아파트와 빌라의 개별 가격 변동성 특성 비교 - AutoML기법을 활용하여

민병철

HF 한국주택금융공사
주택금융연구원

국민의 주거 행복을 책임지는 지속가능 주택금융의 선도기관

분야	주택시장

아파트와 빌라의 개별 가격 변동성 특성 비교 - AutoML 기법을 활용하여

2023. 11

한국주택금융공사 주택금융연구원

■ 저 자 : 민병철 (051-663-8177 / bcmin@hf.go.kr)

■ 본고의 내용은 필자의 개인 의견으로 한국주택금융공사의 공식적인 견해와 다를 수 있습니다.

목 차

1. 연구배경 및 목적 1

2. 선행연구 2

3. 연구방법 및 사용자료 3

4. 분석결과 6

4.1 모형의 설명력 6

4.2 주요 변수의 설명력 비교 10

4.3 매매가, 전세가의 변동률 분포 특성 17

4.4 매매가, 전세가의 결합 분포 특성 분석 27

5. 요약 및 향후 활용 방안 30

참 고 문 헌 33

요 약

- 이 연구에서는 머신러닝 기법을 이용해서 아파트 및 빌라가격을 추정하고, 이를 이용해 가격 변화율 분포를 만들어 그 특성을 분석함

- 거래 정보, 주택 정보, 지리 정보들을 결합하여 모형을 훈련시킨 결과, 아파트 매매의 경우 시험군 MAPE가 4.5%수준이었으며, 아파트 전세 및 빌라 매매의 경우 약 9.3% 수준이었음

- 분석결과, 빌라 매매가격의 변동성보다 아파트 매매가격의 변동성이 더 큰 것으로 나타났으며, 아파트 전세가격의 변화율은 개별 주택간 편차가 큰 편인 것으로 분석됨

- 체결된 전세건에 대하여 전세가율을 분석한 결과, 아파트에 비해 빌라의 전세가율이 높은 것으로 보이며, 특히 2016년과 2021년에 전세가율이 90%를 넘는 주택의 비중이 급등한 것으로 보임

- 전세가율과 향후 2년간 가격 변동 변동률을 살펴보면, 2011-2013년에는 전세가율이 높을수록 향후 2년간 실현된 매매가격 상승률이 높았으나, 그 이후로는 전세가율이 낮을수록 매매가격 상승률이 높았음

  - 다만, 수도권, 특히 서울의 빌라의 경우 대부분의 시기에서 전세가율이 높을수록 매매가격 상승률이 높았음

- 매매가격과 전세가격 간 결합분포를 분석한 결과, 전세가격이 하락할수록 매매가격이 큰 폭 하락할 비중은 급격하게 증가하는 것으로 계산됨

  - 전세가격이 0~10% 하락할 때는 매매가격이 20% 이상 하락하는 비중이 3.18%였지만, 전세가격이 10%~20%하락할 때는 10.31%로 세 배 이상 증가함

- 이 연구의 내용은 빌라 시장에서의 깡통전세 위험 분석에 활용될 수 있으며, 아파트의 경우 역전세 시 깡통전세 발생 위험 계산에 쓰일 수 있음

I. 연구 배경 및 목적

- 아직 거래가 일어나지 않아 가치를 제대로 평가하기 어려운 신축 빌라에서 과도하게 높은 전세보증금을 받아 깡통전세 피해가 발생하는 사례가 최근 빈번하게 발생함

- 또한, 역전세 및 깡통전세 위험을 제대로 평가하기 위해서는 주택가격 변화율 분포가 필요한데, 이를 위해서는 과거 각 시점의 주택 가격을 적절히 평가할 수 있어야 함
  - 관련 내용은 민병철(2021) 참조

- 주택 가격 추정은 감정평가사가 실사 후 비용접근법(cost approach), 시장접근법(market approach), 수익접근법(income approach)등을 통해 산정하는 방법과 대량평가모형(mass appraisal model)을 통해 통계적으로 추정하는 방식이 있음

- 이 중 감정평가사를 통해 주택가격을 추정하는 방식은 인력과 이에 따르는 비용이 수반되므로 간단하게 주택 가치를 추정하기 위해서는 대량평가모형이 용이할 수 있는데, 이에 관련된 대부분의 연구는 아파트에 집중되고 있음

- 특히, 아파트의 경우는 거래가 발생하면 동일 단지, 동일 면적, 그리고 비슷한 층에 있는 집들은 유사한 시장가치를 지닐 것이라고 가정해도 무리가 없기 때문에 가치평가가 손쉬움

- 하지만 빌라의 경우 세대수가 많지 않아 동일 주택, 동일 면적에 대한 거래빈도가 매우 낮으며, 특히 신축 빌라의 경우 이러한 방법을 적용하기가 더더욱 어려움

- 이 연구에서는 실거래 기록이 없는 경우에도 머신러닝 기법을 통해 빌라

가격을 추정하고, 이를 이용해 주택가격 변화율 분포의 특성을 살펴봄

II. 선행연구

□ 전통적으로 사용된 부동산 평가 모형으로는 헤도닉 모형을 꼽을 수 있는데, 이는 주택 가격을 설명변수들의 선형 회귀식으로 표현함

□ 하지만 이는 변수들간의 관계 및 비선형적 영향력, 입지조건 등의 실제 부동산 시장이 갖는 복잡한 본질을 포착하기에는 충분하지 못하다는 지적을 피하기 어려움

□ 이에 머신러닝 기법을 이용한 부동산 평가모형은 다양하게 연구되어 왔으나, 이에 대한 국내 학술논문은 그 수가 많지 않음

- 대표적인 해외 논문으로는 Limsombunchai(2004), Selim(2009), Antipov and Pokryshevskaya(2012), Mu. et. al.(2014)등이 있음
- 국내에서는 민간기업 혹은 한국부동산원 등의 기관에서 자체적으로 모형을 개발하여 활용하고 있으나, 이를 외부에 공개하지는 않고 있음
- 국내의 대표적인 연구로는 이창로, 박기호(2016), 조유나, 김수현, 송규원(2017), Hong. et. al(2020), 홍정의(2021) 등을 꼽을 수 있음

□ 하지만 머신러닝 기법을 이용한 국내 선행연구는 대부분 아파트 시장에 국한되어, 연립·다세대 주택 및 다가구 주택을 포괄하는 빌라 시장에 대한 연구는 찾기 어려움

- 예외적으로, 이창로, 박기호(2016)은 기계학습 모형을 이용하여 단독주택 가격 추정을 시도하였음
- 또한, 김국현(2009), 그리고 김두환 외(2010)는 기계학습 기법은 아니지만 전통적 헤도닉 기법을 이용하여 국내 강남지역의 고급빌라 가격에 영향을 미치는 요인을 분석하였음

- 민간 프롭테크 기업들에서는 자체적으로 빌라 가격 추정모형을 훈련시켜 상업적 목적으로 이용하고 있는 것으로 알려짐

□ 이 연구에서는 실거래 데이터, 건축물대장, 그리고 경도 및 위도 정보만을 결합하되 다양한 머신러닝 모형을 결합한 앙상블 모형을 훈련시킨 후, 이를 이용하여 주택가격 변화율 분포의 특성을 분석함을 목적으로 함

- 역전세 및 깡통전세의 위험을 분석하는데 있어 가격지수가 아닌 개별 주택 단위의 가격 변화율 분포를 사용한 연구로 민병철(2020, 2021)이 있는데, 이 연구는 빌라가 아닌 아파트로 분석대상을 한정하였다는 한계가 있음

III. 연구방법 및 사용자료

1. 연구방법

□ 이 연구에서는 AutoML 툴 중 하나인 Autogluon을 이용하여 아파트, 연립•다세대의 매매가/전세가를 설명하는 모형을 훈련시킴

- AutoML이란 머신러닝 모형 훈련 과정을 자동화해주는 프로그램으로, 다양한 모형의 훈련, 하이퍼파라미터 선택, 배깅(bagging), 스태킹(stacking), 앙상블(ensemble) 모형의 훈련까지 스스로 해 줌

□ 이를 이용하여 시·도별로 가격을 설명하는 모형을 훈련시킴

- 전체 데이터를 한번에 이용하여 훈련시키되 시·도를 나타내는 더미변수를 포함시킨 경우와 비교하여 시·도별로 데이터셋을 나누어서 모형을 훈련시킨 경우의 설명력이 더 높게 나타남
  - 트리 기반 모형의 경우 이론적으로는 시·도별 더미변수를 이용하여 데이터셋을 나누어 모형을 훈련시킬 수 있기 때문에 굳이 데이터셋을 나눌 필요가 없을 수 있으나, 실제로는 과최적화를 방지하는 과정에서 트리의 깊이를 제한하기 때문에 시도별로 데이터셋을 나누는 단계가 추가되는 것이 설명력에 영향을 미칠 수 있음

□ **훈련된 가격예측모형을 이용해, 실제 거래된 건들의 2년 후 가격을 예측하고 이를 이용해 가격 변화율 분포를 추정함**

- 예측모형을 토대로 가격변화율을 추정하기 때문에 실제 거래된 기록이 없더라도 가격변화율을 추정 가능

- 실제 거래가 존재했었던 경우, 해당 표본이 훈련에 사용될 수 있기 때문에 예측치가 좀 더 정확할 가능성 높음

- 실제 거래의 경우 이상 거래라든지 지인 간 거래 등으로 인해 실제 가치보다 높거나 낮은 가격으로 거래될 수 있는데, 예측모형의 예측치를 이용할 경우 보다 실제 가치에 가까운 가격을 얻을 가능성이 있음

□ **전세 거래의 경우, 해당 거래 기록에 대하여 매매가격 및 2년 후 매매가격을 추정하여 전세가율을 추가적으로 추정하고 전세가격과 매매가격의 결합분포를 도출함**

- 편의상 전세 거래라고 표현했지만, 아파트의 경우 표본 확대를 위해 보증금 수준이 240개월치 월세를 초과하는 반전세 전체를 포함시킴
  - 이후 편의상 전세 및 반전세를 합하여 전세라 부름

2. 사용자료

□ **이 연구에서는 국토교통부에서 공개하는 실거래가 자료를 건축물대장, kakao API의 경도•위도 정보와 결합하여 사용함**

- 아파트, 연립•다세대 주택의 매매 및 임차에 대한 실거래 자료를 활용함

- 건축물대장 자료 출처: 건축데이터 민간개방 시스템 (https://open.eais.go.kr)

- 기간: 2011년 1월-2022년 12월

□ **실거래가 자료 사용 정보: 계약년월, 주소, 거래금액, 전용면적, 층, 건축**

년도

- 전월세 거래의 경우 거래금액으로 보증금 값과 월세 값을 사용함
- 최종적으로 설명하고자 하는 변수로는 거래금액이 아닌 거래금액/전용면적을 사용함
  - 전월세 거래의 경우에는 보증금/전용면적 변수를 이용
- 전월세의 경우 월세/전용면적 변수를 추가함

□ **건축물대장 자료 사용 정보: 주소, 세대수, 옥내기계식면적, 옥외기계식면적, 옥내자주식면적, 옥외자주식면적, 구조코드, 기타_구조, 지붕코드, 기타_지붕, 지상층수, (지하층수), 승용승강기수, 비상용승강기수, 부속건축물수, 내진설계여부**

- 옥내기계식면적, 옥외기계식면적, 옥내자주식면적, 옥외자주식면적은 모두 합쳐서 주차장면적이라는 변수를 만듦
- 여러 개의 동으로 이루어져있지만 같은 주소로 되어 있는 경우에는 세대수, 주차장면적, 승강기 수 등을 모두 합침
- 주차장면적, 승강기 수는 총 세대수로 나누어 세대당 주차장면적, 세대당 승강기 수 정보를 사용함
- 지하층수는 연립•다세대의 경우에만 사용

□ **Kakao API 사용정보: 주소, 경도, 위도**

□ **건축물대장 정보 중 일부 부정확한 정보로 이해되는 표본은 제외함**

- 세대수나 층 수가 0이거나, 사용한 변수들의 값이 없는 경우

□ **주소를 기준으로 결합하되, 세 가지 자료 모두 존재하는 표본만 사용함**

□ **시군구·연도별로 가격/전용면적 비율의 상·하위 1%를 제거한 표본을**

분석에 사용함

- 임대차거래의 경우 보증금/전용면적 비율의 상·하위 1% 제거

IV. 분석결과

1. 모형의 설명력

□ 설명력의 기준으로는 MAPE(Mean absolute percentage error)을 이용함

- $MAPE = (\sum_i \frac{|실제값_i - 예측치_i|}{실제값_i})\frac{100}{n}$

□ 아마트 매매시장에 대한 시험군의 MAPE는 4.45%, 아파트 전세에 대한 MAPE는 9.28%, 연립·다세대 매매에 대한 MAPE는 9.25% 수준임

- 이는 지역별로 훈련된 모형들의 예측치들을 모두 모은 전체 데이터에 대한 예측력임

- 일반적으로 쓰이는 빌라라는 표현은 연립·다세대와 다가구 주택을 아우르는 표현이나, 여기부터는 편의상 연립·다세대를 빌라로 통칭함

〈표 1〉 시장별 예측모형의 설명력

*	훈련군 MAPE	시험군 MAPE	표본수
아파트 매매	3.34%	4.45%	5,005,767
아파트 전세	6.83%	9.28%	4,091,363
빌라 매매	5.25%	9.25%	1,387,041

□ 연립·다세대의 전세 가격에 대한 예측 모형은 설명력이 매우 낮아 본 보고서에서는 추가적으로 다루지 않음

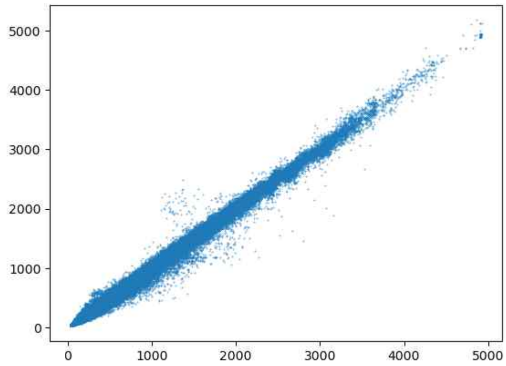
- 훈련군, 시험군을 통합한 데이터셋의 MAPE가 31% 수준으로, 분석 대상으로 삼기 어려움

- **사실상 동일한 변수들을 이용한 아파트의 매매/전세, 연립·다세대의 매매 가격 예측 모형과 비교해 설명력이 매우 낮다는 것은 연립·다세대의 전세 가격은 다른 주택시장부문을 설명하는데 충분했던 변수들만으로는 설명되지 않는다는 것을 의미함**
  - 이는 해당 시장이 다른 시장들에서는 큰 의미가 없었던 다른 변수들에 의해 영향을 크게 받는다는 것을 의미함
  - 또한, 연립·다세대의 전세 중에서 이상거래 및 비합리적인 가격으로 거래된 건들이 많을 때도 이러한 설명력 하락이 발생할 가능성이 있음

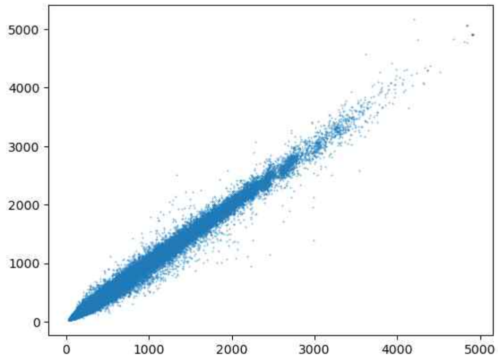
- **<표 2>는 x축에 실제 값, y축에 예측치를 나타낸 그림으로, 대체로 일관성있는 예측 정확도를 보이는 것으로 나타남**

<표 2> 실제 값과 예측치 분포도

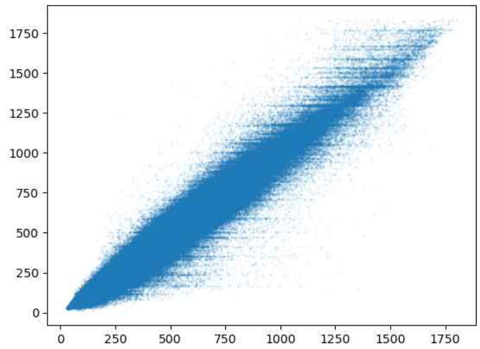
아파트 매매 훈련군



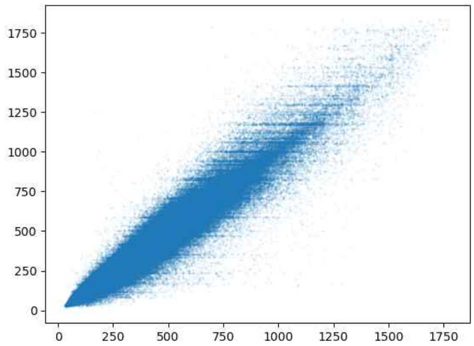
아파트 매매 시험군



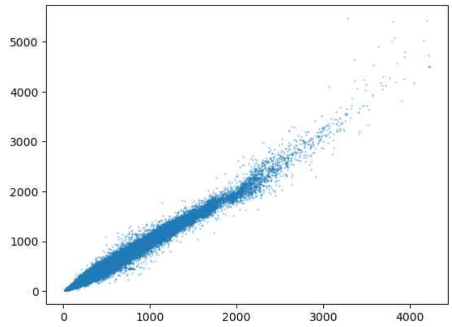
아파트 전세 훈련군



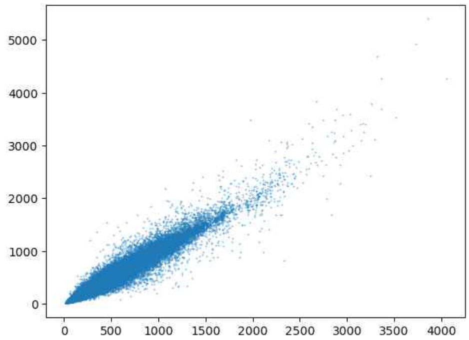
아파트 전세 시험군



빌라 매매 훈련군



빌라 매매 시험군



※ x축은 실제값, y축은 예측치를 나타냄

- **지역별로 보면, 대체로 수도권, 5대광역시에 대한 모형의 설명력이 높은 편이며 도 지역에 대한 모형은 낮은 설명력을 나타냄**
  - 아파트 매매가격 예측모형의 경우, 서울특별시, 경기도, 인천광역시에 대한 모형의 시험군 MAPE는 각각 3.48%, 3.75%, 3.96%로 가장 낮았음
  - 반면 충청북도, 제주특별자치도, 경상북도에 대한 시험군 MAPE는 각각 6.34%, 6.50%, 5.86%로 상대적으로 높았음

- 인천의 경우 특이하게도 아파트 매매가격 모형에 비해 전세가격 및 빌라 매매가격 모형의 설명력이 매우 떨어짐

〈표 3〉 아파트 매매가격 예측모형의 지역별 설명력

*	훈련군 MAPE	시험군 MAPE	표본수
강원도	4.12%	5.41%	215,988
경기도	2.80%	3.75%	956,386
경상남도	3.95%	5.13%	320,195
경상북도	4.43%	5.86%	234,415
광주광역시	3.74%	4.74%	262,270
대구광역시	3.08%	4.13%	352,258
대전광역시	3.13%	4.09%	213,758
부산광역시	3.45%	4.84%	459,291
서울특별시	2.53%	3.48%	799,967
울산광역시	3.65%	5.00%	173,657
인천광역시	2.99%	3.96%	385,912
전라남도	4.54%	5.66%	201,497
전라북도	4.45%	5.79%	120,069
제주특별자치도	4.00%	6.50%	38,963
충청남도	3.58%	4.69%	182,051
충청북도	4.96%	6.34%	89,090

〈표 4〉 아파트 전세가격 예측모형의 지역별 설명력

*	훈련군 MAPE	시험군 MAPE	표본수
강원도	7.26%	9.96%	94,110
경기도	7.96%	9.96%	921,948
경상남도	6.29%	9.08%	158,330
경상북도	5.59%	8.49%	119,280
광주광역시	5.93%	8.44%	117,592
대구광역시	6.32%	8.93%	182,016
대전광역시	6.80%	8.84%	155,558
부산광역시	6.04%	8.94%	278,633
서울특별시	6.29%	8.78%	1,367,807
울산광역시	6.78%	9.83%	76,231
인천광역시	7.53%	10.19%	310,818
전라남도	6.17%	8.43%	99,300
전라북도	6.36%	8.93%	53,481
제주특별자치도	7.70%	10.91%	17,893
충청남도	8.11%	10.65%	98,838
충청북도	7.87%	11.20%	39,801

〈표 5〉 빌라 매매가격 예측모형의 지역별 설명력

*	훈련군 MAPE	시험군 MAPE	표본수
강원도	6.9%	11.9%	19,408
경기도	5.1%	8.6%	265,853
경상남도	5.0%	9.6%	27,740
경상북도	4.3%	9.4%	35,603
광주광역시	3.1%	6.8%	12,312
대구광역시	5.4%	10.7%	41,515
대전광역시	5.8%	11.7%	32,179
부산광역시	4.0%	8.4%	142,424
서울특별시	5.1%	8.5%	468,674
울산광역시	4.9%	10.0%	20,298
인천광역시	6.9%	11.2%	214,846
전라남도	5.7%	10.4%	17,828
전라북도	4.9%	9.9%	11,020
제주특별자치도	4.1%	8.5%	35,834
충청남도	5.3%	10.7%	26,141
충청북도	5.4%	10.3%	15,366

2. 주요 변수의 설명력 비교

□ 지역에 따라 가격을 설명하는데 사용되는 변수들의 중요도가 조금씩 다른 것으로 나타남

- 여기서 중요도는 순열중요도(Permutation Importance)로, 각 변수에 무작위 값을 부여했을 때 하락하는 설명력의 크기를 나타냄

□ 대부분의 지역에서 계약년월 혹은 건축년도가 가장 중요한 변수로 나타나, 시기에 따른 전체적인 가격 수준 및 주택의 노후 정도가 가격 차이에 가장 큰 영향을 주는 것으로 보임

- 대체로 대도시에서는 계약년월이 가장 중요한 변수로 나타나며, 도 지역에서는 건축년도가 가장 중요한 변수로 나타남

□ 시군구코드, 경도, 위도 변수 또한 대부분의 지역에서 중요하게 평가되어, 입지 정보 또한 중요한 변수임을 알 수 있음

□ 특이하게도, 서울의 경우 건축년도가 크게 중요하지 않게 나타나고 계약년월

및 입지정보가 가격에 절대적인 영향을 미치는 것으로 보임

- 전 지역에서 유일하게 건축년도 정보가 5순위 변수에도 들지 못함

□ 경기도의 경우 시군구 정보가 가장 중요한 것으로 나타났는데, 이는 같은 경기도라도 시군구별로 가격 수준 및 가격 특성의 차이가 매우 크다는 점을 시사함

〈표 6〉 아파트 매매가격 예측모형의 지역별 변수 중요도

*	1번		2번		3번		4번		5번	
	변수	중요도	변수	중요도	변수	중요도	변수	중요도	변수	중요도
강원	건축년도	0.15	계약년월	0.14	시군구	0.08	지붕유형	0.06	건축면적	0.04
경기	시군구	0.21	계약년월	0.17	건축년도	0.12	위도	0.12	전용면적	0.06
경남	건축년도	0.18	계약년월	0.09	위도	0.07	시군구	0.06	지붕유형	0.04
경북	건축년도	0.24	계약년월	0.12	위도	0.08	시군구	0.06	지붕유형	0.06
광주	건축년도	0.20	계약년월	0.19	시군구	0.04	경도	0.04	전용면적	0.03
대구	계약년월	0.23	건축년도	0.15	시군구	0.07	위도	0.07	경도	0.05
대전	계약년월	0.15	건축년도	0.09	위도	0.05	시군구	0.05	건축면적	0.05
부산	계약년월	0.15	건축년도	0.13	경도	0.08	시군구	0.07	지상층수	0.07
서울	계약년월	0.30	시군구	0.23	위도	0.11	전용면적	0.08	경도	0.07
울산	건축년도	0.16	계약년월	0.13	시군구	0.10	세대수	0.05	위도	0.05
인천	계약년월	0.16	건축년도	0.14	전용면적	0.06	시군구	0.05	세대수	0.05
전남	건축년도	0.25	시군구	0.12	계약년월	0.11	경도	0.05	지붕유형	0.04
전북	건축년도	0.19	계약년월	0.08	시군구	0.06	지상층수	0.06	지붕유형	0.05
제주	계약년월	0.27	세대수	0.12	건축년도	0.08	지상층수	0.05	지붕유형	0.04
충남	건축년도	0.19	계약년월	0.09	지상층수	0.07	지붕유형	0.05	전용면적	0.05
충북	건축년도	0.21	계약년월	0.10	전용면적	0.05	지상층수	0.04	시군구	0.04

<표 7> 아파트 전세가격 예측모형의 지역별 변수 중요도

*	1번		2번		3번		4번		5번	
	변수	중요도	변수	중요도	변수	중요도	변수	중요도	변수	중요도
강원	계약년월	0.22	건축년도	0.08	지붕유형	0.06	시군구	0.05	전용면적	0.04
경기	계약년월	0.23	위도	0.10	시군구	0.09	전용면적	0.06	건축년도	0.06
경남	계약년월	0.14	건축년도	0.11	위도	0.05	월세비	0.05	지상층수	0.05
경북	건축년도	0.18	계약년월	0.15	지붕유형	0.06	시군구	0.05	지상층수	0.04
광주	건축년도	0.16	계약년월	0.16	월세비	0.06	경도	0.03	지상층수	0.03
대구	계약년월	0.19	건축년도	0.10	위도	0.06	경도	0.05	시군구	0.05
대전	계약년월	0.17	건축년도	0.06	위도	0.05	지상층수	0.04	월세비	0.04
부산	계약년월	0.12	건축년도	0.12	월세비	0.05	지상층수	0.05	전용면적	0.04
서울	계약년월	0.24	시군구	0.16	월세비	0.09	전용면적	0.08	경도	0.08
울산	계약년월	0.13	건축년도	0.11	시군구	0.06	위도	0.05	경도	0.04
인천	계약년월	0.28	건축년도	0.08	전용면적	0.06	시군구	0.05	경도	0.05
전남	건축년도	0.23	계약년월	0.14	시군구	0.08	월세비	0.04	지붕유형	0.04
전북	건축년도	0.12	계약년월	0.12	월세비	0.07	지상층수	0.05	시군구	0.04
제주	계약년월	0.18	경도	0.06	지붕유형	0.06	지상층수	0.05	건축년도	0.04
충남	계약년월	0.15	건축년도	0.09	지상층수	0.07	전용면적	0.06	지붕유형	0.04
충북	건축년도	0.15	계약년월	0.13	지붕유형	0.06	전용면적	0.05	지상층수	0.05

□ 빌라 매매가격의 경우 계약시점보다 건축년도 및 입지요건이 절대적인 요인으로 보임

- 대부분의 지역에서 건축년도 변수가 가장 중요한 설명력을 가짐

<표 8> 빌라 매매가격 예측모형의 지역별 변수 중요도

*	1번		2번		3번		4번		5번	
	변수	중요도	변수	중요도	변수	중요도	변수	중요도	변수	중요도
강원	건축년도	0.32	계약년월	0.10	시군구	0.10	건축면적	0.08	경도	0.07
경기	위도	0.24	건축년도	0.17	시군구	0.16	경도	0.11	계약년월	0.07
경남	건축년도	0.19	위도	0.07	계약년월	0.05	경도	0.04	시군구	0.04
경북	건축년도	0.32	위도	0.08	계약년월	0.07	시군구	0.04	연식	0.04
광주	건축년도	0.28	계약년월	0.28	건축면적	0.10	위도	0.07	연식	0.05
대구	계약년월	0.16	건축년도	0.08	위도	0.06	경도	0.05	시군구	0.04
대전	건축년도	0.18	계약년월	0.12	위도	0.11	경도	0.08	시군구	0.06
부산	건축년도	0.22	계약년월	0.19	경도	0.10	전용면적	0.10	위도	0.10
서울	시군구	0.17	위도	0.14	건축년도	0.14	경도	0.13	계약년월	0.12
울산	건축년도	0.14	계약년월	0.10	시군구	0.08	경도	0.06	위도	0.06
인천	계약년월	0.08	건축년도	0.08	경도	0.04	층	0.04	위도	0.04
전남	건축년도	0.50	시군구	0.06	연식	0.05	계약년월	0.05	경도	0.04
전북	건축년도	0.40	연식	0.06	건축면적	0.05	위도	0.04	계약년월	0.04
제주	건축년도	0.32	계약년월	0.19	경도	0.07	건축면적	0.06	위도	0.03
충남	건축년도	0.29	전용면적	0.05	위도	0.04	시군구	0.04	내진설계 여부	0.04
충북	건축년도	0.26	전용면적	0.04	연식	0.04	내진설계 여부	0.04	계약년월	0.04

□ **머신러닝 기법을 이용했을 때의 장점은 입지가 가격에 미치는 영향을 적은 개수의 변수로(이 연구에서는 시군구코드, 경도, 위도 변수 사용) 표현할 수 있다는 점임**

- 이론적으로는 경도, 위도의 두 가지 변수에 모든 입지 정보가 포함될 수 있으나, 실제로는 시군구코드를 포함시켰을 때 모형의 설명력이 증가함

□ **주요 변수들은 고정시킨 채 시군구코드, 경도, 위도 변수만을 변화시켜서 학습시킨 모형이 입지에 따른 가격 변화를 어떻게 이해하고 있는지를 표현해 볼 수 있음(<그림 1> ~ <그림 5>)**

- 분석기간 동안 거래가 발생한 전국 아파트에 대하여 시각화함

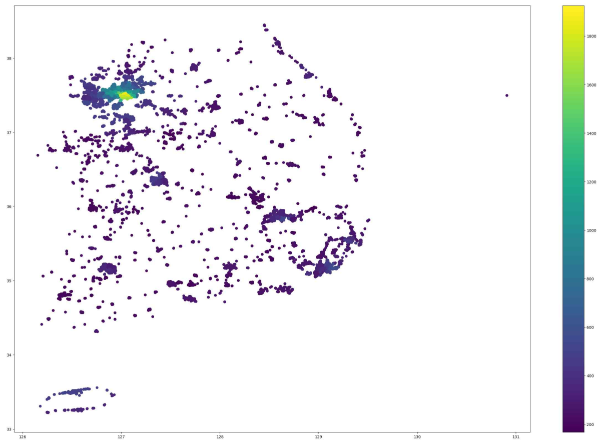
- 지리 정보 외의 특성들은 평균값 혹은 최빈값을 적용하되, 다음 변수들에 대해서는 임의의 값을 적용함
  - 계약년월: 2022년 12월, 전용면적: 55m^2,, 층: 2층, 건축년도: 2002년

- 밝은 색일수록 높은 가격을 나타내며, 짙은 색일수록 낮은 가격을 나타냄

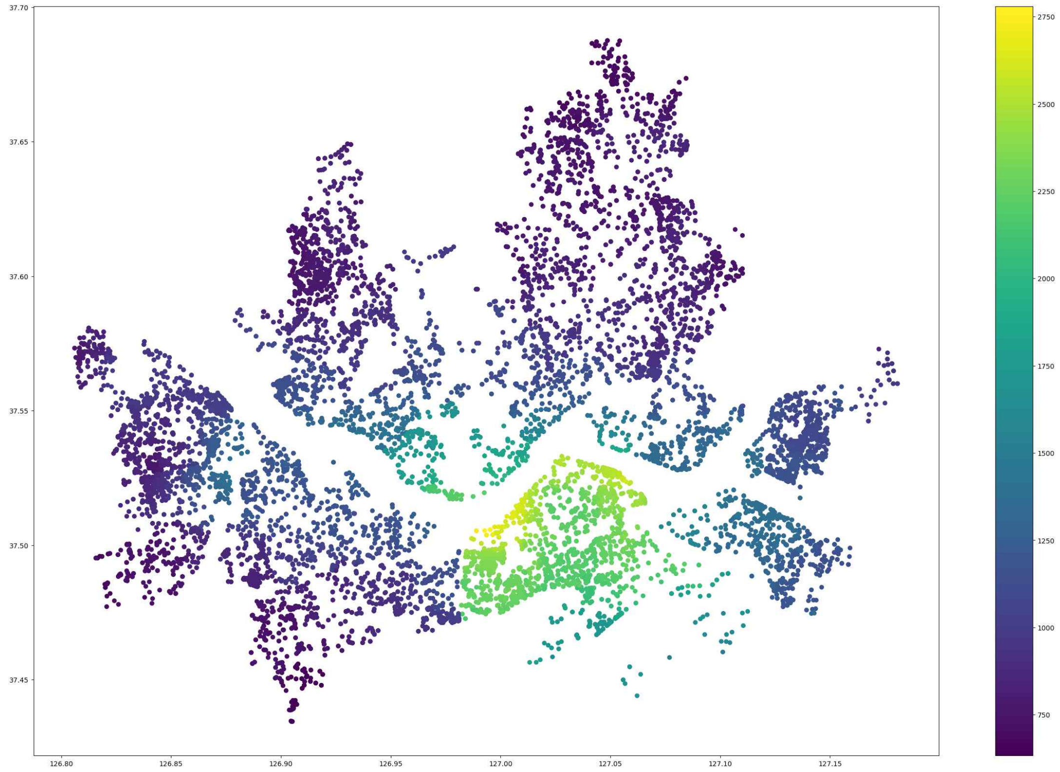
□ **전국적으로 동일 척도로 가격을 나타내면 서울의 가격이 절대적으로 높으며 나머지 지역은 부산을 제외하면 차이가 상대적으로 작은 것으로 보임(<그림 1>)**

- 지도에 빈 곳이 많은데, 국토의 70%가 산지인 특성 상 일부 지역에 주거지가 밀집되어 있는 점에 기인함

〈그림 1〉 지리 정보에 따른 전국 아파트 가격 분포

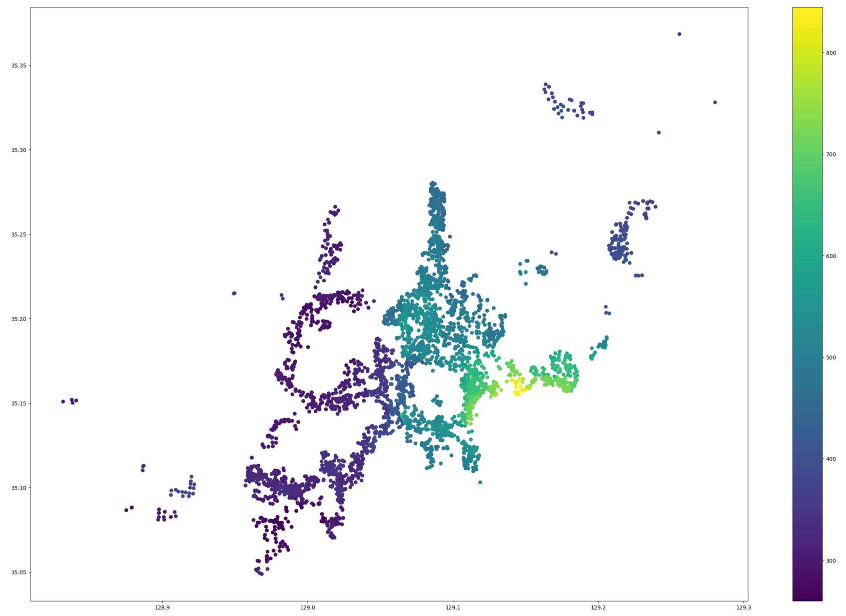


<그림 2> 지리 정보에 따른 서울 아파트 가격 분포



- 서울 지역만 봤을 때는 강남 3구 지역의 가격이 높게 나오되, 한강 지역을 중심으로 상대적으로 높은 가격을 나타냄(<그림 2>)

<그림 3> 지리 정보에 따른 부산 아파트 가격 분포



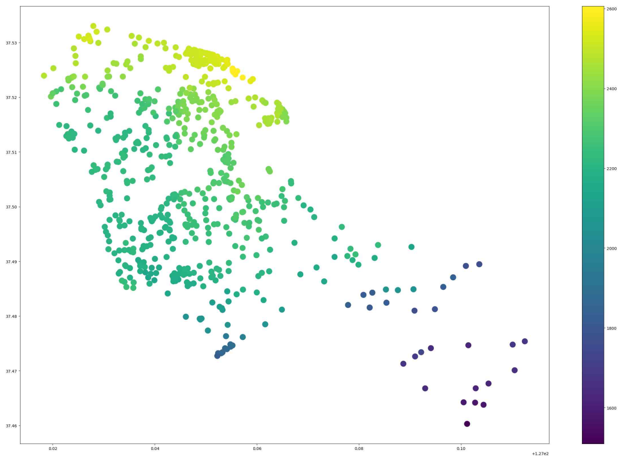
□ 부산의 경우, 해운대 및 수영구 중심으로 높은 가격이 형성되고 서쪽이나 동쪽은 상대적으로 낮은 가격을 나타냄(<그림 3>)

□ 동일 구 내에서도 위도 및 경도 차이로 인해 가격 차이가 관찰됨(<그림 4>: 서울 강남구 , <그림 5>: 부산 해운대구)

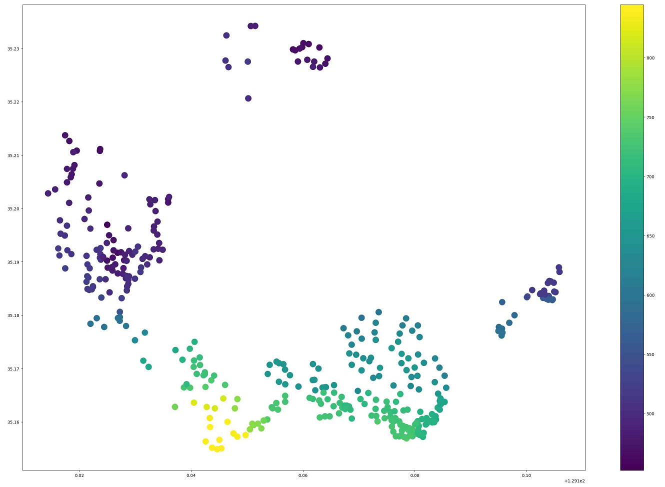
- 위 결과에서 관찰할 수 있는 가격 차이들은 대부분 구별 차이에 의한 것이며, 이는 모형에서 시군구코드를 이용하였기 때문에 나타나는 결과이며 경도 및 위도 변수를 이용한 효과는 보기 어려움

- 〈그림 4〉, 〈그림 5〉는 각각 서울 강남구, 부산 해운대구에 대한 그림을 나타낸 것으로 동일 구 내에서 경도 및 위도 차이가 표현하는 입지 차이에 따른 가격 차이를 볼 수 있음

〈그림 4〉 지리 정보에 따른 서울 강남구 아파트 가격 분포



〈그림 5〉 지리 정보에 따른 부산 해운대구 아파트 가격 분포



3. 매매가, 전세가의 변동률 분포 특성

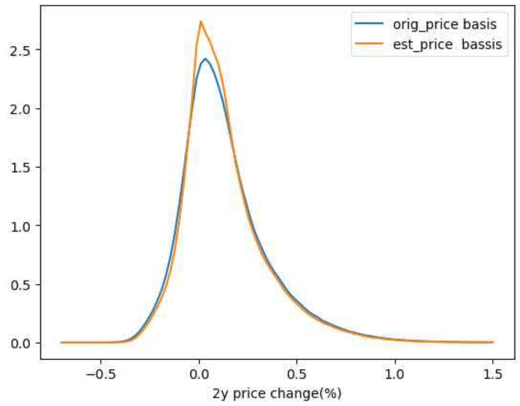
3.1 전체 기간에 대한 변동률 분포

□ 아파트 매매가, 아파트 전세가, 빌라 매매가의 2년 가격변화율 분포에 대한 커널 밀도 함수는 <그림 6>과 같음

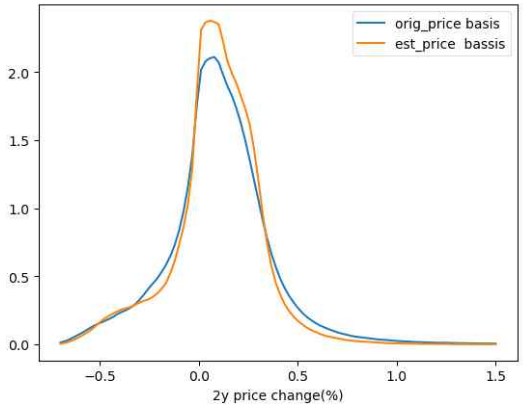
- Min(2022)에서와 같이 히스토그램 형태로 나타낼 수도 있으나, 보다 매끄러운 그래프 형태를 위해 커널 밀도함수를 추정함
  - 파이썬의 scipy패키지를 이용해 가우시안 커널밀도함수를 추정하였음
    - 추정된 가격 변화율의 상·하위 0.01%씩을 이상치로 간주하고 제거함

〈그림 6〉 가격 변화율 분포의 모양

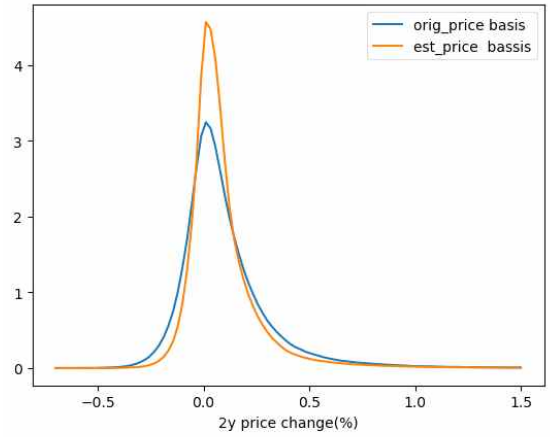
〈아파트 매매〉



〈아파트 전세〉



〈빌라 매매〉



□ **2년 후 가격을 예측모형으로 추정해 2년 가격 변화율을 나타내되, 기준 시점 가격을 실제 가격이 아닌 추정 가격을 이용해 추정한 가격 변화율 분포도 같이 나타냄**

- 모형이 정확하다는 전제 하에, 만약 기준 시점에 거래된 가격이 적정가보다 비싸거나 낮을 경우, 실제 가격을 이용하는 것보다 추정 가격을 이용한 가격 변화율의 주택의 본질적인 가격 변화율에 근접할 수 있음

- 다만, 깡통전세 혹은 역전세의 위험을 추정하기 위해서는 기준 시점 가격으로 반드시 실제 값을 이용해야 함

□ **두 가지 그래프의 모양은 큰 차이를 나타내지 않는데, 이는 실제 거래된 가격 대신 모형이 예측한 가격을 사용하더라도 분포의 모양에는 큰 영향을 끼치지 않는 다는 것을 의미함**

- 두 그래프의 차이는 가격변화율의 식인 $\frac{2\text{년 후 예측치}}{\text{기준값}}-1$ 에서 분모인 기준값이 실제 값인지 모형의 예측치인지의 차이에서 기인하는데, 이 비율의 분포가 대부분 겹친다는 것은 실제 값 대신 추정치를 써도 큰 차이가 없다는 것을 의미함

□ **가격 변화율 그래프는 대체로 좌우 대칭이 아닌 좌측으로 기운 모습을 보이며, 아파트에 비해 빌라의 가격 변화율 폭이 상대적으로 적음**

- 아파트의 가격 변화율 분포에 대해 실거래 데이터를 이용하여 분석한 Min(2022)의 결과와 부합함

〈표 9〉 가격 변화율 분포의 특성

거래유형	분포 기준	평균	표준편차	왜도	첨도
아파트 매매	실제 값 기준	13.56%	22.51%	1.26	2.97
	추정치 기준	13.15%	21.27%	1.31	3.11
아파트 전세	실제 값 기준	10.71%	28.00%	1.63	14.19
	추정치 기준	9.00%	22.34%	0.04	4.16
빌라 매매	실제 값 기준	10.30%	23.10%	2.91	16.92
	추정치 기준	9.12%	18.71%	3.59	23.65

3.2 가격지수 변화율 구간에 대한 조건부 변동률 분포

□ **조건부 분포란, 전체 표본을 이용한 분포가 아닌, 가격지수가 특정 수준 변하였던 시기의 표본만을 모아서 만든 분포를 의미함**

- 예를 들어, 부산광역시 해운대구와 대구광역시 수성구의 가격지수가 2년간 0~5% 상승했던 시기의 표본만을 모아서 분포를 만들 수 있음

- 조건부 분포에 대한 자세한 설명은 민병철()참조

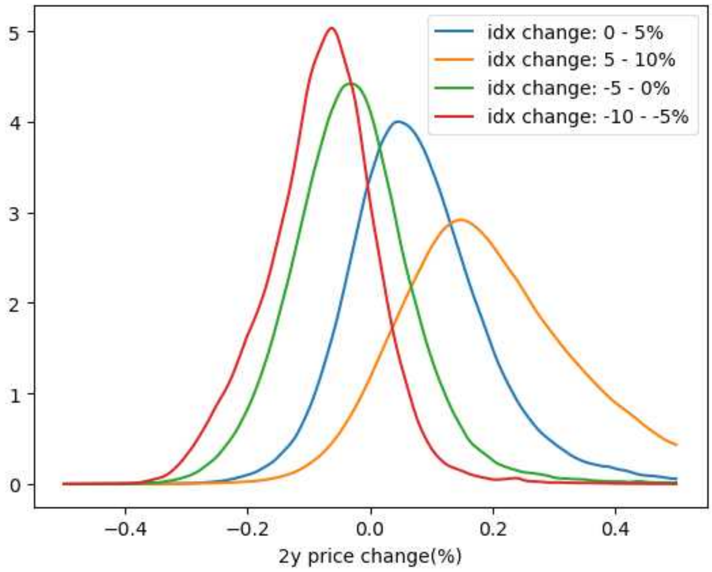
□ **리스크 관리 및 향후 전망을 위해서는 전체 기간에 대한 가격 변화율 분포가 아닌, 가격지수 변화 시나리오에 대한 조건부 분포가 유용하게 쓰일 수 있음**

- 전체 기간에 대한 분포는 가격 상승기 및 하락기를 모두 포함하고 있으므로, 향후 가격하락기가 예상된다면 전체 기간에 대한 분포는 쓰임새가 다소 떨어질 수 있음

□ **아파트 매매가격 변화율에 대한 조건부 분포(<그림 7>)는 Min(2022)에서 분석한 바와 같이, 가격 상승기에는 편차가 커지고 가격 하락기에는 편차가 줄어드는 모습이 나타남**

- 이는 가격 상승기에는 오르는 주택 중심으로 오르고 하락기에는 다 같이 하락한다는 경험적 인식과 일치함

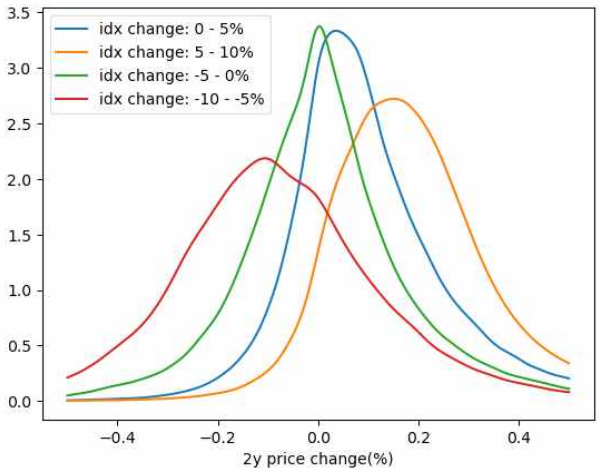
<그림 7> 아파트 매매가격 변화율에 대한 조건부 분포



□ 아파트 전세가격 변화율의 경우, 가격 하락기에 오히려 편차가 커지는 모습을 보여, 가격이 크게 오르거나 하락하는 시기에 변화율 분포의 편차가 큰 모습을 보임

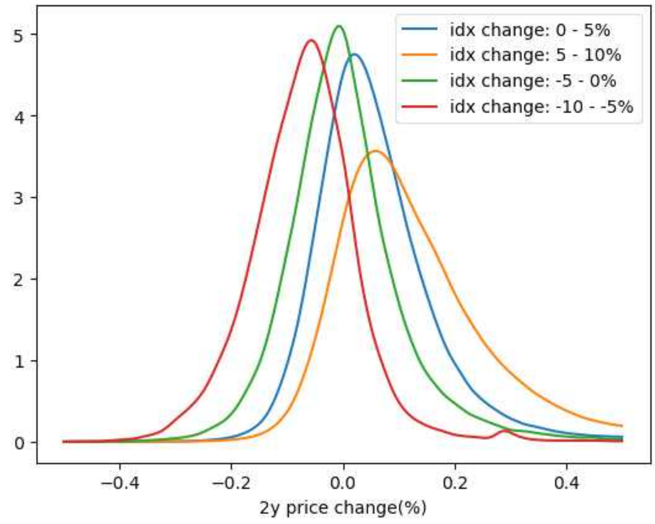
- 매매시장과 달리 가격 하락기에 시장에 대한 인식이 양극화될 수 있음을 의미함

<그림 8> 아파트 전세가격 변화율에 대한 조건부 분포



□ 빌라 매매가격의 경우, 아파트 매매가격과 비교해 가격 변화율의 수준은 작으나, 변화율 간 편차는 오히려 다소 큰 것으로 나타남

〈그림 9〉 빌라 매매가격 변화율에 대한 조건부 분포



〈표 10〉 거래유형별 조건부 가격 변화율 분포의 특성

거래유형	특성	-10% ~ -5%	-5% ~ 0%	0% ~ 5%	5% ~ 10%
아파트 매매	평균	-7.83%	-2.74%	8.03%	20.65%
	표준편차	9.36%	10.52%	12.00%	16.49%
아파트 전세	평균	-6.36%	2.75%	11.93%	20.98%
	표준편차	24.79%	22.06%	21.68%	22.82%
빌라 매매	평균	-6.69%	0.74%	6.34%	15.08%
	표준편차	10.65%	12.64%	14.02%	20.60%

3.3 전세가율 수준별 매매가격 변동률 분포

□ 이론적으로 주택에 대한 향후 기대가격 상승률이 높을수록 전세가율은 낮아지게 됨

- 전세가율 = 전세가격/매매가격 인데, 분자에는 주거가치만 반영되는 반면 분모에는 주거가치에다 투자가치가 반영되기 때문임

□ **하지만 가격상승에 대한 기대 수준이 늘 합리적이지는 않을 수는 있기 때문에, 과연 개별 주택 단위에서 실현된 가격 변화율 또한 전세가율과 역의 관계를 갖는지는 검증이 필요함**

- 평균적으로는 역의 관계를 갖더라도, 개별 주택 단위에서 살펴보면 다른 양상이 나타날 가능성이 있음

□ **또한, 고(高)전세가율 주택에서는 향후 주택가격 하락에 따른 깡통전세 문제가 발생할 수 있기 때문에, 전세가율별로 일정 수준 이상 가격이 하락하는 비중을 분석하는 것도 의미가 있음**

□ **전세 거래 실거래 데이터에 있는 각 거래에 대하여, 앞에서 훈련시킨 모형을 이용하여 매매가격을 추정하여 전세가율을 계산함**

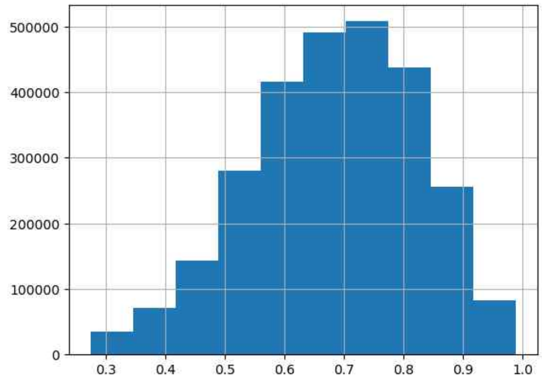
- 또한, 해당 주택의 2년 후 매매가격 또한 추정하여 전세가율별 매매가격 변화율을 계산함

- 이상치 제거를 위해 전세가율, 향후 2년 매매가격 및 전세가격 변화율의 상·하위 1%씩을 제거함

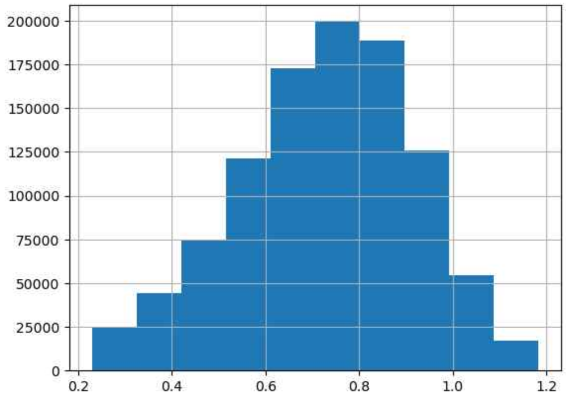
□ **<그림 10>은 전체 분석기간 동안 거래된 전세건들의 전세가율 분포를 나타낸 것인데, 아파트에 비해 빌라에서 높은 전세가율의 비중이 조금 더 높음을 확인할 수 있음**

<그림 10> 체결된 전세거래의 전세가율 분포

<아파트 전세가율 분포>



<빌라 전세가율 분포>



□ **연도별로 살펴보면 약 2015년경부터 높은 전세가율로 체결되는 거래의 비중이 꾸준히 높아지다가, 2020-2021년경부터 그 비중이 다소 줄어듦**

- 여기에 나타난 평균 전세가율은 해당 연도에 거래된 전세 계약건들에 대한 평균 전세가율로, 전체 재고주택의 표본에 대하여 추정되는 공식 전세가율 집계와는 의미가 다소 다름

〈표 11〉 연도별, 주택별 전세가율 수준

*	아파트		빌라	
	평균 전세가율	전세가율>90%	평균 전세가율	전세가율>90%
2011	58.9%	0.9%	57.6%	2.9%
2012	63.2%	1.2%	62.9%	3.8%
2013	67.9%	1.7%	67.0%	5.8%
2014	71.1%	2.2%	69.9%	8.4%
2015	74.5%	4.9%	73.5%	15.5%
2016	76.3%	8.4%	75.6%	20.7%
2017	74.1%	5.8%	76.2%	22.2%
2018	68.7%	3.7%	76.3%	23.9%
2019	67.5%	5.0%	77.6%	26.8%
2020	67.3%	6.9%	79.1%	31.6%
2021	69.4%	12.1%	77.2%	28.4%
2022	64.0%	6.3%	74.2%	21.0%

□ **전세가율 수준별 향후 2년 가격 상승률을 살펴보면, 2014-2016년 이전에는 오히려 전세가율이 높은 집의 가격 상승률이 높았던 것으로 나타남**

- 빌라의 경우 2014년-2016년까지도 높은 전세가율로 거래된 주택의 가격 상승률이 높아, 아파트 시장과 약 1년의 시차가 있었던 것으로 추정됨

□ **보통은 비슷한 전세가율일 때 빌라에 비해 아파트의 가격 상승률이 높았으나, 서울에서 2019-2020년에 높은 전세가율에 거래된 빌라는 비슷한 전세가율 수준의 아파트보다 높은 가격 상승률을 보였음**

- 갭투자에 대한 유인이 더 있었을 것으로 보임

<표 12> 연도별 전세가율 수준별 향후 2년 가격 상승률(전국)

	아파트 전세가율 수준					빌라 전세가율 수준				
	90	80	70	60	50	100	90	80	70	60
'11	21.8%	16.8%	8.5%	1.0%	-4.6%	10.5%	8.2%	4.4%	-2.1%	-5.2%
'12	14.9%	12.9%	7.8%	4.5%	2.4%	6.4%	5.5%	2.7%	0.2%	-1.2%
'13	14.6%	14.9%	13.2%	11.0%	10.3%	9.6%	8.2%	6.1%	5.1%	4.6%
'14	8.9%	13.3%	13.7%	13.0%	13.7%	10.4%	8.9%	8.5%	8.3%	8.4%
'15	9.7%	10.5%	10.8%	11.4%	13.5%	8.6%	7.7%	8.1%	8.7%	9.9%
'16	7.9%	12.9%	12.5%	13.7%	20.1%	5.9%	5.4%	6.3%	7.5%	9.3%
'17	2.2%	9.6%	12.0%	11.0%	13.8%	3.7%	3.8%	4.5%	5.4%	7.3%
'18	4.8%	11.7%	17.2%	21.3%	25.0%	7.2%	7.0%	7.4%	8.6%	10.6%
'19	24.5%	31.8%	37.8%	43.9%	45.5%	16.9%	16.9%	17.4%	19.1%	21.7%
'20	26.1%	29.3%	30.4%	28.8%	24.1%	17.0%	17.6%	19.1%	21.2%	23.6%

<표 13> 연도별 전세가율 수준별 향후 2년 가격 상승률(서울)

	아파트 전세가율 수준					빌라 전세가율 수준				
	90	80	70	60	50	100	90	80	70	60
'11	13.7%	6.6%	-0.9%	-4.0%	-7.1%	-3.0%	-4.3%	-4.6%	-5.3%	-6.0%
'12	11.2%	4.8%	4.9%	3.5%	1.9%	-0.9%	-0.9%	-0.3%	-0.6%	-1.3%
'13	15.8%	14.0%	12.3%	10.5%	11.6%	6.2%	5.4%	5.1%	4.9%	4.8%
'14	18.9%	17.3%	15.4%	15.2%	21.1%	9.6%	8.8%	8.5%	8.7%	9.3%
'15	17.9%	17.9%	17.8%	20.4%	24.3%	8.8%	8.0%	8.5%	9.5%	11.5%
'16	26.0%	30.8%	34.7%	38.1%	40.5%	6.6%	6.1%	7.5%	9.6%	12.9%
'17	22.9%	29.2%	32.5%	33.1%	31.8%	4.8%	4.7%	5.8%	7.9%	11.4%
'18	13.1%	27.9%	34.5%	34.3%	32.1%	8.6%	8.1%	8.8%	10.6%	13.8%
'19	12.5%	26.4%	48.3%	52.3%	49.1%	18.4%	18.3%	18.9%	21.0%	24.5%
'20	9.1%	15.5%	25.4%	27.8%	24.7%	18.7%	18.6%	20.0%	22.8%	26.0%

- 전세가율 수준별로 향후 2년간 10% 이상 하락한 비중을 살펴보면, 2011년(2011-2013)과 2020년(2020-2022)을 제외하면 서울 지역의 하락 비중이 절대적으로 낮음을 볼 수 있음

<표 14> 전세가율 수준별 향후 2년간 10% 이상 하락 비중(전국)

	아파트 전세가율 수준					빌라 전세가율 수준				
	90	80	70	60	50	100	90	80	70	60
'11	0.88%	0.95%	1.49%	2.88%	15.92%	1.62%	1.95%	5.51%	10.19%	16.62%
'12	2.10%	1.46%	0.79%	0.24%	0.43%	1.26%	1.30%	0.71%	0.78%	2.26%
'13	0.54%	0.29%	0.18%	0.09%	0.16%	0.04%	0.21%	0.05%	0.06%	0.10%
'14	5.39%	1.97%	0.71%	0.41%	0.72%	0.76%	0.24%	0.10%	0.07%	0.08%
'15	8.39%	4.47%	3.32%	3.23%	2.08%	1.03%	0.46%	0.47%	0.36%	0.31%
'16	7.40%	5.74%	8.17%	8.09%	5.68%	1.05%	0.76%	0.66%	0.90%	1.13%
'17	17.68%	11.13%	12.58%	13.87%	10.99%	2.62%	1.64%	1.83%	2.99%	3.71%
'18	11.31%	6.33%	4.59%	2.58%	1.49%	1.69%	1.20%	1.31%	2.11%	3.32%
'19	1.89%	0.99%	0.52%	0.30%	0.18%	0.32%	0.16%	0.19%	0.48%	1.14%
'20	1.20%	1.72%	2.10%	3.36%	3.14%	0.24%	0.28%	0.24%	0.44%	0.81%

<표 15> 전세가율 수준별 향후 2년간 10% 이상 하락 비중(서울)

	아파트 전세가율 수준					빌라 전세가율 수준				
	90	80	70	60	50	100	90	80	70	60
'11	9.76%	7.83%	0.62%	1.92%	18.68%	17.20%	8.95%	7.99%	7.65%	11.50%
'12	8.55%	0.03%	0.02%	0.03%	0.19%	2.53%	2.72%	0.46%	0.24%	0.90%
'13	0.00%	0.00%	0.02%	0.01%	0.01%	0.00%	0.16%	0.00%	0.01%	0.03%
'14	0.00%	0.00%	0.00%	0.01%	0.07%	0.00%	0.00%	0.00%	0.01%	0.00%
'15	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.01%	0.00%
'16	0.00%	0.00%	0.00%	0.01%	0.00%	0.00%	0.00%	0.01%	0.02%	0.03%
'17	0.02%	0.00%	0.00%	0.01%	0.00%	0.00%	0.03%	0.05%	0.09%	0.10%
'18	0.45%	0.10%	0.04%	0.00%	0.00%	0.03%	0.02%	0.00%	0.03%	0.04%
'19	1.35%	0.95%	0.19%	0.00%	0.00%	0.00%	0.00%	0.00%	0.01%	0.03%
'20	1.54%	2.18%	3.17%	4.63%	2.95%	0.13%	0.13%	0.10%	0.10%	0.14%

- 이를 토대로 살펴보면, 이론적으로는 높은 전세가율을 나타내는 집은 기대 가격 상승률이 낮으므로 전세 만기 시 깡통전세가 될 확률이 높지만, 실제로는 전세가율이 높은 집의 가격 상승률이 높거나 극단적 하락 비중이 낮은 시기가 대부분이었음

- 경험적으로 이를 알고 있었다면 임차인들은 전세가율이 높을수록 보증금을 잃을 위험 또한 높다는 점을 간과하였을 가능성이 있음

4. 매매가, 전세가의 결합 분포 특성 분석

□ 3.3에서 다룬 내용은 깡통전세의 위험을 분석하는데 필요한 내용이나, 실제 위험인 역전세 시 깡통전세 상태일 위험을 추정하기에는 충분하지 않음

- 역전세: 보증금이 하락하여 발생하는 현상으로, 계약 만료되어 반환해야 할 보증금 수준이 신규 임차인에게 받은 보증금과 임대인의 지불능력을 초과해 보증금을 반환할 수 없는 상황을 말함

- 깡통전세: 주택가격이 하락하여 발생하는 현상으로, 기존에 체결된 임차계약의 보증금 수준보다 주택가격이 낮은 상태를 말함
  - 선순위 주택담보대출이 끼어있다면, 선순위 주택담보대출과 보증금 수준보다 주택가격이 낮은 상태를 말함

□ 깡통전세 상태이더라도 임대인이 주택을 포기하지 않고 보증금을 내어줄 수 있다면 역전세가 발생하지 않으므로, 실제 보증금 미반환 위험을 추정하기 위해서는 전세가격이 하락하여 발생하는 역전세와 매매가격이 하락하여 발생하는 깡통전세가 동시에 일어날 확률을 알아야 함

- 역전세가 발생하였더라도 깡통전세 상태가 아니라면 경매를 통해 임차인은 보증금을 돌려받을 수 있음

- 물론, 이는 개념적으로만 접근하여 상황을 단순화시킨 것이며, 실제로는 이 과정에서 임차인이 겪는 많은 피해는 결코 간과해서는 안될 것임

□ 역전세와 깡통전세가 동시에 일어날 확률을 분석하기 위해서는 동일 주택에 대한 매매가 변화율과 전세가 변화율의 결합 분포를 도출해야 함

□ 이 절에서는 아파트에 대하여 매매가격과 전세가격의 결합분포를 도출하고, 그 특성을 살펴봄

- 빌라의 경우 전세가격 추정모형이 불완전하여 현재 단계에서 결합분포를 추정하는 것이 의미가 없음

- 추정된 전세가격변화율, 매매가격변화율의 상·하위 0.1%씩을 제거하고 분석을 수행함

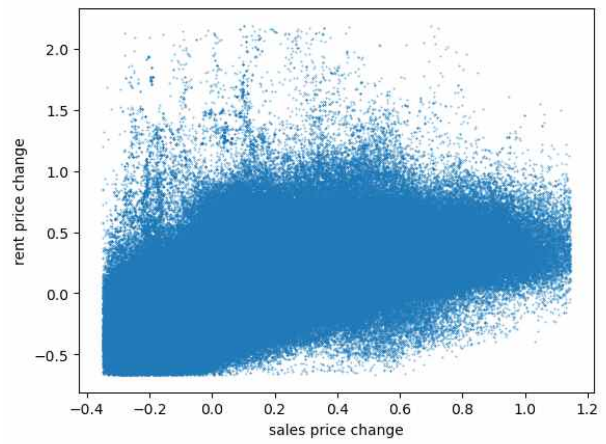
□ 아파트 매매가격과 전세가격의 2년 변화율 간 상관관계는 0.4965 수준으로 계산되며, 그 결합분포의 모양은 <그림 11>과 같음

- 어느 정도 양의 상관관계는 확인할 수 있으나, 그 관계가 아주 강하지는 않음

- 시군구 단위 매매가격지수와 전세가격지수의 2년 변화율 간 상관관계는 약 0.6094로 계산되어, 개별 주택단위에서의 상관관계가 지역 단위에서의 상관관계보다 다소 약한 것으로 분석됨

  ▪ 한국부동산원 가격지수 이용, 분석기간: 2003년 11월-2023년 07월

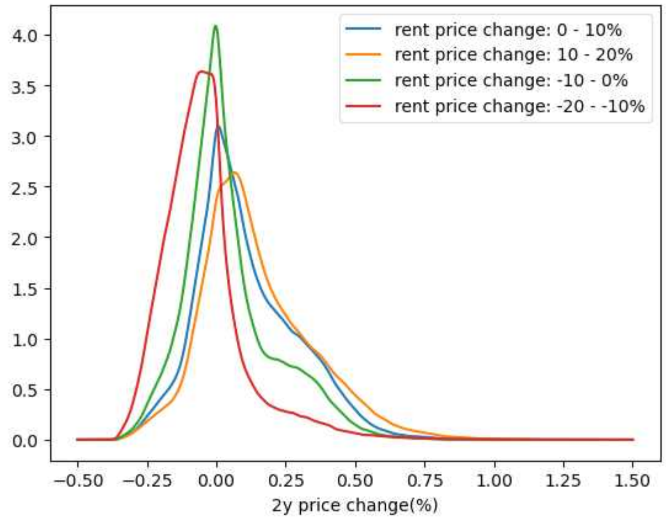
〈그림 11〉 아파트 매매가격과 전세가격 변화율의 결합분포



□ 아파트 전세가격 변화율 구간별로 매매가격 변화율의 분포를 살펴보면, 전세가격 하락 시 매매가격 변화율의 편차가 더 작은 것을 볼 수 있음

- 전세가격이 오르더라도 매매가격은 반드시 오르지 않을 수 있는 반면, 전세가격 하락 시 매매가격 또한 같이 하락하는 비중이 높다는 것을 의미함

〈그림 12〉 아파트 매매가격과 전세가격 변화율의 결합분포



〈표 16〉 전세가격 변화율 조건부 매매가격 변화율 분포의 특성

특성	-20% ~ -10%	-10% ~ 0%	0% ~ 10%	10% ~ 20%
평균	-5.44%	4.79%	11.59%	16.04%
표준편차	14.06%	16.61%	17.88%	19.32%
왜도	1.86	1.03	0.73	0.89
첨도	6.53	1.38	0.59	0.97

□ 전세가격 변화율 구간별로 매매가격이 일정 폭 이상 하락하는 비중을 살펴보면, 전세가격 하락폭이 커질수록 매매가격이 20% 이상 하락하는 비중은 급격히 증가함

- 전세가격이 0~10% 하락할 때는 매매가격이 20% 이상 하락하는 비중이 3.18%였지만, 전세가격이 10%~20%하락할 때는 10.31%로 세 배 이상 증가함

□ <표 15>의 결과는 역전세가 발생할 수 있는 수준의 전세가격 하락 시 깡통전세가 발생할 수 있는 수준의 매매가격 하락이 발생할 수 있는 비중을

나타내나, 실제 임차인의 보증금 손실위험은 경매 낙찰가율에도 영향을 받으므로 이보다 더 클 것임

- 역전세 위험 또한 단순 전세가격 변화율에 의해서만 결정되는 것이 아니고 임대인의 재무상태 및 다주택여부 등에도 영향을 받음

〈표 17〉 전세가격 변화율별 극단적 매매가격 하락 발생 확률

매매가격 변화율	전세가격 변화율			
	-20% ~ -10%	-10% ~ 0%	0% ~ 10%	10% ~ 20%
-20%〉	10.31%	3.18%	1.60%	0.76%
-10%〉	38.35%	14.25%	7.26%	4.15%
0%〉	76.33%	45.80%	27.66%	19.20%

V. 요약 및 향후 활용 방안

□ 이 연구에서는 머신러닝 기법을 이용해서 아파트 및 빌라가격을 추정하고, 이를 이용해 가격 변화율 분포를 만들어 그 특성을 분석함

- AutoML 툴인 Autogluon을 이용하여 다양한 모형을 결합한 앙상블 모형을 추정함

□ 거래 정보, 주택 정보, 지리 정보들을 결합하여 모형을 훈련시킨 결과, 아파트 매매의 경우 시험군 MAPE가 4.5%수준이었으며, 아파트 전세 및 빌라 매매의 경우 약 9.3% 수준이었음

- 대상 기간은 2011년 1월-2022년 12월이었음
- 예측이 목적이 아니고 과거 각 시점의 가격 추정이 목적인만큼 기간 무관하게 무작위로 시험군을 20% 추출함
- 빌라 전세의 경우 성능이 매우 떨어져 이를 이용한 분석은 수행하지 않음

□ **분석결과, 빌라 매매가격의 변동성보다 아파트 매매가격의 변동성이 더 큰 것으로 나타났으며, 아파트 전세가격의 변화율은 개별 주택간 편차가 큰 편인 것으로 분석됨**

- 주택가격 상승기와 하락기 시 가격의 등락폭은 빌라보다 아파트에서 더 큰 것으로 계산됨
- 같은 시기라 할지라도 아파트 전세가격 변동률은 주택별로 차이가 큼

□ **체결된 전세건에 대하여 전세가율을 분석한 결과, 아파트에 비해 빌라의 전세가율이 높은 것으로 보이며, 특히 2016년과 2021년에 전세가율이 90%를 넘는 주택의 비중이 급등한 것으로 보임**

□ **전세가율과 향후 2년간 가격 변동 변동률을 살펴보면, 2011-2013년에는 전세가율이 높을수록 향후 2년간 실현된 매매가격 상승률이 높았으나, 그 이후로는 전세가율이 낮을수록 매매가격 상승률이 높았음**

- 이론적으로는 다른 모든 조건이 동일할 때, 전세가율이 낮을수록 기대가격 상승률이 높다고 봄
- 다만, 수도권, 특히 서울의 빌라의 경우 대부분의 시기에서 전세가율이 높을수록 매매가격 상승률이 높았음

□ **매매가격과 전세가격 간 결합분포를 분석한 결과, 전세가격이 하락할수록 매매가격이 큰 폭 하락할 비중은 급격하게 증가하는 것으로 계산됨**

- 전세가격이 0~10% 하락할 때는 매매가격이 20% 이상 하락하는 비중이 3.18%였지만, 전세가격이 10%~20%하락할 때는 10.31%로 세 배 이상 증가함

□ **이 연구의 내용은 빌라 시장에서의 깡통전세 위험 분석에 활용될 수 있으며, 아파트의 경우 역전세 시 깡통전세 발생 위험 계산에 쓰일 수 있음**

- 빌라의 전세가격 추정 모형을 개발하지 못해 역전세 위험은 추정하지 못함

- 아파트에 대해 기존에 역전세 및 깡통전세 위험을 따로 분석한 연구는 있었지만, 실제 보증금 손실과 보다 직접적인 연관이 있는 역전세 시 깡통전세 발생위험에 대한 연구는 없었음

□ **이 연구에서는 빌라를 구성하는 연립·다세대 주택과 다가구 주택 중 연립·다세대 주택에 대해서만 모형을 추정했으며, 연립·다세대에 대한 전세가격 추정 모형은 도출하지 못한 한계를 지님**

□ **또한, 예측을 목적으로 만든 모형이 아니므로 가격예측 목적으로 이용하려면 다른 형태로 모형을 만들거나 주기적으로 재학습을 하는 것이 적절할 것임**

- 향후 거시 환경 시나리오별로 가격예측을 하려면 모형에 거시변수들을 포함시켜서 학습시켜야 함

- 단지 현 시점의 가격을 추정하기 위해서는 주기적으로 최신 데이터를 추가하여 재학습을 해야 할 것임

참 고 문 헌

민병철. (2021). 전세가격 변동률 분포를 활용한 역전세 위험의 측정. 부동산학연구, 27(2), 63-75.

민병철. (2022). 보증금 미반환 위험의 추정 - 깡통전세 아파트를 중심으로. 주택금융리서치, 2022년 제 28호

이창로, & 박기호. (2016). 단독주택가격 추정을 위한 기계학습 모형의 응용. 대한지리학회지, 51(2), 219-233.

조유나, 김수현, & 송규원. (2017). 헤도닉 가격 모형과 딥러닝을 이용한 주택가격예측 비교: 서울시 아파트를 중심으로. 한국정보과학회 학술발표논문집, 1890-1892.

홍정의. (2021). 랜덤 포레스트 알고리즘을 통한 주택 대량평가모형 연구. 부동산분석, 7(1), 1-28.

Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of Random forest for valuation and a CART-based approach for model diagnostics. Expert Systems with Applications, 39(2), 1772-1778.

Hong, J., Choi, H., & Kim, W. S. (2020). A house price valuation based on the random forest approach: The mass appraisal of residential property in South Korea. International Journal of Strategic Property Management, 24(3), 140-152.

Limsombunchai, V. (2004). House price prediction: hedonic price model vs. artificial neural network.

Min, B. C. (2022). Investigating the housing market volatility through the lens of individual house price distribution. Journal of Housing and Urban Finance, 7(1), 27-47.

Mu, J., Wu, F., & Zhang, A. (2014, August). Housing value forecasting based on machine learning methods. In Abstract and Applied Analysis (Vol. 2014). Hindawi.

Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. Expert systems with Applications, 36(2), 2843-2852.