http://dx.doi.org/10.15576/GLL/2017.1.17

GLL
Geomatics, Landmanagement and Landscape No. 1 • 2017, 17–24

EMPLOYING LINEAR ARTIFICIAL NEURAL NETWORKS IN PROPERTY APPRAISAL AND VALUATION – POSSIBLE APPLICATIONS

Mariusz Górak

Summary

Transactional price is the result of some kind of free market game, and of independent decisions taken by the parties to the transaction. Prices depend on a number of factors, specific to the local real estate market. The impact of some factors is fixed, while others are dependent on the location of the property. Therefore, research into the determination of rules that would describe the relationship between the market price of the real estate, and its market characteristics, remain valid. The article presents the possibilities of applying linear artificial neural networks to real estate valuation. Using a database, the artificial linear neural network is developing a regression model, which produces the results that oscillate close to the market value of the property. The necessary condition is the creation of a database that is representative of the given real estate market.

Keywords

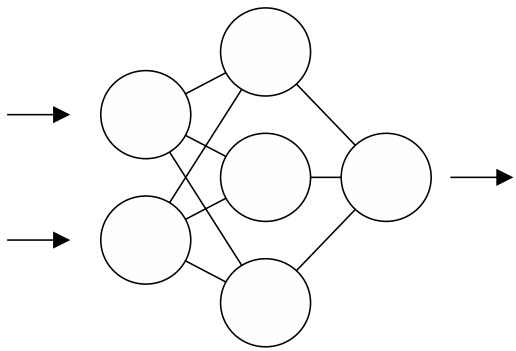
neural network • artificial neurons • training set • property valuation • database

1. Introduction

Artificial neural network operates on the principle of a scheme. In the scheme, by a suitable amount of interconnected elements called neurons, and with an appropriate amount of information (so-called signals), calculations are performed. The result is an approximate solution to the task demanded of the network.

Construction of artificial neural networks should be transparent, so that it can be traced easily and quickly, and if necessary, also easy to control. A sample scheme of an artificial neural network was presented in Figure 1. The artificial neural networks used in practice are flat networks with a regular structure, in which there are layers of neurons with a specific purpose. These neurons are usually joined together according to the "peer-to-peer" principle [Tadeusiewicz et al. 2007].

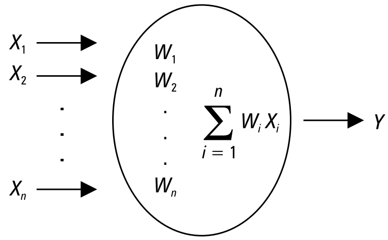
What is characteristic of the neurons, which are the components of the network, is that they have a number of inputs and only one output. Figure 2 shows a sample scheme of a neuron. The input information *Xi* (*i* = 1, 2, ... , *n*) and the output result *Y* should take numerical values from the range of 0 to 1, or in some cases, -–1 to + 1. If we expect certain information or decision from the task demanded of and being solved by the network, we must first assume specific significance for potential output results as well as the input information in our possession. When the obtained values are outside the established range (e.g. between 0 and 1), we apply a scale to the results obtained. The task performed by the essential element (unit) of the artificial neural network, that is, the neuron, is based on processing and modifying the input information *Xi* into the output results of *Y*, using its connections and the knowledge fed to it.



Source: author's study based on [Tadeusiewicz et al. 2007]

Fig. 1. Sample scheme of an artificial neural network

The knowledge of the neurons results from what the given neuron had been taught. For their learning, the neurons utilize the *Wi* coefficients, or the so-called weights. The weights are assigned to each of the neuron's inputs, and may have variable values. The information entering the given input is multiplied by the weight of that input. This modification enhances the information for further calculations, when the weight is greater than 1; or it represses the information, when the weight is less than 1. There exists also the so-called contradictory information. This applies to situations when the weight is of negative value (below zero). The inputs, in which there are negative weights, are commonly called inhibitory inputs, while the inputs with the positive weights are called stimulatory inputs. Inside the artificial neuron, aggregation occurs, which usually involves summing up the information modified by the weights. As a result, we obtain an auxiliary internal signal. The neuron frequently adds an additional ingredient to such an auxiliary internal signal, which is independent of the input information – the so-called threshold. In a situation, where the threshold is taken into account, it is subjected to the process of learning. Therefore it can be assumed that the threshold is an extra weight, associated with the input, to which an internal signal is connected, with a weight equal to 1. In order to obtain the output result, it is necessary to add up the internal signals, previously multiplied by the weights [Tadeusiewicz et al. 2007].



Source: author's study based on [Tadeusiewicz et al. 2007]

Fig. 2. Sample scheme of a neuron

2. Method and area of study

In our studies, we have applied learning networks that were based on the collated database, which was created as a result of the analysis of the local real estate market. The data collected relate to transaction prices of real estate: business premises and residential properties [Siejka 2011].

The proposed learning process is a procedure of network learning "with the teacher." Learning conducted using such a method entails, inter alia, the transfer of examples of proper operation to the network, so that the network imitates these examples later on, during its subsequent operation. The examples should contain specific input information and output result. The presented configurations of the input information and the required output result are analysed by the network in terms of the relationship between the input information and the result. In other words, during the network learning process "with the teacher" there are always two sets, coupled together. The first set contains the data for the task, while the other contains the solutions.

In practice, the "teacher" of the network is a computer program, equipped with a training set, containing properly selected and properly classified data with indications as to which are the input, and which are output data. Inessential data, useless for the purpose of solving the task, should be omitted [Tadeusiewicz et al. 2007].

Table 1 contains corrected transaction prices of 19 selected properties (business premises), with similar characteristics to the property, which is the subject of the value appraisal [Cymerman and Hopfer 2009]. In the presented data, we can distinguish the data we are going to use as input data – i.e. concerning the situation/location/zone, area, wear, standard and floor, and the output data on the corrected price. Therefore, one can assume that Table 1 contains information, based on which the network can be taught. In one of the data sets, features of the real estate market of business premises (offices and similar properties) have been selected, important for the local real estate market, in which the property under evaluation is located. In another data set, we shall find the output result, in the form of corrected prices for selected properties with similar characteristics.

Table 1. Learning set

No.	Situation /location/zone	Area	Wear [%]	Standard W, P, N	Floor	Corrected price [zł/m²]
1	very good	very good	10	W	I	5 508.63 zł
2	good	very good	20	P	III	4 983.31 zł
3	good	very good	10	N	ground	5 300.92 zł
4	very good	very good	30	W	II	5 456.15 zł
5	poor	very good	10	W	I	5 174.14 zł
6	good	good	20	P	III	4 807.01 zł
7	good	very good	40	N	IV	4 983.90 zł
8	very good	very good	30	N	II	5 295.97 zł
9	poor	very good	10	W	I	5 163.69 zł
10	good	good	20	P	III	4 840.07 zł
11	good	very good	40	N	IV	5 014.58 zł
12	very good	very good	30	N	II	5 328.00 zł
13	poor	good	40	N	ground	4 447.52 zł
14	poor	good	40	N	I	4 530.24 zł
15	poor	good	40	N	I	4 509.56 zł
16	very good	very good	30	W	I	5 479.13 zł
17	poor	very good	10	W	II	5 181.64 zł
18	good	very good	10	N	IV	5 292.77 zł
19	good	very good	30	N	IV	5 000.00 zł

Source: author's study

Most data presented in Table 1 is qualitative, that is descriptive or "nominal". This means that the values of the data are names, which must be assigned the appropriate number. In order to properly apply the data, it is necessary to use the "one of *N*" technique, where *N* is the number of different possible values (names), which a given nominal variable can take. For example, one of the characteristics we want to reproduce in the input of the network is the situation in which $N = 3$, because the situation of the property in question might be: poor, good and very good. In order to tell the network the correct value when, for instance, the situation of the property is determined to be good, we give the signal 1, and in other cases, 0 [Tadeusiewicz et al. 2007].

During the process of network learning "with the teacher", the key parameters are the weights, by which we multiply each input signal, and then add that up with other signals. In order to facilitate the start of the learning process, we can use the weights listed in Table 2.

The weights of individual market characteristics, presented in Table 2, have been calculated on the grounds of the database, created as a result of a detailed analysis of the local real estate market, of properties similar to the property which is subjected to appraisal and valuation [Siejka 2010].

Table 2. Weights of various market features

No	Item	Pairs	Calculations	Weight of the feature [%]
1	situation/location/zone	1–9	32.5073	
		1–5	31.5223	
		average =	32.0148	33
2	area	2–10	13.4991	
		2–6	16.6142	
		average =	15.0566	16
3	wear [%]	7–3	29.8760	
		7–18	29.1082	
		average =	29.4921	30
4	standard W, P, N	4–8	15.0960	
		4–12	12.0768	
		average =	13.5864	14
5	floor	13–14	7.7953	
		13–15	5.8464	
		average =	6.8208	7
		sum =	96.9708	100

Source: author's study

During the learning process, the values of the weights change. This is decided independently by each an autonomous neuron, that is to say, every single neuron, based on the appropriately selected rules, independently determines which weights should be changed and by how much. During the learning process, the "teacher" does not have to scrutinize all the details; instead, it is enough to provide the network with a training set, containing the pattern of a correct solution, so that the network can gain knowledge about the error that it has committed. This knowledge will be acquired as a result of the network having compared the solution it provided by itself against the solution stored in the training set as a reference. The appropriate structure of the learning algorithms allows the network to correct the value of the weights based on the error, until it learns to solve problems and tasks from the training set, and on the basis of this knowledge, to solve similar tasks put to it. The input signals Xi ($i = 1,2, ..., n$) and the output signal Y can take the values from an exemplary, delimited range.

When using the scaling function, we assume that $X \in [-1,1]$, for each i, and $Y \in [-1,1]$ while the relation of $Y = f(X_1, X_2, ..., X_n)$. Then, the output signal Y can be determined using the following linear equation:

$$Y = \sum_{i=1}^{n} W_i X_i \tag{1}$$

wherein coefficients W_i, known as synaptic weights, can vary during the learning process. This process is one of the essential traits of neural networks as adaptive information processing systems. Among other things, the element described with the linear equation (1) is capable of recognizing the input signals.

For out explanation, we use vector notation, that is, a set of input signals to the neuron creates a vector (2) stored as a column of n components [Tadeusiewicz 1993].

$$Y = \begin{bmatrix} X_1 \\ X_2 \\ \cdot \\ \cdot \\ \cdot \\ X_n \end{bmatrix} \tag{2}$$

The above vector (2) can be interpreted as a point in the n-dimensional input space of X, which, when described in the form of a transposed matrix, is as follows:

$$X = [X_1, X_2, ..., X_n]^T \tag{3}$$

The set of n weight coefficients can be interpreted as a point in the n-dimensional space of W weights, which, when described in the form of a transposed matrix, is as follows:

$$W = [W_1, W_2, ..., W_n]^T \tag{4}$$

With the above assumptions, we express the equation of the neuron in the form of a scalar product of vectors: of the inputs and the weights.

$$Y = W \cdot X \tag{5}$$

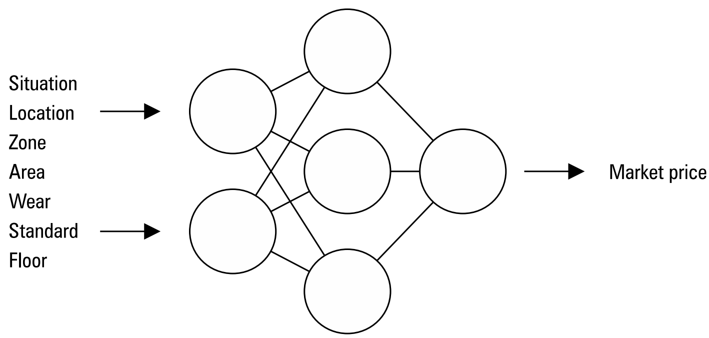
From the properties of the scalar product, it follows that the output signal Y of the neuron is larger when the position of vector X (input) corresponds to the position of vector W (weights). The neuron thus perceives and distinguishes the input signals similar to its own vector of weights, because the components of the input signal X include, for instance, the characteristics of certain objects. At the same time, the output signal Y can be a measure of the similarity to the selected object [Tadeusiewicz 1993].

3. Results and discussion

The result of the network learning process "with the teacher", as presented herein, is the market price of the property being appraised. The result is an approximate value, which should be appropriately scaled. Therefore, for the proper interpretation of the result obtained, it is necessary to interpret the result obtained in the correct way, as artificial neural networks form two types of models: regression models and classification models.

The regression model of the network, shown in Figure 3, is tasked with providing a numerical value. This model can be helpful in assessing the market value of a real estate property (one used for an office or business premises). Then, the input information is provided, which may take the numerical format (e.g. the technical condition/wear given in percentage), as well as information in the nominal format (e.g. the situation/location/zone, designated as for instance "very good"). In turn, at the output we expect the result in the form of a value, which determines the approximate market value of the given property. Based on the information that the network has received for the purpose of learning, provided in the format of an appropriate database (selected as a result of the analysis of the local real estate market) including transaction prices for similar properties and characteristics assigned to them, that artificial neural network is able to perform a regression model that receives the results oscillating close to market value.

In the classification model, the network is designed to allocate the object in question into one of the classes. This model can be helpful in determining whether the profit from the given investment can be classified as, for instance "small", "medium", or perhaps "large". Then, the input information is provided, which may take the numerical format, as well as the information in a nominal format. At the output, we expect a result in the nominal format. Based on the information that the network has received for the purpose of learning, provided in the format of an appropriate database, the artificial neural network is not able to perform a classification model that would produce the precise amount of profit from the given investment [Tadeusiewicz et al. 2007].



Source: author's study based on [Tadeusiewicz et al. 2007]

Fig. 3. Regression network

4. Conclusions

The scheme of operation of the linear artificial neural networks, presented herein, shows that these networks can find their application in real estate appraisal and valuation. Considering the fact that the operation of the linear artificial neural networks is possible to predict, and due to the fact that these networks have a simple mathematical description, they should be deployed for property valuation.

Current technologies make it possible to create systems, which facilitate the use of databases for the valuation of property. Please note that in order to prepare the appropriate databases, a detailed analysis of the local real estate market is required, which must be focused on similarities to the property being appraised.

When using the linear artificial neural networks, we must bear in mind that the results obtained by the network are approximate values. Therefore, we need to rescale them properly for the purpose of interpretation.

Further research will be conducted into the application of the linear artificial neural networks for measurement, using the specific examples of real estate properties.

References

Cymerman R., Hopfer A. 2009. Systemy, zasady i procedury wyceny nieruchomości. PFSRM, Warszawa.

Siejka M. 2010. Próba zastosowania metody AHP do oceny wpływu zmiany cech nieruchomości gruntowych na ich wartość. Infr. Ekol. Ter. Wiej., 12, 93–101.

Siejka M. 2011. Aspekty wykorzystania aktywnych baz danych w wycenie nieruchomości. Infr. Ekol. Ter. Wiej., 3, 235–250.

Tadeusiewicz R., Gąciarz T., Borowik B., Leper B. 2007. Odkrywanie własności sieci neuronowych. Polska Akademia Umiejętności, Międzywydziałowa Komisja Nauk Technicznych, Kraków.

Tadeusiewicz R. 1993. Sieci neuronowe. Akademicka Oficyna Wydawnicza, Warszawa.

Tadeusiewicz R. 2007. Exploration of neural networks properties using specially written programs. Polska Akademia Umiejętności, Międzywydziałowa Komisja Nauk Technicznych, tom II, Kraków, 113–124.

Mgr inż. Mariusz Górak, doktorant
Uniwersytet Rolniczy w Krakowie
Katedra Geodezji
30-198 Kraków, ul. Balicka 253a
e-mail: mariuszgorak@interia.eu