Deep Learning Data Visualization Using Architectural Data

Ji Mok Lee

PAUL MATH SCHOOL

12-11, Dowontongmi-gil, Cheongcheon-myeon, Goesan-gun,

Chungcheongbuk-do, Republic of Korea

(Supervised by teacher **Shin Dong Ho**)

eavatar@hanmail.net

Abstract

Artificial intelligence is human learning ability, reasoning ability, perceptual ability, and other artificially implemented computer programs or computer systems including the same. Recently, as interest in artificial intelligence increases, research related to artificial intelligence is increasing. In the field of architecture, research related to the development of architectural technology using artificial intelligence is increasing too. This Research is understanding deep learning by visualizing deep learning data using architectural data, and through this research, a program was created to visualize the results by using real estate data among architectural data and predicting apartment sales prices. In order to predict the apartment sale price, independent variables were derived using reference to previous studies, web crawling technique, and variable selection method. The deep learning learning was conducted using linear regression analysis using the apartment sales price using the residential area, and the predicted value was extracted using the learned deep learning model. Using a deep learning learning model, the predicted data can be checked by inputting an arbitrary residential area value, and the resulting data is visualized on a map.

keywords

Artificial Intelligence, Machine learning, AI, Architectural Data,

I. Introduction

1. Motivation

Artificial intelligence is human learning ability, reasoning ability, perceptual ability, and other artificially implemented computer programs or computer systems including the same. Recently, as interest in artificial

intelligence increases, research related to artificial intelligence is increasing. In the field of architecture, research related to developing architectural technology using artificial intelligence is increasing too. This research is to understand deep learning by visualizing deep learning data using architectural data.

2. Method

This research creates a program that visualizes the result value by predicting the apartment sale price, which is architectural data, using deep learning. Construction data is collected using public data and web crawling techniques. Deep learning learning proceeds by using simple regression analysis, multiple regression analysis, and artificial neural network-based regression analysis.

II. Body

1. Selection of variables

1.1. Selection of variables through prior research

The results of arranging the independent variables, dependent variables, and significant independent variables used in the research on the factors affecting the apartment sales price and the preceding studies on the estimation of apartment sales prices are as follows.

In the research of Jang and Kim (2020) <A research on the Analysis of Factors Affecting Apartment Price Using Spatial Metering Model-Focused on Suseong-gu, Daegu>, the dependent variable is the apartment complex, the average sale price, and the independent variable is the total number of households, elapsed years, heating method , The highest number of floors, floor area ratio, distance to subway station, elementary, middle and high school distance, apartment builder construction ability evaluation ranking, access road width, school district variables. The significant independent variables were apartment characteristics, subway station distance, school district.

In the research of Byun, Ko, and Choi (2019) <Analysis of determinants of apartment complex price and jeonse price in non-metropolitan areas-focusing on cases in Gwangju, Daejeon, Daegu, and Busan>, the dependent variable is the Apartment charter price, constant term, average acreage, total number , Construction company, construction period, construction time square, educational characteristics, markets, subway entrances, libraries, department stores, parks, and elementary education institutions. The significant independent variables were complex characteristics, educational characteristics, spatial characteristics.

In the research of Jang and Shin (2019) <Analyzing the Impact of External Demand on Housing Prices-Focused on

Apartments in Gwangju>, the dependent variable is the housing price, and the dependent variable is the apartment size, elapsed years, number of transactions, external demand, and internal demand. , Kindergarten, Elementary School, Middle School, and High School. The significant independent variables were apartment characteristics, internal and external demand, spatial characteristics.

In the research of Ko and Kang (2019) <A Comparative research of Real Estate Price Factors and Real Estate Price Increase Rate Factors: Focusing on Apartments in Seoul>, the dependent variables are apartment current price, apartment price, and annual average increase rate, and the dependent variables are floor area ratio, number of households, and subway stations, Number of convenience facilities, proximity to elementary school, proximity to general hospital, Gangnam 4 district, past market price, and significant independent variables were housing complex characteristics, location characteristics, economic characteristics.

1.2. Variable selection method

In this research, to proceed with the variable selection method. The variable selection method selects the variable with the best predictive power from all the given variables, and aims to optimize the number of dimensions and combinations of variables to be used in the model by selecting a meaningful variable or eliminating unnecessary variables. Methods of variable selection include forward selection, backward elimination, and stepwise selection. In this research, a regression analysis is conducted by selecting the optimal variable among the collected independent variables using a stepwise selection method.

1.3. Web crawling

Web crawling is a process in which a web crawler automatically collects users and desired data such as text and images existing on the web according to a set rule. In this research, high-quality data that could not be collected through public data was quickly and easily built through web crawling.

2. Multiple regression analysis based on artificial neural network

2.1. Regression analysis

Regression analysis measures the influence and magnitude of the independent variable on the dependent variable, and is used to predict the value of the dependent variable corresponding to the constant value of the new independent variable by finding the regression plane or regression line for the independent variables. It's the way. In this research, simple regression analysis and multiple regression analysis are used using independent variables collected to predict apartment sales prices.

2.2. Artificial neural network

Artificial neural networks are statistical learning algorithms that mimic human neuronal cells. In this research, the sigmoid function was used among the activation functions, and the number of artificial neural network training was set to 200, and the analysis items used the data of the items that were significant in the multiple regression analysis, The reliability was verified through the comparison of. Also, by checking the loss function, a comparative analysis was performed for accurate prediction.

3. Estimation of real estate prices

3.1. Selection of independent variables

Through prior studies, structural characteristics, complex characteristics, location characteristics, and environmental characteristics were confirmed as significant independent variables. In addition, as a result of analyzing the statistics of factors to consider when choosing a residence in Seoul, the characteristics of traffic factors were added to the independent variable, and data was collected using web crawling. In addition, the final independent variable data was selected using a stepwise selection method. The results of the selected independent variable are as follows.

Table 1. The result of the independent variable

Structural characteristics	residential area, exclusive area, gross floor area
Complex characteristics	number of floors in the building, contract year, month
Location characteristics	Distance to subway station, number of bus stops
Environmental characteristics	number of surrounding cultural facilities, surrounding sports facilities
Traffic factor characteristics	distance to work, transportation cost, time it takes

3.2. Estimation of real estate prices

To estimate real estate prices, to predict numerical data, linear regression analysis and multiple regression analysis based on artificial neural networks are used to predict numerical data. The apartment sale price is selected as the dependent variable, and the structure, complex, location, environment, and traffic factor characteristics of the apartment are selected as independent variables to perform regression analysis. In this research, a total of four regression analyzes were performed to predict apartment sales prices.

First, the apartment sales price is predicted using multiple linear regression analysis. As a result of the prediction, the value of the loss function and the accuracy of the data are checked, and five variables with high correlation among the independent variables are extracted. Next, the artificial neural network-based multiple regression analysis is performed twice by setting the hidden layers to one or two, confirming the loss function value as the prediction result data, and comparing and analyzing the accuracy of the prediction result when compared with the multiple regression analysis.

Next, deep learning learning using linear regression analysis is performed using the apartment sale price using the residential area, and the predicted value is extracted using the learned deep learning model. Using a deep learning learning model, the predicted data can be checked by inputting an arbitrary residential area value, and the resulting data is visualized on a map as shown in the following figure 1.

Figure 1. visualized on a map

III. Results

Through this research, a program was created to visualize the result value by using real estate data among architectural data and predicting the apartment sale price. In order to predict the apartment sale price, independent variables were derived using reference to previous studies, web crawling technique, and variable selection method. The deep learning learning was conducted using linear regression analysis using the apartment sales price using the residential area, and the predicted value was extracted using the learned deep learning model. Using a deep learning learning model, the predicted data can be checked by inputting an arbitrary residential area value, and the resulting data is visualized on a map.

References

Jang Mong-Hyun and Kim Han-Soo. "A Study on Analysis of Factors Affecting Apartment Price Using Spatial Measurement Model-Centering on Suseong-gu, Daegu, *Korean Housing Association   symposium*, 31.1 (2020): 79-86.

Byeon Seil and Go Yeonghwa. "Analysis of determinants of apartment complex price and jeonse price in non-metropolitan area-Gwangju, Daejeon, Daegu, and Busan, *Korean Housing Association Conference, Presentation Proceedings*, 2019.1 (2019): 41-56.

Jang Mijin and Shin Woo-jin. "Analysis of the influence of external demand on house prices- Focused on Gwangju apartment, *Korean Association for Regional Development, Academic conference*, 2019.10 (2019): 477-485.

Go Juhyeong and Gang Myeonggu. "A comparative study of real estate price factor and real estate price increase rate factor : Centering on Seoul apartments, *Korean Association for Regional Development, Academic conference*, 2019.5 (2019): 365-373.