

# 공간적 상관성을 고려한 딥러닝 기반 부동산 가격 예측 방법 제안

김규석\*, 김경민\*\*, 조재우\*\*\*

## A Proposal of Real Estate Valuation Prediction Method using Deep Learning-based Spatial Regression Analysis

Kyuseok Kim\*, and Kyung-Min Kim\*\*, and Jaewoo Cho\*\*\*

본 연구는 한성대학교 학술연구비 지원과제임

### 요 약

부동산은 공간적인 특성을 가지고 있다는 점에서 다른 자산과 차별되는 특성을 갖는다. 공간적인 측면으로 접근하여, 본 연구에서의 데이터는 2013년 1분기부터 2023년 2분기까지의 서울시 3,000세대 이상의 아파트 단지이며, 방법론은 딥러닝 기반의 시계열 분석 기법인 RNN, LSTM, GRU이다. 이를 통해 3,000세대 이상의 아파트 단지들이 특정 지역구의 아파트 단지 가격에 어느 정도 영향을 미치는 지 확인할 수 있다. 모델별 평균 정확도를 MAPE를 통해 산출한 결과, RNN이 10.91, LSTM이 11.44, GRU가 11.12로서 RNN이 가장 우수했고, LSTM과 GRU가 비슷하였다. 추후 본 연구에서 제안하는 모델을 활용하여 높은 정확도의 부동산 가격 예측을 하는데 도움이 될 것으로 기대한다.

### Abstract

Real estate has a characteristic that is differentiated from other assets in that it has spatial characteristics. Approaching from a spatial aspect, the data in this study are apartment complexes with more than 3,000 households in Seoul from the first quarter of 2013 to the second quarter of 2023, and the methodologies are RNN, LSTM, and GRU which are deep learning-based time series analysis techniques. As a result of calculating the average accuracy of each model through MAPE, RNN was 10.91, LSTM was 11.44, and GRU was 11.12. RNN was the best, and LSTM and GRU were similar. With the proposed model, it is expected to be helpful in predicting real estate valuation with higher accuracy in the future.

### Keywords

real estate valuation, spatial regression analysis, time series analysis, deep learning

\* 서울대학교 환경대학원 환경계획학과 박사수료,  
한국폴리텍대학 분당융합기술교육원 데이터융합SW과  
조교수

- ORCID: <https://orcid.org/0000-0001-6613-5125>

\*\* 서울대학교 환경대학원 환경계획학과 교수

- ORCID: <https://orcid.org/0000-0003-0438-7612>

\*\*\* 한성대학교 사회과학부 조교수(교신저자)

- ORCID: <https://orcid.org/0000-0003-3209-5721>

· Received: Sep. 26, 2023, Revised: Nov. 07, 2023, Accepted: Nov. 10, 2023

· Corresponding Author: Jaewoo Cho

School of Social Science, Hansung University

Tel.: +82-2-760-8034, Email: [jaewoocho@hansung.ac.kr](mailto:jaewoocho@hansung.ac.kr)

## I. 서론

부동산 시장은 국가경제의 중심축 중 하나로 개인, 기업, 그리고 정부에 이르기까지 다양한 이해관계자의 관심을 받고 있다. 특히, 부동산 가격의 변화는 경제성장 및 안정에 중요한 영향을 끼치며 더 넓게는 도시계획, 금융, 사회복지 등 다양한 분야에 밀접한 관계를 갖는다. 따라서 부동산 가격을 보다 정확하게 예측하는 모형 개발은 오랜 기간동안 연구자들의 관심을 받아왔다.

그러나 부동산이라는 자산은 공간적인 특성을 가지고 있다는 점에서 다른 자산군과 차별되는 특성을 갖는다. “부동산에서 중요한 세 가지는 입지, 입지, 그리고 입지이다”라는 해럴드 사무엘의 격언은 부동산이 갖는 공간적인 성격을 잘 드러낸다. 나아가 특정 지역의 부동산 가격은 그 지역의 지리적 특성, 주변 지역과의 상호작용, 교통 연결성 등에도 크게 영향을 받는다. 이러한 공간적 상호작용은 부동산 가격의 복잡한 구조를 형성하기에, 보다 신뢰도가 높은 모형을 만들기 위해서는 부동산 시장의 공간적 상관성을 고려해야 한다.

기존의 부동산 가격 예측 모델은 통계적인 접근 방법을 중심으로 하였으나, 최근에는 머신러닝과 딥러닝 기술의 발전으로 예측의 정확도를 높일 수 있는 새로운 방법론이 제시되고 있다. 그러나 인공지능을 활용한 대부분의 기존 모형은 부동산의 가격에 영향을 미치는 공간적 상관성을 무시하거나 충분히 고려하지 않았다는 점에서 한계가 있다.

이에 본 연구는 부동산의 본질적인 공간적 특성을 고려한 딥러닝 기반 부동산 가격 예측 모형을 제시한다. 이 모형은 부동산 시장의 공간적 상관성을 딥러닝에 반영함으로써, 부동산의 복잡한 가격 구조를 더 정확하게 모델링하고 예측하고자 하였다. 이 연구를 통해 부동산 시장 분석, 투자 전략 수립, 도시계획과 정책 결정 등의 분야에서 중요한 시사점을 제공할 수 있으리라 기대한다.

본 연구는 다음과 같이 구성된다. 2장에서는 인공지능 기반의 부동산 가격 예측 연구에 대한 문헌 고찰을 수행하고, 3장에서는 본 연구에서 활용하는 연구 데이터의 범위와 전처리 방법에 대해서 기술한다. 그리고 4장에서는 본 연구에서 활용하는 연구

방법론과 검증 방법론, 5장에서는 연구 결과, 6장에서는 연구의 결론과 추후 연구에 대하여 설명한다.

## II. 관련 연구

부동산 가격 예측모형은 오랜 기간 동안 연구자들의 관심을 받으며 발전해왔다. R. G. Ridker, J. A. Henning이 최초로 헤도닉 모형을 주거용 주택가격 예측에 적용한 이후 예측모형은 비선형성, 동적 평형성, 자기회귀성 등을 포함하며 정교함을 높였다 [1]-[4]. 최근에는 빅데이터와 머신러닝 기술 발전에 힘입어 부동산 가격 예측모형도 전환점을 맞이했다. 비선형적이고 복잡한 관계 탐지에 특화된 머신러닝의 특성은 다양한 요인이 복합적으로 작용하는 부동산 시장에 적합했고 이전 모형보다 향상된 예측 능력을 보였다.

인공신경망을 활용해 부동산 가격을 예측하는 초기 연구로는 A. Q. Do, G. Grudnitski가 있다[5]. 이 연구는 기초적인 구조의 MLP(Multi-Layer Perceptron)를 사용하여 연면적, 건축물연령 등 기본적인 주택 정보를 토대로 미국 샌디에고의 부동산 가격을 예측하였다. 이어진 연구들에서도 투입된 변수에는 다소 차이가 있으나 유사한 구조의 인공신경망 구조를 사용했다[6]-[11]. 일반적으로 인공신경망 예측모형은 이전까지 사용되던 다중회귀분석 모형에 비해 데이터가 커질수록 예측력에서 강점을 갖는 걸로 나타났으나[5][12], 일부 연구에서는 인공신경망이 작은 데이터에서 상대적으로 더 큰 강점이 있다는 주장이 제기되기도 하였다[8][13][14].

2000년대 들어 새롭고 효율적인 인공신경망 알고리즘이 개발되어 딥러닝의 시대가 시작되었고[15], 딥러닝을 활용한 부동산 가격 예측모형도 함께 발전하였다. 특히 경기 변동에 따라 시계열적 특성을 지니는 부동산 시장의 특성상 RNN(Recurrent Neural Network)이나 LSTM(Long Short Term Memory)과 같은 시계열을 반영하는 모형에 대한 연구가 두드러졌다[16]-[19]. 대부분 연구에서 RNN과 LSTM 모형은 부동산 가격 예측에 우수한 예측력을 보였다. 최근에는 GRU(Gated Recurrent Layer)와 같은 새로운 모형에 대한 연구도 진행되고 있으며 다양한 모형을 혼합해 적용하는 방식도 시도되고 있다[20].

국내에서는 2000년대 초반 인공지능망을 활용해 부동산 가격을 예측하는 연구가 진행되었다. T. H. Kim, H. Hong은 서울시 송파구와 도봉구 아파트 단지를 대상으로 회귀모형과 인공지능망 모형을 적용하고 예측력을 비교하였다[21]. 이 연구에서 평형, 방수, 세대수 등을 입력변수로 사용했을 때, 인공지능망 모형은 회귀모형만큼 높은 예측력을 보였다. 이후 연구에서도 평당분양가격, 주택가격지수, 주택 매매실거래가 등을 결과값으로 하여 인공지능망 모형의 예측력을 분석하였는데, 일관되게 우수한 예측력을 보였다[22]-[26]. 그러나 이들 연구는 기초적인 MLP 모델만을 적용했다는 점에서 시계열적 요소를 고려하지 못했다는 한계를 갖는다.

국내 부동산 가격 예측모형에서 LSTM 모델을 최초로 적용한 연구는 S. W. Bae, J. S. Yu이 수행했다 [27]. 이들은 서울 아파트 매매실거래가격지수에 대해 ARIMA(Autoregressive Integrated Moving Average), DNN(Deep Neural Networks), LSTM 모형의 예측력을 비교했는데, 기존 시계열모형인 ARIMA보다 딥러닝 모형인 DNN과 LSTM이 상대적으로 우수한 것으로 나타났다. 딥러닝 모형을 비교했을 때에는 DNN이 LSTM보다 미세하게 우수했다. 이들은 후속 연구에서 VAR(Vector Autoregression), BVAR(Bayesian Vector Autoregressive), SVM(Support Vector Machine), RF(Random Forest), GBRT(Gradient Boosting Decision Tree)를 추가하여 예측력을 비교했는데, 역시 머신러닝이 기존 시계열모형보다 우수한 것으로 나타났다. 특히 외부 요인 등으로 시장이 급변할 때 선형성을 가정하고 예측하는 시계열 모형보다 머신러닝이 시장 추세를 더 잘 예측하는 것으로 나타났다[28]. T. H. Lee, M. J. Cheon은 거시경제지표를 활용해 서울 주아파트 가격지수를 VAR, RNN, LSTM으로 예측했는데, LSTM이 가장 높은 예측력을 갖는 것으로 분석했다[29]. RNN과 LSTM은 예측력에서 큰 차이를 보이지는 않았으나 표준편차에서 LSTM이 작았기 때문에 예측모델로 더 우수하다고 평가했다. 후속 연구에서도 RF, XGBoost(eXtreme Gradient Boosting), LSTM 중에서 LSTM의 예측 정확도가 가장 높은 것으로 나타났다[30].

그러나 같은 예측모형이라도 입력값의 변화에 따라 예측결과가 달라질 수 있다. 특히 시계열 모델에

서는 하나의 입력값, 즉 이전 시점의 부동산 가격만 예측에 활용하는 단변량 모델과 다양한 변수를 함께 고려하는 다변량 모형이 중요한 차이점을 보인다. S. W. Bae, J. S. Yu의 연구에서는 2006-2016년 기간에는 머신러닝방법 중 단변량 LSTM이 가장 예측력이 뛰어났고, 다변량 RF, 다변량 DNN 순으로 예측력이 우수한 것으로 나타났다. 그러나 2006-2008과 같이 외부 충격이 발생한 기간에는 다변량 LSTM이 가장 우수하고 단변량 RF, 단변량 DNN 순으로 예측력이 높았다[28]. 1972년부터 2017년까지의 호주 주택가격을 다양한 모델에 적용하여 비교한 연구에서도 단변량 모델이 전반적으로 예측력이 좋게 나왔는데, 연구자는 그 이유를 다변량 기반 예측모형을 제대로 만들기 위해서는 시간과 노력이 많이 필요하기 때문이라고 설명했다[31]. 즉, 컴퓨팅 자원이 한정되고 시장 흐름이 안정적인 때에는 단변량 모델로 충분할 수 있으나 시장이 불안정할 때에는 충분한 자원을 투입한 다변량 모델이 더 효과적일 수 있다는 의미이다.

또한 RNN과 LSTM은 변수가 시계열로 입력되기 때문에 적절한 시계열을 찾는 것도 중요하다. 전국 50개 아파트 매매가격을 LSTM으로 예측한 연구에서 다양한 입력값의 시간적 범위를 비교한 결과 31개월이 최적이라는 결론을 내렸다[32]. 다시 말해 직전 31개월의 아파트 매매가격과 기타 변수를 입력값으로 사용했을 때, 3개월 뒤의 아파트 매매가격을 가장 잘 예측할 수 있었다. 이와 대조적으로 또 다른 연구에서는 직전 2달 가격을 사용했을 때 예측결과가 가장 우수했는데, 이에 대한 이유로 시장 참여자들이 단기적인 가격에 대한 민감성이 지적되었다[33]. 그러나 앞선 연구와 달리 이 연구는 서울 강남구 소재 아파트만을 대상으로 했기 때문에 직접적인 비교가 어려울 수 있다. 그러나 딥러닝 예측모델의 최적 입력값이 시간적, 공간적 범위에 따라 달라질 수 있다는 점은 확인할 수 있다.

부동산은 본질적으로 공간적 성질을 내포하기 때문에 예측모형에서 공간 요소를 고려하는 것도 중요하다. 이창로·김세형은 강남, 김해, 덕진, 해남 4개 사례지역에 대한 연구에서 지리좌표를 입력값으로 사용했는데, 딥러닝 예측모형에서 지리좌표가 긍정적인 영향을 갖는 것으로 파악했다[26].

12 공간적 상관성을 고려한 딥러닝 기반 부동산 가격 예측 방법 제안

J. S. Kim, K. Kim도 이와 유사하게 딥러닝이 비선형 관계에서 갖는 강점을 위도, 경도, 고도 등 지리좌표와 주택가격과의 관계에 적용했는데, 지리좌표를 포함한 MLP 모델이 전통적인 회귀모형보다 우수한 것으로 나타났다[34].

그러나 예측오차가 지리적 자기상관성을 보였다는 점에서 향후 연구에서 공간적 요소를 섬세하게 적용해야 함을 암시한다.

### III. 연구 데이터

#### 3.1 국토교통부 부동산 실거래가 데이터

본 연구에서 활용하는 데이터는 국토교통부 실거래가 공개 시스템에서 제공하는 아파트 매매 실거래가 데이터이다[36]. 이 데이터들 중 2021년 2월부터 시스템에 추가된 정보인 취소된 계약 건은 제거하였다.

본 연구 데이터의 시간적 범위 시작점은 가장 최근 부동산 상승기의 시작점인 2013년 1분기부터이다[37]. 그리고 시간적 범위의 종료점은 가장 최근까지 업데이트 된 시점인 2023년 2분기이다.

또한, 데이터의 공간적 범위는 표 1과 같다. 2023년 5월 기준, 서울특별시의 아파트 단지 2,582개 중 3,000세대 이상의 대단지 34개를 선별하였다[38]. 이 중 강남구 압구정동의 현대 아파트는 1차부터 14차, 그리고 신현대까지 총 15개의 단지는 위치, 준공년도 및 가격 흐름을 고려하여 1개의 단지로 처리하였다.

이렇게 3,000세대 이상의 아파트 단지가 속해 있는 서울특별시의 지역구는 총 25개 중 강남구, 강동구 등 총 13개이다. 이 아파트 단지들의 최소 세대수는 3,002, 최대 세대수는 9,510이며, 평균은 4,112.5, 표준편차는 1,319.2이다.

#### 3.2 데이터의 전처리

앞에서 수집한 데이터는 표 2와 같이 전용면적 50.0m<sup>2</sup> 이상 75.0m<sup>2</sup> 미만까지의 20평형대와 전용면적 75.0m<sup>2</sup> 이상 100.0m<sup>2</sup> 이하까지의 30평형대 그리고 두 평형대 모두로 구분하였다. 또한 전체 기간,

코로나19 팬데믹 발생 전, 후 등 3가지로 나누어 총 9가지의 샘플군을 만들었다.

표 1. 연구 데이터의 공간적 범위  
Table 1. Spatial ranges of research data

District	Apartment complex name	Number of households
Gangnam-gu	개포주공1단지	5,040
	은마	4,424
	압구정현대	3,130
	도곡렉슬	3,002
Gangdong-gu	고덕그라시움	4,932
	고덕아르테온	4,066
	래미안힐스테이트 고덕	3,658
	강동롯데캐슬퍼스트	3,226
Gangbuk-gu	SK북한산시티	3,830
Gwanak-gu	관악드림타운	3,544
Nowon-gu	월계시영고층	3,930
	중계그린	3,481
	상계보람	3,315
	월계그랑빌	3,003
Dobong-gu	방학신동아1단지	3,169
Mapo-gu	마포래미안푸르지오	3,885
	성산시영	3,710
Seodaemun-gu	DMC파크뷰자이	4,300
	DMC래미안e편한세상	3,293
Seocho-gu	반포본동아파트	3,610
	반포자이	3,410
Seongbuk-gu	돈암동한신한진	3,929
Songpa-gu	헬리오시티	9,510
	잠실파크리오	6,864
	잠실엘스	5,678
	잠실리센츠	5,563
	올림픽선수기자촌	5,540
	올림픽힐러리타운	4,494
	잠실5단지	3,930
	잠실동트리지움	3,696
	신천장미1차2차	3,402
Yangcheon-gu	목동14단지	3,100
	목동 센트럴 아이파크위브	3,045
Jung-gu	신당남산타운	3,116

표 2. 데이터 구분

Table 2. Classification of data

	Size(m <sup>2</sup> )	Period
A1	50.0m <sup>2</sup> ~ 75.0m <sup>2</sup>	2013Q1 ~ 2023Q2
A2		2013Q1 ~ 2019Q4
A3		2020Q1 ~ 2023Q2
B1	75.0m <sup>2</sup> ~ 100.0m <sup>2</sup>	2013Q1 ~ 2023Q2
B2		2013Q1 ~ 2019Q4
B3		2020Q1 ~ 2023Q2
C1	50.0m <sup>2</sup> ~ 100.0m <sup>2</sup>	2013Q1 ~ 2023Q2
C2		2013Q1 ~ 2019Q4
C3		2020Q1 ~ 2023Q2

변수로 사용되는 거래가격은 면적별로 발생할 수 있는 차이를 고려하기 위해 식 (1)과 같이 1m<sup>2</sup> 당 면적별 / 기간별 평균 가격(만원)인 PPSi(Price Per Size)으로 변환하였다.

$$PPS_i = \frac{\sum TransactionPrice_i}{\sum Size_i} \quad (1)$$

PPSi로 정리된 데이터는 공간적 범위가 작기 때문에 분기 마다 1건의 매매 거래도 존재하지 않는 결측치가 발생할 수 있다. 그래서 결측치는 해당 결측치와 가장 가까운 이전, 이후의 평균값으로 한다.

### 3.3 기술 통계

데이터를 전처리 후, 카테고리별 PPS 변수의 기술통계를 산출한 결과는 표 3~5와 같다.

20평형대의 PPS 변수를 시기별로 기술통계를 산출한 결과는 표 3과 같다. 코로나19 팬데믹 이전과 이후인 카테고리 A2와 A3를 구별로 비교하면, PPS 평균값은 강남구가 1.62배로 최소, 노원구가 2.02배로 최대, 구별 평균은 1.82배로서 코로나19 팬데믹 이후의 값이 이전보다 더 높음을 알 수 있다.

30평형대의 PPS 변수를 시기별로 기술통계를 산출한 결과는 표 4와 같다. 코로나19 팬데믹 이전과 이후인 카테고리 B2와 B3를 구별로 비교하면, PPS 평균값은 양천구가 1.25배로 최소, 중구가 1.89배로 최대, 구별 평균은 1.72배로서 코로나19 팬데믹 이후의 값이 이전보다 더 높음을 알 수 있다.

표 3. 카테고리 A의 구별 PPS 기술 통계

Table 3. Descriptive statistics for PPS of category A1(n = 17,694), A2(n = 14,009), and A3(n = 3,685)

	District	Minimum	Maximum	Mean	Std.dev
A1	Gangnam-gu	1000.38	5557.76	2151.72	865.58
	Gangdong-gu	691.78	2559.17	1461.2	510.85
	Gangbuk-gu	316.77	1300.43	660.37	215.8
	Gwanak-gu	417.71	1566.67	751.72	270.09
	Nowon-gu	257.22	1748.25	662.1	283.95
	Dobong-gu	266.21	1091.05	426.57	131.92
	Mapo-gu	519.69	2834.96	1214.4	526.97
	Seodaemun-gu	557.43	2181.96	1158.58	343.38
	Seocho-gu	1333.78	4701.57	2257.07	901.09
	Seongbuk-gu	335.68	1262.29	605.55	214.65
	Songpa-gu	716.29	3652.43	1700.25	633.89
	Yangcheon-gu	481.64	2499.09	1168.2	410.84
	Jung-gu	316.98	2127.13	997.97	376.83
A2	Gangnam-gu	1000.38	5557.76	2080.06	820.92
	Gangdong-gu	691.78	1954.51	1099.14	295.22
	Gangbuk-gu	316.77	816.94	552.61	94.69
	Gwanak-gu	417.71	1044.63	637.61	134.14
	Nowon-gu	257.22	1081.97	559.9	157.84
	Dobong-gu	266.21	526.71	372.46	52.28
A2	Mapo-gu	519.69	2252.06	1013.66	352.25
	Seodaemun-gu	557.43	1650.63	1003.19	211.66
	Seocho-gu	1333.78	3501.17	1980.58	648.01
	Seongbuk-gu	335.68	792.6	505.96	104.7
	Songpa-gu	716.29	2939.21	1493.47	450.69
	Yangcheon-gu	481.64	1862.96	981.09	280.79
	Jung-gu	316.98	1463.13	846.37	223.15
A3	Gangnam-gu	2666.95	5471.1	3375.23	675.44
	Gangdong-gu	1167.44	2559.17	1968.78	251.05
	Gangbuk-gu	691.9	1300.43	984.83	139.56
	Gwanak-gu	980.79	1566.67	1221.23	157.67
	Nowon-gu	591.02	1748.25	1132.63	260.13
	Dobong-gu	410.65	1091.05	614.14	151.49
	Mapo-gu	1009.59	2834.96	2007.91	309.57
	Seodaemun-gu	1350.43	2181.96	1643.89	187.22
	Seocho-gu	3251.08	4701.57	3830.95	352.86
	Seongbuk-gu	587.45	1262.29	945.06	128.93
	Songpa-gu	1670.56	3652.43	2747.22	308.14
	Yangcheon-gu	1162.46	2499.09	1628.23	303.67
	Jung-gu	1251.25	2127.13	1613.31	209.51

14 공간적 상관성을 고려한 딥러닝 기반 부동산 가격 예측 방법 제안

표 4. 카테고리 B의 구별 PPS 기술 통계

Table 4. Descriptive statistics for PPS of category B1(n = 22,333), B2(n = 17,945), B3(n = 4,388)

	District	Minimum	Maximum	Mean	Std.dev
B1	Gangnam-gu	592.21	5090.91	1803.8	667.28
	Gangdong-gu	553.07	2379.65	1228.76	443.19
	Gangbuk-gu	329.72	1048.05	544.89	168.95
	Gwanak-gu	367.23	1359.46	626.16	198.3
	Nowon-gu	295.62	1266.32	575.91	197.53
	Dobong-gu	253.36	736.51	393.55	109.2
	Mapo-gu	407.81	2299.11	1388.73	446.3
	Seodaemun-gu	413.86	1812.32	964.77	282.17
	Seocho-gu	1235.53	4591.31	2183.65	756.61
	Seongbuk-gu	329.92	1160.6	576.49	189.69
	Songpa-gu	646.46	3969.01	1547.99	570.96
	Yangcheon-gu	654.18	2038.41	1131.32	291.73
	Jung-gu	497.17	1873.23	848.92	295.28
B2	Gangnam-gu	592.21	3405.08	1596.49	475.82
	Gangdong-gu	553.07	1703.41	956.19	264.35
	Gangbuk-gu	329.72	684.29	460.13	71.07
	Gwanak-gu	367.23	847.46	555.76	105.03
	Nowon-gu	295.62	849.49	504.44	104.92
	Dobong-gu	253.36	490.45	346.14	43.86
	Mapo-gu	407.81	1950.41	1201.54	323.6
	Seodaemun-gu	413.86	1412.25	851.66	172.11
	Seocho-gu	1235.53	3331.65	1903.18	482.47
	Seongbuk-gu	329.92	774.13	499.84	99.6
	Songpa-gu	646.46	2949.95	1368.13	425.8
	Yangcheon-gu	654.18	1680.47	1015.22	261.85
	Jung-gu	497.17	1396.09	753.62	173.99
B3	Gangnam-gu	1871.37	5090.91	2884.01	441.33
	Gangdong-gu	1012.96	2379.65	1721.71	218.51
	Gangbuk-gu	506.72	1048.05	796.4	116.8
	Gwanak-gu	765.07	1359.46	1003.16	148.74
	Nowon-gu	529.97	1266.32	927.7	167.29
	Dobong-gu	329.8	736.51	555.63	110.34
	Mapo-gu	1646.8	2299.11	1992.22	161.37
	Seodaemun-gu	791.24	1812.32	1411.64	163.86
	Seocho-gu	2872.51	4591.31	3448.24	385.36
	Seongbuk-gu	659.83	1160.6	898.71	130.13
	Songpa-gu	1179.38	3969.01	2399.68	364.06
	Yangcheon-gu	682.59	2038.41	1269.75	264.58
	Jung-gu	1001.41	1873.23	1423.39	213.1

표 5. 카테고리 C의 구별 PPS 기술 통계

Table 5. Descriptive statistics for PPS of category C1(n = 40,027) C2(n = 31,954), C3(n = 8,073)

	District	Minimum	Maximum	Mean	Std.dev
C1	Gangnam-gu	592.21	5557.76	1912	752.04
	Gangdong-gu	553.07	2559.17	1334.52	488.99
	Gangbuk-gu	316.77	1300.43	603.67	202.57
	Gwanak-gu	367.23	1566.67	688.61	244.87
	Nowon-gu	257.22	1748.25	642.3	269.03
	Dobong-gu	253.36	1091.05	416.68	126.43
	Mapo-gu	407.81	2834.96	1250.02	516.26
	Seodaemun-gu	413.86	2181.96	1013.31	310.19
	Seocho-gu	1235.53	4701.57	2205.47	802.46
	Seongbuk-gu	329.92	1262.29	595.36	206.65
	Songpa-gu	646.46	3969.01	1576.98	586.49
	Yangcheon-gu	481.64	2499.09	1159.14	385.21
	Jung-gu	316.98	2127.13	915.19	341.99
C2	Gangnam-gu	592.21	5557.76	1759.44	655.21
	Gangdong-gu	553.07	1954.51	1017.76	286.86
	Gangbuk-gu	316.77	816.94	507.28	95.82
	Gwanak-gu	367.23	1044.63	595.52	126.8
	Nowon-gu	257.22	1081.97	547.04	149.11
	Dobong-gu	253.36	526.71	364.59	51.33
	Mapo-gu	407.81	2252.06	1050.7	354.69
	Seodaemun-gu	413.86	1650.63	888.14	193.53
	Seocho-gu	1235.53	3501.17	1926.82	539.16
	Seongbuk-gu	329.92	792.6	503.75	102.89
	Songpa-gu	646.46	2949.95	1392.22	433.5
	Yangcheon-gu	481.64	1862.96	987.9	277.29
	Jung-gu	316.98	1463.13	793.34	201.75
C3	Gangnam-gu	1871.37	5471.1	2949.96	507.09
	Gangdong-gu	1012.96	2559.17	1843.75	265.55
	Gangbuk-gu	506.72	1300.43	891.8	159.55
	Gwanak-gu	765.07	1566.67	1123.43	188.03
	Nowon-gu	529.97	1748.25	1087.54	257.12
	Dobong-gu	329.8	1091.05	596.47	142.78
	Mapo-gu	1009.59	2834.96	2004.28	282.23
	Seodaemun-gu	791.24	2181.96	1478.13	200.44
	Seocho-gu	2872.51	4701.57	3547.09	411.92
	Seongbuk-gu	587.45	1262.29	930.51	130.91
	Songpa-gu	1179.38	3969.01	2462.93	379
	Yangcheon-gu	682.59	2499.09	1506.55	336.7
	Jung-gu	1001.41	2127.13	1523.39	231.18

전체 평형대의 PPS 변수를 시기별로 기술통계를 산출한 결과는 표 5와 같다. 코로나19 팬데믹 이전과 이후인 카테고리 C2와 C3를 구별로 비교하면, PPS 평균값은 양천구가 1.53배로 최소, 노원구가 1.99배로 최대, 구별 평균은 1.79배로서 코로나19 팬데믹 이후의 값이 이전보다 더 높음을 알 수 있다.

## IV. 연구 방법론

### 4.1 연구 방법론

본 연구에서 활용하는 연구 방법론은 시계열 데이터 예측에 적합한 순환 신경망 기반의 RNN, LSTM, GRU이다[35]. RNN은 시차가 큰 값의 기억력이 소실되는 장기 의존성 문제가 있어 RNN 대비 순환 층이 추가되었다는 차이가 있다[39]. 또한, GRU는 LSTM과 성능은 비슷하며, 그 구조를 단순화 시킨 모형이다[40].

RNN 모형은 과거의 데이터를 기억하여 예측 데이터를 출력하는 방식이다. 그러나 RNN은 짧은 시퀀스를 처리할 경우에 유리하여, 예측하려는 데이터와 과거의 데이터의 거리가 멀어지는 경우에는 학습 능력이 현저하게 떨어지는 장기 의존성 문제(Long-term dependency problem)가 있다. RNN의 장기 의존성 문제를 해결하기 위하여 중간에 기억 소자(Memory cell)를 추가하여 보완한 것이 LSTM 모형이다. 또한, LSTM에서의 출력, 입력, 삭제 게이트를 업데이트, 리셋 게이트로 줄이고, 학습 속도를 개선하면서 비슷한 성능을 보이는 GRU 모형이 있다[39][40][41].

### 4.2 연구 모형 및 관련 설정

본 연구에서는 사용한 딥러닝 모형은 지역구  $i$ 의  $t$  시점에서의 평균 면적당 가격  $PPS_{it}$ 를 출력값(output)으로 설정한다. 입력값은  $i$  지역구를 포함한 13개 지역구의  $t-1 \sim t-3$  시점에서의 PPS로 한다. 특정 지역의 평균 면적당 가격을 예측하기 위해 해당 지역과 주변 지역의 부동산 시장 흐름을 사용하는 것이다.

표 6은 RNN, LSTM, GRU 모형에 공통적으로 적

용된 하이퍼 파라미터 리스트이다. 입력 변수는 12개로 하고, 은닉 레이어는 8, 4로 2개, 학습 횟수인 Epochs는 100, 학습 단위인 Batch size는 1로 한다. 그리고 Sigmoid와 tanh가 갖는 기울기 소멸 문제(Gradient vanishing)을 해결하는 ReLU(Rectified Linear Unit)를 활용한다.

학습 데이터와 검증 데이터는 7 : 3으로 무작위 배분하여 활용한다. 학습과 검증은 5회씩 수행하였으며, 5회 수행의 평균값을 최종 결과값으로 하였다.

표 6. 연구 모형의 하이퍼 파라미터

Table 6. Hyper-parameters on research model

Category	Value	Misc.
Number of input variables	12	The values are common for the RNN, LSTM, GRU models
Number of hidden layers	2	
Epochs	100	
Batch size	1	
Activation function	ReLU	

### 4.3 검증 방법론

본 연구에서는 모형의 정확도를 측정하기 위하여 식 (2)와 같이 MAPE(Mean Absolute Percentage Error)를 활용한다. MAPE는 실제값과 예측값의 차이를 실제값으로 나눈 절대값을 구해 해당 오차의 상대적인 비율을 산출하는데 활용된다[42]. 이 값은 작을수록 오차가 작은 것을 의미하므로 모형의 성능이 좋음을 의미한다.

$$MAPE = \frac{1}{m} \sum_{i=1}^m \left| \frac{Y_i - X_i}{Y_i} \right| \quad (2)$$

## V. 연구 결과

### 5.1 RNN 수행 결과

표 7과 그림 1은 지역구별 RNN 모형을 수행한 결과이다.

MAPE는 8.6에서 13.28 수준이며 평균은 10.91로

16 공간적 상관성을 고려한 딥러닝 기반 부동산 가격 예측 방법 제안

예측의 정확도가 우수한 수준임을 알 수 있었다 [43]. 카테고리별로 살펴보면 코로나19 팬데믹 이전인 2013~2019년만을 사용한 예측결과가 전체 기간을 대상으로 한 예측결과보다 우수한 것으로 나타

났다. 또한, 코로나19 팬데믹 이후를 대상으로 예측한 결과는 전체 또는 코로나 팬데믹 19 이전에 비해 상대적으로 낮은 15.38에서 20.01 수준임을 알 수 있었다.

표 7. RNN 수행 결과  
Table 7. Result on RNN

	A <sub>1</sub>	A <sub>2</sub>	A <sub>3</sub>	B <sub>1</sub>	B <sub>2</sub>	B <sub>3</sub>	C <sub>1</sub>	C <sub>2</sub>	C <sub>3</sub>	MAPE Ave.
Gangnam-gu	10.28	9.64	14.52	12.53	10.59	22.31	7.45	5.25	26.96	13.28
Gangdong-gu	12.62	9.58	13.33	10.22	9.43	24.76	7.01	5.66	23.53	12.9
Gangbuk-gu	15.17	8.91	12.4	10.41	8.27	23.41	8.47	6.17	24.94	13.13
Gwanak-gu	14.18	8.63	14.86	9.3	9.06	21.72	8.31	5.93	21.82	12.65
Nowon-gu	13.45	9.33	15.65	9.76	9.47	20.86	9.23	5.8	21.54	12.79
Dobong-gu	4.95	8.51	21.9	6.12	8.77	20.17	6.03	8.22	5.26	9.99
Mapo-gu	5.01	9.28	15.56	6.71	8.14	20.73	6.29	8.67	7.57	9.77
Seodaemun-gu	5.8	8.67	14.09	6.6	7.73	19.2	6.5	7.48	10.38	9.61
Seocho-gu	6.41	9.1	14.69	6.74	8.96	18.18	6.3	7.74	11.78	9.99
Seongbuk-gu	7.15	9.06	13.46	6.94	8.64	19.31	6.12	7.47	13.65	10.2
Songpa-gu	7.69	4.48	16.69	6.46	3.8	13.76	7.23	7.45	9.8	8.6
Yangcheon-gu	8.16	4.69	15.49	6.57	3.7	18.19	6.99	7.68	13.16	9.4
Jung-gu	8.88	5.33	17.32	6.56	3.69	17.53	6.92	6.39	13.44	9.56
Mean	9.21	8.09	15.38	8.07	7.71	20.01	7.14	6.92	15.68	10.91

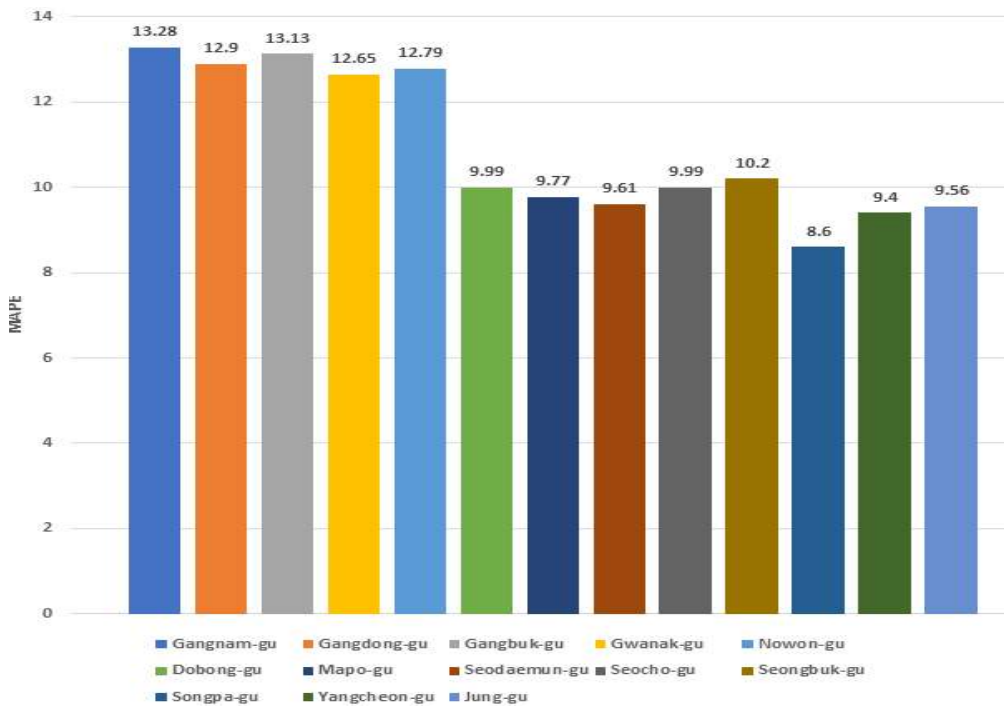


그림 1. RNN 수행 결과  
Fig. 1. Result on RNN



## 5.2 LSTM 수행 결과

표 8과 그림 2는 지역구별 LSTM 모형을 수행한 결과이다. MAPE는 8.3에서 17.06 수준이며 평균은 12.49로 예측의 정확도가 우수한 수준임을 알 수 있

었다[43]. 그리고 RNN 모형과 마찬가지로 코로나19 팬데믹 이후인 카테고리 A3, B3, C3의 성능은 전체 또는 코로나 팬데믹 19 이전에 비해 상대적으로 낮은 17.29에서 22.88 수준임을 알 수 있었다.

표 8. LSTM 수행 결과  
Table 8. Result on LSTM

	A1	A2	A3	B1	B2	B3	C1	C2	C3	MAPE Ave.
Gangnam-gu	7.69	7.25	34.82	9.72	12.35	13.75	18.2	10.08	39.64	17.06
Gangdong-gu	10.15	7.53	21.39	10.58	12.52	16.03	15.59	8.26	40.71	15.86
Gangbuk-gu	10.99	7.91	17.79	9.61	10.46	15.08	12.39	8.16	34.75	14.13
Gwanak-gu	11.61	8.58	14.63	10.12	10.21	22.13	11.76	7.77	28.49	13.92
Nowon-gu	11.6	8.74	19.36	10.35	9.72	21.76	11.19	7.56	28.83	14.35
Dobong-gu	7.6	9.71	4.92	8.77	7.75	4.82	10.21	5.04	15.9	8.3
Mapo-gu	8.47	8.84	7.38	9.89	8.33	7.85	8.36	5.98	14.62	8.86
Seodaemun-gu	9.31	8.97	11.74	9.72	7.93	10.85	9.22	7.33	18.15	10.36
Seocho-gu	8.53	9.3	14.15	8.91	7.89	11.79	10.42	7.89	15.71	10.51
Seongbuk-gu	8.36	8.51	16.7	9.11	8.64	11.52	10.13	8.13	15.17	10.7
Songpa-gu	7.65	5.37	21.55	15.78	3.63	40.67	6.71	3.35	14.72	13.27
Yangcheon-gu	8.87	4.29	22.17	12.82	4.87	36.48	6.53	3.82	14.88	12.75
Jung-gu	10.03	5.67	18.16	12.47	5.23	33.2	6.73	3.22	15.86	12.29
Mean	9.30	7.74	17.29	10.60	8.43	18.92	10.57	6.66	22.88	12.49

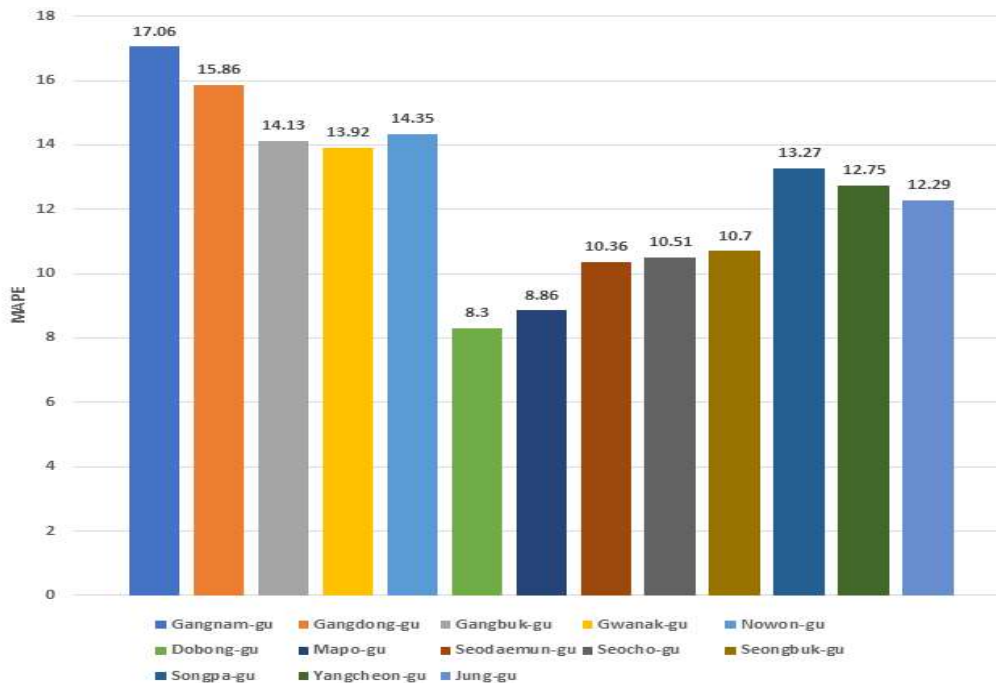


그림 2. LSTM 수행 결과  
Fig. 2. Result on LSTM

5.3 GRU 수행 결과

표 9과 그림 3은 지역구별 GRU 모형을 수행한 결과이다. MAPE는 10.78에서 13.98 수준이며 평균은 12.07로 예측의 정확도가 우수한 수준임을 알 수 있

었다[43]. 그리고 RNN, LSTM 모형과 마찬가지로 코로나19 팬데믹 이후인 카테고리 A3, B3, C3의 성능은 전체 또는 코로나 팬데믹 19 이전에 비해 상대적으로 낮은 17.45에서 23.39 수준임을 알 수 있었다.

표 9. GRU 수행 결과  
Table 9. Result on GRU

	A1	A2	A3	B1	B2	B3	C1	C2	C3	MAPE Ave.
Gangnam-gu	12.92	6.88	23.2	9.72	8.84	25.52	11.97	6.9	19.91	13.98
Gangdong-gu	10.34	8.56	13.71	9.62	11.39	19.47	9.36	7.96	20.3	12.3
Gangbuk-gu	10.33	8.34	12.68	8.95	10.39	25.06	10	7.33	22.02	12.79
Gwanak-gu	10.58	8.42	14.85	9.28	10.26	23.44	10.46	6.95	23.34	13.06
Nowon-gu	10.53	8.11	14.26	8.94	10.41	25.2	10.03	7.64	25.26	13.38
Dobong-gu	5.57	6.49	15.23	5.4	13.9	19.37	5.87	15.69	9.51	10.78
Mapo-gu	5.67	7.8	18.11	5.36	11.66	19.52	6.23	10.81	14.63	11.09
Seodaemun-gu	6.02	8.6	17.54	5.73	10.54	21.96	8.33	10.12	15.31	11.57
Seocho-gu	5.84	8.04	17.63	5.71	10.94	22.2	8.32	9.08	14.16	11.32
Seongbuk-gu	5.84	7.8	19.02	5.75	11.03	21.32	7.89	8.77	13.77	11.24
Songpa-gu	8.68	8.78	23.64	7.08	4.07	28.88	6.25	9.02	10.75	11.91
Yangcheon-gu	7.91	5.95	19.21	7.24	7.85	30.07	6.34	6.41	19.63	12.29
Jung-gu	7.77	5.01	17.77	7.08	7	22.08	7.17	5.34	21.25	11.16
Mean	8.31	7.60	17.45	7.37	9.87	23.39	8.32	8.62	17.68	12.07

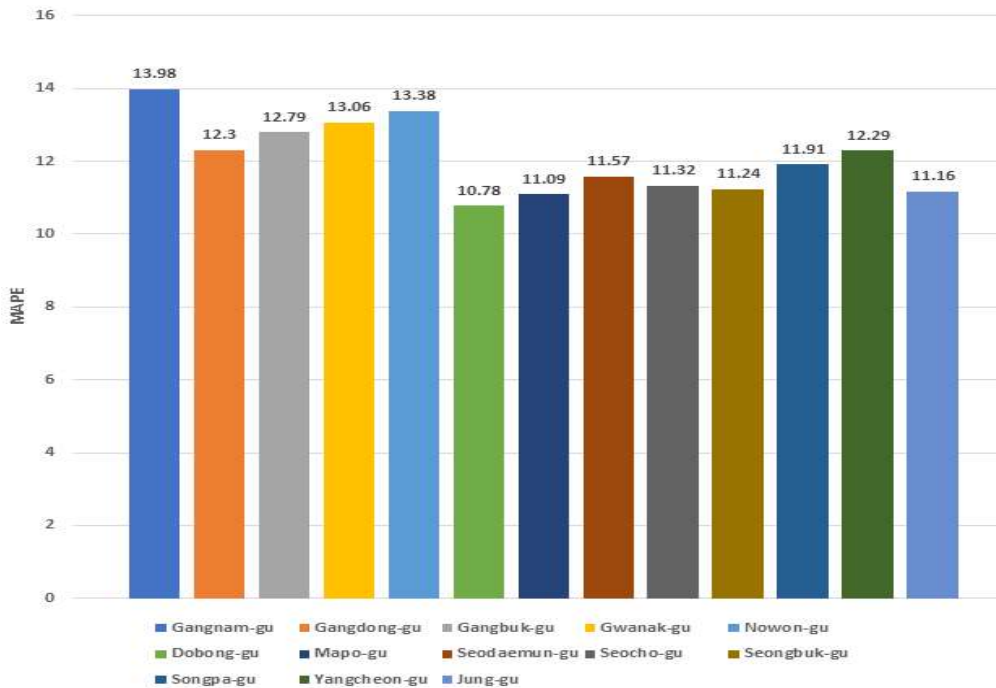


그림 3. GRU 수행 결과  
Fig. 3. Result on GRU

## VI. 결론 및 향후 과제

본 연구에서는 부동산의 공간적 특성을 고려한 딥러닝 기반 부동산 가격 예측 모형을 제시하였다. 이 모형은 부동산 시장의 공간적 상관성을 딥러닝에 반영함으로써 딥러닝 모형에서 부동산 하위시장 간의 역학관계를 반영하고자 시도하였다.

연구 결과는 다음과 같이 정리될 수 있다.

첫째, 본 연구의 모형은 특정 지역구의 평당가를 예측함에 있어 다른 지역의 평당가를 활용하는 것에 대한 가능성을 보여주었음을 알 수 있다.

둘째, 모형과 샘플별로 차이는 있으나 전반적으로 강남구에 대한 예측오차(MAPE)가 크게 나타났다. 이는 강남구 부동산 가격을 설명하는 데 있어 다른 지역의 부동산 하위시장의 영향력이 낮다는 의미이며, 강남구 부동산 시장이 다른 하위시장들보다 선형한다고 해석될 수도 있다. 강동구와 노원구도 상대적으로 높게 나타났는데 강남구만큼은 아니지만 독립적으로 선형해서 움직이는 시장이라 해석될 수 있다. 반면에 도봉구, 마포구는 전반적으로 예측오차가 낮게 나타났는데 다른 시장에 영향을 받아 후행적으로 움직이는 성향이 강하기 때문에 나온 결과라고 볼 수 있다.

셋째, 코로나19 팬데믹을 기준으로 해서 전후의 부동산 하위시장간 공간적 역학관계는 다르게 나타났다. 샘플 수의 차이로 인한 절대적인 예측오차 변화를 감안하더라도 코로나 이전에는 도봉구의 예측오차가 높았으나, 코로나 이후에는 강남구의 예측오차가 급격히 증가해 가장 높은 수치를 보였다. 이는 팬데믹 이전에는 적당한 영향력을 주고받으며 안정적으로 움직이던 하위시장들이 팬데믹 이후 역동성이 강화되면서 독립적으로 선형해 움직이는 시장과 그렇지 않은 시장으로 극명히 구분되었기 때문에 나타나는 현상으로 보인다.

넷째, 모델별 평균 정확도는 RNN이 10.91, LSTM이 11.44, GRU가 11.12로 RNN이 가장 우수했고, LSTM과 GRU가 비슷한 수준이었다. 장기 의존성 문제를 해결한 LSTM, GRU와 그렇지 않은 RNN과의 결과 차이가 크지 않다는 점은 부동산 시장 가격을 예측하는데 있어서는 장기 의존성 문제가 중요하지

않는다는 것을 의미한다. 즉, 부동산 시장의 가격은 최근의 가격이 중요하고, 시차가 큰 오래된 가격은 상대적으로 중요도가 떨어진다는 것을 의미한다. 그러나 일부 하위 시장에서는 RNN과 LSTM, GRU의 차이가 크게 나타나 지역별로 장기 의존성 문제가 다르게 나타날 수도 있다는 점을 시사한다.

그러나 하위시장간 구체적인 연결관계를 밝혀내지 못했다는 점은 본 연구의 한계점이다. 시장을 주도하는 특정 지역으로부터 각각 다른 하위시장으로 퍼져나가는 영향력이 얼마나 강한지, 그 시차는 어떻게 다른지 등에 대한 연구가 추가적으로 진행될 필요가 있다. 부동산 하위시장 간의 연결관계가 보다 구체적으로 밝혀질수록 가격 예측모형의 정확도는 지금보다 개선될 수 있을 것이다.

또한 본 연구는 면적당 평균 가격만을 변수로 사용했다는 점에서 단변량 모형에 가깝다고 할 수 있다. 그러나 후속 연구에서는 다변량 모형을 고려해 볼 수 있을 것이다. 금리, 경제성장률 등 거시경제적 변수뿐 아니라 단지별 접근성, 공공 인프라, 노후도 등 미시적 변수까지 활용한다면 보다 정교하게 각 시장의 흐름을 예측할 수 있을 것이다.

마지막으로 공간적 단위를 나누는 방법을 개선할 수 있다. 본 연구는 임의적으로 자치구별로 하위시장을 구분했지만 향후에는 동별 또는 단지별로 보다 세밀하게 시장을 나눌 수도 있다. 그러나 이 경우 데이터의 결측치를 보완할 수 있는 더 적절한 방법을 함께 고안할 필요가 있다.

## References

- [1] R. G. Ridker and J. A. Henning, "The determinants of residential property values with special reference to air pollution", *The review of Economics and Statistics*, Vol. 49, No. 2, pp. 246-257, May 1967. <https://doi.org/10.2307/1928231>.
- [2] D. M. Grether and P. Mieszkowski, "Determinants of real estate values", *Journal of Urban Economics*, Vol. 1, No. 2, pp. 127-145, Jan. 1974. [https://doi.org/10.1016/0094-1190\(74\)90013-8](https://doi.org/10.1016/0094-1190(74)90013-8).
- [3] A. C. Goodman, "Hedonic prices, price indices

- and housing markets", *Journal of urban economics*, Vol. 5, No. 4, pp. 471-484, Oct. 1978. [https://doi.org/10.1016/0094-1190\(78\)90004-9](https://doi.org/10.1016/0094-1190(78)90004-9).
- [4] R. F. Engle, D. M. Lilien, and M. Watson, "A dynamic model of housing price determination", *Journal of Econometrics*, Vol. 28, No. 3, pp. 307-326, Jun. 1985. [https://doi.org/10.1016/0304-4076\(85\)90003-X](https://doi.org/10.1016/0304-4076(85)90003-X).
- [5] A. Q. Do and G. Grudnitski, "A neural network approach to residential property appraisal", *The Real Estate Appraiser*, Vol. 58, No. 3, pp. 38-45, Jan. 1992.
- [6] A. Evans, H. James, and A. Collins, "Artificial Neural Networks: An Application to Residential Valuation in the UK", University of Portsmouth, Department of Economics, 1992.
- [7] R. A. Borst, "Artificial neural networks in mass appraisal", *Journal of Property Tax Assessment & Administration*, Vol. 1, No. 2, pp. 5-15, 1995.
- [8] P. Rossini, "Artificial neural networks versus multiple regression in the valuation of residential property", *Australian Land Economics Review*, Vol. 3, No. 1, pp. 1-12, Jan. 1997.
- [9] S. McGreal, A. Adair, D. McBurney, and D. Patterson, "Neural networks: the prediction of residential values", *Journal of Property Valuation and Investment*, Vol. 16, No. 1, pp. 57-70, Mar. 1998. <https://doi.org/10.1108/14635789810205128>.
- [10] J. G. Mora-Esperanza, "Artificial intelligence applied to real estate valuation: An example for the appraisal of Madrid", *Catastro*, Vol. 1, pp. 255-265, Apr. 2004.
- [11] A. Khalafallah, "Neural network based model for predicting housing market performance", *Tsinghua Science and Technology*, Vol. 13, No. S1, pp. 325-328, Oct. 2008. [https://doi.org/10.1016/S1007-0214\(08\)70169-X](https://doi.org/10.1016/S1007-0214(08)70169-X).
- [12] N. Nguyen and A. Cripps, "Predicting housing value: A comparison of multiple regression analysis and artificial neural networks", *Journal of real estate research*, Vol. 22, No. 3, pp. 313-336, Jun. 2020. <https://doi.org/10.1080/10835547.2001.12091068>.
- [13] H. James, "The reliability of artificial neural networks for property data analysis", *European Real Estate Society*, No. eres1996\_157, 1996.
- [14] A. B. Khamis and N. K. Kamarudin, "Comparative study on estimate house price using statistical and neural network model", *International Journal of scientific & technology research*, Vol. 3, No. 12, pp. 126-131, Dec. 2014.
- [15] G. E. Hinton, S. Osindero, and Y. W. The, "A fast learning algorithm for deep belief nets", *Neural computation*, Vol. 18, No. 7, pp. 1527-1554, Jul. 2006. <https://doi.org/10.1162/neco.2006.18.7.1527>.
- [16] X. Chen, L. Wei, and J. Xu, "House price prediction using LSTM", *arXiv preprint arXiv:1709.08432*, Sep. 2017. <https://doi.org/10.48550/arXiv.1709.08432>.
- [17] L. Yu, C. Jiao, H. Xin, Y. Wang, and K. Wang, "Prediction on housing price based on deep learning", *International Journal of Computer and Information Engineering*, Vol. 12, No. 2, pp. 90-99, 2018.
- [18] A. S. Temur, M. Akgün, and G. Temur, "Predicting housing sales in Turkey using ARIMA, LSTM and hybrid models", *Journal of Business Economics and Management*. Vol. 20, No. 5, pp. 920-938, Jul. 2019. <https://doi.org/10.3846/jbem.2019.10190>.
- [19] B. Afonso, L. Melo, W. Oliveira, S. Sousa, and L. Berton, "Housing prices prediction with a deep learning and random forest ensemble", *InAnais do XVI Encontro Nacional de Inteligência Artificial e Computacional*, pp. 389-400, Oct. 2019.
- [20] H. Ragb, A. Muntaser, E. Jera, A. Saide, and I. Elwarfalli, "Hybrid GRU-LSTM Recurrent Neural Network-Based Model for Real Estate Price Prediction", *TechRxiv*, Jan. 2023. <https://doi.org/10.48550/arXiv.2301.00000>.

- 36227/techrxiv.21929760.
- [21] T. H. Kim and H. Hong, "A Study on Apartment Price Models Using Regression Model and Neural Network Model", *The Korea Spatial Planning Review*, Vol. 43, pp. 193-200, Dec. 2004.
- [22] Y. W. Nam and J. M. Lee, "A Study on the Applicability of Neural Network Model for Prediction of the Apartment Market", *Korean Journal of Construction Engineering and Management*, Vol. 7, No. 2, pp. 162-170, Jan. 2006.
- [23] W. Chung and S. Lee, "A Study on the Forecasting of the Apartment Price Index Using Artificial Neural Networks", *Housing Studies*, Vol. 14, No. 3, pp. 39-64, Jan. 2007.
- [24] S. W. Min, "A study on the forecasting of housing price using deep learning : focusing on apartment price index in Seoul", Dissertation, Gangnam University, 2016.
- [25] S. Bae and J. Yu, "Estimation of the Apartment Housing Price Using the Machine Learning Methods: The Case of Gangnam-gu, Seoul", *Journal of the Korea Real Estate Analysts Association*, Vol. 24, No. 1, pp. 69-85, Mar. 2018. <https://doi.org/10.19172/KREAA.24.1.5>.
- [26] C. Lee and S. H. Kim, "The Deep Learning Approach to Property Valuation: An Application of a Multilayer Neural Net Model for Estimating House", *Journal of The Korean Regional Development Association*, Vol. 30, No. 4, pp. 179-201, Nov. 2018.
- [27] S. W. Bae and J. S. Yu, "Predicting the Real Estate Price Index Using Deep Learning", *Korea Real Estate Review*, Vol. 27, No. 3, pp. 71-86, Sep. 2017.
- [28] S. W. Bae and J. S. Yu, "Predicting the Real Estate Price Index Using Machine Learning Methods and Time Series Analysis Model", *Housing Studies*, Vol. 26, No. 1, pp. 107-133, Feb. 2018. <https://doi.org/10.24957/hsr.2018.26.1.107>.
- [29] T. H. Lee and M. J. Cheon, "Prediction of Seoul House Price Index Using Deep Learning Algorithms with Multivariate Time Series Data", *SH Urban Research & Insight*, Vol. 8, No. 2, pp. 39-56, Aug. 2018. <https://doi.org/10.26700/shuri.2018.08.8.2.39>.
- [30] J. M. Lee, S. H. Park, S. H. Cho, and J. H. Kim, "Comparison of Models to Forecast Real Estates Index Introducing Machine Learning", *Journal of the Architectural Institute of Korea*, Vol. 37, No. 1, pp. 191-199, Mar. 2021. <https://doi.org/10.5659/JAIK.2021.37.1.191>.
- [31] G. Milunovich, "Forecasting Australia's real house price index: A comparison of time series and machine learning methods", *Journal of Forecasting*, Vol. 39, No. 7, pp. 1098-1118, Feb. 2020. <https://doi.org/10.1002/for.2678>.
- [32] C. W. Kim, D. S. Kim, T. Y. Kim, and J. H. Kim, "A Prediction of the Apartment Sales Price Using Deep Learning", *proc. of 2020 Korean Institute of Communications and Information Sciences*, pp. 463-464, Nov. 2020.
- [33] K. Kim, K. S. Kim, and D. S. Nam, "A Study on the Index Estimation of Missing Real Estate Transaction Cases Using Machine Learning", *Journal of the Economic Geographical Society of Korea*, Vol. 25, No. 1, pp. 171-181, Mar. 2022. <https://doi.org/10.23841/egsk.2022.25.1.171>.
- [34] J. S. Kim and K. Kim, "How the Pattern Recognition Ability of Deep Learning Enhances Housing Price Estimation", *Journal of the Economic Geographical Society of Korea*, Vol. 25, No. 1, pp. 183-201, Mar. 2022. <https://doi.org/10.23841/egsk.2022.25.1.183>.
- [35] dProgrammer lopez, <https://dprogrammer.org/rnn-lstm-gru/> [accessed: Jul. 1, 2023]
- [36] Transaction Price Open System, Ministry of Land, Infrastructure and Transport, <https://rt.molit.go.kr/> [accessed: May 1, 2023]

- [37] S. J. Han and S. C. Jeong, "A Study on the Factors Affecting Apartment Price in Seoul and Provinces", Korea Real Estate Academy Review, Vol. 78, No. 3, pp. 92-103, Aug. 2019. <https://doi.org/10.31303/KREAR.2019.78.92>.
- [38] K-apt, <http://www.k-apt.go.kr/> [accessed: May 1, 2023]
- [39] Deep Learning Bible, <https://wikidocs.net/book/7100> [accessed: May 1, 2023]
- [40] H. J. Chun and H. S. Yang, "A Study on Prediction of Housing Price Using Deep Learning", Journal of the Residential Environment Institute of Korea, Vol. 17, No. 2, pp. 37-49, Jul. 2019. <https://doi.org/10.22313/reik.2019.17.2.37>.
- [41] H. Kim, G. Joo, and H. Im, "Product Category Classification using Word Embedding and GRUs", Journal of Korean Institute of Information Technology, Vol. 19, No. 4, pp. 11-18, Apr. 2021. <https://doi.org/10.14801/jkiit.2021.19.4.11>.
- [42] D. Chicco, M. J. Warrens, and G. Jurman, "The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation", PeerJ. Computer Science, Vol. 7, pp. 1-24, Jul. 2021. <https://doi.org/10.7717/peerj-cs.623>.
- [43] C. D. Lewis, "Industrial and business forecasting methods: a practical guide to exponential smoothing and curve fitting", Butterworth Scientific, pp. 40, 1982.

## 저자소개

### 김 규 석 (Kyuseok Kim)



2011년 2월 : 한국항공대학교

정보통신공학(공학사)

2019년 8월 : 아주대학교

정보통신공학(공학석사)

2019년 6월 : LG전자(주)

선임연구원

2020년 2월 : (주)LG유플러스 책임

2019년 9월 ~ 현재 : 서울대학교 환경계획학과 박사수료

2020년 ~ 현재 : 한국폴리텍대학 분당융합기술교육원

데이터융합SW과 조교수

관심분야 : 데이터 수집, 데이터 분석, 머신러닝, 딥러닝,

도시계획, 근거리 무선통신, 상황인지

### 김 경 민 (Kyuseok Kim)



1995년 2월 : 서울대학교 지리학  
(학사)

2002년 5월 : UC Berkeley School  
of Information 정보 시스템 전공  
(석사)

2008년 6월 : Harvard University  
도시계획 · 부동산 전공(박사)

2002년 5월 : Oracle, US 소프트웨어 엔지니어

2001년 1월 : Hyosung Data Systems 소프트웨어  
엔지니어

2009년 ~ 현재 : 서울대학교 환경대학원 환경계획학과  
교수

관심분야 : 도시계획, 부동산 시장 분석, 인공지능,  
머신러닝

### 조 재 우 (Jaewoo Cho)



2005년 2월 : 서울대학교 조경학  
(학사)

2009년 2월 : 서울대학교

환경대학원 환경계획학(석사)

2018년 3월 : UC Irvine School of  
Social Science 도시계획학(박사)

2018년 4월 ~ 2019년 2월 : KAIST

4차산업혁명정책연구센터 Post-doc

2019년 ~ 현재 : 한성대학교 사회과학부 조교수

관심분야 : 스마트도시, 블록체인, 딥러닝, 도시 에너지