Intercontinental Geoinformation Days

http://igd.mersin.edu.tr/2020/

A House Valuation with Multiple Regression Analysis and Artificial Neural Networks

Mehmet Emin Tabar*[1], Aslan Cihat Başara [1], Yasemin Şişman [2]

[1]*Ondokuz Mayis University, Institute of Graduate Studies, Department of Geomatics Engineering, Samsun, Turkey*
[2]*Ondokuz Mayis University, Faculty of Engineering, Department of Geomatics Engineering, Samsun, Turkey*

Keywords
Real estate appraisal
Multiple regression analysis
Artificial neural network
Artificial intelligence
Matlab

ABSTRACT

Valuation, in its simplest form, is the determination of the amount that a property will be processed at a certain date. Valuation can be done for many purposes. These; can be listed as buying and selling, transfer, tax assessment, expropriation, inheritance distribution, investment, financing and credit. There are various methods of valuation. These methods are examined under 3 main groups as traditional, statistical and modern valuation methods. The aim of the article is to provide an overview of regression analysis, one of the statistical valuation methods, and artificial neural networks, one of the modern valuation methods, and to compare the accuracy values. Matlab software was used for artificial neural network modeling and Minitab software was used for regression analysis. The accuracies of the obtained values were determined by the average absolute percent error (MAPE) formula.

1. INTRODUCTION

Valuation of real estate and the reflection of these values to tax is one of the most important economic foundations of developed societies. With the development of real estate investment trusts, construction companies, education, technology and professionalism, the real estate market has gained a positive momentum. These developments have also brought professional real estate appraisals to the agenda in order to make the right investments (Atik, KÖSE, Yilmaz, & ERBAŞ, 2015).Real estate is independent and permanent rights registered on a separate page in the land register and independent parts of the property ownership register, which give the owner the right to use as she wishes, except for the restrictions developed for the benefit of the public (Açlar & Çağdaş, 2002). The concept of housing is the place where the consumer ultimately lives with his wife and children (Özdamar, 2004).

The concept of value is defined as the abstract measure used to determine the importance of something, the value of something, the value of something that can be measured with money, price (Yomralıoğlu, Nişancı, Çete, & Candaş, 2011). Real estate valuation is the provision for trading or corporate transactions that varies according to needs, wishes and financial capacity (Ring & Dasso, 1977).

According to another definition, it is the process of determining the provision of the seller according to the properties of the property for investment or long-term use (Brown, 1965). Real estate valuation is done in many different ways. However, for professional real estate valuation, a mathematical model should be mentioned rather than subjective value estimates. Statistical and modern methods involve less initiative as they depend on a mathematical model.

Traditional methods, on the other hand, are far from being objective compared to statistical and modern valuation methods since they do not depend on a mathematical model. Because in traditional valuation methods, valuation experts try to make value calculations by only estimating an exchange price. However, one of the most important issues in the valuation area is the need to ensure that the information provided to the customer is clear and clear. (Pagourtzi, Assimakopoulos, Hatzichristos, & French, 2003). When the methods used in the mathematical model are examined; Many methods such as fuzzy logic, artificial neural networks, spatial analysis, support vector machines, regression analysis are reached. Real estate valuation methods are shown in Table 1.

* **Corresponding Author**

*(mehmetemintabar@gmail.com) ORCID ID 0000 – 0002 – 3234 – 5340
(aslancihatbasara@gmail.com) ORCID ID 0000 – 0001 – 6644 – 6097
(ysisman@omu.edu.tr) ORCID ID 0000 – 0002 – 6600 – 0623

Cite this study

Tabar M E, Basara A C & Sisman Y (2020). A House Valuation with Multiple Regression Analysis and Artificial Neural Networks. Intercontinental Geoinformation Days (IGD), 96-99, Mersin, Turkey

Table 1. Real estate valuation methods

Traditional methods	Statistical methods	Modern methods
Comparison method	Nominal method	Fuzzy logic
Income method	Multiple regression method	Artificial neural networks
Cost method	Hedonic pricing method	Support vector machine

In this study, regression analysis, one of the statistical valuation methods, and artificial neural networks, one of the modern valuation methods, were used. Accuracy values obtained were calculated with the formula of average absolute percent error (MAPE).

$$MAPE = 100\frac{\sum_{i=1}^{n}\frac{|A_i - F_i|}{A_i}}{n}$$

The calculated accuracy values are compared and given in a table.

2. MATERIAL AND METHOD

Housing data were obtained from a residential sales site. While choosing the residences, care was taken to choose those with a facade to the tramway street. Housing values are valid from August to October 2020. Housing values were determined by modeling the obtained data with multiple regression analysis and artificial neural networks.

2.1. Multiple Regression Analysis

Regression analysis consists of dependent and independent variables. The number of dependent variables is single, but the number of arguments can be more than one. If there is only one independent variable, it is called simple linear regression, if there are two or more independent variables, it is called multiple linear regression. In regression analysis, it is aimed to explain the relationship between variables functionally and to define this relationship with a model (Chatterjee & Hadi, 2015).

It is based on the principle of examining how the other variable changes according to this level by keeping one of the variables or the category of the variable at predetermined levels. Regression is also interpreted as finding the unknown with the help of what is known in modern statistics (Akış, 2013). In the prediction, when the correlation between variables is zero, it is concluded that there is no relation about X, Y. However, if this correlation is not zero, the result will be less erroneous. If the correlation is ± 1.00, then the prediction probability is high and successful. When this correlation moves away from ± 1.00, the amount of error increases (Zeng & Zhou, 2001).

2.2. Artificial Neural Networks

Artificial neural networks emerged by artificially imitating the way the human brain works. It can be thought of as a complex system that occurs as a result of connecting many nerve cells in the brain with different levels of influence. In artificial neural networks, the system first performs the learning process by analyzing input data and output data (Öztürk & Şahin, 2018). It gives approximate outputs of new input data after learning process as a result of iterations. Artificial neural networks are especially used in engineering applications. Engineering problems that are difficult to solve with classical methods have gained a different dimension with artificial neural networks and have created an effective alternative (Yegnanarayana, 2009). Although the human brain is limited in mathematical operations such as division, multiplication, addition, subtraction, it is more successful than machines in many processes such as learning, remembering, and predicting. The main features of artificial neural networks are nonlinearity, learning, parallel working, generalization, working with missing data, using a large number of variables and parameters, applicability, fault tolerance and flexibility.

Artificial neural networks consist of 3 main components. These; architectural structure, learning algorithm and activation function. When we examine the architectural structure, the input layer consists of the hidden layer and the output layer. In the learning algorithm, the weights in the whole network should take optimal values. In fact, training the net is to find the best value of the weights (Graupe, 2013). The activation function provides the match between input and output layers.

Artificial neural networks learn by making mistakes. Basically, artificial neural networks learn in 3 stages. In the first step, outputs are calculated. In the second step, it compares the outputs with the target outputs and calculates the error. In the last stage, the process repeats by changing the weights (Livingstone, 2008). Artificial neural networks architecture is shown in Fig 1 as input layer, hidden layer and exit layer.

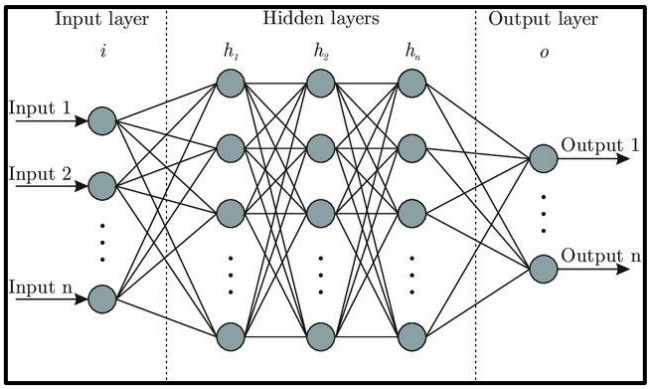

Figure 1. Artificial neural network architecture (Bre, Gimenez, & Fachinotti, 2018)

Artificial neural networks consist of multi-layer computational units. The data received from the external environment is applied to the input layer and processed at the entrance, the information is transmitted to the middle layer without any change in the flow direction. The information entering the process in these layers is transmitted forward to the output layer. Artificial neural networks based on the working principle of transmitting information in the forward direction from input to output are called forward feed artificial neural networks (Canan, 2006).

3. RESULTS

In Samsun, Atakum, Yenimahalle Quarter, 200 data were collected about the properties and values of the houses facing the tramway and these values were transformed into tables and normalized with maximum minimum normalization.

Normalized data were defined in Minitab program and analyzed by multiple regression. The result of the analysis was taken as a mathematical model and compared with the actual values of the test data.

For the other model, Matlab software was used. In the artificial neural network module in Matlab software, residential data are defined as input, output and test data. A feed forward network was created using this module. For the test data, the outputs of the Matlab software are taken and compared with the real residential values.

A comparison was made by calculating the accuracy values of the data obtained from both models. Real estate data are generally shown in Table 2.

Table 2. Real estate data

Area	Number of rooms	Building age (Years)	Floor/Number of floors	Number of bathrooms	Balcony	Furnished	Value (TRY)
45	1+1	0	0/4	1	Yes	Yes	205 000
125	3+1	21-25	0/4	2	Yes	No	300 000
160	4+1	11-15	8/8	2	Yes	No	375 000
85	2+1	0	-1/4	1	Yes	No	260 000
130	3+1	5-10	5/8	2	Yes	No	435 000
…	…	…	…	…	…	…	…
130	3+1	4	1/5	1	Yes	No	349 000
145	3+1	0	1/6	2	Yes	No	475 000
135	3+1	0	1/6	2	Yes	No	495 000
145	3+1	16-20	6/6	1	Yes	No	290 000
95	2+1	11-15	0/7	2	Yes	Yes	260 000

3.1. Multiple Regression Analysis Application

Normalized real estate data were defined in Minitab software and regression equation was obtained. Housing values were calculated with the regression equation obtained.

Regression Equation = (-0.612 + 0.405 area + 0.307 room + 0.4227 b.age + 0.1027 floor + 0.1712 bathroom + 0.2276 balcony - 0.0229 furnished)

The calculated normalized value has been converted into the real house value with the maximum-minimum formula.

$$Normalized\ data = \frac{(x - \min)}{max - min}$$

Normal probability plot created according to the value of the house is shown in Fig 2.

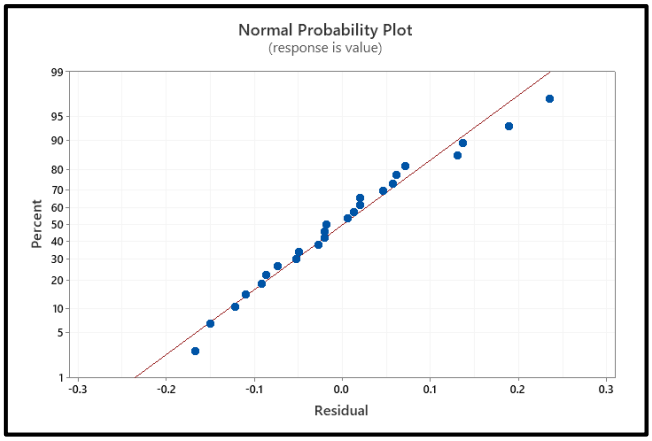

Figure 2. Normal probability plot

The values calculated by multiple regression analysis are shown in Table 3 together with the real house prices.

Table 3. Multiple regression analysis results

MRA Value (TRY)	Resale Value (TRY)	Accuracy (%)
219361,667	205000,000	92,994
294014,667	300000,000	98,004
410175,000	375000,000	90,620
286537,000	260000,000	89,793
380271,000	435000,000	87,418
…	…	…
331190,000	349000,000	94,896
437010,667	475000,000	92,002
426800,667	495000,000	86,222
295148,667	290000,000	98,224
267976,333	260000,000	96,932

3.2. Artificial Neural Networks Application

Normalized housing data are defined in Matlab software. The training of the network was carried out with 8 neurons using experimental data with the created feed forward artificial neural network.

The training process was repeated a few times to make the learning process more accurate. The maximum failure value was entered as 500, and the iteration amount was determined as 1000.

The accuracy and consistency values of the data were examined by looking at the regression chart after the training. The trained network is simulated with test data.

Housing values were determined by applying the maximum-minimum normalization reversed to the values obtained. Artificial neural network training regression is shown in Fig 3. The results are shown in Table 4.

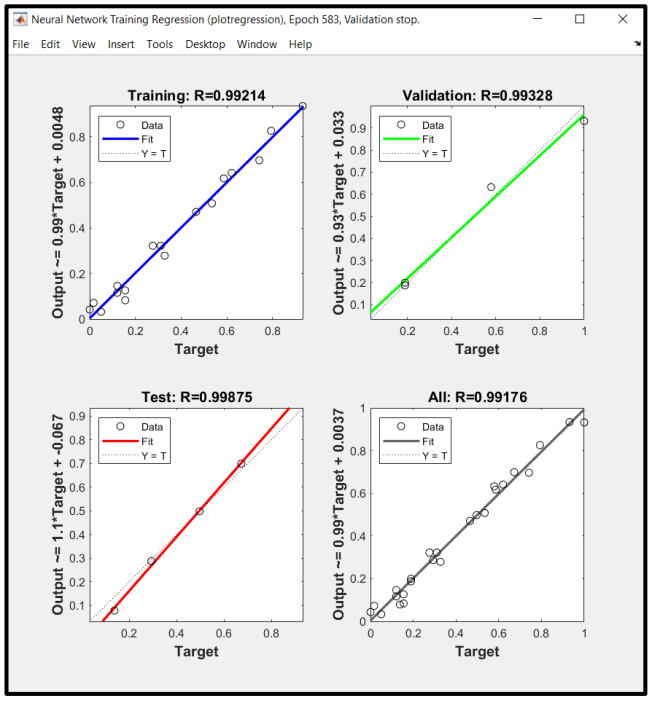

Figure 3. Artificial neural network training regression

Table 4. Artificial neural network results

ANN Value (TRY)	Resale Value (TRY)	Accuracy (%)
222862,080	205000,000	91,286
293771,863	300000,000	97,923
376021,579	375000,000	99,727
257108,193	260000,000	98,887
454902,955	435000,000	95,424
…	…	…
320026,201	349000,000	91,698
471619,035	475000,000	99,288
471124,399	495000,000	95,177
291098,107	290000,000	99,621
259925,188	260000,000	99,971

4. DISCUSSION

In the study, the values obtained in the house valuation model created using multiple regression analysis and artificial neural networks were compared with the real values of the houses and the accuracy values were determined. The houses used in the study were obtained from Samsun province Atakum district Yenimahalle neighborhood. In determining the residences, the ones facing the tramway street were chosen as the location.

The total accuracy value obtained in the multiple regression analysis was determined as 93.232%. The total accuracy value obtained in artificial neural networks was determined as 97.090%.

5. CONCLUSION

Considering the accuracy values obtained in the study, it is seen that the housing valuation model created with artificial neural networks gives a higher accuracy value. Multiple regression analysis is more suitable in terms of ease of application and understandability of the model. In artificial neural networks, there may be some results that vary with the change of the number of neurons and the selection of functions. The authors recommend the use of statistics and modern methods in frequent applications such as real estate appraisal.

REFERENCES

Açlar, A., & Çağdaş, V. (2002). Taşınmaz (gayrimenkul) değerlemesi. *TMMOB Harita ve Kadastro Mühendisleri Odası, Ankara*.

Akış, B. (2013). *İstatistiki yöntemlerle değer belirleme ve değer haritası üretimi Selçuklu örneği.* Selçuk Üniversitesi Fen Bilimleri Enstitüsü.

Atik, M., KÖSE, Y., Yilmaz, B., & ERBAŞ, M. (2015). Şehirlerin İlerleme Yönlerinin Gayrimenkul Değerleri Üzerindeki Etkisinin Ölçülmesi. *Çankırı Karatekin Üniversitesi, 5*, 443-458.

Bre, F., Gimenez, J. M., & Fachinotti, V. D. (2018). Prediction of wind pressure coefficients on building surfaces using artificial neural networks. *Energy and Buildings, 158*, 1429-1441.

Brown, R. K. (1965). *Real estate economics: an introduction to urban land use*: Houghton Mifflin.

Canan, S. (2006). *Yapay sinir ağları ile GPS destekli navigasyon sistemi.* Selçuk Üniversitesi Fen Bilimleri Enstitüsü

Chatterjee, S., & Hadi, A. S. (2015). *Regression analysis by example*: John Wiley & Sons.

Graupe, D. (2013). *Principles of artificial neural networks* (Vol. 7): World Scientific.

Livingstone, D. J. (2008). *Artificial neural networks: methods and applications*: Springer.

Özdamar, N. (2004). 4822 Sayılı Yasa İle Değişik 4077 Sayılı Yasa'da Tanımlanan Konut Nedir. *Türkiye Barolar Birliği Dergisi, Yıl, 17*, 317-331.

Öztürk, K., & Şahin, M. E. (2018). Yapay Sinir Ağları ve Yapay Zekâ'ya Genel Bir Bakış. *Takvim-i Vekayî, 6*(2), 25-36.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: a review of valuation methods. *Journal of Property Investment & Finance*.

Ring, A. A., & Dasso, J. J. (1977). *Real estate principles and practices*: Prentice Hall.

Yegnanarayana, B. (2009). *Artificial neural networks*: PHI Learning Pvt. Ltd.

Yomralıoğlu, T., Nişancı, R., Çete, M., & Candaş, E. (2011). DÜNYA'DA VE TÜRKİYE'DE TAŞINMAZ DEĞERLEMESİ.

Zeng, T. Q., & Zhou, Q. (2001). Optimal spatial decision making using GIS: a prototype of a real estate geographical information system (REGIS). *International Journal of Geographical Information Science, 15*(4), 307-321.