

부동산 가격 정밀 예측을 위한 후처리 어댑터 연구

김지현, 박현석, 김수현, 이남경, 김희원, 이건희*
에이치디씨랩스

{jh991219, hyunskki, soohyun.kim, nk_lee, ive2go, Gunhee_Lee}@hdc-labs.com

Post-processing Adapter Methods for Precise Real Estate Price Forecasting

Jihyeon Kim, Hyunseok Park, Soohyun Kim, Namkyeong LEE, Heewon Kim, Gunhee Lee*
HDC LABS

요약

본 논문은 지역, 가격대, 거시 시장 환경 변화에 따라 이질적인 변동성을 보이는 부동산 시장의 시세를 정밀하게 예측하기 위한 어댑터 기법을 제안한다. 변동하는 시장의 국지적인 패턴을 즉시 반영하기 위해서 백본(Backbone) 예측 모델에 도시별, 가격대별, 글로벌 어댑터를 부착하고, 최적의 어댑터 조합을 통해 시세 예측 후보정에 사용하여 최신 시장 상황에 적용한다. 2024년 한국 부동산원 실거래 데이터 실험 결과, 기존 모델 대비 평균 절대오차(MAE)를 최대 5.21%, 평균 제곱오차(MSE)를 최대 9.01% 감소시키는 결과를 확인했다.

I. 서론

어떻게 하면 다양한 요인의 영향을 받는 부동산 시장의 시세를 정밀하게 예측할 수 있을까? 데이터 처리 기술의 발전으로 다양한 분야의 데이터를 수집하고, 인공지능 모델을 적용하여 가치를 창출하려는 시도들이 이루어지고 있다. 그 중 시계열 데이터는 시간의 흐름에 따라 변화되는 데이터로, 제조, 운송 등 다양한 분야에서 수집 및 활용되고 있다.

최근 인공지능을 활용하여 복잡한 시계열 데이터를 처리하고 예측하는 수요가 늘어남에 따라 주식, 부동산 시세 예측에 대한 연구들이 진행되고 있다. 그러나 부동산 거래는 연속적으로 일어나지 않아 결측 데이터가 다수 존재하고[1], 도시별 데이터 수집 규모와 사람들의 수요의 불균형, 부동산 가격대별 투자 성향의 차이, 금리 등 전역적인 환경 요인에 따라서 시세에 복잡하게 작용한다.

기존 연구들은 전역적인 시세 예측을 하는 단일 모델을 개발하는데 그쳤다[2]. 또한 최근에는 아파트 브랜드의 가치, 주변 상권 차이를 반영하여 평형 타입 단위의 시세를 예측하는 연구가 진행되었다[1]. 하지만 도시별로 사람들의 수요는 다르게 작용하고[3], 데이터 수집 규모가 다르며, 부동산 가격대별 가격 움직임 또한 다르다. 또한 시장과 정책에 따라서 급격하게 변화하는 시세를 인공지능 모델에 빠르게 반영하는 연구는 부재한 상황이다.

본 연구에서는 부동산 실거래 가격을 예측하는 모델을 좀 더 정밀한 시세 예측을 수행할 수 있으며 지속적으로 추적하고 현재 시장 상황을 반영할 수 있는 후처리 어댑터 기법을 제안한다.

결과적으로 본 연구의 기여는 다음과 같다.

1. 기존 단일 모델이 포착하지 못했던 지역별 수요

편차, 가격대별 투자 성향의 차이, 거시적 시장 변화 등 다양한 시세 결정 요인을 반영하기 위해 도시별, 가격대별, 전역 어댑터를 설계하고 이를 조합하여 정밀 예측을 수행한다.

2. 시간 흐름에 따라 변화하는 시장 특성을 반영하기 위해, 어댑터들을 주기적으로 최적의 조합을 찾아 선택하고 반영한다.

3. 실제 한국부동산원 실거래 데이터를 기반으로 실험을 진행하고 어댑터 조합의 효과를 검증한다.

II. 본론

본 논문에서는 최근 8 주간의 실제 부동산 실거래 데이터와 사전 학습된 예측 모델 결과 간의 차이를 매주 추적하고, 이를 보정하기 위한 어댑터 기반 후처리 기법을 제안한다. 제안 기법은 크게 세 가지 어댑터로 구성되며, 이들은 각각 전역적인 추세, 가격대별 편향, 지역별 가격 및 수요 차이를 보완하는 역할을 한다. 각 어댑터는 독립적으로 보정값을 계산한 뒤, 최종 예측값은 8 주간의 실거래가를 가장 잘 보완하는 조합으로 산출된다.

부동산 시장은 금리, 정책, 경제 상황 등 다양한 거시 환경 요인의 영향을 받으며[4], 이러한 변화는 예측 모델의 전반적인 예측값의 스케일과 실제 시세 간의 괴리를 유발할 수 있다. 글로벌 어댑터는 이러한 전역적인 scale mismatch 를 보정하기 위한 보정계수를 추정한다. 이를 위해 예측값과 최근 실거래 시세 간의 관계를 로그 변환 후 Huber 회귀를 통해 선형적으로 회귀하며, 이상치에 덜 민감하게 전체적인 스케일을 재조정한다. 보정은 다음의 식으로 이루어진다:

$$\log(P_{real}) = \beta_0 + \beta_1 * \log(P_{pred})$$

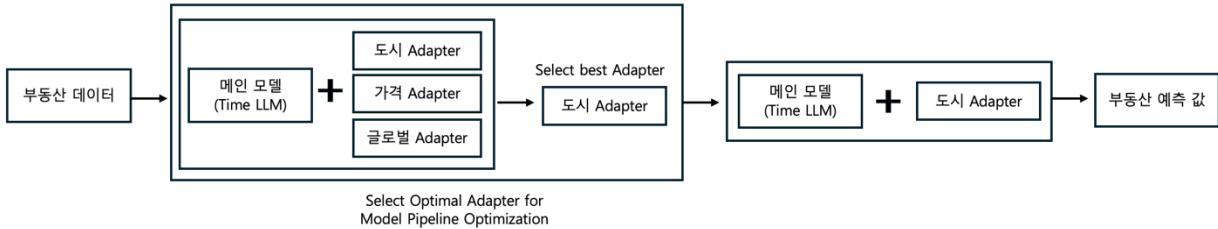


그림 1. 제안 방법의 구조

이를 통해 전체 예측값을 재 스케일링하여, 전역적인 시세 흐름과의 괴리를 줄인다.

또한 특정 가격 구간(예: 저가/고가 아파트)에 대한 예측 편향을 줄이기 위해 가격대별 어댑터를 적용한다. 이는 예측값 분포를 분위수 구간(Quantile Bin)으로 나누고, 각 분위수 구간별로 실제 실거래값 대비 예측값의 평균 비율을 계산하여 보정계수를 산출한다. 예를 들어, 0~20% 분위 구간(저가 가격대)에서 예측값이 일관되게 실제보다 낮게 나올 경우, 해당 구간에 보정계수를 곱하여 상향 보정하는 방식으로 적용된다. 이 방식은 국지적인 왜곡을 줄이고 전 구간에서 예측 정확도를 높이는 데 기여한다.

도시마다 부동산 시장의 수요, 공급, 거래 빈도가 다르기 때문에 지역별로 편차가 발생할 수 있다. 도시별 어댑터는 이러한 편차를 반영하여 예측값을 지역별로 조정한다. 우선 각 도시별로 예측값과 실제값의 비율을 계산한 후, 그 비율에서 1을 뺀 잔차를 추정한다. 이 평균 잔차를 도시별 보정계수로 사용하며, 과하게 벗어난 지역에 선별적으로 적용된다. 이 방식은 도시별로 발생하는 예측 편향을 보완하는 데 효과적이다.

각 어댑터는 개별적으로 동작하며, 8 주간의 실거래 데이터를 기반으로 가장 잘 보완하는 어댑터 조합을 찾아서 적용한다. 이를 통해 변화하는 시장 환경에 따라 유연하게 대응할 수 있다. 그림 1은 도시 어댑터가 최적의 어댑터로 선택되었을 때의 예시로, 본 논문의 전체적인 구조를 설명한다.

	기존 예측값	어댑터 적용 예측값	개선율
MAE(6/20)	2642.87	2505.10	5.21%
MSE(6/20)	42,124,930	38,328,210	9.01%
MAE(6/27)	2744.47	2708.43	1.31%
MSE(6/27)	46,775,710	45,073,750	3.64%

표 1. 어댑터 적용 실험 결과

제안하는 방법의 성능을 평가하기 위해 25년 6월 20일과 27일에 발생한 부동산 거래에 대해 사전 학습된 예측모델에 적용했을 때 시세에 대한 정확도를 평가한다. 평가는 MAE, MSE 두 가지로 평가하며 표 1. 어댑터 적용 실험 결과에서 보여준다. 6월 20일은 MAE는 5.21% 개선되었으며, MSE는 9.01% 개선된다. 그 다음주인 6월 27일은 MAE는 1.31% 개선되었으며, MSE는 3.64% 개선된다. 이 결과를 통해 제안하는 어댑터 방법이 예측 모델의 성능을 개선하는데 도움이 된다는 것을 입증한다.

III. 결론

본 논문에서는 다양한 요인의 영향을 받는 부동산 시장의 시세를 더 정밀하게 예측하고, 변동성을 반영하기 위한 후처리 어댑터 기반 시세 예측 보정 기법을 제안한다. 제안하는 기법은 전역적인 시장 어댑터(Global Adapter), 가격대별 어댑터(Price Adapter), 지역별 어댑터(City Adapter)를 최근 8 주간 실거래 데이터를 기반으로 추정하고, 최적의 조합을 동적으로 적용함으로써 기존 예측값의 정밀도를 높인다. 실험 결과 제한한 어댑터는 2025년 한국 부동산원 실거래 데이터를 기반으로 평균절대오차(MAE)를 최대 5.21%, 평균제곱오차(MSE)를 최대 9.01%까지 개선하며 기존 사전 학습 모델의 성능을 유의미하게 향상시키는 것을 확인했다. 특히, 본 기법은 시장의 변화나 국지적 수용 변동에 따른 모델의 예측 불확실성을 줄이는 데 효과적이며, 실시간 대응력을 강화할 수 있다는 점에서 실용성이 높다.

향후 연구로는 지속적인 학습을 사용하는 인공지능 기반 어댑터, 경제 지표 및 정책 변수를 반영할 수 있는 어댑터로의 고도화를 고려하고 있다. 본 연구는 AI 기반 부동산 시세 예측의 정밀도 향상에 기여할 뿐만 아니라, 다양한 어댑터 모듈들의 개발로의 확장성을 제시한다.

참고 문헌

- [1] 이다연, 이선빈, 박현석, 김희원, 이건희, T-FOREST: 상호 보완적 결측치 처리를 통한 부동산 시세의 시계열 예측 방법, KDBC 2024 학술대회 논문집, pp. 77–86, 2024.
- [2] 남상현, 한태호, 김이주 and 이은지. (2020). 머신러닝 기법을 통한 대한민국 부동산 가격 변동 예측. 한국인터넷방송통신학회 논문지, 20(6), 15–20.
- [3] 홍정의, 최한재, 송민재 and 홍시온. (2019). 지역별 장기주택수요의 예측과 비교 : Mankiew-Weil 모형에 기반하여. 부동산연구, 29(3), 93–103.
- [4] 송영오, 김재환, 심희철. (2023). 부동산정책 효과 분석 및 시사점 – 수도권 아파트 실거래 평균매매가격을 중심으로. 아시아태평양융합연구교류논문지, 9(1), 125–140. <http://dx.doi.org/10.47116/apjcri.2023.01.11>