International Journal of Strategic Property Management
ISSN: 1648-715X / eISSN: 1648-9179
2022 Volume 26 Issue 2: 86–105
https://doi.org/10.3846/ijspm.2022.15962

REAL ESTATE VALUATION MODELS PERFORMANCE IN PRICE PREDICTION

Adela DEACONU [1,*], Anuța BUIGA [2], Helga TOTHĂZAN [1]

[1]Department of Accounting and Auditing, Faculty of Economic Sciences and Business Administration, Babeș-Bolyai University, Cluj-Napoca, Romania
[2]Department of Mathematics and Statistics, Faculty of Economic Sciences and Business Administration, Babeș-Bolyai University, Cluj-Napoca, Romania

Received 18 March 2021; accepted 12 October 2021

Abstract. Using a sample of 900 apartments from Cluj-Napoca, Romania, containing selling transactions for the second semester of 2019, and data for 33 locational, physical and neighbourhood-related attributes (socio-cultural, environmental, and urbanism related), our research objective is to test the performance in price prediction, and hence the utility, of the Artificial Neural Networking (ANN), as artificial intelligence model versus the Generalized Linear Model (GLM), as a regression model. By contributing to an ongoing debate, our empirical findings confirm the results of a predominant group of earlier studies, namely the superiority of ANN. Precisely, we found that ANN can better predict selling prices and provides stability of results. Additionally, we addressed the critiques related to the transparency of results, showing that ANN also has the ability to illustrate the significance of the different attributes of real estate, if appropriate statistical indicators are used. These findings can serve the different real estate valuation purposes, including that of the review of valuation reports.

Keywords: real estate, valuation reports review, artificial neural networking, price prediction.

Introduction

The value of real estate reflects the influence of factors such as the property attributes, market evolution, and participants' perceptions in a specific context. In the real estate typology, residential property valuation plays an important role for households, investors, financial institutions, regulators, and the public policy. Therefore, the accuracy of an estimate for the value of a given property, as well as assuring appropriate tools for the review of such estimates, aiming to avoid arbitrary measurements, is critical. Our research objective is a comparative analysis of the performance of real estate valuation models in price prediction, subsumed to market approach, intended to provide a reviewing tool for valuation reports. The models we compare are a regression model, *i.e.* the Generalized Linear Model (GLM) which we found appropriate for residential properties and for our dataset and an artificial intelligence model (machine learning), *i.e.* the artificial neural networking model (ANN). GLM approach allows us to build a linear relationship between the response and the predictors, even though their underlying relationship is not linear. We opted for ANN from its family of applications, given that it is the most used today for real estate value predictions and also because it is controversial in terms of superiority over classic hedonic models. Thus, we were encouraged to test it for a new context, in order to add to the literature new results for or against ANN. We do not believe that ANN replaces traditional valuation approaches, however it facilitates the application of the market approach by eliminating some of the subjectivism valuators might experience when using the market grid. It also helps verify evaluations by using a model (ANN) developed by researchers, and there is no need to redo the evaluation of the professional who prepared the evaluation report. The research question we formulated is whether ANN is superior to GLM in terms of transparency, predictive ability and stability of output.

The investigation was conducted on a segment of the housing market in the Romanian urban area and used a significant number of properties' attributes, both quantitative and qualitative. The real estate market investigated is located in Cluj-Napoca, an important city from Romania.

*Corresponding author. E-mail: *adela.deaconu@econ.ubbcluj.ro*

Our model used 22 variables as predictors of the selling price, almost half of which were qualitative variables (for example, the existence of relaxing places in the area). We are thus trying to cover a gap in the literature, noting that qualitative variables are less used in articles that have tested ANN's performance, although the interest of real estate market participants in these price predictors is growing.

There is mixed evidence of the superiority of ANN over the regression models for real estate prices prediction and analysis, according to a body of literature spanning for the last three decades (*e.g.* Tai & Ho, 1991; Din et al., 2001; Worzala et al., 1995; Nguyen & Cripps, 2001; Curry et al., 2002; McCluskey et al., 2012; Núñez Tabales et al., 2013; Bogin & Shui, 2020). Furthermore, real estate markets and prices vary geographically, and the findings of such studies are dependent on the economic context under analysis. For emergent contexts we noticed the scarcity of similar studies (*i.e.* Cechin et al., 2000 for Brazil; Selim, 2009 for Turkey; Lai, 2011 for Taiwan; Sampathkumar et al., 2015 for India; Hu et al., 2019 for China).

Nowadays, ANN (as a version of AVM – Automated Valuation Model) is not restricted only to researchers, but it is used by practitioners to estimate value for various purposes, mainly for mortgage lending, and tax purposes (Glumac & Des Rosiers, 2021b). We could also add the review of real estate valuation reports, which is a relevant issue for the valuation profession, as an emerging specialization. In professional associations, valuation reviewing has evolved in the last decades in countries with a tradition for market estimations (Isakson, 1998; Vascu, 2015), more prominently in countries like USA, UK and Australia. In Europe, although there was limited interest in this area in the past, nowadays it is more pronounced in professional standards prescriptions and valuation practice (Scheurwater, 2017). Generally, there are difficulties in valuation reviews triggered by the inherent subjectivism of valuators' opinions. Isakson (1998) asserted that the usual techniques consist of a check of the valuation report for math errors and deviations from the standards of practice. Besides these, few independent techniques have been developed by practitioners or researchers (for example for the American market, Roulac, 1986; Mathieson & Dryer, 1993). We therefore plead for the use of ANN in property valuation, including for reviewing purposes.

Our empirical results suggest that ANN is superior to GLM in terms of real estate selling prices prediction ability. Additionally, we rank real estate attributes in terms of significance for the selling price – that prove to be confirmed by GLM –, in order to counteract the criticism regarding ANN's lack of transparency. Finally, we argue that ANN results are stable as regards the analysis criteria of the model utility.

The research and professional gap we intend to cover is linked to the goal of our research – improving valuation reviews, and the geographical context investigated. Firstly, our results are a new attempt to apply ANN as a research methodology, additional to the traditional regression models. Secondly, our study brings practical/professional contributions. Therefore, we believe that the research output could be useful to valuators specializing in real estate review that are now aware of the utility of an ANN model to substantiate the accuracy of valuations. This could be promoted and even developed and periodically updated for different geographical areas, including by the Romanian valuation professional association, the Association of National Valuators of Romania (thereafter ANEVAR). It is worth mentioning that ANEVAR has recently introduced valuation reviews in its standards, based on the International Valuation Standards (IVS) and the European ones (EVS). Also, even if the model developed by us proves its usefulness for a confined geographical area, Cluj county, we believe we have confirmed its utility for the entire Romanian emergent context. The peculiarities of this type of market and the economic transition factors justify a distinct analysis for the space and the economic context proposed by us.

In the following, Section 1 highlights the core of market approach as preferred method in the residential real estate valuation, respectively the relevant models used and the content of valuation review, as area of application of valuation models under analysis. Section 2 reveals our research design, describing the variables and the sample. In section 3 we conduct an empirical investigation in order to test models' performance. The last section concludes on the most suited model in general and in the described context, as a grounded support for valuation reviews.

1. Literature review

1.1. Regression models and ANN

There are several methods proposed in literature to best estimate the selling price of a property. The work coordinated by Kauko and d'Amato (2008) examines the appraisal practices from an international perspective, including comparable sales approach, automated valuation models, fuzzy logic techniques, hedonic price modelling, spatial analysis. Three approaches have been developed in the professional standards (*e.g.* IVS) and guides, namely the market, income and cost approaches, from which several valuation methods derived.

The market approach is widely considered the most appropriate for valuing residential real estate in an active market (Worzala et al., 1995; American Society of Appraisers, 2004; Glumac & Des Rosiers, 2021a). Moreover, this approach is best suited to, and is the most reliable in, the appraisal of single-family homes. It relies on market efficiency and the similarity of a specific property to another, a recently traded one.

As market approach related methods, hedonic models have been used to complete the sales comparison grid and to establish a hedonic price index for a property with given attributes (DiPasquale & Wheaton, 1994; Mayer & Somerville, 2000). Curry et al. (2002) assert that this approach has the advantage of using observed prices and attributes rather than subjective valuations or statements of intent. Based on regressions, these models estimate the marginal effect of real estate attributes on property sale trading price (Rosen, 1974). The link between property price P and observable attributes of the property $X_1, X_2, ...X_k$ is given by the hedonic price function f.

First of all, we discuss the linear regression models, which are frequently used for property valuation given the large number of price predictors involved, as we intend to compare a hedonic model and ANN, included in machine learning methodology. The least squares model (OLS) is the most common instrument used in the valuation process, a traditional one according to Benjamin et al. (2004). Its use was apparently initiated in 1980s, a decade associated with the beginning of the rise of information systems (Kuburić et al., 2012). OLS is part of the Generalized Linear Model (GLM) approach and it is based on several independent (explanatory) variables able to predict the real estate price, and assumes a normal distribution for the dependent variable, that the independent variables' coefficients have linear effects on the dependent variable, and also the hypothesis of homoscedasticity of errors. This classic linear model was criticized for the non-linearity of the input variables observed in practice (Brunson et al., 1994; Do & Grudnitski, 1992); the heteroscedasticity of errors that leads to biased estimations and therefore to biased inference on the coefficients (Stevenson, 2004; Helbich et al., 2014); and the multicollinearity of the independent variables and the inclusion of outlier properties in the sample (Worzala et al., 1995). In this context of non-normal distribution of prices, of heteroscedasticity issues, and of the nonlinearity of a predictor, GLMs could be used as they provide flexibility for house price distributions and variance assumptions (McCulloch & Searle, 2001). Another concern raised by some researchers (Do & Grudnitski, 1992; Brunson et al., 1994) about a hedonic model is its inadequacy for a market that requires precise and fast responses, but this critique is not applicable to the post-analysis of valuation reports, the valuation aim to which we refer in this paper. For this verification, a post-calculated model that suggests an average evolution of real estate attributes and prices is suitable.

Besides the hedonic regression, we also propose another model for valuation reviews purposes – ANN, pertaining to the machine-learning family. It is described as a flexible nonlinear model that enables a universal approximation (Do & Grudnitski, 1992; Nguyen & Cripps, 2001; Curry et al., 2002; Jahanshiri et al., 2011). Artificial intelligence systems, mainly artificial neural networks (ANN), are viewed as predictive systems which replicate the activity of the human brain, based on learning processes and providing solutions to complex issues (Worzala et al., 1995; Kuburić et al., 2012; Ho et al., 2020). The very beginning of this model is in 1943, when it was introduced as an alternative to algorithmic programming (Núñez Tabales et al., 2013), and since then it has continuously evolved as application refinements and work speed. The content of ANN is based on the fact that the learning process undertaken by the human brain occurs and reoccurs through the repetition of the input stimuli and the output response. The method uses a regression having as dependent variable a sales price generated by a software containing mathematical algorithms that allow the repetition of tests, so as the regression error minimizes. More precisely, ANN uses a first layer of inputs (the independent variables, namely the real estate attributes), a "hidden" layer which contains weights (coefficients) allocated by the software to different independent variables, and the output layer (the dependent variable). In order to conclude on the ANN prediction accuracy the set of weights must be found and assigned by a nonlinear transfer function to all the independent variables which can minimize the prediction error, the error between the neural network output and the current sales price observed on a sample of real estate transactions.

ANN has been reported as appropriate for real estate valuation since 1990s (*e.g.* Do & Grudnitski, 1992; Tay & Ho, 1992; Worzala et al., 1995) and widely considered useful for mass appraisals (Nguyen & Cripps, 2001; Mora-Esperanza, 2004; McClukey et al., 2012; Yacim & Boshoff, 2018).

Generally, these studies compared ANN with the regression models in terms of the ability to predict an accurate selling price for properties, usually residential ones, and standardized to a certain number of rooms. Although in most cases ANN turned out to be superior, there are many sceptical researchers. This mixed evidence was revealed in the first period of research interest in ANN, by Worzala et al. (1995). The review undertaken mentioned several studies that attested ANN superiority, *i.e.* Borst (1991), Do and Grudnitski (1992), Tay and Ho (1992), and Evans et al. (1992). However, Worzala et al. (1995) referred to other studies that did not report successful results of ANN application in finance generally, as in Allen and Zumwalt (1994) literature review. Worzala et al. (1995) also revealed this scepticism in their study. The authors invite to prudence in using ANN in the real estate field due to differences in the results obtained by running the two software packages, reduced speed in data processing, and finally, disparity between the results of the same software package with repeated tests. McCluskey et al. (2012) who revisited the ANN topic refer to the findings of Worzala et al. (1995) and McCluskey and Borst (1997) who were not in favour of ANN in order to test themselves the predictive accuracy of ANN. They expressed significant reserves towards ANN superiority, especially for mass appraisals. Their restraints are related to transparency, stability of output, predictive ability and defensibility, as factors not provided by ANN. However, this study asserts the ANN potential for predictive modelling in a quick and cost-effective way.

Table 1 presents a summary of the literature review from above, focusing on the three qualitative criteria which we set out to test (transparency, predictive ability and stability of output).

Table 1. ANN versus regression models: pros and cons

Transparency of the underlying model structure
Cons for ANN: The black box nature of the model is criticized: no final model structure/evidence is provided to suggest the individual contribution of variables to the predicted value (McCluskey et al., 2012) Presents weaknesses due to the black box nature and the lack of interpretation of the produced output (Abidoye & Chan, 2017) Presents a limitation due to the black box nature of the model: the valuator cannot know with certainty what values and forms the variables assume in the learning processes (Valier & Micelli, 2020)
Predictive power/accuracy
Pros for ANN: Is superior as predictive power (Tay & Ho, 1992; Do & Grudnitski, 1992) Provides highly precise outputs (Mora-Esperanza, 2004) Offers more realistic pricing of individual properties (Din et al., 2001) Generates less pricing errors and has greater pricing precision out-of-sample (Peterson & Flanagan, 2009) Provides more realistic marginal prices (Núñez Tabales et al., 2013) Offers a better goodness of fit with the usual statistical measures, e.g. determination coefficient (Núñez Tabales et al., 2013) Is an improved alternative for prediction (Selim, 2009) The proven predictive power could be enhanced by optimization algorithms (Vo et al., 2014) Produces more accurate and reliable estimates (Abidoye & Chan, 2017) Provides a higher accuracy of prediction (Valier & Micelli, 2020)
Cons for ANN: Its superiority as predictive power or accuracy is not fully proven (Worzalla et al., 2005; McGreal et al., 1998)
Results extrapolation/replicability and stability
Pros for ANN: Extrapolates better from more volatile pricing environments (Peterson & Flanagan, 2009) Has generalization capabilities (Xie & Hu, 2007)
Cons for ANN: The non-repeatability of model design (different results each time the model is run) impacts the reliability of the output (McCluskey et al., 2012)

An analysis of the literature published in recent decades shows that ANN, although used in real estate prediction processes since the 1980s, is still relevant. It proves relevant and it is tested in more and more geographical contexts due to the availability of high volume data. ANN is even more current in this age of big data and deep learning trends. What has changed in recent years is the proposal of new optimization algorithms that should be considered for ANN training (Vo et al., 2014). The ANN performance analysis characteristics have not changed. Of the literature findings presented above, we have deemed that the predictive power, the transparency, and the stability of results are useful to test in our research.

1.2. The valuation review as one of property valuation aim

This section attempts to offer a synthesis of valuation review regarding the content, objectives, users and standards that guide it, given that valuation review is at the core of our research. The presentation is based on professional standards and guides, because we did not find relevant scientific literature about this purpose of property valuation and the methods to do the review. In fact, this is why we are preoccupied by ANN, as a tool to review the valuation reports.

A valuation review is a review of a valuator's work undertaken by another valuator exercising impartial judgement (International Valuation Standards Committee [IVSC], 2003). Even if this International Valuation Guidance Note no. 11 "Reviewing Valuations" is obsolete and the current IVS no longer contains a document dedicated to this activity, the definition provided earlier is the nutshell of this specific work. The valuation reviewer may perform certain valuation procedures and/or provide an opinion of value (International Valuation Standards Council [IVSC], 2017). In this case, specific professional competencies must be provided, for which reason the valuation associations, including ANEVAR in Romania, developed a distinct specialization for its members (Vascu, 2015).

Valuation reviews are performed for multiple reasons, at the request of a third party, the most significant in terms of effects being in our opinion the valuations for legal proceedings/circumstances (potential litigations), respectively for mortgage lending processes. The usefulness of a valuation review for the courts is linked to the value "approval" between several valuation reports; therefore, it is critical to provide to courts supplementary data that enable correct legal judgments. The effect of non-compliant valuations on banks also has obvious consequences; therefore, the need to prevent the risks related to miscalculated mortgage values is critical and requires internal valuation. In addition to these two aims, an interesting issue is raised by TEGoVA in its EVS in relation to volatile or illiquid markets, which are supposed to require valuation reviews more often (TEGoVA, 2020).

Valuation associations around the globe distinguish between various types of reviews, reflected in valuation review reports. Considering our aim, we limited this classification to two types, such as a simple or extended objective (Association of the Romanian Valuators [ANEVAR], 2020). In the first case, the valuator examines the conformity of the analysed valuation report with the valuation standards in place and with other specific regulations; and in the second case, the reviewer's opinion is required on the value estimated in the investigated valuation report. This second case suits our research aim as the reviewing valuator should proceed with his own valuation process, often a retrospective one. After that, a quantitative threshold is sometimes advanced, in order to decide on the initial valuation report quality. The Romanian valuation standards, the Standards of Valuation of Properties (SEV), 2020 edition discuss a 20% limit as difference between the initial value and the reviewer's estimated value for the same property; even if this limit is not considered a priori a non-compliance with the valuation standards before it is argued in the spirit of the professional standards in force (ANEVAR, 2020). RICS appreciate as negligence in evaluation an error margin of 10–15%, this quantitative threshold being used in courts. As characteristics of the reports reviewed, several criteria were evoked in different texts, *e.g.* correctness, consistency, reasonableness and completeness (IVSC, 2003), accuracy of valuation (TEGoVA, 2020).

2. Research methodology and data

2.1. Research design

The mixed evidence described in the previous section prompted us to test the two models, a hedonic regression and ANN, in the Romanian context and to provide an opinion on ANN usefulness for valuation reports review purposes. Therefore, our research question is if ANN could be considered the optimal model, and if so, how it could be used in further valuation endeavours in this specific context.

Given the fact that critics addressed linear models with estimators of OLS type (*e.g.* heteroscedasticity issues) in literature, and taking into account the characteristics of the data collected (many predictors are dummy variables; and there are nonlinear relationships between the predictors and the dependent variable), we tested several specifications of GLM in order to estimate the hedonic equation. We applied GLM in case the dependent variable is assumed to have a normal distribution and also in case the dependent variable follows a gamma distribution. Next, we excluded as predictors the variables without statistical significance (*F-test* or *T-test*) and those that were correlated with each other, in order to eliminate the multicollinearity effect.

Thus, our research objective is to test the performance of a GLM we will select, and ANN, in predicting the selling price. The usefulness of the two models will be tested using three qualitative criteria we deem relevant, *i.e.* transparency, predictive ability and stability of the results. This list is built around the ANN weaknesses revealed in the literature and reviewed in Table 1, section 1.1. Also, these criteria take into consideration the requirements of valuation reports review prescribed by the professional standards, namely IVS and the Romanian SEV standards (2020 edition), which combine the international (IVS, 2017 edition) (IVSC, 2017) and the European Valuation Standards (EVS, 2020 edition) (TEGoVA, 2020), adapted to the specific context of the Romanian market (see section 1.2).

In order to select the performance measures for GLM and ANN, we provide a review of several studies on this subject matter and other useful details (the context of the analysis and the conclusion on the potential superiority of ANN). As a general idea, the models' performance is observed by calculating a benchmark for fair market value, which is afterwards compared with the real prices of the transactions. Attention must be paid to the model which is closest to those prices or which provided the fewest errors. The bulk of studies analysed residential, single-family properties. The presentation of the real estate attributes used in the studies was adjusted in Table 2 in order to assure comparability for further analysis. The property type variable covers slightly different contents, such as basic apartment or one with some improvements, detached or semi-detached, purely residential or combined with commercial premises. There are studies that included the temporal variable, namely the selling date as control variable or adjusted selling price. The resulting variables in the table could be different to the ones listed in the studies due to categories used inside the same variable or due to the multiple descriptions of a variable, not detailed in our table. Finally, Table 2 provides a selection of the studies in order to observe at least two papers for several timeframes of publication, due to ANN methodology evolution and hence some temporal differences in terms of performance. Also, the table provides evidence for economically developed as well as emergent contexts, observe if there are differences in terms of models application results, and analyse the type of variables used, not only the physical and locational ones, but also qualitative (neighbourhoods).

In Table 2, a variety of items can be observed for the type of explanatory variables (predictors) included in the models. These emerged from different analysis criteria (concerning the transactions and assets) of the valuated property and its comparables selected from the market, which differ according to the practice in each jurisdiction and the nature of the property. The valuation standards (we refer mainly to IVS) do not provide an exhaustive list of real estate analysis criteria. There is instead an impressive body of literature and professional guidelines that prescribe and test these predictors, which will be listed in the following iteration. By way of example only, professional associations and researchers (*e.g.*, Appraisal Institute, 2001 or Kokinis-Graves, 2006) show that the adjustment to selling prices must be made according to the following general and specific criteria: property rights, financing terms, conditions of sale, expenditures immediately after purchase, market conditions, location, or physical

Table 2. Description, methodology and findings of the selected studies in terms of ANN versus MRA models predictive ability

Research	Sample size	Period	Location	Performance measures and models specification (designations provided by the authors of the studies)	Findings in terms of predictive ability
Tay and Ho (1992)	1,055	1989	Singapore	(1) mean absolute error (2) std. dev. of percentage error (3) percent error in excess Models compared to ANN: Traditional multiple regression analysis (MRA)	support ANN superiority
Worzalla et al. (1995)	288	1993–1994	US, Colorado	(1) mean absolute error between the predicted sales price and the current prices in the sample (2) percentage of properties whose absolute error was less than 5% of the current sales price Models compared to ANN: Hedonic pricing (multiple regression) models	do not support ANN superiority
Nguyen and Cripps (2001)	3,906	1993–1994	US, Tennesse	(1) mean absolute error (2) percentage of properties less than 5%, 15% Models compared to ANN: Multiple regression analysis (MRA)	support ANN superiority
Din et al. (2001)	285	1978–1992	Switzerland, Geneva	(1) simple comparison between predicted prices Models compared to ANN: Standard linear regression model using Geographic Information System – GIS)	difference across scenarios much more pronounced for ANN
Selim (2009)	5,741	2004	Turkey, all regions	(1) mean absolute error (2) mean squared error (3) root mean squared error Models compared to ANN: Hedonic regression model – semi-log	support ANN superiority
McCluskey et al. (2012)	2,694	2002–2004	UK, Northern Ireland	(1) R^2 (2) percentage of properties less than 10%, 15% and 20% Models compared to ANN: three regression models [an OLS regression model (baseline model) and two non-linear multiple regression models (semi-log and log-log)]	ANN out-performed by the non-linear models
Abidoye and Chan (2018)	321	2010–2016	Nigeria, Lagos	(1) R^2 (2) mean absolute error (3) root mean squared error (4) mean absolute percentage error Models compared to ANN: Hedonic pricing model	support ANN superiority
Yacim and Boshoff (2018)	3,526	–	South Africa, Cape Town	(1) root mean squared error (2) mean absolute error Models compared to ANN: three multiple regression models (linear, semi-log and log-log)	ANN out-performed by hedonic models; if a search algorithm is used to train ANN, same results as for hedonic models

attributes. There are opinions that geographic variables should also be taken into account besides typology, economic and demographic general attributes (Abraham & Hendershott, 1996; Lamont & Stein, 1999; Ghysels et al., 2007). Earlier literature offers a more complex picture of the analysis criteria.

Besides the basic criteria of judgment for accurate property valuation, such as the locational and physical attributes, it turned out that residential real estate selling price is also sensitive to another category of items. We will call this third analysis criterion the *neighbourhood*. It reflects some qualitative attributes of the property derived from the area in which the property is located and the proximity to various facilities, in other words from the characteristics of the zone. For this reason, some researchers include neighbourhood and location variables in the same category. Some examples of neighbourhood attributes of real estate we extracted as pertaining to this class are the following: according to Din et al. (2001) – quality of neighbourhood, and of location, quietness, public transport, view; according to Ibeas et al. (2012) – quality of the environment, accessibility and other local land-use attributes; according to Chiarazzo et al. (2014) – proximity from the beach, rural zone, number of bus lines and other items linked to the transport system, items linked to pollution; according to Hu et al. (2019) – education facilities, health care facilities, natural amenities, commercial facilities, public transportation.

Our research investigates the location-related physical (technical) attributes, as well as the neighbourhood (qualitative) analysis criteria for the residential real estate subject to analysis. We decided to include neighbourhood related items due to the complexity of market participants' current perception of the real estate and of the market *per se*. It seems that the real estate analysis is more precise if qualitative criteria (social, cultural, ethical, psychological, religious, demographic) are additionally considered (Kaklauskas et al., 2010). We selected from this description some social, cultural and religious items. In addition, we intend to follow the new direction of the latest studies on ANN that incorporated such influential factors, thinking of the scarcity of such studies, given that the majority of ANN research used mostly locational and physical analysis criteria. Finally, the ANN seems better suited to reflect the neighbourhood criteria than a traditional MRA (Chiarazzo et al., 2014), therefore it is worth testing this assertion.

Combining the choices observed in earlier studies (Table 2) and our own judgment, as model performance measures we will use the Mean Standard Error (MSE) between the predicted sales price and the current prices in the sample. We preferred this measure instead of the Mean Absolute Error (MAE), which is more difficult to interpret for the two tested models. Additionally, we will compute the Root Mean Squared Error (RMSE) and the Root Mean Square Percentage Error (RMSPE). It is recommended to use the second indicator since RMSE is sensitive to outliers. R^2 values (coefficient of determination) for the entire data set and raw residuals statistics will also be observed, as measures of the model quality.

Our sample contains 900 observations for Cluj-Napoca city in Romania, and for the period July – December 2019. The sample will be described in the next section.

2.2. Independent variables and the data

We are aware of the potential differences between regions (cities) and sub-markets in the city at the level of a country, considering the prices, rents and factors of influence for a variety of real estate properties. In this regard, Blackley et al. (1986) offer strong empirical evidence of the heterogeneity of interurban pricing. The existence of housing sub-markets in the same city is explained by Jones (2002) through factors such as searching costs, transaction costs, imperfect information and inelastic supply. Therefore, we controlled such housing market potential differences of the prices using standardized houses (apartments) located in one important Romanian city. The unit of the study are apartments with 1 to 4 rooms; this choice was also influenced by the availability of data since transactions with apartments are more frequent. Then, we controlled the potential differences inside the city by differentiating the sample by zone (location with a specific price area).

Taking into account earlier literature suggestions about standard housing attributes, the Romanian appraisal practice most common choices and the data availability, we selected the attributes described in Table 3, 33 in total, out of which 2 are location variables, 14 are variables regarding the physical attributes of the apartments and the other 17 variables reflect qualitative attributes related to the neighbourhood (culture, environment, and urbanism). We took into account all the physical attributes (along with the location ones) traditionally used in the formation/prediction of the housing price, correlated with their availability in our database. In addition, we included qualitative variables in the model, given the growing interest of buyers for them in recent years, as well as their insufficient use in articles that analysed ANN performance.

Some details on the variables' meanings are provided in the followings. The first category of variables is that of location. For the variable *Zone (distance from the city centre)* we used Google Maps and we took into account the shortest walking option. The variable *Zone (neighbourhoods)* is an alternative to the properties localisation inside the town, based on their appurtenance to a specific neighbourhood, in its turn located at a greater or lesser distance from downtown area. Therefore, besides downtown and semi-central zone (zones 0 and 1), we considered the others in ascending order, proportional to the distance from the centre or the proximity to the periphery.

From the second category of real estate attributes, the physical ones, *Construction type* represents the building age, denoting a new or old building. The apartments with the year of construction under 2000 were considered as belonging to an old building and those built after 2000 to a new building. This choice is in line with the housing characteristics evolution in Romania, which have begun to integrate elements of modernity such as construction materials, finishing, or design in the last three decades and especially in the last two (after the fall of the communism rule and the transition to a new approach in the construction sector). In addition, observing the selling prices dimension for different periods and talking with several valuators and real estate developers, we set 2000 as the threshold for analysis, to which prices seem to be sensitive. For the apartments that did not have the year of construction mentioned, we used the information available in Argus database where it was mentioned as a newly built or old construction or we searched for them by address to appraise directly the type of building. The degree of *Finishing* was retrieved from the Argus database, which provided information from a minimal finishing, with basic materials of walls, and with thermal insulation and electrical installations (which we titled semi-finished), reaching a complete and qualitative finishing of the walls and endowment with modern electric and sanitary installations (which we titled modern finished or ultra-finished, depending on the quality and price of the finishes).

For the majority of the attributes pertaining to the third category of variables that are neighbourhood-related, we observed the existence or the nonexistence of the items in the area. Then, for the description of the *Cultural-social level* in the area (evaluated as low, moderate, high) we took into account, for example, the position in the central area or towards the periphery (denoting economic and social status), the education level of the population (including the observed social behaviour or proximity to the best

Table 3. Residential real estate attributes

Attributes	Analysis elements within the attribute	Statistical category of the variable
Zone (distance from the city centre)	Km	Continuous
Zone (neighbourhoods)	10 zones surrounding the city centre (Downtown, Semi-central, 4 historical neighbourhoods, and 4 groups each one covering 2 smaller zones)	Categorical
Useful area	Number of square meters	Continuous
Rooms	Number of (1, 2, 3 or 4)	Categorical
Parking	Existence: Yes/No	Categorical (dummy)
Bathrooms	One, More than one	Categorical
Balcony	Without, One, More than one	Categorical
Finishing	Semi-finished, Finished, Modern finished, Ultra-finished	Categorical
Partitioning	Not detached, Semi-detached, Detached	Categorical
Construction type	Old or new building	Categorical (dummy)
Elevator	Existence: Yes/No	Categorical (dummy)
Insulated windows	Existence: Yes/No	Categorical (dummy)
Metal door	Existence: Yes/No	Categorical (dummy)
Own central heating	Existence: Yes/No	Categorical (dummy)
Storage room or attic	Existence: Yes/No	Categorical (dummy)
Thermally insulated	For the building: Yes/No	Categorical (dummy)
Cultural-social level	Low, moderate, high	Categorical
Environmental pollution	Low, less, high	Categorical
Urban density	Low, moderate, high, very high	Categorical
Business centres	Existence: Yes/No	Categorical (dummy)
Farmers' markets	Existence: Yes/No	Categorical (dummy)
Financial and banking institutions	Existence: Yes/No	Categorical (dummy)
Green area	Existence: Yes/No	Categorical (dummy)
Health institutions (hospitals, clinics, pharmacies)	Existence: Yes/No	Categorical (dummy)
Hotels, resorts	Existence: Yes/No	Categorical (dummy)
Hypermarkets	Existence: Yes/No	Categorical (dummy)
Religious institutions	Existence: Yes/No	Categorical (dummy)
Relaxing places (restaurants, clubs, pubs)	Existence: Yes/No	Categorical (dummy)
Schools, kindergartens	Existence: Yes/No	Categorical (dummy)
Shopping centres	Existence: Yes/No	Categorical (dummy)
Sport centres	Existence: Yes/No	Categorical (dummy)
Transportation lines	Existence: Yes/No	Categorical (dummy)
Universities	Existence: Yes/No	Categorical (dummy)

listed kindergartens and schools) or degree of cleanliness observed. For the variable *Environmental pollution,* the apartments located in the city centre and in areas without parks were considered as belonging to a highly polluted area. The apartments that have parks in the area, are less circulated were considered as belonging to an area with a moderate level of pollution, and the apartments built on the outskirts of the city with lots of greenery (in principle these details were mentioned in the Argus database) were considered the least polluted. To assess *Urban density* in the area we used Google Maps and, based on the address, we grouped the properties in four categories as follows: very high (usually, the buildings in the centre of the city), high (old buildings with many apartments such as the Mehedinti area), moderate (new constructions, more peripheral) and low concentration (peripheral area of the city, as the area Colina).

The sample technique used was the non-probability sampling, the data set being collected from a platform dedicated to real estate professionals, Argus Property Resources[1]. Argus collects data on real estate transactions

[1] https://www.mediapress.ro/argus-reo-imobiliare.php

that take place throughout the country and allows filtering of the elements of interest on various types of properties. We used this source of information to collect the location, physical and partially neighbourhood variables of the apartments offered for sale in the Cluj-Napoca city area. We checked our sample through interviews with real estate agents who made public the sales offers to assure data reliability.

Cluj-Napoca is located in the northwest part of Romania, in Transylvania province, and has over 300,000 inhabitants. We argue the choice of Cluj-Napoca city given its importance in Romania as economic development and hence dynamism of the market and relevance of the real estate transactions. Cluj-Napoca is the country's second largest city by population at the time (2011 Census), after the country's capital (Bucharest and its adjacent area), and the third as volume of real estate transactions in 2019, respectively the second in the first semester of 2020, after Bucharest, according to the data registered in ANCPI – the National Agency for Cadastre and Real Estate Advertising[2]. It is also a university city and an IT hub, with a high degree of development and attractiveness for people who want to settle here and find a home.

We chose a very short period of time for the sales, being aware of the time sensitivity of real estate market and in order not to have to adjust selling prices, to inflation for example, or to include the date of sale (offer) in the model. As Fabozzi et al. (2010) show, participants' expectations regarding price on the local real estate market are strongly influenced by the most recent series of prices. The database covers the second half of 2019 and contains transactions (offer) prices designed in our research as selling prices[3]. Initially the sample included a number of 927 apartments. Due to extravagant or too low prices, the extreme values were eliminated and the database contains 900 subjects in the final version.

2.3. Descriptive statistics and other statistical tests

The Table 4 presents our results in terms of values or frequencies of the variables, after processing the database we created.

Table 4. Descriptive statistics for the variables

PANEL A Continuous variables				
Variables	Mean	Min.	Max.	Std. Dev.
Dependent				
Selling price (Euro)	95,952	16,500	247,000	38,887
Independent (Predictors)				
Zone (distance from the city centre)	3.7	0.0	10.0	1.8
Useful area	57.9	11.0	197.0	23.8
Rooms[a]	2.4	1.0	4.0	0.9

PANEL B Categorical variables			
Independent (Predictors)	Predictor typology	Frequency (%)	Mean price per predictor type (Euro)
Zone (neighbourhoods)	Zone 0 (Downtown)	3.58	116,822
	Zone 1 (Semi-central)	2.68	97,708
	Zone 2 (Bună ziua+Andrei Mureșanu)	5.14	111,626
	Zone 3 (Grigorescu)	5.03	103,160
	Zone 4 (Gheorgheni)	12.51	104,909
	Zone 5 (Mărăști)	22.01	92,229
	Zone 6 (Mănăștur)	27.71	90,542
	Zone 7 (Zorilor+Europa)	9.39	109,966
	Zone 8 (Gruia+Gară+Dâmbul rotund)	4.02	98,732
	Zone 9 (Iris+Someșeni)	7.93	66,587
Parking	Yes	64.0	113,248
	No	36.0	86,223
Bathrooms[b]	One	76.6	86,654
	More than one	23.4	126,314
Balcony	Without	43.0	85,731
	One	44.0	98,357
	More than one	13.0	121,621

[2] http://ancpi.ro/index.php/presa-3/statistici

[3] during this period, the prices remained relatively stable on the market.

Continue of Table 4

Independent (Predictors)	Predictor typology	Frequency (%)	Mean price per predictor type (Euro)
Finishing[c]	Semi-finished	6.8	86,168
	Finished	60.2	95,629
	Modern finished	26.3	93,521
	Ultra-finished	6.7	118,427
Partitioning[b]	Not detached	2.9	61,219
	Semi-detached	25.2	94,677
	Detached	71.9	97,795
Construction type	Old	74.7	93,479
	New	25.3	103,242
Elevator	Yes	12.7	100,481
	No	87.3	95,295
Insulated windows	Yes	75.4	92,124
	No	24.6	107,713
Metal door[a]	Yes	47.4	79,625
	No	52.6	110,692
Own central heating	Yes	61.7	99,897
	No	38.3	89,607
Storage room or attic	Yes	30.4	112,258
	No	69.6	88,815
Thermally insulated[b]	Yes	43.6	99,318
	No	56.4	93,355
Cultural-social level	Low	34.9	80,381
	Moderate	24.3	100,710
	High	40.8	106,436
Environmental pollution[b]	Low	6.3	62,343
	Moderate	25.9	88,747
	High	67.8	101,845
Urban density[b]	Low	1.6	105,350
	Moderate	26.1	93,625
	High	68.9	97,042
	Very high	3.4	87,561
Business centres[b]	Yes	20.3	95,840
	No	79.7	95,981
Farmers' markets	Yes	35.8	89,242
	No	64.2	99,690
Financial and banking institutions	Yes	70.7	101,766
	No	29.3	81,947
Green area	Yes	61.0	93,548
	No	39.0	99,713
Health institutions (hospitals, clinics, pharmacies)	Yes	67.6	104,601
	No	32.4	77,944
Hotels, resorts[b]	Yes	77.2	96,072
	No	22.8	95,545
Hypermarkets	Yes	92.6	96,840
	No	7.4	84,916
Religious institutions[b]	Yes	77.9	94,950
	No	22.1	99,484

End of Table 4

Independent (Predictors)	Predictor typology	Frequency (%)	Mean price per predictor type (Euro)
Relaxing places (restaurants, clubs, pubs)	Yes	86.9	100,398
	No	13.1	66,491
Schools, kindergartens[b]	Yes	79.7	94,947
	No	20.3	99,890
Shopping centres[b]	Yes	22.3	99,892
	No	77.7	94,819
Sport centres[b]	Yes	55.3	94,874
	No	44.7	97,287
Transportation lines	Yes	94.1	98,864
	No	5.9	49,417
Universities	Yes	47.6	102,450
	No	52.4	90,060

Notes: [a] the variable metal doors was eliminated from the predictors, being correlated with insulated windows (*phi* coefficient = 0.480, *p-value* = 0.000); [b] variables that will be eliminated for the GLM and ANN model processing due to their lack of statistical significance according with *F-test* from ANOVA; [c] in Argus database we observed the following states and numbers for apartments finishings: unfinished (16 apartments), semi-finished (47), finished (526), with classic finishing (19), modern (239) and ultramodern (61). Due to the uneven number of finishing types we applied ANOVA test to check the impact of the imbalance in the number of finishing types in the sample and noticed that a regrouping or removal from the sample of subjects with classic finish, respectively of unfinished apartments was needed. Our choice was a regrouping, thus the group of unfinished apartments is included in semi-finished group. Another group contains the finished apartments and the classic finished ones; the apartments with modern and ultramodern finishing remained unchanged as number of observations. Finally, we created four stages of finishing, as shown in Table 3.

Panel A of Table 4 indicates, for the dependent variable Selling price, a mean price of 95,952 Euro, and a standard deviation of 38,887 Euro, for an average useful area of 57.9 square meters. In addition, our analysis reveals that this variable follows a distribution that tends towards the normal distribution. Next, we will identify the factors that explain the variation of this variable using the two proposed models, GLM and ANN. The data was processed using SPSS, version 22.

We performed a comparative analysis between several GLM specifications, under some common assumptions regarding price distribution, such as normal and Gamma distributions. As we noticed a strong association between the useful area and the number of rooms, as characteristics of the apartments' size, we considered the following four specifications for the hedonic model: M1– normal distribution with identity link for prices (this is the traditional linear model) with the variable area, M2 – normal distribution with identity link for prices with the variable rooms, M3 – Gamma distribution with log link with the variable area and M4 – Gamma distribution with log link with the variable rooms. The statistical results are provided in the Table 5.

Table 5. Regression model choice

Regression models indicators	M1	M2	M3	M4
R^2	0.7485	0.6668	0.5880	0.6633
Likelihood Ratio Chi-Square	1,242.4	989.4	1,011.2	1,149.1
MAE	13,933.4	16,753.8	14,822.5	16,642.2

For the choice of the regression model to use in our comparison with ANN, we looked at the indicators' values. The Likelihood Ratio Chi-Square, R^2 and MAE suggest that the M1 performs better in terms of goodness-of-fit and prediction accuracy. Consequently, for this model, hereinafter referred to as GLM, the estimated parameters of the predictors will be presented.

3. Results and discussion

3.1. Results for the Generalized Linear Model (GLM)

We applied several statistical measures which converged in confirming the accuracy of the model's effectiveness. For example, the Omnibus test indicates a value of 1,242 for Likelihood Ratio Chi-Square, *p-value* = 0.000, and this justified proceeding with the model variation investigation.

Of the variables whose values were collected in the database (see Table 3) only the variables shown in Table 4 were kept for running in GLM (and also in ANN). These denote those predictors that are significant in explaining selling price variation, the dependent variable, as ANOVA and correlation tests revealed.

The test of model effects revealed the statistical significance of the predictors. Hence, untabulated statistics indicate 11 of the 22 variables that qualified in the previous step as being significant, for a *p-value* <0.1. The regression results for the variables of the GLM that enable us to comment and rank the significant ones are presented in Table 6.

Table 6. Regression results and variables significance according to GLM

Variable	Coefficient (B)	Wald Chi-Square[b]	Ranking of the variables relative importance[c]
Constant	28,323.05	115.09	
Zone (neighbourhoods)[a] – Downtown		32.88	III
Zone 1	–764.51		
Zone 2	–5,257.10		
Zone 3	–4,680.16		
Zone 4	2,590.18		
Zone 5	–1,631.51		
Zone 6	–8,541.76**		
Zone 7	–2,272.53		
Zone 8	–8,833.58*		
Zone 9	–11,335.29**		
Zone (distance from the city centre)	–2,930.23***	24.33	V
Useful area	1,033.00***	659.38	I
Parking[a] – No			IV
Yes	7,645.47***	26.18	
Bathrooms[a] – One			
More than one	2,638.10	2.03	
Balcony[a] – Without		17.35	VI
One	6,090.26***		
More than one	2,626.62		
Finishing[a] – Semi-finished		43.03	II
Finished	5,510.84*		
Modern finished	10,733.70***		
Ultra-finished	20,385.40***		
Partitioning[a] – Not detached		1.25	
Semi-detached	4,768.16		
Detached	4,453.06		
Construction type[a] – Old building		0.068	
New building	–471.57		
Insulated windows[a] – No			
Yes	–2,463.64	1.79	
Own central heating[a] – No		0.628	
Yes	–1,260.90		
Storage room or attic[a] – No			VII
Yes	4,228.56***	7.48	
Thermally insulated[a] – No			
Yes	2,066.35	1.90	
Cultural-social level[a] – Low		1.70	
Moderate	–1,425.75		
High	1,140.99		
Farmers' markets[a] – No			
Yes	–2.326.88	2.35	
Financial and banking institutions[a] – No			
Yes	1,429.20	0.63	
Green area[a] – No			VIII
Yes	–4,095.49***	7.39	
Health institutions[a] – No			
Yes	2,882.27	2.65	

End of Table 6

Variable	Coefficient (B)	Wald Chi-Square[b]	Ranking of the variables relative importance[c]
Hypermarkets[a] – No			X
Yes	–5,474.49**	3.90	
Relaxing places[a] – No			XI
Yes	4,072.53*	2.88	
Transportation lines[a] – No			IX
Yes	8,048.46**	6.43	
Universities[a] – No			
Yes	–827.35	0.29	
$R^2 = 0.7485$[d]			
F-statistic = 74.44			
$F(p\text{-}value) = 0.000$			

Notes: [a] reference value; ***,**,* significant at 1%, 5%, respectively 10%; [b] according to joint Wald Chi-Square in the Test of model effects; [c] only the variables that proved statistical significance were ranked, knowing that for GLM the variables that are not statistically significant have no effect on selling price; [d] in the framework of maximum likelihood method, the coefficient $R^2 = 1 - D_1/D_0$, where D_1 is the residual deviance for the estimated model, and D_0 residual deviance for the model including only the intercept, measures the proportion of the null deviance accounted for by the model (Fox, 2008); the value of R^2 suggest the goodness-of-fit for the model.

The coefficient signs for the GLM predictors that proved to be statistically significant according to Table 6 are, in general, consistent with our expectations. A negative influence on selling prices is observed for other areas than the downtown, with one exception, a neighbourhood in full development, in high demand during the analysed period. Therefore, a greater distance from the city centre decreases the price. A positive influence is caused by the rest of the variables, as follows: the useful area is directly related to the price; an apartment with balcony is favourably appreciated, one balcony being enough, even if several balconies have a positive influence; the degree of finishing obviously influences the selling price positively, to an increasing extent from lower to upper; the existence of a parking space, of a storage room or neighbourhood characteristics as relaxing places and transportation lines are important for the real estate transactions. Contrary to our expectations, the existence of green area and hypermarkets nearby does not cause a positive influence on the selling price. Some explanations may be related to the existence of sufficient green spaces and hypermarkets all around the town, so that these items affect the price, the larger congestion induced by the proximity of a hypermarket could penalize the price, or that the green spaces are found especially in the marginal areas of the city, where prices are lower.

3.2. Results for the Artificial Neural Network (ANN)

We used a feed forward/backward propagation neural network software package, inside SPSS version 22, Multilayer Perceptron in order to construct ANN model. ANN architecture is the following: one hidden layer, activation function-hyperbolic tangent and output layer-activation function-identity.

To ensure a good out-of-sample generalization performance, we used a cross validation technique to choose the best network structure. According to the literature that prescribe a split into 2 or 3 subsamples, our database of 900 selling prices was divided into 2 subsamples, *i.e.* a training set of 71.6% and a testing set of 28.4% of the transactions. This is the best result offered by the network algorithm that looked for different divisions of data in 2 subsamples around the partition we established, *i.e.* 70/30%.

The training set is very well sized, larger than prescribed by earlier studies (*e.g.* Nguyen & Cripps, 2001 pleaded for a percentage of 13–39% for the relevance of the ANN results). We are in line with Curry et al. (2002), who worked with a training set of 75% in order to combat the risk of overfitting.

In our ANN analysis, we have 22 input nodes (2 continuous and 20 categorical variables), as this is the number of the predictors left after application of ANOVA and correlation tests, 1 hidden layer with 8 nodes and 1 output node, the estimated selling price. Because there is no theoretical ground for the number of hidden layers (Tay & Ho, 1992), we finally used 1 layer, bearing in mind not to provoke an overfitting of the network and also knowing that normalized importance of the variables is viewed as relevant when one hidden layer is used. Previously, we ran the model for two hidden layers, but the results were inferior to those obtained by using a single hidden layer. The number of nodes, 8, in the hidden layer was automatically provided by the backpropagation ANN model, which aimed to obtain the best architecture, as the error decreases in relation to the number of iterations of the program, with different learning rates and momentum as default parameters. The number of nodes depends on the number of variables included in the model (inputs), which in our case represents 22 (n)

and which generates 45 as a maximum ($2n+1$), according to Tay and Ho (1992). However, we stopped at a number of 8 hidden nodes, aiming to obtain a model with the fewest possible nodes, which would increase the network training speed and which would at the same time indicate a good fit for out-of-sample price behavior (Din et al., 2001). As type of functions, we tested several activation functions and we finally used a hyperbolic one for the hidden layer, which provided the best results.

For the inputs and outputs, we calculated the normalized values using the minimum and maximum of their values (within the range 0 to 1), to which we added one bias node. Hence, the independent variable importance on the selling price variation is presented in Table 7, being useful – if ANN model has one hidden layer – for a comparison of the variables' significance and hence their salience, previously obtained with GLM.

The data was run until we obtained the smallest relative error for the 2 samples, of 0.18, and 0.24 respectively. At the same time, we chose these values to be as close as possible. Another summary for the estimation errors of applying ANN to the 2 sets of data is the following: sum of squares error of 57.48 for the training set, and of 30.11 for the testing set.

In order to generate a model after running the best solution for ANN, trying to imitate GLM, which provides coefficients useful to other samples, we obtained a matrix with the connections between the input and hidden layers, respectively the hidden and output layers. This is the "optimal combination", meaning that after the network generated randomly initial connection weights, it self-adjusted the next weights it produced, learning from the estimation errors committed after each iteration. The optimal matrix is shown in Table 8 and presents the correlations between the input layers, hidden layers and the output as connection weights.

Table 7. Variables significance according to ANN

Variable	Normalized importance (%)	Ranking of the variables according to the normalized importance
Zone (neighbourhoods)	13.0	IV
Zone (distance from the city centre)	28.1	II
Useful area	100.0	I
Parking	4.2	XIV
Bathrooms	4.1	XVI
Balco	8.6	V
Finishing	18.4	III
Partitioning	8.3	VI
Construction type	5.8	VIII
Insulated windows	3.8	XVIII
Own central heating	4.3	XIV
Storage room or attic	4.8	XII
Thermally insulated	5.0	XI
Cultural-social level	6.7	VII
Farmers' markets	4.2	XV
Financial and banking institutions	3.0	XXI
Green area	3.2	XX
Health institutions (hospitals, clinics, pharmacies)	4.4	XIII
Hypermarkets	3.5	XIX
Relaxing places (restaurants, clubs, pubs)	3.9	XVII
Transportation lines	8.3	VI
Universities	5.3	X

Table 8. ANN connection weights between layers

PANEL A Weights between inputs layer and hidden layer nodes								
Inputs layer	Hidden layer nodes							
	1	2	3	4	5	6	7	8
(Bias)	–0.007	–0.065	–0.118	0.429	0.180	–0.322	0.030	–0.266
Zone (neighbourhoods)								
Zone 0	–0.382	–0.170	–0.079	–0.026	–0.302	–0.175	0.441	–0.423
Zone 1	0.439	0.198	–0.080	0.024	–0.098	–0.467	–0.470	0.177
Zone 2	–0.053	–0.406	0.189	–0.399	0.429	0.127	0.478	0.471
Zone 3	–0.014	0.367	0.330	–0.281	0.304	–0.172	–0.149	0.364
Zone 4	0.077	0.212	0.007	–0.167	0.124	–0.506	–0.374	0.473
Zone 5	–0.059	0.346	0.312	–0.051	–0.323	–0.052	–0.423	–0.212
Zone 6	–0.209	–0.043	0.116	0.247	–0.127	0.591	–0.247	0.284
Zone 7	–0.398	–0.416	0.258	0.286	0.208	0.095	0.283	0.140
Zone 8	0.435	–0.365	0.128	0.080	0.026	0.068	0.205	0.310
Zone 9	0.269	–0.185	0.052	–0.067	0.220	0.039	0.047	–0.342

Continue of Table 8

Inputs layer	Hidden layer nodes							
	1	2	3	4	5	6	7	8
Useful area	0.091	0.210	1.063	–0.097	–0.867	–0.345	0.129	0.810
Parking								
No	0.528	–0.067	0.508	–0.113	–0.146	0.043	0.373	0.268
Yes	0.122	–0.321	0.334	–0.087	–0.069	–0.301	0.517	0.297
Bathrooms								
One	0.450	0.295	–0.008	–0.312	–0.387	–0.120	0.408	–0.464
More than one	0.443	0.221	–0.155	–0.497	0.242	–0.325	–0.264	0.062
Balcony								
Without	–0.326	0.179	0.259	–0.111	0.270	0.203	–0.073	0.137
One	0.032	–0.468	0.119	0.010	0.050	–0.391	0.259	–0.216
More than one	–0.419	0.254	0.259	–0.253	–0.448	0.232	0.085	–0.464
Finishing								
Semi-finished	0.209	0.071	–0.608	0.505	–0.095	0.346	0.199	0.098
Finished	–0.369	0.374	–0.618	0.663	0.046	–0.087	–0.106	0.151
Modern finished	0.550	–0.080	–0.361	–0.370	–0.380	0.231	0.415	–0.106
Ultra-finished	–0.334	–0.441	–0.186	0.373	–0.321	–0.074	–0.414	0.461
Partitioning								
Not detached	0.540	–0.213	–0.165	–0.063	–0.306	0.006	–0.071	0.070
Semi-detached	–0.270	0.087	0.437	–0.267	–0.061	–0.206	–0.486	–0.330
Detached	0.071	–0.275	0.116	–0.124	–0.425	0.239	–0.162	0.197
Construction type								
Old	–0.197	–0.146	0.117	0.170	–0.390	0.512	0.009	0.306
New	0.010	0.249	0.356	–0.144	0.364	–0.296	0.419	–0.020
Insulated windows								
No	–0.374	0.417	–0.259	0.287	0.157	0.173	–0.425	0.625
Yes	0.147	0.090	–0.447	0.486	–0.476	0.270	–0.023	–0.052
Own central heating								
No	0.162	–0.324	0.091	–0.413	–0.193	–0.139	0.273	–0.297
Yes	0.271	–0.129	0.599	–0.219	0.126	0.094	0.187	–0.481
Storage room or attic								
No	–0.038	–0.128	0.373	0.646	0.192	0.256	–0.392	–0.518
Yes	–0.351	–0.431	–0.096	0.113	0.493	–0.112	0.156	–0.037
Thermally insulated								
No	0.019	0.030	0.378	0.345	0.257	–0.394	0.359	0.245
Yes	0.156	0.226	0.131	0.551	–0.588	–0.320	0.438	0.488
Cultural-social level of the area								
Low	0.470	–0.477	–0.116	–0.060	0.289	–0.344	0.179	0.404
Moderate	–0.238	0.505	0.300	–0.152	–0.339	0.259	0.122	0.108
High	–0.382	–0.137	0.001	0.077	0.387	–0.543	0.309	–0.174
Farmers' markets								
No	0.384	0.233	–0.021	0.489	–0.200	0.496	–0.260	–0.056
Yes	0.179	–0.306	–0.686	0.494	–0.048	0.130	0.368	0.440
Financial and banking institutions								
No	–0.054	0.456	–0.268	0.074	0.240	–0.085	0.453	0.371
Yes	–0.155	–0.203	–0.181	0.024	0.438	0.216	0.366	0.317
Green area								
No	0.164	–0.083	–0.154	0.063	–0.072	0.223	–0.378	–0.592
Yes	–0.202	0.340	–0.260	0.071	–0.175	0.543	–0.003	–0.098

End of Table 8

Inputs layer	Hidden layer nodes							
	1	2	3	4	5	6	7	8
No	0.465	–0.406	–0.188	–0.399	–0.145	0.769	0.538	0.144
Yes	–0.469	0.366	–0.364	–0.142	0.056	0.235	–0.086	0.517
Hypermarkets								
No	0.055	–0.208	–0.324	0.108	0.293	–0.229	0.070	0.375
Yes	0.143	–0.204	–0.338	0.146	0.285	–0.163	–0.161	–0.241
Relaxing places								
No	–0.075	–0.382	–0.521	–0.143	0.352	0.031	0.263	0.424
Yes	–0.356	0.200	–0.021	–0.422	0.068	0.413	0.245	–0.404
Transportation lines								
No	0.417	0.170	0.474	–0.104	–0.369	0.342	0.132	0.250
Yes	–0.228	–0.458	0.488	0.441	–0.403	–0.179	0.438	0.126
Universities								
No	0.442	–0.184	–0.326	–0.185	0.060	–0.161	0.487	–0.150
Yes	–0.282	0.320	0.369	0.211	0.073	0.300	–0.135	–0.494

PANEL B Weights between hidden layer nodes and output layer	
Hidden layers	Output
Bias	0.622
1	–0.261
2	–0.349
3	0.637
4	–0.521
5	–0.525
6	–0.606
7	–0.131
8	0.246

These final weights optimally emulate the valuation function for the sample we used, but in a black-box model. However, the matrix could be associated with a software and generate a solution to run other tests with new samples.

To test the stability of the model we modified the sample structure on the subsamples representing the training set and the testing set, for the variables kept in the model (30 iterations with random extraction) and for the total of the initial variables (20 iterations). For each iteration, we tracked relative error and manipulated the number of units contained in the hidden layer, keeping the type of the initial function, hyperbolic tangent. If we refer only to the tests we applied on the significant variables kept in our model, we can say that the relative error ranged from 0.178 to 0.243 for the training sample and from 0.210 to 0.315 for the testing sample. The number of units in the hidden layer varied from 6 to 13. We concluded that the predictive power of ANN did not change much, in other words, we did not get better results than initially reported, thus demonstrating the stability of the model.

3.3. Discussion of empirical results

The results for GLM and the optimal version of the applied algorithm for ANN are presented in Figure 1. A first observation is that the ANN model has better prediction power as the points wrap better on the diagonal of the graph.

The verifiable arguments are listed in Table 9.

MSE suggests that ANN provides a slightly better agreement between the predicted and the current sale prices in the sample than GLM. The results suggested by RMSE and RMSPE also prove that ANN is better, as it causes fewer errors that GLM.

R^2 indicates a slightly better quality of ANN. Also, the raw residuals, standardized, as the difference between real selling prices and the predictive ones confirm, as mean values, the superiority of ANN over GLM, even if the distance between the models is not extremely large.

In order to provide a better comparability between GLM and ANN, completing the results provided by the performance measures, we performed a comparison in terms of variables salience. As a reference for ranking, we

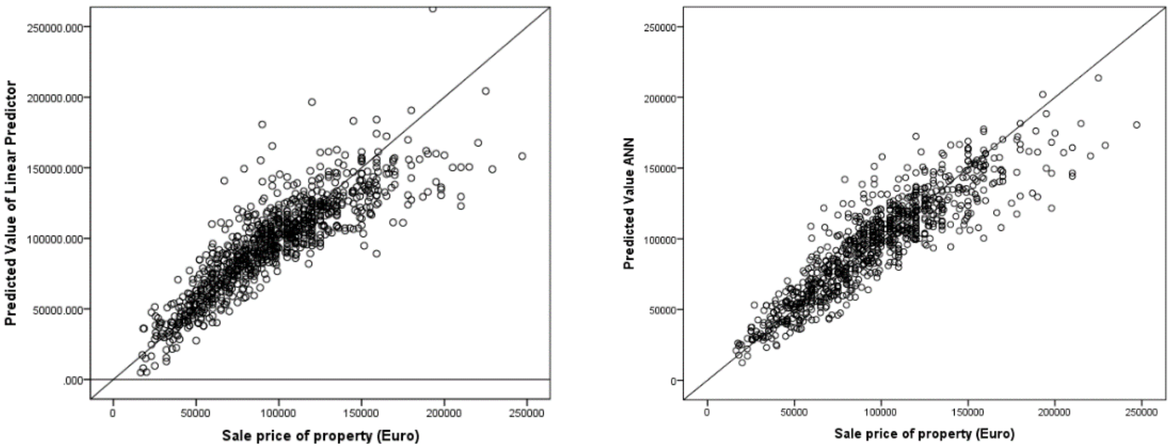

Figure 1. GLM and ANN prediction compared to observed selling price

Table 9. Measures for GLM and ANN performance

Performance measures	GLM	ANN
Mean Standard Error (MSE)	650.01	572.28
Root Mean Squared Error (RMSE)	19,489.47	17,162.74
Root Mean Square Percentage Error (RMSPE)	2.2271	1.8844
R^2	0.7485	0.8051
Mean raw residuals	–3.4954	–2.2697

used Wald Chi-square for GLM and normalized importance for ANN. The comparative analysis between the two models reveals that the first four places as impact on selling price are identical, namely the variables useful area, zone in its two declinations and finishing (the two location variables and other two physical). Then, for GLM, are listed the other physical variables, the neighbourhood ones being all confined to the second set of variables as importance. For ANN, the rest of the variables, apart from the first six, are a mix between physical and neighbourhood variables.

Conclusions

According to our research objective, this section contains our findings for the comparison between two models for real estate selling price prediction, GLM and ANN, and the implications of these results. For this research, we used a sample of 900 apartments from Cluj-Napoca, Romania, containing selling transactions for the second semester of 2019, and data for 33 locational, physical and neighbourhood related attributes (socio-cultural, environmental, and urbanism related). Our research question was if ANN is superior to GLM in predicting selling price in terms of transparency, predictive ability and stability of the results.

Thus, firstly, according to the specific literature and our own judgements, five performance measures were used in order to observe which model is better. All these measures indicated ANN superiority, our results being in line with other studies reviewed and synthesized in Table 2.

Secondly, we report the results of predictors' importance for GLM, as well as for ANN, based on the ranking of variables. We conclude that the two models also rank very closely from the perspective of illustrating the significance of the different attributes of real estate, and that ANN enable the knowledge of these details, if appropriate statistical indicators are used.

Now, if we compare the classical hedonic regression and the neural network program, there is a rich literature that highlighted the strengths and weaknesses of each one (see Table 1). In short, the traditional model of MRA we used, GLM, is viewed as presenting a good quality of estimators that are used and which can be statistically validated and easily interpreted, to the advantage of the professional environment in terms of accessibility. In the case of ANN, the predictions are better but the interpretation of the weights assigned to the input variables is more difficult to understand. In this research, we filter the two models according to the criteria we established, based on Table 1 review, namely transparency, predictive ability and stability of the results, based on our empirical results. Aside from the results *per se*, we sustained the occurrence of errors in predicting the selling price, and that it implicitly denotes the prediction ability. We therefore assert that the ANN model provides a good stability in terms of results, the last of our analysis criteria. It is to mention here that in the case of results stability a comparison with GLM is not relevant as the regression methodology does not allow predictions for various iterations on the sample. Therefore, ANN checks two of the three criteria compared to GLM. The transparency of the model was the main critique concerning neural networks applications. These are black-boxes that do not allow the observation of the whole process of selling prices estimation, especially its finality, the optimal model obtained. However, even the lack of ANN transparency can be counteracted if appropriate measures are used. Hence, we used the normalized importance to rank the variables, and that is equivalent to observing each real estate characteristic impact on the selling price (even if not the direction of the influence), as the coefficients in a hedonic regression show.

Apparently, using an ANN program, setting the parameters, applying specific tests are skilled tasks in statistical terms, and require a general substantial effort of implementing ANN. However, if statistical experts handle ANN, it can be viewed as a valuable tool for valuators who review other professionals' valuation reports. This is a contribution to the valuation practice and professionals, in Romania and elsewhere.

We believe that our study brings more input about the predictive power of ANN by integrating a consistent number of qualitative variables (referred to as neighbourhood related attributes) into the model. Even if they did not prove to be the prime influencers of selling prices, the qualitative variables enriched the quality of the ANN model. We have proved that this type of variables can be introduced in the model and generate a good prediction. In addition, we have shown that, regardless of the type of economic context, emerging or not, the functionality of the real estate market does not differ much, since ANN proved superior to the classical hedonic model, even in the emerging context studied by us. This is a contribution to the specific literature and ANN methodology application.

Our results should be interpreted in the light of certain limitations, which can encourage further developments. A possible limitation is the subjective nature of the real estate qualitative characteristics, which induced certain choices in their measurement. Other researchers may have other solutions. Another limitation is related to the studied context, a single country/city, this narrowing the generalization of our findings (for example, ANN weights), even if we are aware of the market volatility (in time and space) in properties' valuation.

References

Abidoye, R. B., & Chan, A. P. C. (2017). Artificial neural network in property valuation: application framework and research trend. *Property Management*, *35*(5), 554–571. https://doi.org/10.1108/PM-06-2016-0027

Abidoye, R. B., & Chan, A. P. C. (2018). Improving property valuation accuracy: a comparison of hedonic pricing model and artificial neural network. *Pacific Rim Property Research Journal*, *24*, 71–83. https://doi.org/10.1080/14445921.2018.1436306

Abraham, J. M., & Hendershott, P. H. (1996). Bubbles in metropolitan housing markets. *Journal of Housing Research*, *7*(2), 191–207. https://doi.org/10.3386/w4774

Allen, W. C., & Zumwalt, J. K. (1994). *Neural networks: a word of caution* (Working Paper). Colorado State University.

American Society of Appraisers. (2004). *Valuing machinery and equipment*. ANEVAR.

Appraisal Institute. (2001). *The appraisal of real estate* (12th ed.). Chicago.

Association of the Romanian Valuators. (2020). *Standards for the goods valuation (SEV)*. ANEVAR. http://site2.anevar.ro/sites/default/files/page-files/standarde_2020_dupa_cn_27_iulie_final_31.07.2020.pdf

Benjamin, J. D., Guttery, R. S., & Sirmans, C. F. (2004). Mass Appraisal: an introduction to multiple regression analysis for real estate valuation. *Journal of Real Estate Practice and Education*, *7*(1), 65–78. https://doi.org/10.1080/10835547.2004.12091602

Blackley, D. M., Follain, J. R., & Lee, H. (1986). An evaluation of hedonic price indexes for thirty-four large SMSAs. *American Real Estate and Urban Economics Association Journal*, *14*(2), 179–205. https://doi.org/10.1111/1540-6229.00382

Bogin, A. N., & Shui, J. (2020). Appraisal accuracy and automated valuation models in rural areas. *The Journal of Real Estate Finance and Economics*, *60*, 40–52. https://doi.org/10.1007/s11146-019-09712-0

Borst, R. A. (1991). Artificial neural networks: the next modelling/calibration technology for the assessment community. *Property Tax Journal*, *10*(1), 69–94.

Brunson, A., Buttimer, R. J., & Rutherford, R. (1994). *Neural networks, nonlinear specifications, and industrial property values* (Working paper series). University of Texas at Arlington, Arlington, TX.

Cechin, A., Souto, A., & Gonzales, A. M. (2000). Real estate value at Porto Alegre city using artificial neural networks. In *Proceedings. Vol. 1. Sixth Brazilian Symposium on Neural Networks* (pp. 237–242). IEEE. https://doi.org/10.1109/SBRN.2000.889745

Chiarazzo, V., Caggiani, L., Marinelli, M., & Ottomanelli, M. (2014). A neural network based model for real estate price estimation considering environmental quality of property location. *Transportation Research Procedia*, *3*, 810–817. https://doi.org/10.1016/j.trpro.2014.10.067

Curry, B., Morgan, P., & Silver, M. (2002). Neural networks and non-linear statistical methods: an application to the modelling of price–quality relationships. *Computers & Operations Research*, *29*(8), 951–969. https://doi.org/10.1016/S0305-0548(00)00096-4

Din, A., Hoesli, M., & Bender, A. (2001). Environmental variables and real estate prices. *Urban Studies*, *38*(11), 1989–2000. https://doi.org/10.1080/00420980120080899

DiPasquale, D., & Wheaton, W. C. (1994). Housing market dynamics and the future of housing prices. *Journal of Urban Economics, 35*(1), 1–27. https://doi.org/10.1006/juec.1994.1001

Do, A. Q., & Grudnitski, G. (1992). A neural network approach to residential property appraisal. *The Real Estate Appraiser*, *58*(3), 38–45.

Evans, A., James, H., & Collins, A. (1992). Artificial neural networks: an application to residential valuation in the UK. *Journal of Property Valuation and Investment*, *11*, 195–203.

Fabozzi, F. J., Shiller, R. J., & Tunaru, R. S. (2010). Property derivatives for managing European real-estate risk. *European Financial Management*, *16*(1), 8–26. https://doi.org/10.1111/j.1468-036X.2009.00528.x

Fox, G. (2008). *Applied regression analysis and generalized linear models*. Sage Publications.

Ghysels, E., Plazzi, A., & Valkanov, R. (2007). Valuation in US commercial real estate. *European Financial Management*, *13*(3), 472–497.
https://doi.org/10.1111/j.1468-036X.2007.00369.x

Glumac, B., & Des Rosiers, F. (2021a). Practice briefing – Automated valuation models (AVMs): their role, their advantages and their limitations. *Journal of Property Investments & Finance*, *39*(5), 481–491.
https://doi.org/10.1108/JPIF-07-2020-0086

Glumac, B., & Des Rosiers, F. (2021b). Towards a taxonomy for real estate and land automated valuation systems. *Journal of Property Investments & Finance*, *39*(5), 450–463.
https://doi.org/10.1108/JPIF-07-2020-0087

Helbich, M., Brunauer, W., Vaz, E., & Nijkamp, P. (2014). Spatial heterogeneity in hedonic house price models: the case of Austria. *Urban Study*, *51*, 390–411.
https://doi.org/10.1177/0042098013492234

Ho, W. K. O., Tang, B. S., & Wong, S. W. (2020). Predicting property prices with machine learning algorithms. *Journal of Property Research*, *38*(1), 48–70.
https://doi.org/10.1080/09599916.2020.1832558

Hu, L., He, S., Han, Z., Xiao, H., Su S., Weng, M., & Cai, Z. (2019). Monitoring housing rental prices based on social media: an integrated approach of machine-learning algorithms and hedonic modelling to inform equitable housing policies. *Land Use Policy*, *82*, 657–673.
https://doi.org/10.1016/j.landusepol.2018.12.030

Ibeas, T., Cordera, R., dell'Olio, L., Coppola, P., & Dominquez, A. (2012). Modelling transport and real-estate values interactions in urban systems. *Journal of Transport Geography*, *24*, 370–382. https://doi.org/10.1016/j.jtrangeo.2012.04.012

International Valuation Standards Committee. (2003). *International Valuation Guidance Note No. 11: reviewing valuations*.
http://www.romacor.ro/legislatie/25-gn11.pdf

International Valuation Standards Council. (2017). *International Valuation Standards*. https://www.ivsc.org/standards/international-valuation-standards

Isakson, H. R. (1998). The review of real estate appraisals using multiple regression analysis. *Journal of Real Estate Research*, *15*(1/2), 177–190.
https://doi.org/10.1080/10835547.1998.12090922

Jahanshiri, E., Buyong, T., & Shariff, A. R. M. (2011). A review of property mass valuation models. *Pertanika Journal of Science and Technology*, *19*(1), 23–30.

Jones, C. (2002). The definition of housing market area and strategic planning. *Urban Studies*, *39*(3), 549–564.
https://doi.org/10.1080/00420980220112829

Kaklauskas, A., Zavadskas, E. K., Bagdonavicius, A., Kelpsiene, L., Bardauskiene, D., & Kutut, V. (2010). Conceptual modelling of construction and real estate crisis with emphasis on comparative qualitative aspects descriptions. *Transformations in Business and Economics*, *9*(1/19), 42–61.

Kauko, T., & d'Amato, M. (2008). *Mass appraisal methods: an international perspective for property valuers*. John Wiley & Sons. https://doi.org/10.1002/9781444301021

Kokinis-Graves, C. (2006). Use of the cost, income and sales-comparison approaches in the valuation of real estate. *Journal of State Taxation*, 23–32.

Kuburić, M., Tomić, H., & Mastelić Ivić, S. (2012). Use of multicriteria valuation of spatial units in a system of mass real estate valuation. *KiG*, *11*(17), 58–74. https://hrcak.srce.hr/85966

Lai, P. Y. (2011). Analysis of the mass appraisal model by using artificial neural network in Kaohsiung city. *Journal of Modern Accounting and Auditing*, *7*(10), 1080–1089.

Lamont, O., & Stein, J. (1999). Leverage and house price dynamics in US cities. *RAND Journal of Economics*, *30*(3), 498–514.

Mathieson, K., & Dryer, B. J. (1993). Improving the effectiveness and efficiency of appraisal reviews: an information systems approach. *Appraisal Journal*, *61*(3), 57–63.

Mayer, C. J., & Somerville, C. T. (2000). Land use regulation and new construction. *Regional Science and Urban Economics*, *30*(6), 639–662.
https://doi.org/10.1016/S0166-0462(00)00055-7

McCluskey, N., Davis, P., Haran, M., McCord, M., & McIlhatton, D. (2012). The potential of artificial neural networks in mass appraisal: the case revisited. *Journal of Financial Management Property and Construction*, *17*(3), 274–292.
https://doi.org/10.1108/13664381211274371

McCluskey, W. J., & Borst, R. (1997). An evaluation of MRA, comparable sales analysis and ANNs for the mass appraisal of residential properties in Northern Ireland. *Assesment Journal*, *4*(1), 47–55.

McCulloch, C. E., & Searle, S. R. (2001). *Generalized, linear, and mixed models*. John Wiley & Sons.
https://doi.org/10.1002/9780470057339.vag009

Mora-Esperanza, J. G. (2004). Artifical intelligence applied to real estate valuation: an example for the appraisal of Madrid. *CATASTRO*, 255–274.

Nguyen, N., & Cripps, A. (2001). Predicting housing value: a comparison of multiple regression analysis and artificial neural networks. *Journal of Real Estate Research*, *22*(3), 313–336.
https://doi.org/10.1080/10835547.2001.12091068

Núñez Tabales, J. M., Caridad y Ocerin, J. M., & Rey Carmona, F. J. (2013). Artificial neural networks for predicting real estate prices. *Revista de Metodos Cuantitativos para la Economia y la Empresa*, *15*(1), 29–44.

Peterson, S., & Flanagan, A. B. (2009). Neural network hedonic pricing models in mass real estate appraisal. *Journal of Real Estate Research*, *31*(2), 147–164.
https://doi.org/10.1080/10835547.2009.12091245

Rosen, S. (1974). Hedonic prices and implicit markets: product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Roulac, S. E. (1986). *Real estate as a strategic resource* (pp. 317–321). Chief Financial Officer International.

Sampathkumar, V., Santhi, M. H., & Vanjinathan, J. (2015). Evaluation of the trend of land price using regression and neural network models. *Asian Journal of Scientific Research*, *8*(2), 182–194. https://doi.org/10.3923/ajsr.2015.182.194

Scheurwater, S. (2017). *The future of valuations: the relevance of real estate valuations for institutional investors and banks – view from a European Expert Group*. RICS.

Selim, H. (2009). Determinants of house prices in Turkey: hedonic regression versus artificial neural network. *Expert Systems with Applications*, *36*(2/2), 2843–2852.
https://doi.org/10.1016/j.eswa.2008.01.044

Stevenson, S. (2004). New empirical evidence on heteroskedasticity in hedonic housing models. *Journal of Housing Economics*, *13*(2), 136–153. https://doi.org/10.1016/j.jhe.2004.04.004

Tay, D. P. H., & Ho, D. K. H. (1992). Artificial intelligence and the mass appraisal of residential apartments. *Journal of Property Valuation and Investment*, *10*(2), 525–540.
https://doi.org/10.1108/14635789210031181

TEGoVA. (2020). *European Valuation Standards (EVS)*. https://tegova.org/european-valuation-standards-evs

Valier, A., & Micelli, E. (2020). Automated models for value prediction: a critical review of the debate. *Valori e Valutazioni*, *24*, 151–162.

Vascu, A. (2015). *About valuation and valuation review* [Despre evaluare şi verificarea evaluării]. Hamangiu / IROVAL.

Vo, N., Shi, H., & Szajman, J. (2014). Optimisation to ANN inputs in automated property valuation model with Encog 3 and winGamma. *Applied Mechanics and Materials*, *462–463*, 1081–1086. https://doi.org/10.4028/www.scientific.net/AMM.462-463.1081

Worzala, E., Lenk, M., & Silva, A. (1995). An exploration of neural networks and its application to real estate valuation. *Journal of Real Estate Research*, *10*(2), 185–201. https://doi.org/10.1080/10835547.1995.12090782

Xie, X., & Hu, G. (2007). A comparison of Shanghai housing price index forecasting. In *3rd International Conference on Natural Computation (ICNC)* (pp. 221–225), Haikou. https://doi.org/10.1109/ICNC.2007.14

Yacim, J. A., & Boshoff, D. G. B. (2018). Combiniang BP with PSO algorithms in weights optimisation and ANNs training for mass appraisal of properties. *International Journal of Housing Markets and Analysis*, *11*(2), 290–314. https://doi.org/10.1108/IJHMA-02-2017-0021