 - 용인 동백택지개발지구를 사례 -

구 분	요 인	비 고
개별적 특성	대지지분	규모를 측정하기 위한 변수로서 다중공선성 관계에 있기 때문에 둘 중 하나만을 이용
	평형	
	층	주택의 해당 층
	각도(향)	정남향에 근접한 정도
단지적 특성	평균층수	단지 내 주거동의 평균층수
	세대수	단지 내 총 주택수
	브랜드 더미	LH=0 민영=1
접근성	중심업무지구까지의 거리	동백지구에서는 다중공선성 관계에 있기 때문에 둘 중 하나만을 이용
	중앙공원까지의 거리	
	경전철역까지의 거리	-
	초등학교까지의 거리	-
	고속도로IC까지의 거리	-
	복합행정타운까지의 거리	-

1. 기존 연구를 참고하여 경제 지표와 주거 특성이 주택 전세가격에 영향을 준다는 것을 파악하여 변수로 채택
2. 기존 연구 외에 주택 전세가격에 영향을 미칠 것으로 예상되는 변수들로서 범죄율과 인구수를 채택

# 부동산 전세가격 예측·전세가율 분석

## 데이터 전처리

### Raw Data

부동산 데이터

시계열 데이터

위치 데이터

경제 데이터

인구 데이터



### Process

데이터 분할 (아파트, 오피스텔, 연립다세대)

결측치 & 이상치 전처리

파생변수 생성(건물연식, 최단거리 ...)

주소 데이터로 좌표 추출

전세 데이터에 나머지 데이터 병합



### Results

데이터 정리



# 부동산 전세가격 예측·전세가율 분석

## 데이터 전처리 - Coordinates

### Problem



추출할 좌표  
데이터가  
너무 많은  
문제..☹



### Solution



1. drop\_duplicates 함수를 활용, 공통된  
위도 & 경도의 주소 데이터를  
일괄적으로 축약



2. 주소 데이터에 Google Maps API를  
사용하여 좌표 추출



3. 기존 데이터와 '주소' 컬럼을 기준으로  
병합



### Results



모든 데이터의  
위도 & 경도를 찾아냄

자세한 내용은 해당 Github 링크 참조:

[https://github.com/HSYhrae/team3\\_semi\\_project/blob/data/%EC%A2%8C%ED%91%9C\\_%EC%B6%94%EC%B6%9C.ipynb](https://github.com/HSYhrae/team3_semi_project/blob/data/%EC%A2%8C%ED%91%9C_%EC%B6%94%EC%B6%9C.ipynb)

## 데이터 전처리 - Missing Values

Before

	Year	
	Month	
	YearMonth	
	Region_Name	
	Floor	
	Building_Age	
	Building_Use	
	JS_BA	
	Population	
	IR	→ IR (기준금리): 3.5로 대체
	UR	
	LC_index	
	CA_index	→ XGBoost 활용
	TC_index	
	SDT_index	
	HSP_index	
	Crime_Rates	→ Crime_Rates (범죄율): 연도별 & 자치구별 평균값 매치
	latitude	
	longitude	
	Shortest_Distance_to_Subway	
	Shortest_Distance_to_University	→ XGBoost 활용
	Shortest_Distance_to_School	
	Sell_Price	
	JS_Price	

Preprocess outliers

After

	Year
	Month
	YearMonth
	Region_Name
	Floor
	Building_Age
	Building_Use
	JS_BA
	Population
	IR
	UR
	LC_index
	CA_index
	TC_index
	SDT_index
	HSP_index
	Crime_Rates
	latitude
	longitude
	Shortest_Distance_to_Subway
	Shortest_Distance_to_University
	Shortest_Distance_to_School
	Sell_Price
	JS_Price

## 데이터 전처리 - Outliers

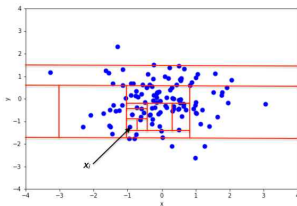
### Isolation Forest

$$c(m) = \begin{cases} 2H(m-1) - \frac{2(m-1)}{n} & \text{for } m > 2 \\ 1 & \text{for } m = 2 \\ 0 & \text{otherwise} \end{cases}$$

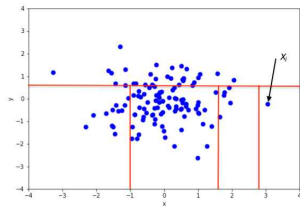
$$H(i) = \ln(i) + \gamma \quad \gamma = 0.5772156649$$

$$s(x, m) = 2^{\frac{-E(h(x))}{c(m)}}$$

- if  $s$  is close to 1 then  $x$  is very likely to be an anomaly
- if  $s$  is smaller than 0.5 then  $x$  is likely to be a normal value



Non-anomalous points



Anomalous points

### Modified Z-Score vs. Isolation Forest 비교 결과

#### Modified Z-Score

평균 절대 오차(MAPE) : 8.356902429380652

결정계수(R^2) : 0.9029853021864345

MAPE: 약 1.65 감소 (24% 개선)

R-Squared: 약 0.028 증가 (3.1% 개선)

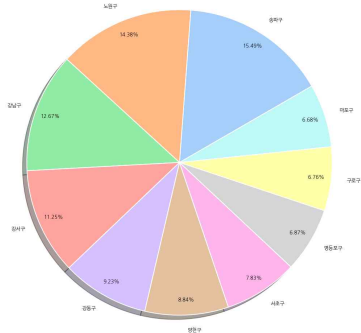
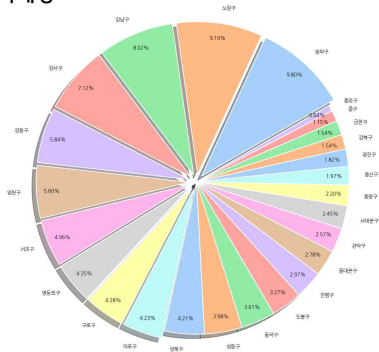
#### Isolation Forest

Mean Absolute Percentage Error (MAPE): 6.704866679030175

R-Squared: 0.9297818877725617

# 부동산 전세가격 예측·전세가율 분석

## EDA - 거래량



서울 자치구별 부동산 거래 데이터 비율

- 자치구별 거래량 차이가 많이 나는 것을 확인
- '종로구, 중구, 금천구'가 하위 3개 자치구임을 확인

서울 자치구별 부동산 거래 데이터 비율 Top 10

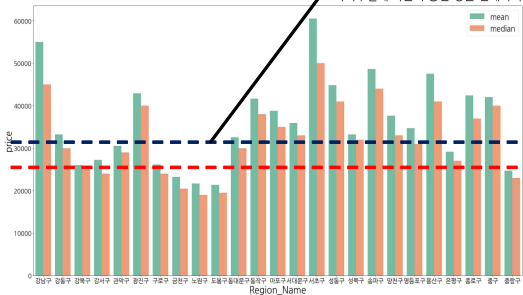
- '송파구, 노원구, 강남구' 순서로 거래량이 많은 것을 확인
- 거래량이 활발한 자치구에서 다양한 매물이 있을 것으로 예상

# 부동산 전세가격 예측·전세가율 분석

## EDA - Metrics

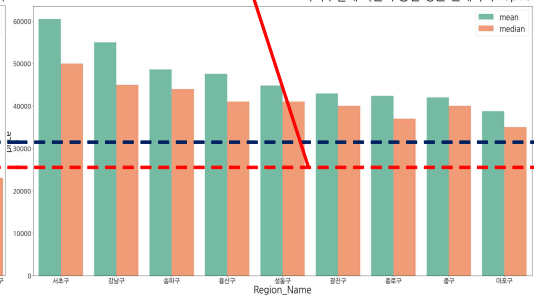
모든 자치구별 부동산 평균 거래가  
전세가격 평균 : 37005 (만원)

자치구별에 따른 부동산 평균 전세가격



모든 자치구별 부동산 중위 거래가  
전세가격 평균 : 31000 (만원)

자치구별에 따른 부동산 평균 전세가격 Top10



### 서울 자치구별 부동산 평균/중위 전세가격

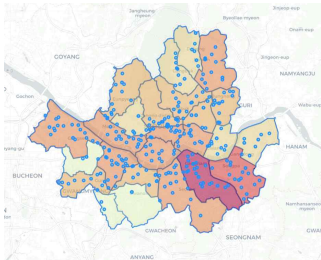
- 자치구별로 거래량의 차이가 큰 것을 확인
- '종로구, 중구, 금천구'가 하위 3개 자치구임을 확인

### 서울 자치구별 부동산 평균/중위 전세가격 Top 10

- '서초구, 강남구, 송파구' 순서로 전세가격이 높은 것으로 확인
- '강남구'와 '송파구' -> 거래량과 전세가격 간의 상관관계가 있을 것으로 추정

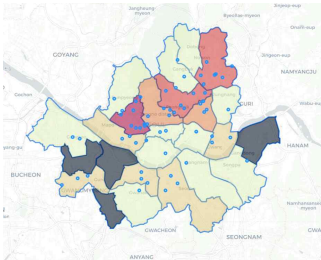
# 부동산 전세가격 예측·전세가율 분석

## EDA - Infrastructure



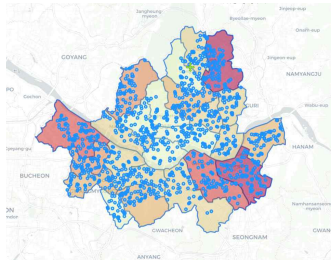
자치구별 지하철 수

강남구	⇒	33개
송파구	⇒	23개
마포구	⇒	22개
중구	⇒	21개
강서구	⇒	20개



자치구별 대학교 수

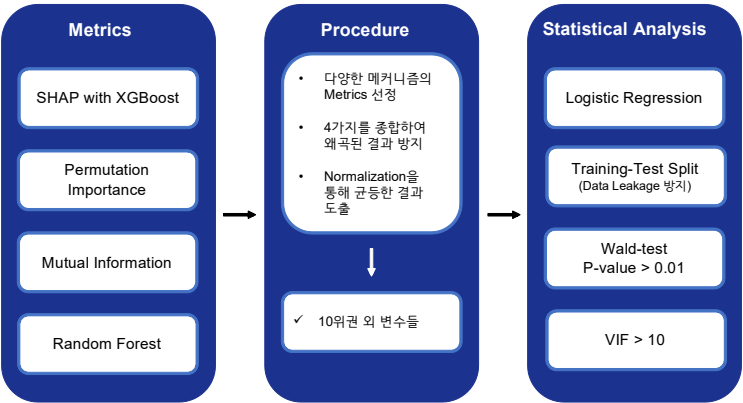
서대문구	⇒	8개
노원구	⇒	7개
성북구	⇒	7개
종로구	⇒	8개
동대문구	⇒	5개



자치구별 초/중/고등학교 수

노원구	⇒	93개
송파구	⇒	89개
강서구	⇒	77개
은평구	⇒	66개
양천구	⇒	64개

## 통계 분석 – Feature Importance



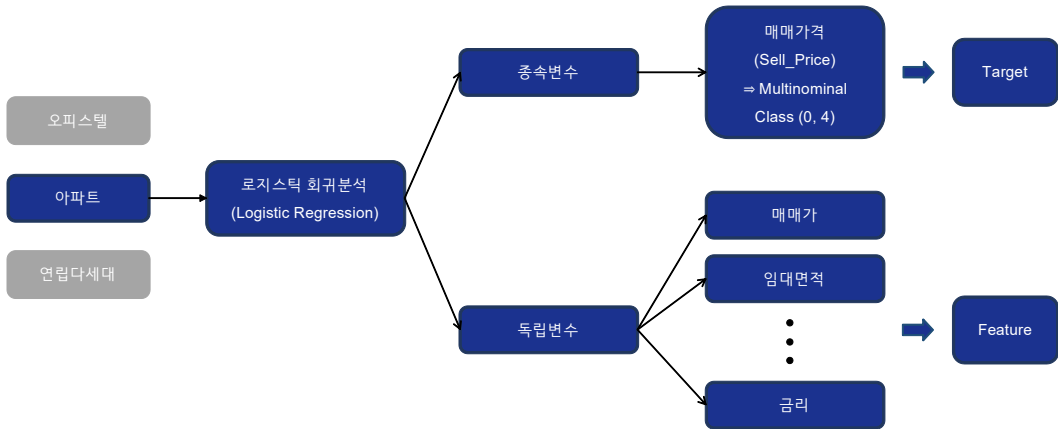
## Features Ranking

	Feature	Combined_Ranking
12	JS_Price	0.986482
7	JS_BA	0.684255
0	Building_Age	0.251346
19	latitude	0.189964
20	longitude	0.181056
10	Region_Name	0.145718
9	Population	0.136831
15	Shortest_Distance_to_University	0.124809
13	Shortest_Distance_to_School	0.123984
14	Shortest_Distance_to_Subway	0.121152
8	LC_index	0.106573
18	YearMonth	0.102640
3	Crime_Rates	0.093738
2	CA_index	0.052059
5	HSP_index	0.051452
6	IR	0.041633
16	TC_index	0.039424
11	SDT_index	0.038430
4	Floor	0.022902
1	Building_Use	0.020333
17	UR	0.005931

10위권 외 변수들

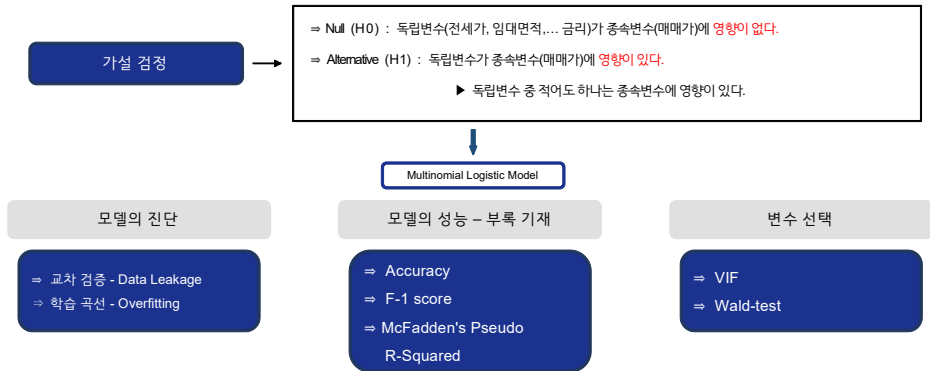
# 부동산 전세가격 예측·전세가율 분석

## 통계 분석 – Logistic Regression





## 통계 분석 – Logistic Regression



# 부동산 전세가격 예측·전세가율 분석

## 통계 분석 – Data Leakage & Overfitting

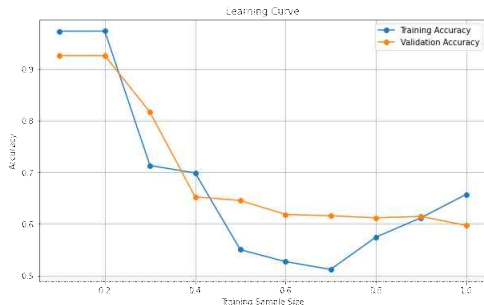
⇒ 교차 검증 - Data Leakage 여부 파악

Fold	Fold-wise Accuracy	Mean Accuracy
1	0.595264467	0.595206678
2	0.595778152	
3	0.595023678	
4	0.595216309	
5	0.594750783	

폴드별 정확도에 큰 차이가 없음

Data Leakage 가능성 ↓

⇒ 학습 곡선 - 모델의 Overfitting 여부 파악



Training Accuracy < Validation Accuracy

Overfitting 가능성 ↓

# 부동산 전세가격 예측·전세가율 분석

## 변수 설정

아파트

1차 제거 VIF > 10

독립변수	VIF
TC_index	86.497
CA_index	79.168

2차 제거 P-value > 0.01

전세가 Class	독립변수	p (유의확률)
1	HSP_index	0.247
	UR	0.506



UR, HSP\_index, TC\_index, CA\_index 제거

오피스텔

독립변수	VIF
CA_index	48.400
TC_index	47.602

전세가 Class	독립변수	p (유의확률)
1	IR	0.083



IR, TC\_index, CA\_index 제거

연립다세대

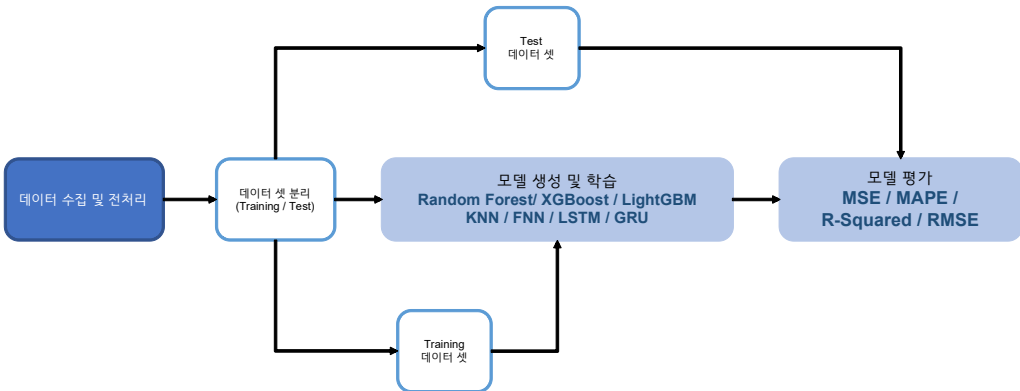
독립변수	VIF
CA_index	97.289
TC_index	89.541



TC\_index, CA\_index 제거

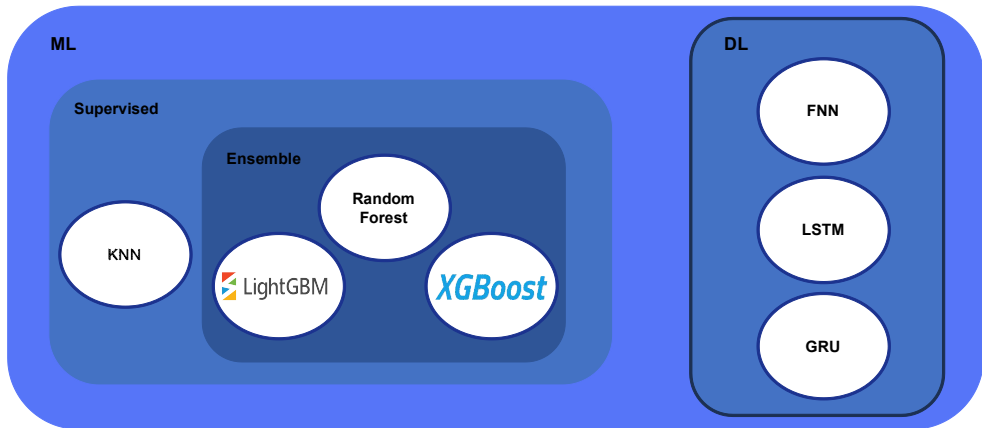
# 부동산 전세가격 예측·전세가율 분석

## 분석 모델 설정 – Training / Test Flowchart



# 부동산 전세가격 예측·전세가율 분석

## 분석 모델 설정 - 모델 생성 및 학습



# 부동산 전세가격 예측·전세가율 분석

## 분석 모델 설정 - 모델 평가 지표

### MAE

실제 값과 예측 값 간의 절대 오차를 나타내는 지표

MAE가 작을수록 모델의 예측이 더 정확

$$MAE = \frac{\sum |y - \hat{y}|}{n}$$

### RMSE

오차 제곱 값의 평균을 측정하고 루트를 씌운 지표

RMSE가 작을수록 모델의 예측이 더 정확

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y - \hat{y})^2}{n}}$$

### MAPE

예측 오차를 백분율로 표현한 지표

MAPE는 예측 성능을 백분율로 이해하기 용이

$$MAPE = \frac{100}{n} \sum_{i=1}^n \left| \frac{y - \hat{y}}{y} \right|$$

### R-Squared

회귀 모델이 종속 변수의 분산을 얼마나 잘 설명하는지 측정하는 지표

R-Squared가 1에 가까울수록 모델이 더 좋은 성능을 가짐

$$R^2 = \frac{\sum (\hat{y}_i - \bar{y})^2}{\sum (y_i - \bar{y})^2}$$

## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)  
주제 선정 배경  
프로젝트 수행 방향  
프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집 & 전처리  
EDA  
통계 분석  
변수 설정  
분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가  
Streamlit 구축 & 배포

### 04. 최종 결론

기대 효과  
한계점

### 05. Document

부록  
참고문헌  
출처

# 부동산 전세가격 예측·전세가율 분석

## 성능 평가

	LightGBM	XGBoost	LSTM	GRU	Random Forest	FNN	KNN
MAE	3654.68	3760.784	3886.538	4018.997	4006.445	4232.778	6279.989
RMSE	6053.86	7135.28	6223.073	6502.211	6645.358	6623.921	9522.973
MAPE	0.128	0.166	0.132	0.139	0.139	0.147	0.218
R-Squared	0.899	0.87	0.893	0.887	0.879	0.879	0.879
Rank	1	2	3	4	5	6	7



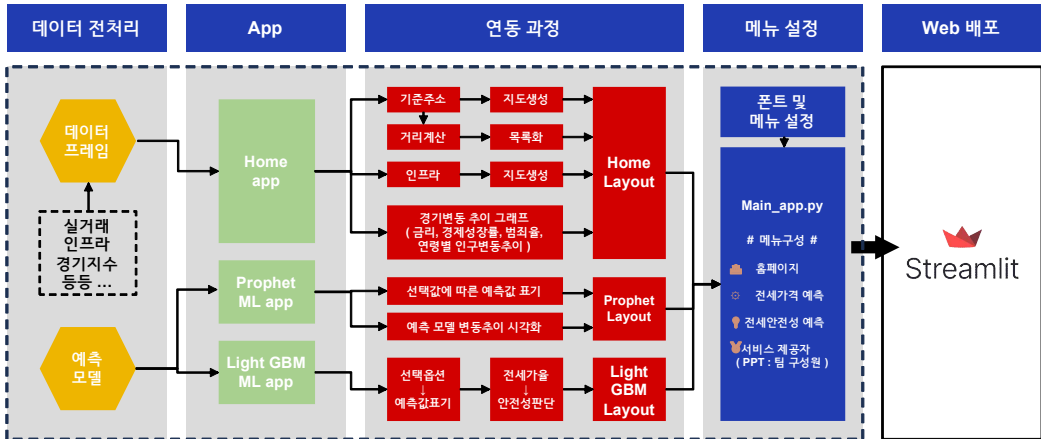
# 부동산 전세가격 예측·전세가율 분석

## 성능 평가 (Normalization)

	LightGBM	XGBoost	LSTM	GRU	Random Forest	FNN	KNN
MAE	1	0.959584186	0.911683539	0.861229	0.86601	0.77979811	0
RMSE	1	0.91581508	0.951222978	0.870759	0.829496	0.83567529	0
MAPE	1	0.977777778	0.955555556	0.877778	0.877778	0.78888889	0
R-Squared	1	0.93877551	0.959183673	0.918367	0.863946	0.86394558	0
Total	1	0.947988139	0.944411437	0.882033	0.859307	0.81707697	0
Rank	1	2	3	4	5	6	7

# 부동산 전세가격 예측·전세가율 분석

## Streamlit 구축



### 메뉴 구성



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

⚙️ 전세가격 예측

💡 전세 안전성 예측

🏆 서비스 제공자

### Streamlit / 전세 안전성 예측 / 예측 모델 구현

주변 인프라를 확인할 기준 도로명 주소를 입력 후 연락을 누르세요.

서울시 강남구 개포동 1194-1

예측하고 싶은 날짜를 선택하세요.

2023/10/01

건물 연식을 입력하세요.

6

계약하시는 물건의 전세 가격을 입력하세요. (단위: 만원)

60000

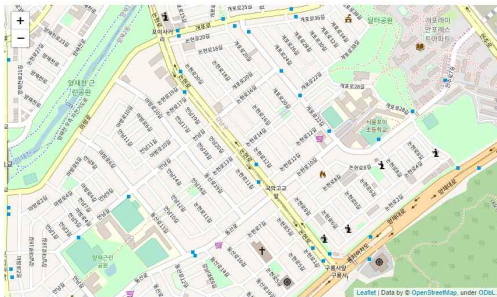
임대 면적을 입력하세요. (단위: 제곱미터)

28

임대 건물의 층수를 입력하세요.

1

예측 결과 값 : 74,352만원 / 전세가율 : 80% / 안전성 : 위험



## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)  
주제 선정 배경  
프로젝트 수행 방향  
프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집 & 전처리  
EDA  
통계 분석  
변수 설정  
분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가  
Streamlit 구축 & 배포

### 04. 최종 결론

기대 효과  
한계점

### 05. Document

부록  
참고문헌  
출처

# 부동산 전세가격 예측·전세가율 분석

## 기대 효과 & 한계점

### 기대 효과



사회초년생의 부동산 전세 계약 의사결정에 기여



전세가율을 통해 위험 매물 판단 가능



희망하는 부동산 근처의 인프라 확인 가능

### 한계점



서울시 이외의 지역에 대한 서비스 미제공



클래스별, 자치구별로 각 모델을 생성할 필요성



더욱 다양한 인프라 데이터 활용

## INDEX

### 01. 프로젝트 개요

프로젝트 조직(구성원 및 역할)  
주제 선정 배경  
프로젝트 수행 방향  
프로젝트 추진 일정

### 02. 프로젝트 수행 과정

데이터 수집 & 전처리  
EDA  
통계 분석  
변수 설정  
분석 모델 설정

### 03. 프로젝트 수행 결과

성능 평가  
Streamlit 구축 & 배포

### 04. 최종 결론

기대 효과  
한계점

### 05. Document

부록  
참고문헌  
출처

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 데이터 정의서

Num	Column Name	Column Name(KOR)	Dtype	etc
1	Year	연도	int64	
2	Month	월	int64	
3	YearMonth	연도 & 월	int64	
4	Region_Name	자치구명	object	
5	Floor	층	float64	
6	Building_Age	건물 연식	int64	계약년도 - 건축년도
7	Building_Use	건물 용도	object	아파트, 오피스텔, 연립다세대
8	JS_BA	전세 임대 면적	float64	

## 부록 - 데이터 정의서

Num	Column Name	Column Name(KOR)	Dtype	etc
9	Population	인구수	int64	
10	IR	기준 금리	float64	
11	UR	실업률	float64	
12	LC_index	선행종합지수	float64	
13	CA_index	동행종합지수	float64	
14	TC_index	후행종합지수	float64	
15	SDT_index	전세수급동향지수	float64	
16	HSP_index	공동주택 매매 실거래가격지수	float64	



# 부동산 전세가격 예측·전세가율 분석

## 부록 - 데이터 정의서

Num	Column Name	Column Name(KOR)	Dtype	etc
17	Crime_Rates	범죄율	float64	(발생횟수/인구수)*100
18	latitude	위도	float64	
19	longitude	경도	float64	
20	Sell_Price_mean	평균 매매가격	float64	연도 & 월별, 구별, 유형별, 주소별
21	Shortest_Distance_to_Subway	지하철역까지의 최소거리	float64	
22	Shortest_Distance_to_University	대학까지의 최소거리	float64	
23	Shortest_Distance_to_School	학교(초, 중, 고등학교)까지의 최소거리	float64	
24	JS_Price	전세가격	int64	

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 모델의 성능

⇒ 아파트 데이터에 대한 정확도, F1-score,  $R^2$

평가지표	정확도	F1-score	McFadden's Pseudo $R^2$
지수	0.63075	0.56866	0.37286

⇒ 오피스텔 데이터에 대한 정확도, F1-score,  $R^2$

평가지표	정확도	F1-score	McFadden's Pseudo $R^2$
지수	0.39132	0.38637	0.14111

⇒ 연립다세대 데이터에 대한 정확도, F1-score,  $R^2$

평가지표	정확도	F1-score	McFadden's Pseudo $R^2$
지수	0.68601	0.58466	NaN



Features\_Ranking의 상위 10가지 변수를 제외하고 만든 모델이기 때문에 낮은 성능을 보임

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 오즈비

$$\begin{aligned} \checkmark \text{ Odds} &= \frac{\text{전세가 클래스에 해당}}{\text{전세가 클래스에 해당하지 않음}} \\ &= \frac{p}{1-p} = \frac{e^{\beta_0 + \beta_1 X_1 + \dots + \beta_i X_i}}{1} \end{aligned}$$

$$\text{오즈비 (OR)} = \frac{\frac{A}{B}}{\frac{C}{D}} = \frac{AD}{BC}$$

구분		전세가클래스(예측)		
		YES	NO	합계
전세가클래스 (실제)	YES	A	B	A+B
	NO	C	D	C+D
합계		A+C	B+D	A+B C+D

- Odds = 해당 사건이 발생할 확률 / 발생하지 않을 확률
- OR>1 : 사건 발생 확률이 오즈비만큼 더 크게 나타남
- OR<1 : 사건 발생 확률이 오즈비만큼 더 작게 나타남
- OR=1 : 1에 가까울수록 사건이 발생할 확률에 유의미한 영향을 준다고 볼 수 없다고 판단

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

APT 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
1	Population	-0.0418	0.007	-6.393	0	0.959	-0.059	-0.025
	Crime_Rates	-0.1136	0.007	-17.465	0	0.893	-0.13	-0.097
	YearMonth_encoded	0.942	0.014	68.842	0	2.565	0.907	0.977
	HSP_index	-0.016	0.014	-1.157	0.247	0.984	-0.052	0.02
	SDT_index	-0.1018	0.011	-9.413	0	0.903	-0.13	-0.074
	IR	-0.1471	0.013	-11.724	0	0.863	-0.179	-0.115
	UR	-0.0051	0.008	-0.666	0.506	0.995	-0.025	0.015
	Floor	0.0202	0.007	3.02	0.003	1.02	0.003	0.037
2	Population	-0.1233	0.007	-17.351	0	0.884	-0.142	-0.105
	Crime_Rates	-0.2613	0.007	-35.716	0	0.77	-0.28	-0.242
	YearMonth_encoded	1.6035	0.016	102.374	0	4.97	1.563	1.644

## 부록 - 통계 분석

APT 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
2	HSP_index	-0.2258	0.015	-15.209	0	0.798	-0.264	-0.188
	SDT_index	-0.388	0.012	-32.641	0	0.678	-0.419	-0.357
	IR	-0.7147	0.014	-49.307	0	0.489	-0.752	-0.677
	UR	-0.4233	0.008	-50.491	0	0.655	-0.445	-0.402
	Floor	0.089	0.007	12.042	0	1.093	0.07	0.108
3	Population	0.1161	0.008	14.81	0	1.123	0.096	0.136
	Crime_Rates	-1.079	0.011	-97.512	0	0.34	-1.108	-1.051
	YearMonth_encoded	2.9254	0.018	161.788	0	18.641	2.879	2.972
	HSP_index	-0.5844	0.016	-35.723	0	0.557	-0.627	-0.542
	SDT_index	0.1711	0.015	11.508	0	1.187	0.133	0.209
	IR	0.4112	0.016	26.41	0	1.509	0.371	0.451

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

APT 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
3	UR	-0.2894	0.009	-31.192	0	0.749	-0.313	-0.266
	Floor	-0.727	0.009	-77.312	0	0.483	-0.751	-0.703
4	Population	-2.9085	0.02	-143.339	0	0.003	-2.961	-2.856
	Crime_Rates	-3.1043	0.023	-136.49	0	0.055	-3.163	-3.046
	YearMonth_encoded	4.6526	0.04	115.655	0	0.045	4.549	4.756
	HSP_index	-0.8101	0.039	-20.993	0	104.855	-0.909	-0.711
	SDT_index	-5.0977	0.038	-132.662	0	0.445	-5.197	-4.999
	IR	-6.3749	0.051	-125.013	0	0.006	-6.506	-6.244
	UR	-3.6309	0.026	-137.941	0	0.002	-3.699	-3.563
	Floor	1.4547	0.016	89.271	0	0.026	1.413	1.497

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

Officetel 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
1	Population	-0.1366	0.007	-18.611	0	0.872	-0.156	-0.118
	Crime_Rates	0.0291	0.007	4.11	0	1.03	0.011	0.047
	YearMonth_encoded	0.4018	0.009	44.858	0	1.495	0.379	0.425
	HSP_index	-0.4685	0.008	-55.121	0	0.626	-0.49	-0.447
	SDT_index	0.088	0.009	9.475	0	1.092	0.064	0.112
	IR	-0.0199	0.011	-1.735	0.083	0.98	-0.05	0.01
	UR	-0.0391	0.008	-5.167	0	0.962	-0.059	-0.02
	Floor	0.0414	0.006	6.397	0	1.042	0.025	0.058
2	Population	-0.252	0.008	-33.144	0	0.777	-0.272	-0.232
	Crime_Rates	-0.2329	0.008	-28.603	0	0.792	-0.254	-0.212
	YearMonth_encoded	0.6472	0.01	61.978	0	1.91	0.62	0.674

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

Officetel 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
2	HSP_index	-1.0736	0.01	-107.96	0	0.342	-1.099	-1.048
	SDT_index	-0.2494	0.01	-25.804	0	0.779	-0.274	-0.224
	IR	-0.7263	0.014	-51.548	0	0.484	-0.763	-0.69
	UR	-0.0367	0.008	-4.599	0	0.964	-0.057	-0.016
	Floor	-0.0858	0.007	-12.129	0	0.918	-0.104	-0.068
3	Population	0.2088	0.008	27.218	0	1.232	0.189	0.229
	Crime_Rates	-0.0909	0.009	-9.943	0	0.913	-0.114	-0.067
	YearMonth_encoded	1.0304	0.011	92.489	0	2.802	1.002	1.059
	HSP_index	-0.3933	0.009	-43.591	0	0.675	-0.416	-0.37
	SDT_index	0.0395	0.01	3.903	0	1.04	0.013	0.066
	IR	-0.2354	0.013	-17.718	0	0.79	-0.27	-0.201



# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

Officetel 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
3	UR	0.1105	0.008	14.271	0	1.117	0.091	0.13
	Floor	0.0197	0.007	2.892	0.004	1.02	0.002	0.037
4	Population	-0.202	0.009	-22.203	0	0.817	-0.225	-0.179
	Crime_Rates	-1.3403	0.016	-84.006	0	0.262	-1.381	-1.299
	YearMonth_encoded	1.7522	0.015	120.505	0	5.767	1.715	1.79
	HSP_index	-0.7354	0.013	-58.641	0	0.479	-0.768	-0.703
	SDT_index	0.377	0.013	28.665	0	1.458	0.343	0.411
	IR	-0.2796	0.02	-13.64	0	0.756	-0.332	-0.227
	UR	0.4174	0.009	45.522	0	1.518	0.394	0.441
	Floor	-0.5651	0.008	-67.73	0	0.568	-0.587	-0.544

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

Townhouse 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
1	Population	0.6509	0.01	65.185	0	1.917	0.625	0.677
	Crime_Rates	0.4193	0.01	44.092	0	1.521	0.395	0.444
	YearMonth_encoded	1.7948	0.023	76.829	0	6.018	1.735	1.855
	HSP_index	-1.0262	0.023	-45.488	0	0.358	-1.084	-0.968
	SDT_index	0.2602	0.013	19.897	0	1.297	0.227	0.294
	IR	0.5075	0.014	35.172	0	1.661	0.47	0.545
	UR	0.3526	0.009	37.37	0	1.423	0.328	0.377
	Floor	0.0725	0.008	9.024	0	1.075	0.052	0.093
2	Population	-7.1568	0.408	-17.539	0	0.001	-8.208	-6.106
	Crime_Rates	6.3095	0.161	39.301	0	549.753	5.896	6.723
	YearMonth_encoded	-20.6099	3.231	-6.379	0	0	-28.933	-12.287

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

Townhouse 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
2	HSP_index	362.7142	17.218	21.065	0	3.35E+157	318.362	407.066
	SDT_index	31.5697	1.11	28.447	0	5.13531E+13	28.711	34.428
	IR	87.375	2.667	32.766	0	8.84E+37	80.506	94.244
	UR	12.3919	0.864	14.342	0	240841.239	10.166	14.617
	Floor	2.6607	0.135	19.763	0	14.307	2.314	3.008
3	Population	-0.6846	0.1	-6.851	0	0.504	-0.942	-0.427
	Crime_Rates	4.4764	0.121	37.011	0	87.917	4.165	4.788
	YearMonth_encoded	117.8526	3.575	32.968	0	1.52E+51	108.645	127.061
	HSP_index	-38.8477	1.708	-22.742	0	0	-43.248	-34.448
	SDT_index	23.8996	0.884	27.044	0	2.39587E+10	21.623	26.176
	IR	10.8796	1.064	10.222	0	53081.064	8.138	13.621

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 통계 분석

Townhouse 전세가별 독립변수들의 통계량

전세가 Class	독립변수	회귀계수	S.E.	Wald	p (유의확률)	OR (오즈비)	99% 신뢰구간	
							하한	상한
3	UR	10.8816	0.681	15.989	0	53190.779	9.129	12.635
	Floor	-12.2244	0.23	-53.041	0	0	-12.818	-11.631

# 부동산 전세가격 예측·전세가율 분석

## 부록 - 모델 생성 및 학습

### Random Forest

- ✓ 다수의 결정 트리를 생성 데이터 일부를 무작위로 선택하여 트리를 구성
- ✓ 부트스트랩 샘플링을 사용하여 각 트리를 학습 모든 트리의 예측을 결합하여 최종 예측을 수행

### XGBoost

- ✓ 그래디언트 부스팅 알고리즘의 한 종류
- ✓ 결정 트리를 기반으로 한 모델
- ✓ 목적 함수를 최소화, 모델의 과적합을 피하도록 도움

### LightGBM

- ✓ 리프 중심(Leaf-Wise) 방식으로 수행
- ✓ 결정 트리를 기반으로 한 모델
- ✓ 조기 종료 기능을 지원
- ✓ 분산 학습 및 병렬 처리를 지원

### Ensemble Model

- ✓ 여러 개의 기본 모델을 조합하여 예측 모델을 만드는 머신 러닝 기법
- ✓ 다수결 원칙에 기반으로 다른 모델을 합쳐서 예측을 수행, 각 모델의 예측 결과를 조합, 최종 예측을 만들
- ✓ 대표적인 앙상블 모델로 배깅(Bagging), 부스팅(Boosting), 스택킹(Stacking)

### KNN

- ✓ 근접 이웃 기반 예측
- ✓ 주어진 데이터 포인트와 이웃 데이터 포인트 간의 거리를 측정, 가장 가까운 이웃을 찾음
- ✓ 이웃의 수(K)를 조절하여 모델의 성능을 조절

## 부록 - 모델 생성 및 학습

### LSTM

- ✓ 순차적으로 나열된 시퀀스 데이터를 처리하는 데 사용.
- ✓ 장기 및 단기 의존성을 효과적으로 학습
- ✓ 게이트라고 불리는 특별한 구조를 사용

### GRU

- ✓ RNN의 장기 의존성 문제를 해결하고자 고안된 순환 신경망 아키텍처
- ✓ LSTM보다 더 간단한 구조
- ✓ 시계열 데이터, 자연어 처리 등 다양한 시퀀스 데이터 처리 작업에 사용

### FNN

- ✓ 뉴런 모델을 기반
- ✓ 핵심은 뉴런 간의 연결 가중치와 활성화 함수
- ✓ 학습 데이터를 사용하여 가중치를 조정하여 원하는 작업을 수행

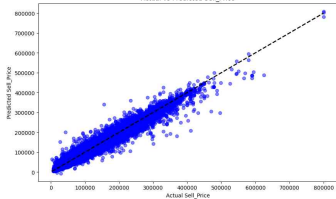
# 부동산 전세가격 예측·전세가율 분석

## 성능 평가 - 모델 성능 검증

### 판매 가격 예측

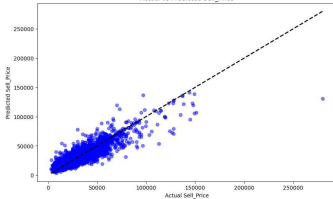
아파트

Actual vs Predicted Sell\_Price



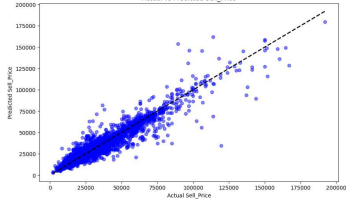
오피스텔

Actual vs Predicted Sell\_Price



연립다세대

Actual vs Predicted Sell\_Price



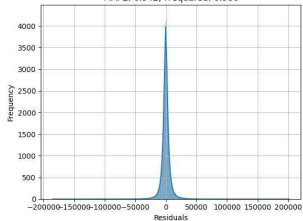
# 부동산 전세가격 예측·전세가율 분석

## 성능 평가 - 모델 성능 검증

### Residuals

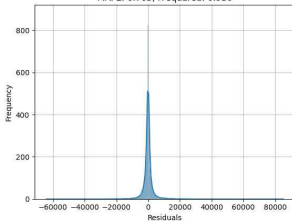
#### 아파트

Histogram of Residuals  
MAPE: 6.942, R-squared: 0.980



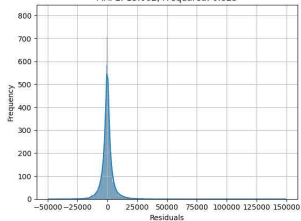
#### 오피스텔

Histogram of Residuals  
MAPE: 6.705, R-squared: 0.930



#### 연립다세대

Histogram of Residuals  
MAPE: 13.062, R-squared: 0.825





# 부동산 전세가격 예측·전세가율 분석

## 부록 - 전세가율 분석

성주한. (2023). 아파트 전세가율의 변동이 아파트 매매가격에 미치는 영향에 관한 연구. Journal of Real Estate Analysis, 9(2), 95-120.

변수		모형 2
		수도권(FE) 모형 Coef.
부동산 특성	L0. 아파트 전세가격	1.019***
	L24. 아파트 전세가격	-0.116***
	아파트 미분양	0.000***
거시경제 특성	회사채수익률	-0.326***
	M2	0.002***
	소비자 물가지수	0.414***
전세가율 변화 특성	L0. 아파트 전세가율	-0.580***
	L24. 아파트 전세가율	0.069***
	L0. 아파트 전세가율 70%대	0.508***
	L24. 아파트 전세가율 70%대	-0.850***
	L0. 아파트 전세가율 80%대	1.237***
	L24. 아파트 전세가율 80%대	-0.145

전세가율 변화 특성에서 수도권은 아파트 매매 가격이 낮고 아파트 전세가격이 높은 상황에서, 현재 아파트 전세가율이 높은 경우 현재 아파트 매매가격이 상승함에 따라 갭투자 가능성은 있는 것으로 파악되었다.

거시경제특성에서 금리인 회사채수익률은 급격히 하락한 시기로서, 수도권에서 금리가 약간 증가하여도 아파트 매매가격에 대해 정(+)의 영향을 미치는 것으로 파악되었다.

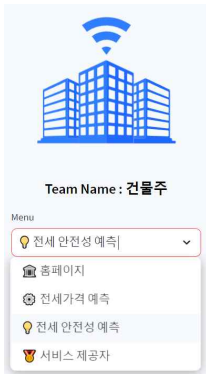
수도권의 경우 현재 아파트 전세가율 70%대인 경우에는 24개월 후 아파트 매매가격이 하락함에 따라 역전세의 가능성이 있다.

# 부동산 전세가격 예측·전세가율 분석

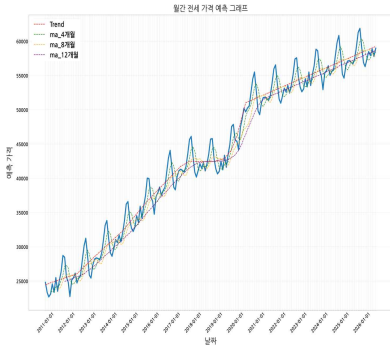
## 부록 - Streamlit 서비스 구현

주요 기능

### Main\_app.py



### Prophet 활용한 전세가격 예측통계



### LightGBM 활용한 전세계약 안전성 평가

주변 인프라를 확인할 기존 도로와 주소를 입력 후 결과를 출력하세요.

서울시 강남구 테헤란로 1194-1

예측하고 싶은 날짜를 입력하세요.

2023/10/01

건물 안전성을 입력하세요.

5

계약하는 월의 월간 전세 가격을 입력하세요. (단위: 원)

60000

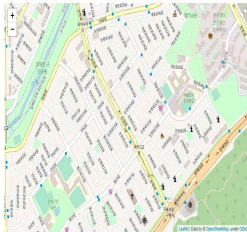
임대 만약을 입력하세요. (단위: 개월)

28

임대 건물의 층수를 입력하세요.

1

예측결과값: 74,352만원 / 전세가율: 80% / 안전성: 위험



# 부동산 전세가격 예측·전세가율 분석

부록 - Streamlit 서비스 구현

홈페이지

가격 예측

안전성 평가

제공자

## 메뉴 구성



Team Name : 건물주

Menu

전세 안전성 예측

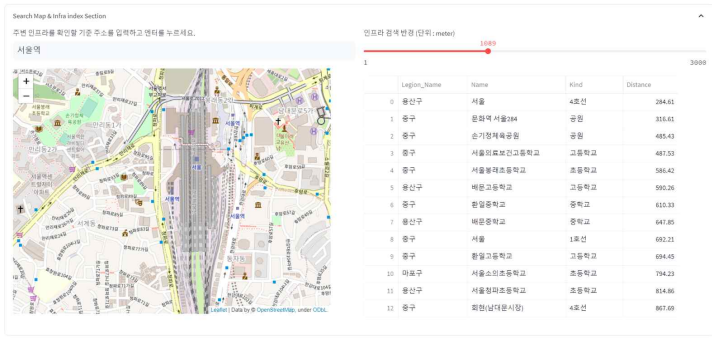
홈페이지

전세가격 예측

전세 안전성 예측

서비스 제공자

## Streamlit / Home / 주변 인프라 탐색 기능 구현



### 메뉴 구성



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

⚙️ 전세가격 예측

💡 전세 안전성 예측

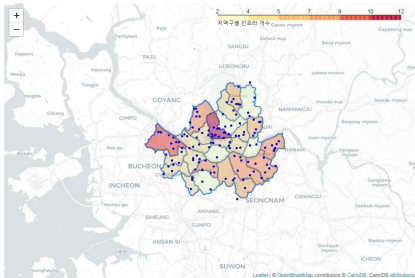
🏆 서비스 제공자

### Streamlit / Home / 자치구 별 인프라 현황 시각화

주변 인프라 탐색 구역

확인할 인프라 지도 탭을 선택하세요.

공원 위치정보



Infra Map Section

확인할 인프라 지도 탭을 선택하세요.

공원 위치정보

공원 위치정보

대학 위치정보

지하철 위치정보

학교 위치정보

### 메뉴 구성



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

📊 전세가격 예측

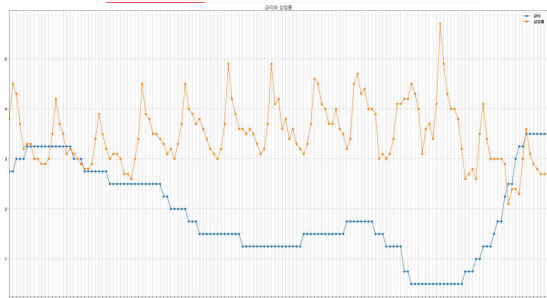
💡 전세 안전성 예측

👤 서비스 제공자

### Streamlit / Home / 경기 변동 추이 시각화

경기변동통량합계구역

경제성장률 및 금리 변동추이   경기종합지수 변동추이   범죄율 변동추이   연립별 인구수 변동추이



### 메뉴 구성



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

⚙️ 전세가격 예측

💡 전세 안전성 예측

🏆 서비스 제공자

### Streamlit / 전세 가격 예측 / 예측 모델 구현

ML\_Predictions\_Section

사용 방법 안내

예측하실 날짜를 조정하기 위해 슬라이더를 이용하세요.

ML to Apartment ML to Office ML to Townhouse

아파트 전세 가격 예측

모델 결과 확인

Option : Month

1

선택한 날짜 값(2011-01-01)의 예측 가격: 24,805 만원

1

189

Model\_Dataframe\_Section

ML to Apartment ML to Office ML to Townhouse

	ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper	additive_terms	additive_terms_lower	additive_terms_upper	yearly	yearly_lower	yearly_upper	multiplicative
0	2011-01-01 08:00:00	24,422.106	-5,488.6188	54,982.1675	24,422.336	24,422.106	383.2341	383.2341	383.2341	383.2341	383.2341	383.2341	
1	2011-02-01 08:00:00	24,537.927	-5,787.8793	52,489.8466	24,537.927	24,537.927	-1,198.8749	-1,198.8749	-1,198.8749	-1,198.8749	-1,198.8749	-1,198.8749	
2	2011-03-01 08:00:00	24,642.5395	-7,570.5183	51,931.0887	24,642.5395	24,642.5395	-1,946.7947	-1,946.7947	-1,946.7947	-1,946.7947	-1,946.7947	-1,946.7947	
3	2011-04-01 08:00:00	24,758.3605	-6,066.9645	53,331.8698	24,758.3605	24,758.3605	-1,651.3739	-1,651.3739	-1,651.3739	-1,651.3739	-1,651.3739	-1,651.3739	
4	2011-05-01 08:00:00	24,843.1774	-3,413.0246	53,353.8796	24,843.1774	24,843.1774	-265.7604	-265.7604	-265.7604	-265.7604	-265.7604	-265.7604	
5	2011-06-01 08:00:00	24,938.8215	-7,471.4863	51,556.5244	24,938.8215	24,938.8215	-1,543.662	-1,543.662	-1,543.662	-1,543.662	-1,543.662	-1,543.662	
6	2011-07-01 08:00:00	25,035.6383	-3,238.6509	53,458.4446	25,035.6383	25,035.6383	477.1782	477.1782	477.1782	477.1782	477.1782	477.1782	
7	2011-08-01 08:00:00	25,163.2824	-4,466.9288	54,313.258	25,163.2824	25,163.2824	-1,587.6736	-1,587.6736	-1,587.6736	-1,587.6736	-1,587.6736	-1,587.6736	
8	2011-09-01 08:00:00	25,199.9265	-6,295.9915	53,561.9239	25,199.9265	25,199.9265	-68.7279	-68.7279	-68.7279	-68.7279	-68.7279	-68.7279	
9	2011-10-01 08:00:00	25,276.5214	-5,628.3326	54,133.7796	25,276.5214	25,276.5214	1,086.3031	1,086.3031	1,086.3031	1,086.3031	1,086.3031	1,086.3031	

### 메뉴 구성



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

⚙️ 전세가격 예측

💡 전세 안전성 예측

🏆 서비스 제공자

### Streamlit / 전세 가격 예측 / 예측 모델 구현

ML\_Predictions\_Section

사용 방법 안내

예측하실 날짜를 조정하기 위해 슬라이드를 이동시키세요.

ML to Apartment ML to OfficeTel ML to Townhouse

아파트 전세 가격 예측

모델 결과 확인

Option : Month

1

1

189

선택한 날짜 값(2011-01-01)의 예측 가격: 24,805 만원

# 부동산 전세가격 예측·전세가율 분석

부록 - Streamlit 서비스 구현

홈페이지

가격 예측

안전성 평가

제공자

## 메뉴 구성

## Streamlit / 전세 가격 예측 / 예측 모델 데이터 프레임화



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

⚙️ 전세가격 예측

💡 전세 안전성 예측

🏆 서비스 제공자

Model\_DataFrame\_Section

ML to Apart\_DF ML to Official\_DF ML to TownHouse\_DF

ds	trend	yhat_lower	yhat_upper	trend_lower	trend_upper	additive_terms	additive_terms_lower	additive_terms_upper	yearly	yearly_lower	yearly_upper	multiplicative_terms	multiplicative_terms_lower	multiplicative_terms_upper	yhat
8 2011-09-01 00:00:00	25,190.9265	-8,295.9425	53,501.9219	25,190.9265	25,190.9265	-65.7279	-65.7279	-65.7279	-65.7279	-65.7279	-65.7279	0	0	0	25,125.1986
9 2011-10-01 00:00:00	25,276.5214	-5,620.3326	54,130.7796	25,276.5214	25,276.5214	1,086.3031	1,086.3031	1,086.3031	1,086.3031	1,086.3031	1,086.3031	0	0	0	26,362.8244
10 2011-11-01 00:00:00	25,364.9594	1,895.4833	58,724.4513	25,364.9594	25,364.9594	3,362.4364	3,362.4364	3,362.4364	3,362.4364	3,362.4364	3,362.4364	0	0	0	28,727.4108
11 2011-12-01 00:00:00	25,450.5643	-3,324.7086	57,238.9949	25,450.5643	25,450.5643	3,022.6953	3,022.6953	3,022.6953	3,022.6953	3,022.6953	3,022.6953	0	0	0	28,473.2337
12 2011-01-01 00:00:00	25,539.0124	-1,450.519	55,580.9123	25,539.0124	25,539.0124	53.5487	53.5487	53.5487	53.5487	53.5487	53.5487	0	0	0	25,592.5811
13 2011-02-01 00:00:00	25,627.4804	-2,714.4206	54,733.2205	25,627.4804	25,627.4804	-779.0732	-779.0732	-779.0732	-779.0732	-779.0732	-779.0732	0	0	0	24,848.3872
14 2011-03-01 00:00:00	25,711.7363	-7,322.8562	52,674.4506	25,711.7363	25,711.7363	-2,965.6463	-2,965.6463	-2,965.6463	-2,965.6463	-2,965.6463	-2,965.6463	0	0	0	22,746.09
15 2011-04-01 00:00:00	25,801.8243	-4,151.5319	52,415.4012	25,801.8243	25,801.8243	-647.7115	-647.7115	-647.7115	-647.7115	-647.7115	-647.7115	0	0	0	25,354.1129
16 2011-05-01 00:00:00	25,889.0963	-3,330.5909	54,428.203	25,889.0963	25,889.0963	-426.2127	-426.2127	-426.2127	-426.2127	-426.2127	-426.2127	0	0	0	25,462.7936
17 2011-06-01 00:00:00	25,970.0943	-1,422.3107	55,812.2295	25,970.0943	25,970.0943	164.8019	164.8019	164.8019	164.8019	164.8019	164.8019	0	0	0	26,143.8952
18 2011-07-01 00:00:00	26,054.7363	-5,000.856	53,400.9164	26,054.7363	26,054.7363	-1,130.7655	-1,130.7655	-1,130.7655	-1,130.7655	-1,130.7655	-1,130.7655	0	0	0	24,738.6307



### 메뉴 구성



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

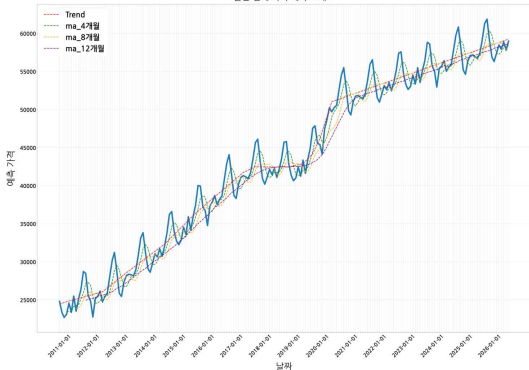
⚙️ 전세가격 예측

💡 전세 안전성 예측

🏆 서비스 제공자

### Streamlit / 전세 가격 예측 / 예측 모델 시각화

월간 전세 가격 예측 그래프



### 메뉴 구성



Team Name : 건물주

Menu

💡 전세 안전성 예측

🏠 홈페이지

⚙️ 전세가격 예측

💡 전세 안전성 예측

🏆 서비스 제공자

### Streamlit / 전세 안전성 예측 / 전세사기 예방 가이드라인

#### 전세사기 예방 가이드라인

	사이트 이름	가이드	사이트 주소
0	건축행정시스템 세움터	계약 전 → 주택상태(불법 무허가주택 여부)확인	<a href="http://cloud.eais.go.kr">http://cloud.eais.go.kr</a>
1	대법원인터넷 등기소	계약 전 → 선순위 권리관계 확인, 근저당 등 권리관계 확인	<a href="http://www.iros.go.kr">http://www.iros.go.kr</a>
2	홈택스	계약 중 → 임대인 세금 체납 여부 확인	<a href="https://www.hometax.go.kr/">https://www.hometax.go.kr/</a>
3	위택스	계약 중 → 임대인 세금 체납 여부 확인	<a href="https://www.wetax.go.kr">https://www.wetax.go.kr</a>
4	부동산거래관리시스템	계약 중 → 주택임대차 표준계약서 확인, 주택임대차 계약 여부 신고하기	<a href="https://rtms.molit.go.kr/">https://rtms.molit.go.kr/</a>
5	대법원인터넷 등기소	계약 중 → 등기 상 권리관계 재확인	<a href="http://www.iros.go.kr">http://www.iros.go.kr</a>
6	정부24	계약 후 → 전입 신고하기	<a href="http://gov.kr">http://gov.kr</a>
7	HUG(주택도시보증공사)	계약 후 → 전세보증금반환보증 보험가입	<a href="https://www.khug.or.kr/index.jsp">https://www.khug.or.kr/index.jsp</a>
8	HF(한국주택금융공사)	계약 후 → 전세보증금반환보증 보험가입	<a href="https://www.hf.go.kr/ko/index.do">https://www.hf.go.kr/ko/index.do</a>
9	SGI(서울보증)	계약 후 → 전세보증금반환보증 보험가입	<a href="https://www.sglc.co.kr/chp/main.mvc">https://www.sglc.co.kr/chp/main.mvc</a>

# 부동산 전세가격 예측·전세가율 분석

## 참고문헌

### 뉴스기사

- 윤종은(2023.07.18) “전세사기 막을 보완 대책 시급”(오마이뉴스) <https://omn.kr/24uhx>
- 심민규(2023.08.02) 무자본 갭투자로 126채 산 뒤 전세보증금 250억 편취한 일당(연합뉴스) <https://www.yna.co.kr/view/AKR20230802045200060>
- 박형기(2023.08.28) 파월은 물론 연준 간부들 잇달아 추가 금리인상 시사(뉴스1) <https://www.news1.kr/articles/5152515>
- 우경희(2023.08.29) 中 부동산위기 진앙지 형태, 상반기만 6조 손실... “주식 거래재개”(머니투데이) <https://news.mt.co.kr/mtview.php?no=2023082809170892576>

### 논문

- Cihan Çilgin, Hadi Gökçen. (2023). Machine learning methods for prediction real estate sales prices in Turkey. Journal of Construction, 22(1), 163-177.
- Aswin Sivam Ravikumar. (2017). Real Estate Price Prediction Using Machine Learning. National College of Ireland.
- Moritz Stang. (2023). Real Estate Valuation in the Age of Artificial Intelligence – Modern Machine Learning Algorithms and their Application in Property Appraisal. IREBS.
- Ping-Feng Pai and Wen-Chang Wang. (2020). Using Machine Learning Models and Actual Transaction Data for Predicting Real Estate Prices. MDPI, 10, 5832.
- 성주한. (2023). 아파트 전세가율의 변동이 아파트 매매가격에 미치는 영향에 관한 연구. Journal of Real Estate Analysis, 9(2), 95-120.
- 김학현, 유한규, and 오하영. (2023). 딥러닝과 머신러닝을 이용한 아파트 실거래가 예측. 정보처리학회논문지. 소프트웨어 및 데이터 공학, 12(2), 59-76.
- Yadav, S., Dhanda, N., Sahai, A., Verma, R., Pandey, S. (2023). Real Estate Price Prediction Using Machine Learning. In: Sarkar, D.K., Sadhu, P.K., Bhunia, S., Samanta, J., Paul, S. (eds) Proceedings of the 4th International Conference on Communication, Devices and Computing. ICCDC 2023. Lecture Notes in Electrical Engineering, vol 1046. Springer, Singapore. [https://doi.org/10.1007/978-981-99-2710-4\\_9](https://doi.org/10.1007/978-981-99-2710-4_9)
- 고주형 and 강명규. (2019). 부동산 가격 요인과 가격상승률 요인 비교 연구: 서울시 재건축 아파트를 중심으로. 부동산학연구, 25(2), 7-22.
- 최남진. (2023). 초과 유통성과 전국 아파트 가격 간의 관계에 대한 연구. 부동산분석, 9(1), 195-210.
- 김은미. (2022). 거시경제변수가 지역 별 아파트 전세가격에 미치는 영향 및 예측모델 구축에 관한 연구. 지적과 국토정보, 52(2), 211-231.
- M Sri Rangan, W. (2023) *Creating accurate valuation models for real estate properties*.
- Prof. Dr. Sven Bienrt 등 16명 (2023) Real Estate Valuation in the Age of Artificial Intelligence - Modern Machine Learning Algorithms and their Application in Property Appraisal

# 부동산 전세가격 예측·전세가율 분석

## 출처

- 남녀 픽토그램 [https://kr.freepik.com/free-vector/flat-design-male-female-symbols\\_22340731.htm#query=%EB%82%A8%EB%85%80%20%ED%94%BD%ED%86%A0%EA%B7%B8%EB%9F%A8&position=15&from\\_view=keyword&track=ais](https://kr.freepik.com/free-vector/flat-design-male-female-symbols_22340731.htm#query=%EB%82%A8%EB%85%80%20%ED%94%BD%ED%86%A0%EA%B7%B8%EB%9F%A8&position=15&from_view=keyword&track=ais)
- 많은 문제 픽토그램 <https://pixabay.com/ko/vectors/%EC%84%B8%EC%85%98-%EA%B3%BC%ED%95%99-%ED%94%BD%ED%86%A0%EA%B7%B8%EB%9F%A8-%ED%94%BC%EB%A1%9C-1989711/>
- 종이 자르기 픽토그램 [https://www.flaticon.com/kr/free-icon/cut-paper\\_4224992?term=%EC%A2%85%EC%9D%B4+%EC%9F%90%EB%A5%B4%EA%B8%B0&page=1&position=2&origin=search&related\\_id=4224992](https://www.flaticon.com/kr/free-icon/cut-paper_4224992?term=%EC%A2%85%EC%9D%B4+%EC%9F%90%EB%A5%B4%EA%B8%B0&page=1&position=2&origin=search&related_id=4224992)
- 주소 픽토그램 [https://www.flaticon.com/kr/free-icon/location\\_4985945?term=%EC%A3%BC%EC%86%8C&page=1&position=18&origin=search&related\\_id=4985945](https://www.flaticon.com/kr/free-icon/location_4985945?term=%EC%A3%BC%EC%86%8C&page=1&position=18&origin=search&related_id=4985945)
- 병합 픽토그램 [https://www.flaticon.com/kr/free-icon/merge\\_1831973?term=%EB%B3%91%ED%95%A9&page=1&position=1&origin=search&related\\_id=1831973](https://www.flaticon.com/kr/free-icon/merge_1831973?term=%EB%B3%91%ED%95%A9&page=1&position=1&origin=search&related_id=1831973)
- 예측 픽토그램 [https://www.flaticon.com/kr/free-icon/prediction\\_5856929?term=%EC%98%88%EC%B8%A1&related\\_id=5856929](https://www.flaticon.com/kr/free-icon/prediction_5856929?term=%EC%98%88%EC%B8%A1&related_id=5856929)
- 모델링 픽토그램 [https://www.flaticon.com/kr/free-icon/3d-modeling\\_3112363?term=%EB%AA%A8%EB%8D%88%EB%A7%81&page=1&position=2&origin=search&related\\_id=3112363](https://www.flaticon.com/kr/free-icon/3d-modeling_3112363?term=%EB%AA%A8%EB%8D%88%EB%A7%81&page=1&position=2&origin=search&related_id=3112363)
- 인프라 픽토그램 [https://www.flaticon.com/kr/free-icon/infrastructures\\_11930238?term=%EC%9D%B8%ED%94%84%EB%9D%BC&page=1&position=12&origin=search&related\\_id=11930238](https://www.flaticon.com/kr/free-icon/infrastructures_11930238?term=%EC%9D%B8%ED%94%84%EB%9D%BC&page=1&position=12&origin=search&related_id=11930238)
- ppt 템플릿 <http://pptbizcam.co.kr/?p=6897>
- 서울 전세 실거래: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/OA-21276/S/1/datasetView.do>)
- 서울 매매 실거래: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/OA-21275/S/1/datasetView.do>)
- 기준 금리: 한국은행 경제통계시스템 (<https://ecos.bok.or.kr/#/SearchStat>)
- 실업률:서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/59/S/2/datasetView.do>)
- 전세수급동향: 공공데이터포털 (<https://www.data.go.kr/data/15102751/fileData.do>)
- 지하철역 위치: 레일포털 (<https://data.kric.go.kr/rips/>)
- 학교(초, 중, 고등학교) 위치: 학구도안내서비스 (<https://schoolzone.emac.kr/publicData/publicDataView.do>)
- 대학 위치: 공공데이터포털 (<https://www.data.go.kr/data/15052369/fileData.do>)
- 인구수: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/10043/S/2/datasetView.do?stcSrl=10043>)
- 경기종합지수: 국가통계포털 ([https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT\\_1C8015&vw\\_cd=MT\\_ZTITLE&list\\_id=J1\\_1&seqNo=8&lang\\_mode=ko&language=kor&obj\\_var\\_id=&itm\\_id=&conn\\_path=MT\\_ZTITLE](https://kosis.kr/statHtml/statHtml.do?orgId=101&tblId=DT_1C8015&vw_cd=MT_ZTITLE&list_id=J1_1&seqNo=8&lang_mode=ko&language=kor&obj_var_id=&itm_id=&conn_path=MT_ZTITLE))
- 서울 5대 범죄: 서울 열린 데이터 광장 (<https://data.seoul.go.kr/dataList/316/S/2/datasetView.do>)