

대중의 부동산 관심에 영향을 미치는 정책 및 환경 예측을 위한 머신러닝 모형 제안

서예진*, 김규석**

Development of an Artificial Intelligence Model for Predicting the Policy and the Environment Affecting the Public Interest on the Real Estate

Yejin Seo*, Kyuseok Kim**

요 약

2017년 6월부터 2021년 2월까지 약 44개월 동안 25번의 부동산 정책이 수립, 시행되었다. 그러나 부동산 정책이 수립, 시행된 44개월 중 아파트 매매 가격의 하락은 2019년 1월부터 4월까지 0.7%, 2020년 5월 0.2%에 그쳤다. 반면에 전체기간인 약 4년 동안 아파트 매매 가격은 약 78%가 상승하였다. 본 논문에서는 부동산에 대한 대중의 관심에 영향을 미치는 정책과 환경을 예측하기 위한 회귀분석 기반의 머신러닝 모형을 제안한다. 연구 데이터는 Naver에서의 부동산 키워드의 검색량, 국토교통부의 부동산 대책, 통계청과 KOSIS의 인구통계학적 정보 등이다. 연구 결과, 회귀 분석 모형의 설명력인 Adj. R²는 0.749이었으며 머신러닝 모형의 MAPE는 약 2.2976%였다. 본 방법론을 이용하여 부동산 심리지수를 예측하는데 도움이 될 것으로 기대된다.

Abstract

25 of the real estate policies were established and implemented for about 44 months from June 2017 to February 2021. However, among the 44 months of real estate policies being established, the apartment sales prices fell only for 5 months, 0.7% from January to April 2019, and 0.2% in May 2020. On the other hand, the apartment sales prices increased by about 78% over the entire period of about 4 years. Public interest in policies can influence the policy formulation and implementation. In this paper, a machine learning model based on regression analysis for predicting the policy and the environment affecting public interest on the real estate is proposed. The research data includes the real estate related keyword search volumes on Naver, the policies for the real estate from the Ministry of Land, Infrastructure and Transport, demographic indicators from Statistics Korea and KOSIS, and so on. As a result of the study, the explanatory power of the regression analysis model, Adj. R² was 0.749 and the MAPE (Mean Absolute Percentage Error) of the machine learning model was about 2.2976%. With the methodology, it is expected to be helpful to predict the real estate sentiment index.

Keywords

real estate policy, artificial intelligence, machine learning, regression analysis

* 한국폴리텍대학 구미캠퍼스 전기과 (교신저자) · Received: Sep. 03, 2021, Revised: Nov. 18, 2021, Accepted: Nov. 21, 2021
- ORCID: <https://orcid.org/0000-0003-4823-5047> · Corresponding Author: Yejin Seo
** 한국폴리텍대학 분당융합기술교육원 데이터융합SW과/ Dept. of electricity, Korea Polytechnics, 84, Suchul-daero 3-gil, Gumi-si, Gyeongsangbuk-do, Republic of Korea, 서울대학교 환경대학원 환경계획학과 Tel.: +82-54-468-5244, Email: syj1122@kopo.ac.kr
- ORCID: <https://orcid.org/0000-0001-6613-5125>

1. 서 론

2021년을 1월부터 6월까지 수도권 아파트 매매 가격의 상승률은 12.97%로 19년 내 최고를 기록하였다[1]. 2021년 1월부터 6월까지 전국 아파트 매매 가격의 상승률은 9.97%로 지난 한 해 동안의 상승률인 9.65%를 이미 넘겼다[1]. 또한, 2021년 5월 기준, 최근 4년 동안의 서울, 수도권, 전국 아파트 매매가의 상승률은 각 46.44%, 35.05%, 21.52%에 이르는 것으로 파악되었다[1].

2017년 6월부터 2021년 2월까지 약 44개월 동안, 총 25번의 부동산 대책이 수립, 시행되었다. 이 대책들의 목적은 부동산 가격 상승을 억제하고 하락을 유도하여 많은 실수요자들이 혜택을 받을 수 있도록 하는 것이다[2]. 이를 위하여 부동산 갭 투자를 통한 부동산 투기 수요를 억제하고 다주택자들에게 추가적인 세금을 부과하는 등의 대책을 통해 부동산 시장의 억제를 유도하였다. 그러나 2019년 1월부터 4월까지 4개월 간 약 0.7%, 2020년 5월 1개월 간 약 0.2%의 부동산 매매가 하락이 있었던 시기를 제외한 기간은 모두 더 큰 폭으로 상승하였다. 또한, 이렇게 규제의 범위를 넓히고 강도를 높였음에도 불구하고 버티기, 풍선효과 등 역효과를 양산하는 결과를 낳는 등의 부작용도 발생하였다.

본 연구에서는 2017년 6월부터 2021년 2월까지 약 44개월 간 수립되어 시행된 25번의 부동산 대책에 대하여 대중의 부동산 관심에 영향을 미치는 요인에 대하여 분석을 하였다. 부동산에 대한 대중의 관심이 인터넷 검색 활동과 연계된다는 점을 반영하여 부동산에 대한 관심의 척도로는 종속변수로서 국내 최대 포털사이트인 Naver 검색 엔진에서의 “부동산” 키워드의 검색량으로 하였다[3][4]. 이 종속변수와 인과성을 판단하기 위하여 부동산 정책의 종류, 기준금리, 고용률, 부동산 대책의 반응도, 부동산 심리지수 등의 독립변수를 활용하여 다중 선형 회귀분석을 1차적으로 수행하였다. 이렇게 다중 선형 회귀분석을 통해 설명력이 확보된 모형을 머신러닝을 통해 예측 오차율을 측정함으로써 대중의 부동산 관심에 영향을 미치는 정책 및 환경 예측을 위한 머신러닝 모형을 제안하였다.

본 논문의 구성은 다음과 같다. 2장에서는 부동

산 지수와 관심도 등을 분석 또는 예측하는 기준 문헌을 고찰하며 3장에서는 본 연구에서 활용하는 데이터와 방법론을 설정한다. 4장에서는 연구 결과에 대하여 분석한다. 마지막으로 5장에서는 결론과 함께 향후 연구 방향에 대하여 제안한다.

II. 문헌 고찰 및 이론적 배경

2.1 부동산 지수 예측을 위한 인공지능 모형 연구

이주미 외(2021)는 아파트 매매지수, 지가지수, 전세가격지수, 부동산 심리지수 등의 부동산 지수를 종속변수로 하여 머신러닝을 활용한 예측 모델 비교 분석을 하였다[3]. 머신러닝 모형으로는 RF(Random Forest), XGBoost(eXtreme Gradient Boosting), LSTM(Long Short-Term Memory)을 활용하였다. 여러 개의 표본을 생성하고 의사결정 나무 모델(Decision tree model)을 적용하여 결과를 종합하는 앙상블 알고리즘의 일종인 RF 기반의 모형은 아파트 매매지수를 예측했을 때 가장 높은 정확도를 보였다. 그리고 RF와 같이 다수의 결정 트리를 결합시킨 앙상블 알고리즘 중 하나인 XGBoost를 활용했을 때에는 아파트 매매지수, 장기 패턴을 학습할 때 발생하는 기울기 소멸 문제를 극복하는 RNN(Recurrent Neural Network)의 일종인 LSTM을 활용했을 때에는 전세 가격지수와 지가지수의 예측 정확도가 가장 높았다.

남상현 외(2020)는 부동산 가격 예측을 위한 머신러닝 모형을 제안하였다[5]. 2012년 1월부터 2019년 9월까지의 지역별 부동산 매매 실거래가를 회귀 분석 모형의 머신러닝에 학습시켜 수행하였다. 그러나 해당 머신러닝 모형의 정확도 평가를 한 결과는 없었기에 예측의 정확도 판단이 되지 않는 한계점이 존재하였다.

2.2 언론 및 대중의 부동산 관심도 연구

카카오(2020)는 연령대별 부동산 관심도의 순위를 분석하여 40대, 30대, 50대, 60대, 20대 순으로 정리하였다[6]. 채영길 외(2018)는 정부가 발표한 2017년 8월 2일 부동산 대책에 대한 언론 보도 경

향을 비교 분석하였다[7]. 김대원 외(2014)는 인터넷 검색 활동과 주택 가격 및 거래량 간의 유의미한 관계성을 파악하는 연구를 진행하였다[4]. Naver에서 제공하는 인터넷 검색활동의 자료와 아파트 매매가 지수, 아파트 매매 거래량 지수 간의 동적 관계를 분석하였다. 분석결과, 인터넷 검색 활동은 잠재 주택 수요자들의 구매 수요를 반영한 결과로 판단되었고, 이 활동이 주택 가격 및 거래량에 유의미한 영향력을 미치고 있음을 알 수 있었다.

2.3 본 연구의 차별점

본 연구는 다음과 같은 차별점이 존재한다.

첫째, 대중의 부동산 관심도에 영향을 주는 독립변수로 부동산 관련 정책의 성격을 분류하였다. 지난 2017년 6월부터 2021년 2월까지 25회의 부동산 정책에 따른 대중의 부동산 관심도를 분석하는 것이다.

둘째, 기준 금리, 취업률 등 거시 경제 변수를 독립변수로 활용하였다. 이 모형을 다중 선형 회귀분석을 통해 설명력을 입증한 후, 딥러닝을 수행하여 데이터를 학습, 예측한다.

셋째, 부동산 관심도의 척도로 국내 최대 포털 사이트인 Naver에서의 “부동산” 키워드의 검색량을 종속변수로 한 연구이다. 이는 수치적 근거를 기반으로 관련 정책 수립에 도움이 될 수 있을 것이다.

따라서 대중의 부동산 관심의 척도를 예측할 수 있는 실증적 연구로서 관련 정책 수립 시 참고할 수 있는 지표가 될 수 있을 것으로 기대한다.

III. 제안하는 연구 모형

본 연구에서는 다중 선형 회귀 분석을 통해 설명력을 갖춘 모형을 회귀 분석 기반의 딥러닝 모형에 대입하여 그 결과를 예측한다.

3.1 연구 데이터

2017년 6월 19일 첫 부동산 정책이 발표된 후로부터 2021년 2월 4일까지 25번의 부동산 대책이 발표되었다. 이는 1,326일간 약 55.25일 간격으로 한번씩 부동산 정책이 발표, 시행된 것이다.

이와 관련하여 본 연구에서 활용하는 데이터의 공간적 범위와 시간적 범위는 표 1과 같다. 시간적 범위의 시작일은 첫 부동산 정책 발표일보다 41일 전이면서 대통령 선거일인 2017년 5월 9일로 하였다. 25번째 부동산 정책 발표일인 2021년 2월 4일부터 81일 후이자 26번째 부동산 정책 발표일인 2021년 7월 28일의 중간에 해당하면서 해당 주 월요일인 2021년 4월 26일로 하였다. 또한, 공간적 범위는 전국으로 설정하였다.

본 연구에서 변수로 활용하는 데이터와 출처는 표 2와 같다. 부동산 검색 키워드량은 Naver 데이터랩에서 추출하였다[8]. 그리고 기준 금리 및 고용률은 나라지표 사이트에서 추출하였으며 부동산 정책과 관련된 내용은 국토교통부, 부동산 소비심리지수는 국가통계포털에서 추출하였다[9]-[11].

본 연구에서 설정한 변수는 표 3과 같다.

표 1. 연구 데이터의 시공간적 범위

Table 1. Spatial and temporal range for research data

Spatial range	Temporal range
National wide	2017. 05. 09. ~ 2021. 04. 26

표 2. 연구 데이터와 출처

Table 2. Research data and source

Data	Source
Search keyword volume of “Real Estate”	datalab.naver.com
Base rate	index.go.kr
Employment rate	index.go.kr
Policy on the real estate	rt.molit.go.kr
Real estate sentiment index	kosis.or.kr

표 3. 변수

Table 3. Variable

Category	Variable
Dependent variable	Search Keyword Volume of “Real Estate”(SKV)
	Employment Rate(ER)
Independent variable	Base Rate(BR)
	Reaction to Real estate from survey result (RTR)
	Real Estate Sentiment index(RES)
	Reinforcement of Law(RL, Dummy variable)
	Tax Increase(TI, Dummy variable)
	Loan Reduction(LR, Dummy variable)
	Supply Expansion(SE, Dummy variable)

종속변수로는 Naver에서의 “부동산” 키워드의 검색량(SKV)이다. 독립변수로는 고용률(ER), 기준금리(BR), 부동산 소비심리지수(RES)이다. 그리고 더미 변수로는 부동산 정책의 종류를 구분하는 법령강화(RL), 세금상승(TI), 대출축소(LR), 공급증대(SE)로 설정하였다.

독립변수들 중 설문조사에서 부동산에 대한 반응도(RTR)는 긍정의 비율에서 부정의 비율을 뺀 값으로 식 (1)과 같이 계산되어 진다. 예를 들어, 설문조사 결과, 부동산에 대한 긍정의 비율이 5%이고 부정의 비율이 7%라면 RTR은 2가 된다.

$$RTR_i = PPR_i - PNR_i \tag{1}$$

PPR : Proportion of Positive Reaction

PNR : Proportion of Negative Reaction

3.2 연구 방법

본 논문에서의 연구 순서도는 그림 1과 같다. 모델을 설정하고 그에 맞는 변수들의 데이터를 모두 수집한다. 데이터를 수집 후 다중 회귀분석을 통해 설명력을 갖춘 결과가 나왔다면 이를 머신러닝 모형에 학습시켜 결과를 얻는다. 만약, 설명력을 갖추지 못했다면 모형을 다시 설계하여 회귀분석을 진행한다. 그리고 머신러닝을 통해 얻은 결과들의 정확도를 MAPE를 활용하여 계산한다.

본 연구에서는 식 (2)를 기반으로 다중 선형 회귀분석을 통해 변수들 간의 인과성을 파악한다. 설명력이 확보된 모형을 머신러닝에 대입하여 학습시킨 후 해당 모형의 예측 정확도를 통해 모형을 검증한다. 회귀분석과 회귀분석 기반의 인공지능 모형의 예측 정확도를 판단하는 방법으로는 MAPE (Mean Absolute Percentage Error)가 있다[12]-[14]. MAPE는 정확도의 오차를 백분율로 표시하여 이는 식 (2)와 같이 계산되어 진다.

$$SKV_i = \beta_0 + \beta_1 ER_i + \beta_2 BR_i + \beta_3 RTR_i + \beta_4 RES_i + \beta_5 RL_i + \beta_6 TI_i + \beta_7 LR_i + \beta_8 SE_i + \epsilon \tag{2}$$

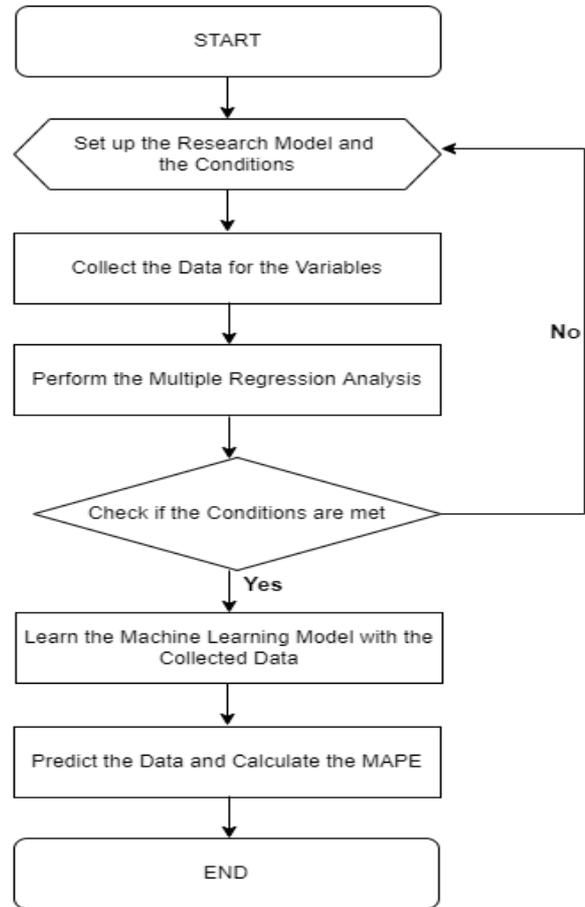


그림 1. 연구 순서도
Fig. 1. Flow chart for research

위 모형을 파이썬(Python) 기반의 텐서플로우(Tensorflow) 모듈에 대입한다. 배치 사이즈(Batch size)는 1이며 에포크(Epoch)는 1부터 10씩 늘려서 해당 모형의 예측 오차율인 MAPE를 평가한다. MAPE는 식 (3)과 같이 계산 되어지며 그 값이 작을수록 평균 예측 오차가 작다는 것으로 그만큼 예측 정확도가 높다는 것을 의미한다.

$$MAPE[\%] = \frac{100}{n} \sum \frac{|Prediction Value - Actual Value|}{Actual Value} \tag{3}$$

IV. 연구 결과

4.1 기술 통계 분석

본 연구에서 활용하는 데이터의 소수점 둘째자리

까지 표현한 기술통계량은 표 4와 같다. “부동산” 키워드의 검색량인 SKV 값은 해당 키워드 검색량이 최대치인 날을 100으로 하여 일별 상대 값을 나타내는 것이다. 그러나 한 번에 1년 치만을 추출할 수 있어 중첩되는 날짜를 기준으로 총 4개년 치를 상대 값으로 계산하였다. 수행 결과, SKV 값의 최대값은 72.45였으며, 최소값은 26.28, 평균값은 42.70이었다. 고용률인 ER 값의 평균은 66.41이었으며, 최대값은 67.4, 최소값은 64.3이었다. 기준금리인 BR 값의 평균은 1.12, 최대값은 1.75, 최소값은 0.5였다. 국토연구원에서 제공하는 부동산 소비심리지수인 RES는 0 ~ 200의 값으로 표현되며 총 9단계, 하강, 보합, 상승의 3개의 구간으로 나뉜다. 평균값인 108.86은 6단계인 105 ~ 114 구간의 값으로 보합의 최고 단계이다. 최대값인 131.7은 7단계이자 상승의 첫 단계이며, 최소값인 87.7은 3단계이자 하락의 마지막 단계이다.

또한, 부동산 정책의 종류를 구분하기 위하여 독립변수로 활용하는 RL, TI, LR, SE 변수는 각 정책별 중복 할당될 수 있다. RL의 평균값이 0.45로 가장 높았고, LR의 평균값인 0.11로 가장 낮아 법령강화 정책이 상대적으로 많았고, 대출축소 정책이 상대적으로 적었음을 알 수 있다. 이는 정부가 대출범위를 축소하여 은행권까지 그 영향이 미치는 것은 피하는 것으로 풀이된다.

표 4. 기술통계량 (n = 207)

Table 4. Descriptive statistics (n = 207)

Variable	Mean	SD	Min.	Max.
SKV	42.70	9.70	26.28	72.45
ER	66.41	0.69	64.3	67.4
BR	1.12	0.46	0.5	1.75
RES	108.86	11.02	87.7	131.7
RL	0.45	0.50	0	1
TI	0.22	0.41	0	1
LR	0.11	0.41	0	1
SE	0.25	0.43	0	1

4.2 다중 선형 회귀분석 결과

식 (1)의 모형을 기반으로 다중 선형 회귀분석을 수행한 결과는 표 5와 같다.

표 5. 회귀 분석 결과 (n = 207)

Table 5. Result of regression analysis (n = 207)

Variable	Coef.	t-test	VIF
Constant	264.3913***	5.894	-
ER	-4.6139***	-6.771	1.939741
BR	3.8135*	2.558	4.185282
RES	0.7003***	14.248	2.586781
RL	4.1365***	3.580	2.910216
TI	4.3956***	3.348	2.583577
LR	0.0742	0.061	1.223483
SE	5.8132***	4.028	3.406837
R ² : 0.757 / Adj. R ² : 0.749 F-statistics(P-value) : 88.79(9.49 X 10 ⁻⁵⁸)			

본 모형의 설명력인 Adj. R²의 값은 0.749로서 약 74.9%의 설명력을 갖는다고 할 수 있다. F-statistics의 P-value는 9.49 X 10⁻⁵⁸로 모형 변수들이 전체적으로 유의함을 알 수 있었다. 또한, 모든 독립변수들의 VIF 값은 5 미만으로 한 개의 변수에 의존성이 높은 다중공선성이 없음을 알 수 있었다.

세부적으로 분석하면, 종속변수인 “부동산” 키워드 검색량에 영향을 미치는 유의한 독립변수는 ER, BR, RES, RL, TI, SE로서 LE 더미변수를 제외한 전체였다. 또한, 독립변수 ER, BR, RL, TI, SE의 계수의 절대값은 약 3에서 5사이로 1 미만인 RES 변수의 계수에 비해 그 영향력이 큼을 알 수 있었다.

다중 회귀분석 결과를 해석하면 다음과 같다. 고용률(ER)이 1만큼 올라갈수록 “부동산” 키워드 검색량은 약 -4.6139만큼 감소하고 기준금리(BR)나 부동산 소비심리지수(RES)가 높아지거나 최근에 시행된 부동산 정책이 법령강화(RL), 세금상승(TI), 공급증대(SE)의 정책이라면 검색량 각 0.7003에서 5.8132 만큼 높아지는 것으로 나타났다.

4.3 머신러닝 수행 결과

0.749의 설명력을 갖춘 식 (1)의 모형을 파이썬의 텐서플로우 모듈에 대입하여 수행한 결과는 표 6과 같았다. 수행 결과, 배치 사이즈가 1로 동일할 때, 에포크가 커질수록 MAPE 값이 작아지는 것을 알 수 있었다. 배치사이즈를 1로하고 에포크를 1로 했을 때의 MAPE는 56.14038059072763이었으나 에포크를 10으로 했을 때의 MAPE는 5.219994519198487로 감소하였다.

표 6. 인공지능 수행 결과

Table 6. Results from artificial intelligence model

No.	Epoch	Batch size	MAPE(%)
1	1	1	56.14038059072763
2	10	1	5.219994519198487
3	100	1	3.15674510280238
4	1,000	1	2.8932316184783764
5	10,000	1	2.59358538013991
6	100,000	1	2.5607275289080775
7	1,000,000	1	2.37420687409587
8	10,000,000	1	2.2976134538341537

이는 에포크가 1이었을 때의 예측 정확도의 오차는 약 56.14%였으나 에포크가 10이었을 때의 예측 정확도가 약 5.22%로 줄어든다는 것을 의미한다. 또한, 에포크 사이즈를 10,000,000으로 늘렸을 때의 MAPE는 2.2976134538341537로 줄어드는 것을 알 수 있었다. 이는 인공지능 학습의 횟수가 늘어날수록 예측의 정확도가 높아지는 것을 의미한다.

4.4 소결

본 연구에서는 머신러닝을 활용한 대중의 부동산 관심도를 예측할 수 있는 모형을 제안하였다. 해당 머신러닝 모형을 정의하기에 앞서 부동산 정책을 법령강화, 세금상승, 대출축소, 공급증대 등 4가지(중복 선택)로 세분화 하였다. 이렇게 세분화된 부동산 정책과 각종 지표들을 변수로 하여 다중 선형 회귀분석으로 검증을 하였다.

분석 결과, 본 연구에서 제안하는 모형의 선형 회귀분석 설명력은 0.749였으며, 머신러닝을 통해 10,000,000회 학습 후 예측한 결과와 실제의 차이는 약 2.2976%임을 알 수 있었다.

회귀분석의 설명력을 의미하는 R2는 종속변수와 독립변수에 의해 설명되는 정도를 나타내는 수치이다. 학문에 따라 그 기준은 다르지만 일반적으로 R2 값은 0.6 이상이면 높다고 판단하며, 0.4 이상이어도 설명력이 있다고 판단한다. 본 논문의 연구 모형에서의 R2는 0.749로 약 74.9%가 설명이 된다는 의미이다.

V. 결 론

우리나라의 2017년 6월부터 2021년 상반기까지 25회의 부동산 정책은 부동산 가격을 안정시키기보다 상승세를 꾸준히 이어가는데 일조하는 경향을 보였다. 이에 따라 부동산 실수요자들을 위한 집값 상승을 억제하고 하락을 유도하려 했던 정책들이 버티기와 풍선효과와 같은 역효과의 범위와 강도를 높이고 있었다.

본 연구에서는 지난 25번의 부동산 대책에 대하여 대중의 부동산 관심에 영향을 미치는 요인에 대한 분석을 하였다. 부동산에 대한 관심의 척도로는 국내 최대 포털 사이트인 Naver에서의 30 ~ 50대의 “부동산” 키워드의 검색량으로 하였다. 이와 관련된 독립 변수를 설정하여 다중 선형 회귀분석을 통해 1차적 검증을 하고 딥러닝을 통해 예측 정확도가 높은 인공지능 모형을 제안하였다.

본 연구에서 제안하는 모형에 대한 수치적 검증을 위하여 다중 선형 회귀분석을 수행하였고 0.749의 설명력을 보임을 알 수 있었다. 이를 기반으로 파이썬 텐서플로우를 통해 10,000,000회의 학습을 한 결과 MAPE가 약 2.2976%를 보였다.

추후 연구에서는 본 논문에서 제안하는 모형에 코로나19 등의 환경적인 독립변수를 추가하고 다중 선형 회귀분석의 설명력을 더 높일 수 있을 것이다. 그 후 텐서플로우를 통해 부동산 시장에 대한 대중의 관심도, 심리 지수 등을 예측할 수 있을 것이다. 이렇게 구성된 인공지능 모형을 통해 부동산 정책의 종류와 시기에 따른 그 효과를 예측하는데 활용될 수 있을 것으로 기대한다.

References

- [1] KBLAND, <https://kbland.kr>. [accessed: Aug. 26, 2021]
- [2] M. J. Jang, J. J. Kim, D. W. Yang, and W. J. Shin, "Comparing the Effect of External Demands on a Local Housing Market before and after a Real-Estate Policy Change", *Journal of Korea Planning Association*, Vol. 55, No. 6, pp. 90-98, Nov.

2020. <https://doi.org/10.17208/jkpa.2020.11.55.6.90>.

[3] 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, Jan. 2021. <http://dx.doi.org/10.5659/JAIK.2021.37.1.191>.

[4] D. W. Kim and J. S. Yu, "A Dynamic Relationship Between Internet Search Activity, Housing Price, and Trading Volume", Korea Real Estate Review, Vol. 24, No. 2, pp. 125-140, Jun. 2014.

[5] S. H. Nam, T. H. Han, L. J. Kim, and E. J. Lee, "Real-Estate Price Prediction in South Korea via Machine Learning Modeling", The Journal of the Institute of Internet, Broadcasting and Communication, Vol. 20, No. 6, pp. 15-20, Dec. 2020. <http://dx.doi.org/10.7236/IIBC.2020.20.6.15>.

[6] The Korea Economic Daily, <https://www.hankyung.com/it/article/202007291177g>. [accessed: May 16, 2021]

[7] Y. G. Chae and S. Y. Jang, "An analysis of Newspaper reports on Government Real Estate Reform Policy in Korea", Journal of the Korea Contents Association, Vol. 18, No. 8, pp. 446-458, Aug. 2018. <http://dx.doi.org/10.5392/JKCA.2018.18.08.446>.

[8] NAVER DataLab. <https://datalab.naver.com>. [accessed: May 16, 2021]

[9] e-Country Index. <https://www.index.go.kr>. [accessed: May 16, 2021]

[10] Ministry of Land, Infrastructure and Transport. <http://rt.molit.go.kr>. [accessed: May 16, 2021]

[11] National Statistical Portal. <https://kosis.kr>. [accessed: May 16, 2021]

[12] Y. S. Hwang, "A Hybrid Forecasting Framework based on Case-Based Reasoning and Artificial Neural Network", Journal of Intelligence and Information Systems, Vol. 18, No. 4, pp. 43-57, Dec. 2012. <https://doi.org/10.13088/jiis.2012.18.4.043>.

[13] A. de Myttenaere, B. Golden, B. Le Grand, and F. E. Rossi, "Mean Absolute Percentage Error for

Regression models", Neurocomputing (Amsterdam), Vol. 192, pp. 38-48, Jun. 2016. <https://doi.org/10.1016/j.neucom.2015.12.114>.

[14] D. Chicco, M. Warrents, and G. Jurman, "The coefficient of determination R-squared in more informative than SMAPE, MAE, MAPE and RMSE in regression analysis evaluation", PeerJ. Computer science, Vol. 7, pp. 1-24, Jul. 2021. <https://doi.org/10.7717/peerj-cs.623>.

저자소개

서 예 진 (Yejin Seo)



2002년 2월 : 울산대학교
전기전자자동화공학부(공학사)
2004년 2월 : 울산대학교
전기전자정보시스템공학부
(공학석사)
2015년 2월 : 울산대학교
전기전자컴퓨터공학과(공학박사)

2020년 2월 ~ 현재 : 한국폴리텍대학 구미캠퍼스 전기과
조교수

관심분야 : 시스템제어, 신호처리, 머신러닝, 딥러닝

김 규 석 (Kyuseok Kim)



2011년 2월 : 한국항공대학교
정보통신공학(공학사)
2019년 6월 : LG전자(주)
선임연구원
2019년 8월 : 아주대학교
정보통신공학(공학석사)
2020년 2월 : (주)LG유플러스 책임

2019년 9월 ~ 현재 : 서울대학교 환경계획학과
(박사수료)

2020년 2월 ~ 현재 : 한국폴리텍대학 분당융합기술교육원
데이터융합SW과 조교수

관심분야 : ITS, UTM, UAV, Context-awareness, 데이터
분석