

Artificial intelligence algorithms to predict Italian real estate market prices

Luca Rampini and Fulvio Re Cecconi

Department of Architecture, Built Environment and Construction Engineering,
Politecnico di Milano, Milan, Italy

Abstract

Purpose – The assessment of the Real Estate (RE) prices depends on multiple factors that traditional evaluation methods often struggle to fully understand. Housing prices, in particular, are the foundations for a better knowledge of the Built Environment and its characteristics. Recently, Machine Learning (ML) techniques, which are a subset of Artificial Intelligence, are gaining momentum in solving complex, non-linear problems like house price forecasting. Hence, this study deployed three popular ML techniques to predict dwelling prices in two cities in Italy.

Design/methodology/approach – An extensive dataset about house prices is collected through API protocol in two cities in North Italy, namely Brescia and Varese. This data is used to train and test three most popular ML models, i.e. ElasticNet, XGBoost and Artificial Neural Network, in order to predict house prices with six different features.

Findings – The models' performance was evaluated using the Mean Absolute Error (MAE) score. The results showed that the artificial neural network performed better than the others in predicting house prices, with a MAE 5% lower than the second-best model (which was the XGBoost).

Research limitations/implications – All the models had an accuracy drop in forecasting the most expensive cases, probably due to a lack of data.

Practical implications – The accessibility and easiness of the proposed model will allow future users to predict house prices with different datasets. Alternatively, further research may implement a different model using neural networks, knowing that they work better for this kind of task.

Originality/value – To date, this is the first comparison of the three most popular ML models that are usually employed when predicting house prices.

Keywords Artificial intelligence, House market price, Real estate, Machine learning, Deep learning

Paper type Research paper

1. Introduction

Traditionally, in the Real Estate (RE) market, the value of assets is measured using manual evaluation methods, e.g. comparative, investment, residual and so on (Pagourtzi *et al.*, 2003). However, these methods often fail in capturing the complexity and the variety of the assets in the current RE market, thus the assessments are often characterised by a high inaccuracy (Zurada *et al.*, 2006). Various RE stakeholders such as house owners, buyers, investors and agents use house price prediction frameworks usually depending on multiple factors; some of them are linked to the geometrical aspects of the physical assets, such as the volume and the area, whereas some other are related to the geographical position of the buildings within the Built Environment (Gao *et al.*, 2019). Moreover, the general physical conditions and the assets' year of construction can further influence the final prices (Rahadi *et al.*, 2015).

© Luca Rampini and Fulvio Re Cecconi. Published by Emerald Publishing Limited. This article is published under the Creative Commons Attribution (CC BY 4.0) licence. Anyone may reproduce, distribute, translate and create derivative works of this article (for both commercial and non-commercial purposes), subject to full attribution to the original publication and authors. The full terms of this licence may be seen at <http://creativecommons.org/licenses/by/4.0/legalcode>

Compliance with Ethical Standards: The authors declare no conflict of interest.



Lately, fostered by the increasing amount of data available and the advancement in Information Technology, Artificial Intelligence techniques are widely deployed to solve complex, non-linear problems. AI algorithms can provide evaluation methods that are more accurate and efficient than the traditional ones (Taffese, 2006; Abidoye and Chan, 2017). To date, there is no agreed definition of what constitutes AI, although, according to Baldini *et al.* (2019), AI usually refers to “machines or agents that are capable of observing their environment and, experience gained, taking intelligent action or proposing decisions”. Among AI methods, Machine Learning (ML) is one of the cutting-edge techniques that can be used to identify, interpret and analyse tabular data. ML algorithms represent a paradigmatic shift in computer programming. Traditionally, computer scientists would code rules and data as inputs and results as output. However, in ML, computers receive data and results, whereas the algorithm produces the rules. Therefore, a ML system is trained rather than explicitly programmed. Recently, Artificial Neural Networks (ANNs), a subset of ML, are gaining momentum for their efficiency and easy implementation. ANNs, also referred to as Deep Learning (DL), are characterised by a pipeline of trainable modules called layers, which do not require any parameters’ tuning if not the size of the network (LeCun *et al.*, 2015).

This study provides a comparison among three popular ML techniques, namely ElasticNet, XGBoost, and ANN, to assess the more accessible and reliable model that can predict house stock market prices. Indeed, the knowledge of housing prices offered a sound basis for a better understanding of some Built Environment’s aspects such as the impacts of urban renewal (Lee *et al.*, 2017), the effects of environmental sustainability (Huang and Yin, 2015), and the attractiveness of ecological factors (Luttk, 2000). *Inter alia*, an efficient prices prediction model may mitigate the effects of the price variations caused by bubbles, economic slowdowns, or bankruptcy (Abidoye *et al.*, 2019), helping clients and customers of the RE market. A prediction model is characterised by several input parameters and one or more output values, which, in this specific article, is the price of a dwelling. The predictions’ accuracy measures the goodness of a model, and its usability can be related to the number and type of input parameters. The more parameters are required, the more difficult it is to retrieve these parameters, the less useable the forecasting model will be and therefore the more it will be directed to a limited number of users. In this research the features provided by the commerce chamber of Lombardy are six, as discussed in Section 3. Finally, the models are tested with two datasets from Brescia and Varese, which are two cities in Northern Italy.

This paper is organised as follows. Section 2 shows the related works on the topic. Section 3 describes the methodology used to collect and clean the datasets as well as models’ descriptions. The review of data features and entries made by exploratory data analysis is described in Section 4, whereas the results and the comparison of the models are presented in Section 5. Finally, the last section concludes the paper.

2. Related work

The most popular evaluation method of properties is the hedonic pricing model (HPM). HPM was introduced by Lancaster (1966) and then has been extended to the residential market by Rosen (1974) to assess the effects of social, environmental and urban characteristics on property values. Since then, this model has been widely utilised to correlate house prices and house characteristics (Adair *et al.*, 1996; Keskin, 2008). Several applications of hedonic price modelling have enabled the identification of correlations that contradict empirical evidence: for instance (Espey and Lopez, 2000), analysing the effect of proximity to the airport on residential property values, highlighted that this location could be perceived as an amenity rather than a detractor.

However, the recent financial and economic events have caused a great deal of turbulence in the field of valuation theories and techniques, not only at the academic level, where new

approaches to value creation have been developed, but also at the operational level, due to the apparent ambiguity and approximation of the results obtained using traditional techniques ([Tajani et al., 2018](#)).

Recently, data-driven AI algorithms have increased dramatically ([Lu, 2019](#)) in all industrial sectors. Hence, several studies implemented ML algorithms to predict housing prices ([Park and Kwon Bae, 2015; Rafiei and Adeli, 2016; Abidoye and Chan, 2017](#)). [Mohd et al. \(2019\)](#) compared several ML algorithms such as Random Forest, Ridge Regression and Lasso to see which technique performs better. Their results concluded that Random Forest (RF) was the best in terms of accuracy. The same conclusion was reached by [Wu and Wang \(2018\)](#), where RF outperformed Regression models in predicting house prices data from Virginia (US).

Recently, gradient boosted classifiers have been used to win many data science competitions in [Kaggle \(2019\)](#). These optimisation techniques use a combination of decision tree models that can exceed RF performance. [Ho et al. \(2020\)](#) compared Support Vector Machine (SVM), RF and gradient boosting machine (GBM) to examine 40,000 housing transactions in Hong Kong. Their results confirmed the better accuracy of the GBM model over the others. Moreover, [Mrsic et al. \(2020\)](#) demonstrated that the XGBoost algorithm outperformed RF and Gradient Boosting AdaBoost, resulting in the best ensemble method for this task.

Besides, the use of ANN, thanks to the increased computational power of standards computers and the availability of open-source datasets, is gaining momentum. Currently, Neural Networks are widely adopted in different fields, such as Healthcare ([Jiang et al., 2017](#)) Finance ([Guresen et al., 2011](#)) and Agriculture ([Abiodun et al., 2018](#)). To evaluate dwelling prices, [García-Magariño et al. \(2020\)](#) compared the Multi-Layer Perceptron (MLP) neural network with other ML techniques. They found that MLP achieved the lowest error and presented no outliers in the predictions. Finally, [Zhou \(2020\)](#) implemented a back-propagation (BP) neural network to assess China's property prices. He concluded that the application of the model in RE price evaluation is technically feasible and credible.

Finally, there are a variety of publications in the literature that compare property price estimates using HPM and ANN models. The findings are contradictory: in various studies, the advantage of ANN is to capture, automatically from data, the non-linear relations between explanatory variables and prices ([Limsombunc et al., 2004; Selim, 2009; Lin and Mohan, 2011](#)). On the other hand, other studies ([Worzala et al., 1995; Lenk et al., 1997](#)) claims that the ANN is a "black box" that generates solutions rather than a plain functional link between the input and output values. The results are not consistent, although they do improve as the sample size grows. The marginal prices computed using ANN are more realistic than the traditional hedonic pricing, however, they come with significant computational difficulties. Due to over-parameterisation, an excessive number of neurons might result in a lack of predicting power.

3. Method and tools

Recently, new techniques for predicting RE values have appeared on the scientific research scene, either alongside or replacing traditional techniques. However, advances in computational tools, particularly those in artificial intelligence, make it necessary to update methods for estimating RE prices frequently. This paragraph shows how the objective of comparing a modern RE price estimation system, based on an artificial neural network, with two other methods widely used in scientific literature has been achieved. The proposed methodology can be summarised in the following steps:

- (1) *Data collection.* The market value data for residential buildings were retrieved from the local Chamber of Commerce, which maintains and updates those data periodically.

-
- (2) *Exploratory Data Analysis (EDA)*. EDA is an approach for data analysis that employs various techniques, most of them graphical, to (i) maximise insight into a dataset, (ii) uncover underlying structure, (iii) extract essential variables and (iv) detect outliers and anomalies. EDA is further described in [Section 4](#).
 - (3) *Data cleaning*. After the EDA, incorrect records and outliers have been removed from the dataset. The raw data contained some records likely to be incorrect. These records were identified according to some rules, described in [Section 4](#), and deleted.
 - (4) *Models selection and parameters tuning*. The models selected for this study are three: ElasticNet, XGBoost and ANN. Each model presented different parameters to tune in order to achieve the best accuracy. Models and the tuning process are described in detail in [Section 3.1](#). The models are implemented inside a Python environment using most popular ML libraries such as SKlearn (for ElasticNet), XGBoost and Keras (for the ANN).
 - (5) *Dataset preparation*. Traditionally, the dataset used in ML research is split into two groups: training and test sets. The model uses the training set to tune the weights, whereas the test set is used to check if the algorithm can generalise the problem properly, so it can also be used with data that do not come from the original dataset. This study used 80% of the original data to train the algorithm and 20% to test it.
 - (6) *Comparison of the predictions*. The evaluations of the models' predictions are compared to establish which model reached the best accuracy. The results are shown in [Section 5](#).

3.1 Models description

This study compared the performance of three different ML models that adopt different philosophies while solving a problem. The three models used in this research are as follows:

- (1) *ElasticNet*. The elastic-net regularisation theme has been proposed for the first time by [Zou and Hastie \(2005\)](#). Regularisation in regression algorithms is the process of introducing additional information to the model in order to prevent overfitting ([Ghojogh and Crowley, 2019](#)). Traditionally, in ML the models are regularised by adding some constraints to the loss function. ElasticNet adopts a combination of two methods: (i) Ridge regression, also called L2 regularisation, penalises feature variance to reduce model complexity and (ii) Lasso, also known as L1 regularisation, which penalises some features and sometimes excluding them from the model's training. To evaluate the model's performance in making predictions on unseen data, k -fold cross-validation is conducted. This technique splits the training dataset into k groups, or folds, that have approximately the same size, and then treat the first fold as a validation set, whereas the algorithm is trained in the remaining $k-1$ folds ([Arlot and Celisse, 2010](#)). This study implemented 10-fold cross-validation.
- (2) *XGBoost*. Gradient boosting is a ML algorithm that produces a prediction model in the form of an ensemble of weak prediction models, usually decision trees ([Friedman, 2001](#)). Hence, an ensemble model is a combination of simple individual models that together create a more powerful one. XGBoost starts by fitting an initial decision tree to the data. Then, a second model focuses on accurately predicting the cases where the first model performs poorly. This process of boosting is repeated many times and each successive model attempts to correct for the shortcomings of the combined boosted ensemble of all previous models. Unlike the ElasticNet model, which had few parameters to tune, XGBRegressor presented more than ten factors to optimise.

Thus, the hyperparameter tuning phase was considerably expensive compared to ElasticNet. Nonetheless, the size of the datasets did not imply unacceptable training time; therefore, also for XGBRegressor a 10-folds cross-validation was used. However, with bigger datasets that contain hundreds of thousands of records, it is necessary to assess if this option is viable.

- (3) *Artificial Neural Network (ANN)*. ANNs are a branch of AI developed since the 1940s. The concept beyond ANNs is to mimic the behaviour of human brains by using a set of connected units or nodes, called neurons, arranged in multiple computational layers (Grosan and Abraham, 2011). The most significant advantage of ANNs is that they do not require much human intervention in tuning the model; therefore, ANNs are relatively easy to implement. On the other hand, the major concern about these models is that they act like a “black box” and the weights that these algorithms are learning are not always controllable (Castelvecchi, 2016). The structure of an ANN is composed of a series of layers: in the first, called input layer, the data are fed into the model, whereas in the last, called output layer, the expected results are given, usually as a prediction value of the target variable. In the middle, several layers, called hidden layers, learn the correlations among features to make the best forecasting. In this study, an ANN with 3 hidden layers has been implemented (see Table 1).

The whole dataset is typically divided into three sets: training, validation and test. The weights are calculated on the training set, while the model is evaluated on the validation set. This is done to prevent overfitting and to be sure that the model is generalising (Chollet, 2018). However, the MAE score of the model is evaluated on the test set, since the model has not had access to any information about it. To date, there is not a standard and accepted method to choose the best number of layers in a neural network, therefore, the typical approach is to try different configurations and see which performs better (Olson *et al.*, 2018). In this case, configurations with 2, 3, 4 hidden layers have been tested. The configuration with 3 layers outperforms the one with 2 layers, whereas the model with 4 layers had a drop in the validation accuracy, meaning that it was overfitting.

4. Exploratory data analysis

This section describes in detail the composition of the two datasets from Brescia and Varese. Both datasets had six numerical features that describe each housing unit: the geographic coordinates (WGS X is the latitude, while WGS Y is the longitude), the unit's surface, the number of car's spot, the construction year and the energy performance index EPh measured in [$\text{kWh m}^{-2} \text{y}^{-1}$]. Although this reduced number of features might affect the performance of the models, it is interesting testing the models because (1) the reduced number of factors can make new databases creation more expeditious and (2) the

Table 1.
ANN architecture

Layer (type)	Output shape	Parameters
Normalisation	(None, 6)	13
Dense_1	(None, 64)	448
Dense_2	(None, 128)	8,320
Dense_3	(None, 128)	16,512
Dense_4	(None, 1)	129
Total parameters: 25,422		

ANN's black-box nature will not necessarily benefit from features traditionally used to determine housing prices. Besides, raw data retrieved from the local chamber of commerce were cleaned and all incorrect records and outliers were removed. The adjustments made from the original data comprehend the following:

- (1) Removing data with incorrect geographical location, i.e. outside the target city;
- (2) Removing data not associated with a single residential housing unit;
- (3) Removing incorrect data, i.e. record with (i) more than 10 parking lots, (ii) too big surface (more than 700 sqm), (iii) too low price for squared metre (less than 600 Euro/sqm), and (iv) huge EPh consumption (more than 525 kWh/m²).

The process of data exploration and data cleaning is iterative, but, in the end, it was possible to show the characteristics of the dataset without errors and outliers.

4.1 EDA of Brescia's dataset

The first EDA has been performed on the city of Brescia (approx. 197,000 inhabitants), which is the second largest city in the region and the fourth of northwest Italy. The cleaned dataset contained 1,203 entries evenly distributed throughout the municipal territory, with a slight prevalence for the central area ([Figure 1](#)).

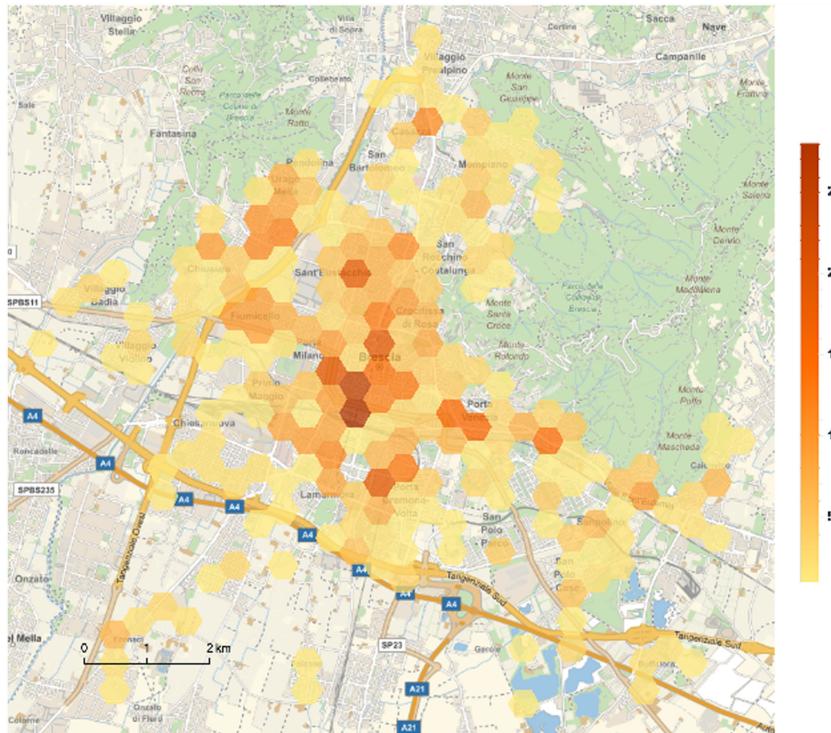


Figure 1.
Geographical
distribution of the
Brescia dataset

Note(s): Each hexagon is coloured according to the number of records pertaining to it

The dwelling units in Brescia's dataset had an average price equal to 291,486 Euro and a standard deviation of 266,465 Euro. A complete statistical description of the dataset is shown in [Table 2](#).

Almost 80% of the dwellings were under 400,000 Euro, whereas half were under 200,000 Euro. The distribution of the prices is shown in [Figure 2](#). The price distribution is skewed right where a very high price characterised few assets.

EDA also aimed to identify the most important features from the dataset. Typically, a heat map that contains Pearson correlation's values represents the relationship between features and the target variable ([Figure 3](#)). The hue channel facilitates the reading of the results where red blocks represent a negative correlation, i.e. an inverse relationship between the features, whereas green blocks represent a high correlation. Albeit Pearson's coefficient may vary between -1 and $+1$, when dealing with a topic that embeds social aspects and human behaviour, the fluctuations of the variables limit the possibility to reach high values of correlation. Social aspects influence housing prices, therefore, correlation coefficients between -0.20 and 0.20 are generally considered weak, between 0.20 and 0.50 (positive or negative) are considered moderate, and above 0.50 (positive or negative) are considered strong ([Spiegelhalter, 2019](#)).

Table 2.
Statistic description of
Brescia's dataset

	Surface	Box or parking lots	Construction year	EPh	Price/sqm	Price
Count	1,203	1,200	1116	1,092	1,203	1,203
Mean	143	1	1969	162	1,946	291,486
Std	95	0.83	84	96	820.5	266,465
Min	25	0	700	1	611	35,000
25%	86	0	1964	100	1,346	129,000
50%	116	1	1977	175	1,785	200,000
75%	165	2	2008	200	2,360	352,500
Max	818	4	2020	512	6,923	2,040,000

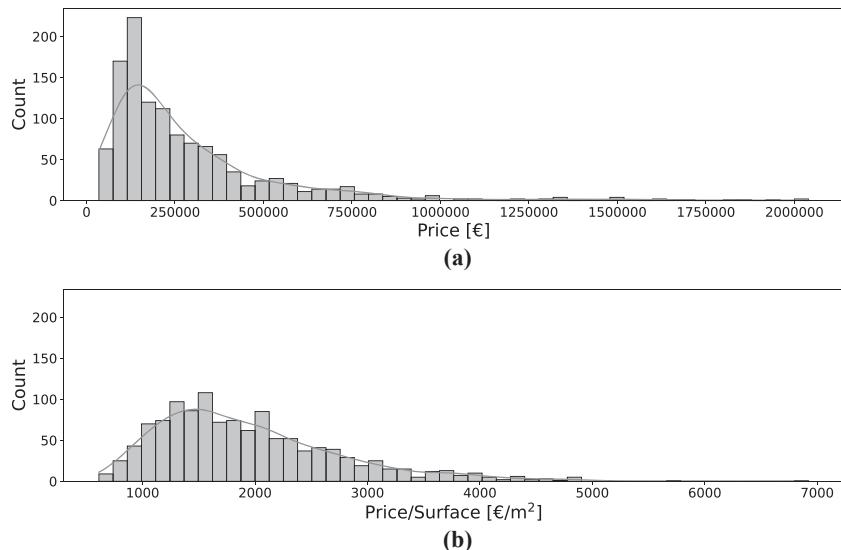


Figure 2.
Statistical distribution
of the price (a) and price
per square metre (b) of
the estates contained in
the dataset pertaining
to the city of Brescia

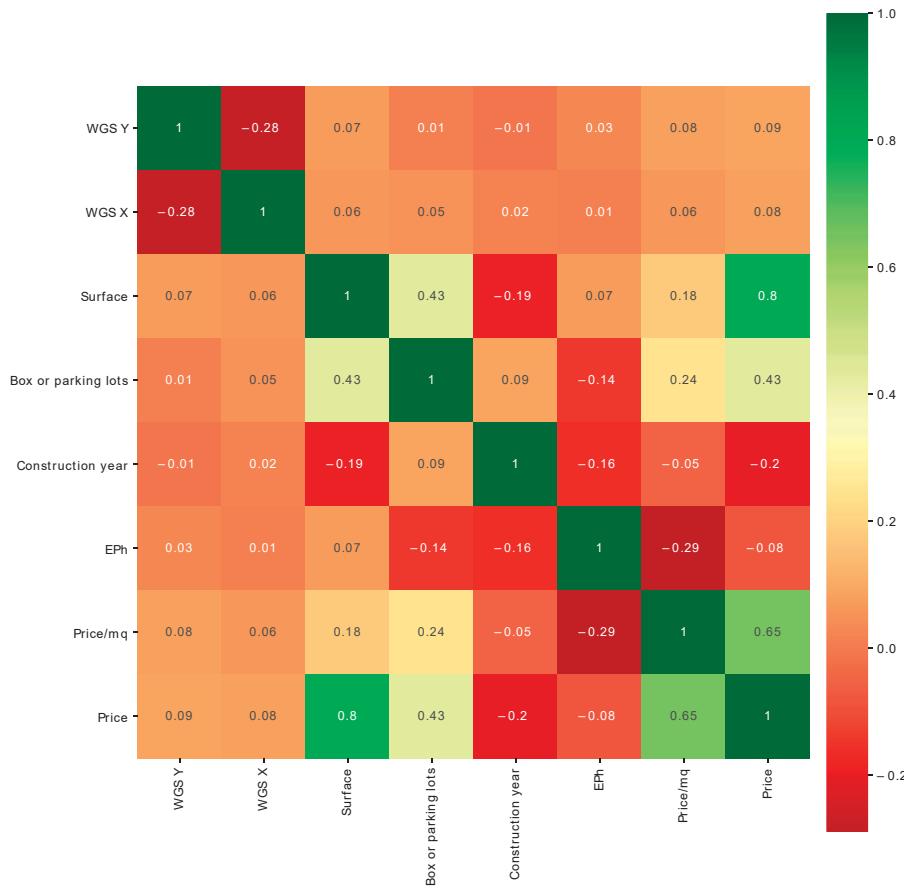


Figure 3.
Correlation among
features in the dataset
of the city of Brescia

The heat map suggested the following deductions:

- (1) Although the location is known to be the most important characteristic affecting the price of a house, the geographical information did not show a high value of correlation with the Price since the Pearson coefficient was calculated for every single coordinate; hence, longitude or latitude alone cannot be well correlated with prices.
- (2) The Surface was the feature with the highest correlation to the Price (Figure 4b), being the Pearson's correlation coefficient equal to 0.8; this was predictable given the above considerations about how the location was treated in the correlation coefficient computation.
- (3) EPh was more related to the Price/Surface (Pearson's correlation coefficient equal to -0.29) parameter than the Price (Pearson's correlation coefficient equal to -0.08)

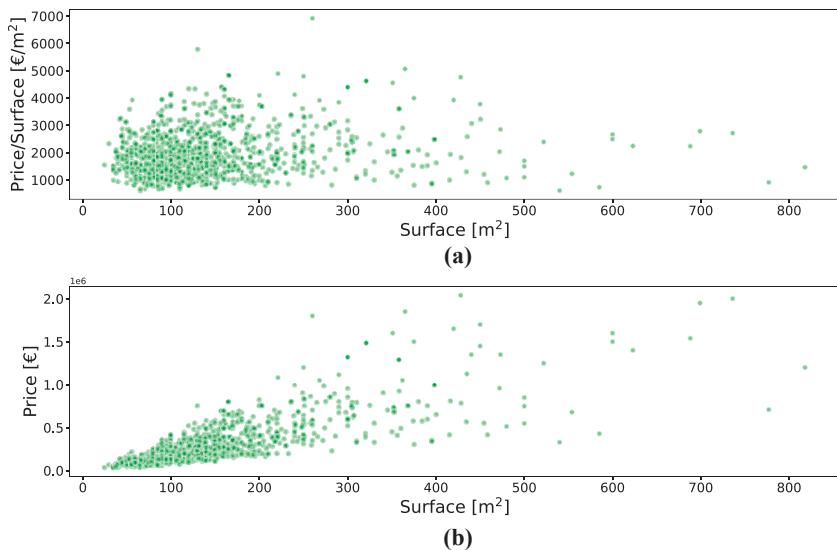


Figure 4.
Correlation between Surface and (a) Price per square metre; (b) Price for the city of Brescia

Note(s): Each point represents a dwelling unit

itself. This was reasonable since the energy consumption of an asset also depends on its dimensions. The market perceived a higher value of EPh, i.e. a higher energy demand per square metre, for large surface houses.

- (4) The good correlation between box or parking lots and prices was noteworthy, and the inverse correlation between price and construction years. These correlations may result from the peculiar price of houses in the city's historic centre: a costly location characterised by ancient buildings where parking lots are few and very sought-after.

The data visualisation techniques presented for the Brescia database can be transposed to Varese as well. Therefore, in the next paragraph, a brief description of Varese's dataset is presented.

4.2 EDA Varese's dataset

The records for the city of Varese (approx. 81,000 inhabitants) were 1,228. The cleaned entries were distributed throughout the municipal territory, albeit some clusters were in the city centre (Figure 5). The housing units in Varese had an average price equal to 234,175 Euro and a standard deviation of 217,138 Euro. A complete statistical description of the dataset is shown in Table 3.

The prices in Varese are slightly lower than Brescia, and the Price per metre square also demonstrates this. The distribution of prices is shown in Figure 6.

The heat map of the correlation for Varese's dataset is shown in Figure 7.

In this dataset, the correlation between the construction year and price was weaker probably because the price difference between buildings located in the historical city centre (i.e. with the older age) and those located in surrounding areas was less marked than in Brescia. The correlation between Surface Area and Price (Figure 8), conversely, was as high as in the case of the city of Brescia, being the Pearson's correlation coefficient for this city equal to 0.81.

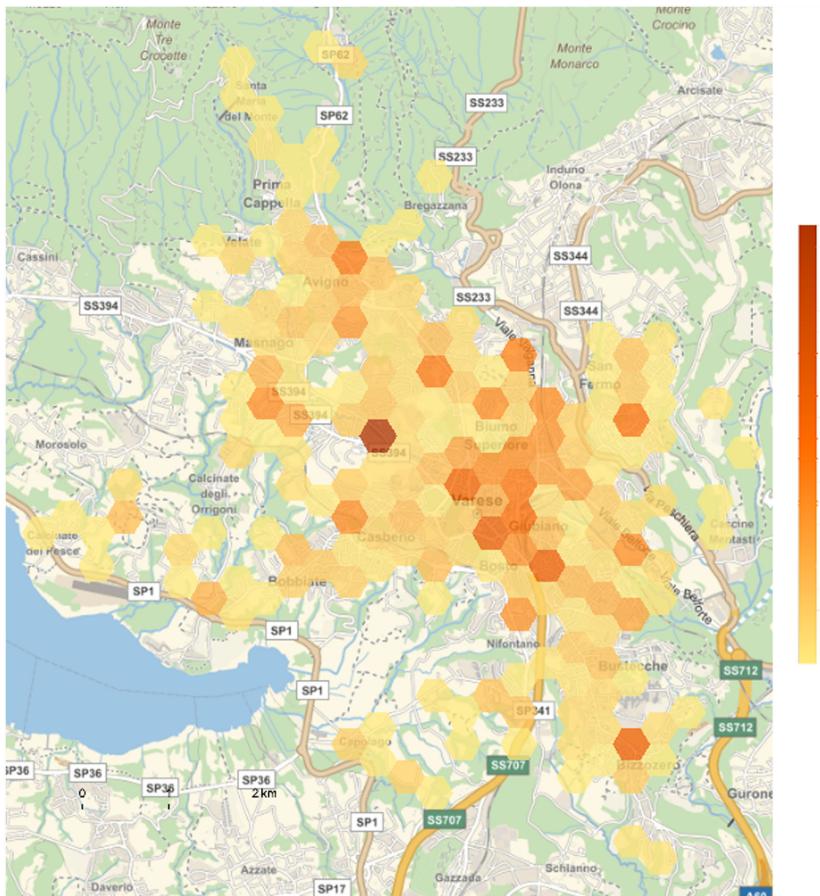


Figure 5.
Geographical distribution of the
Varese dataset

Note(s): Each hexagon is colored according to the number of records pertaining to it

	Surface	Box or parking lots	Construction year	EPh	Price/sqm	Price
Count	1,228	835	878	1,026	1,228	1,228
Mean	129	1	1972	223	1,729	234,175
Std	92	0.5	43	109	707	217,138
Min	21	1	1100	5	400	25,000
25%	75.75	1	1960	150	1,239	100,000
50%	100	1	1970	222	1,588	160,000
75%	145	2	1995	292	2,090	281,250
Max	680	4	2018	518	5,750	1,900,000

Table 3.
Statistical description
of Varese's dataset

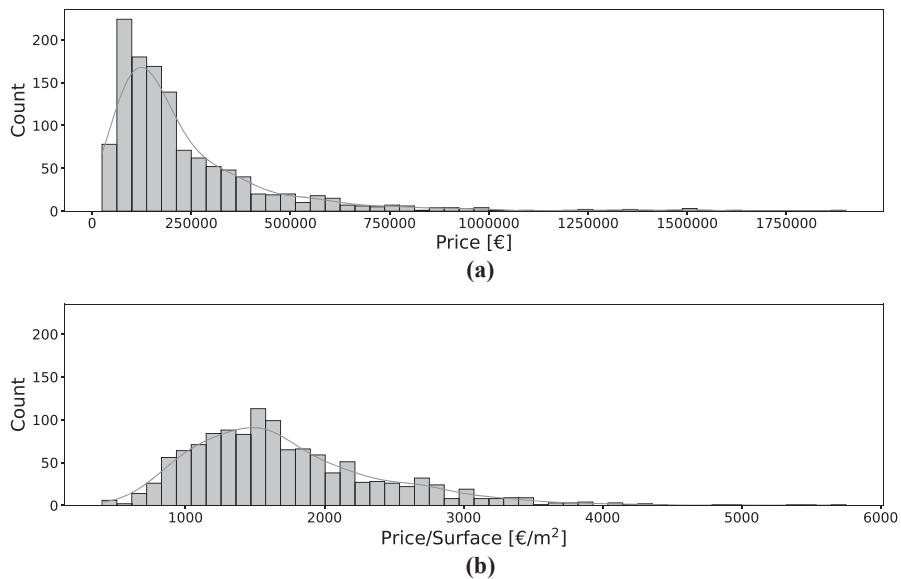


Figure 6.
Statistical distribution of the price (a) and price per square metre (b) of the estates contained in the dataset pertaining to the city of Varese

5. Comparison of the prediction models

With the expansion of computer methods, researchers are more and more frequently faced with the problem of evaluating the accuracy of a particular prediction algorithm (Baldi *et al.*, 2000). Both the root mean square error (RMSE) and the mean absolute error (MAE) are regularly employed in model evaluation studies (Chai and Draxler, 2014) but MAE is often preferred because it is a more natural measure of average error (Willmott and Matsuura, 2005). MAE is defined as follows:

$$\text{MAE} = \frac{1}{n} \sum_{j=1}^n |y_j - \hat{y}_j| \quad (1)$$

where:

y_j = predicted value

\hat{y}_j = real value

n = number of data

Therefore, a larger value of MAE means that the accuracy of the algorithm is decreasing, while the best possible accuracy, which coincides with the real value, happens when the MAE is equal to zero. Moreover, MAE can also be used to determine a “confidence interval” where the predicted prices may vary. Indeed, the little fluctuations that usually characterise data affected by social implications might be included within MAE. Finally, it is expected that the MAE will decline if more data are collected, so the range of possible house prices will also be narrowed.

The MAE scores are calculated on the test dataset, accounting for 20% of the whole dataset. The ElasticNet method scores a MAE equal to 92,988 € for Brescia and 62,921 € for Varese. The predictions against the real values are plotted in Figure 9. The algorithm predicts the first half of the datasets quite well, i.e. those with price less than circa 200,000 Euro,

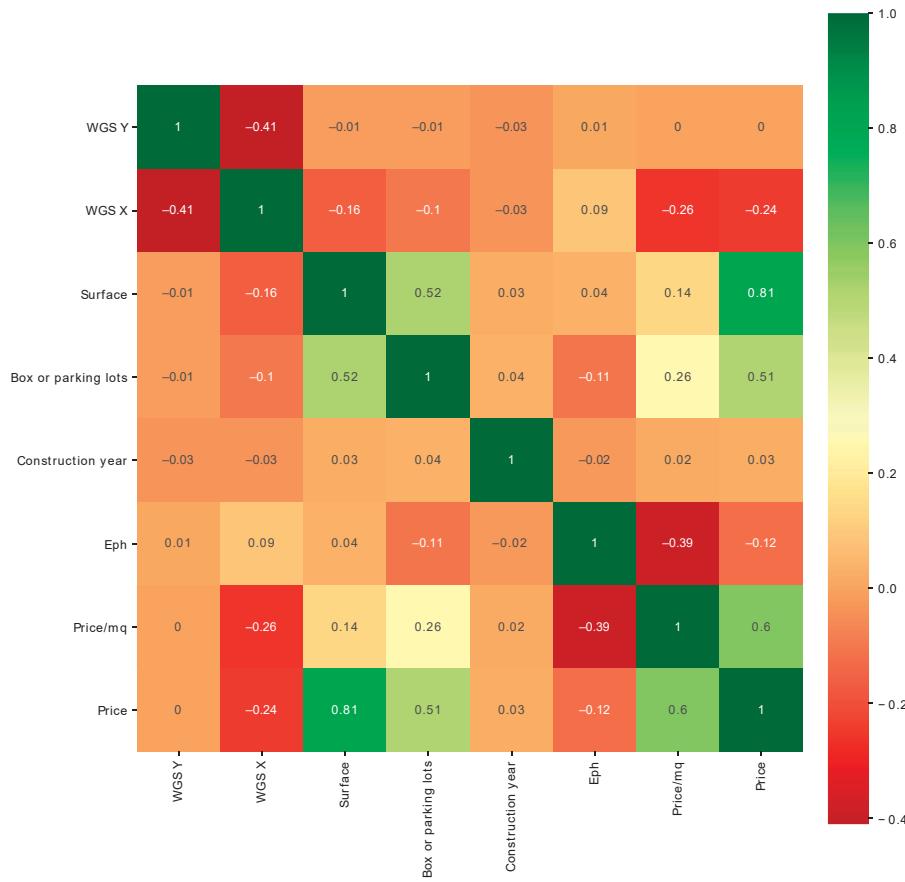


Figure 7.
Correlation among
features in the dataset
of the city of Varese

Note(s): Each record shows the Pearson correlation coefficient: green colors represent a positive correlation, while red colors indicates a negative correlation

whereas the most expensive properties are not correctly predicted, with errors up to 1 million (see Figure 10), probably due to the low number of records.

XGBoost algorithm performs better than ElasticNet with MAE equal to 81,025 € (Brescia) and 58,990 € (Varese). The reduced error is due to the better performance in the first half of the datasets (Figure 11), which decreased the MAE, even though the value of the maximum absolute error is greater than that produced by ElasticNet. Even in this case, predicting the most expensive property values was inaccurate, and the most significant error occurred in predicting the most expensive dwelling units (Figure 12).

Finally, it has been assessed if the ANN model generalises properly during the training phase, avoiding overfitting. Overfitting occurs when the algorithm becomes too good in evaluating training data without generalising the problem properly (Dietterich, 1995). Therefore, it is crucial to measure the loss and accuracy of the training set, as well as the ones of the validation set and control that are decreasing together. In Figure 13, both the training

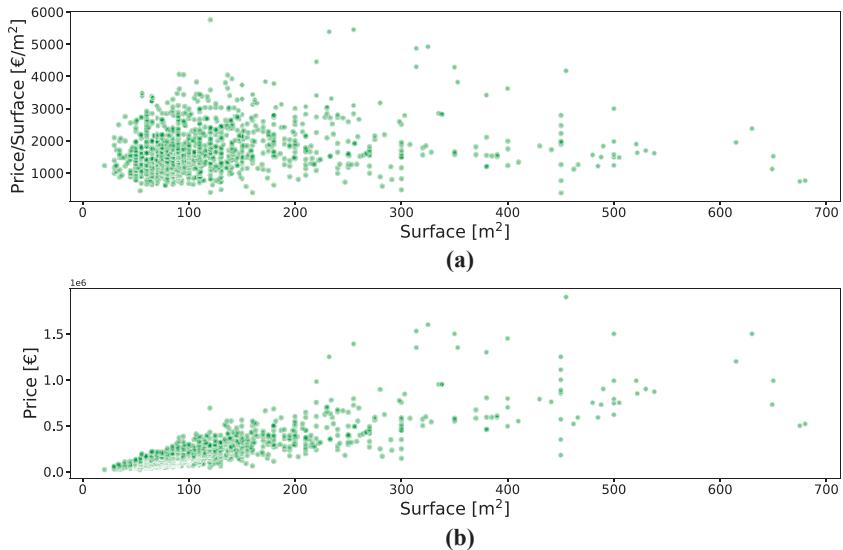


Figure 8.
Correlation between
Price and Surface for
the city of Varese

Note(s): Each point represents a dwelling unit

loss and the validation loss are plotted, showing the close behaviour of the two losses, which means the model is not overfitting.

The MAE scores for ANN are equal to 77,015 € (Brescia) and 56,128 € (Varese). In [Figure 14](#), it is reported the correlation between true and predicted values formulated by the ANN model. The algorithm can predict data quite well within the third quartile, whereas the higher prices have the biggest errors, even the highest for the most expensive housing units. Nonetheless, the distributions of the errors are more concentrated around zero ([Figure 15](#)) than the ones of the previous models, which overall gives better MAE scores.

Summing up, the results of the model's performances are showed in [Table 4](#), where all the MAE scores are listed as well as the relative difference among them computed as follows:

$$(\text{MAE}_i - \text{MAE}_{i+1}) / (\text{MAE}_i)$$

where:

MAE_i is the MAE of the first or second model

MAE_{i+1} is the MAE of the second (if MAE_i is the first) or third model (if MAE_i is the second)

The ElasticNet model has the worst results, although it can offer good speed of execution and ease of use. XGBoost reduces the ElasticNet errors by 13 and 6% for Brescia and Varese respectively. However, the number of parameters to tune can be a hurdle for non-expert users. Finally, the ANN is the best model for both datasets and reduces the errors by 5% towards XGBoost performances.

The MAE scores are strongly influenced by the errors made in predicting the most expensive properties. These errors are expected because the data range within the

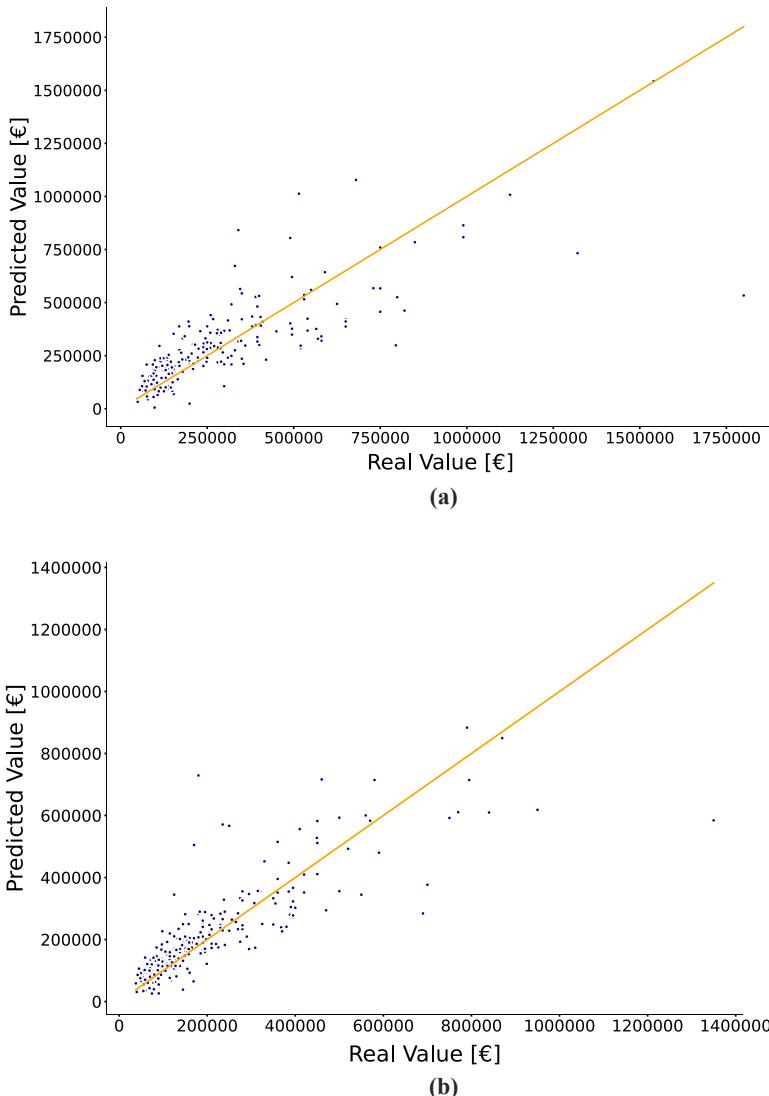


Figure 9.
Regression plot for
ElasticNet algorithm's
predictions for Brescia
(a) and Varese (b)

third quartile is much lower than that in the fourth quartile (Table 5); hence, the models had few scattered data to train on for the most expensive housing units.

Noteworthy, if housing units with prices higher than 800,000 € are removed from the test dataset, i.e. deleting only 3% of the dwellings from Brescia and 2% from Varese, using the identical ANN, trained on the same training set, the MAE score drops more than 10% in both datasets (Table 6).

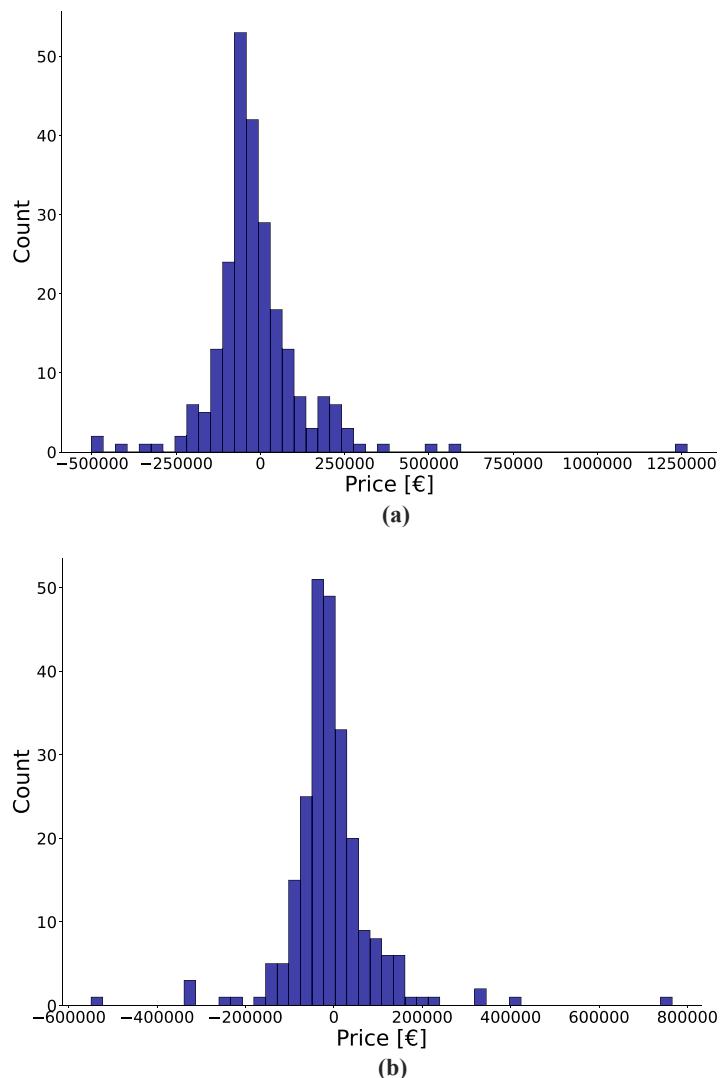


Figure 10.
Distribution of errors
for Brescia (a) and
Varese (b), produced
by ElasticNet

6. Discussion and conclusions

The current house stock's complexity and variety required new instruments to assess asset prices beyond traditional evaluation methods. This study uses modern AI techniques to predict RE values from data with multiple features. To detect the best algorithm for this problem, a comparison of three different techniques was conducted.

The results showed that ElasticNet had the worst performance. This conclusion was expected since ElasticNet is a linear regression algorithm, therefore it could not understand market evaluations in RE, particularly for the most expensive cases where the non-linearity of the problem is even more evident.

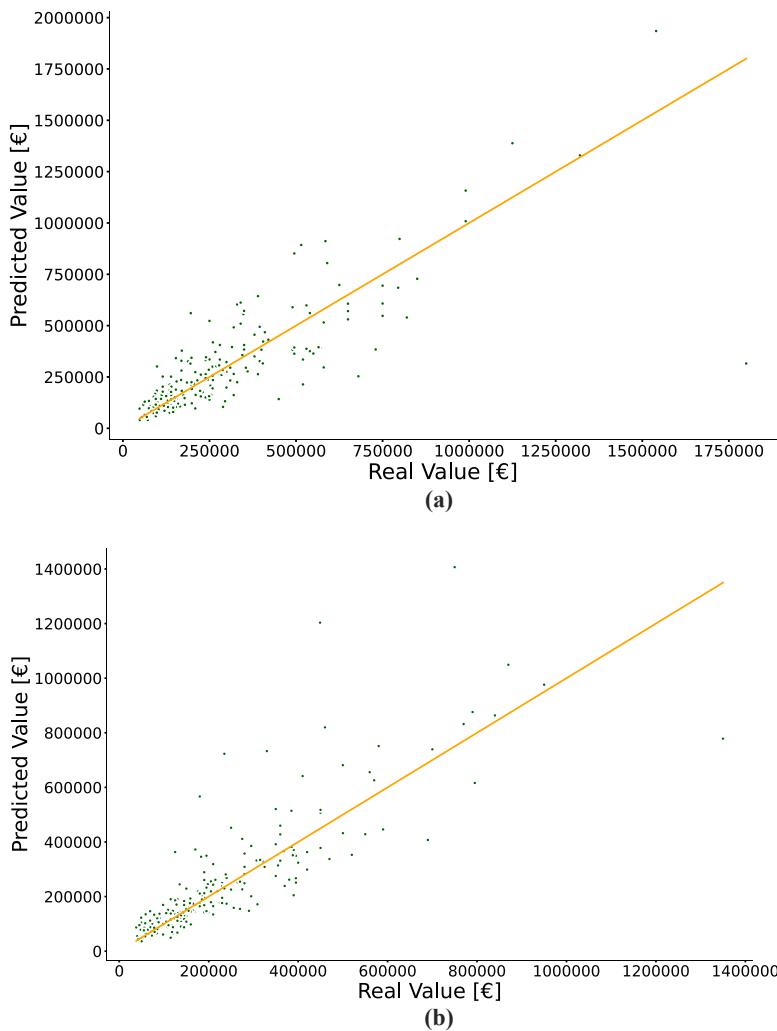


Figure 11.
Regression plot for
XGBoost algorithm's
predictions for Brescia
(a) and Varese
(b)

The second algorithm considered in this paper is XGBoost, which is a popular ML technique that uses gradient boosting machines to improve the performance of multiple regression tree models. XGBoost is widely used in many scientific fields for its execution speed and performance and has been already found as one of the best prediction algorithms for the RE market ([Mrsic et al., 2020](#)). The results obtained were better than the ElasticNet, however, the models produced the highest absolute error, which means that the robustness of the model should be tested with more data.

The last technique adopted was an ANN with a total of 5 layers. The results showed that this model outperformed the others even with a small architecture. For most of the dataset, the predictions made by the ANN were incredibly precise, with errors within ten thousand euros. The main concern about ANN models is the interpretability of the associations made by the algorithms, which may mean they end up using features that we would normally think

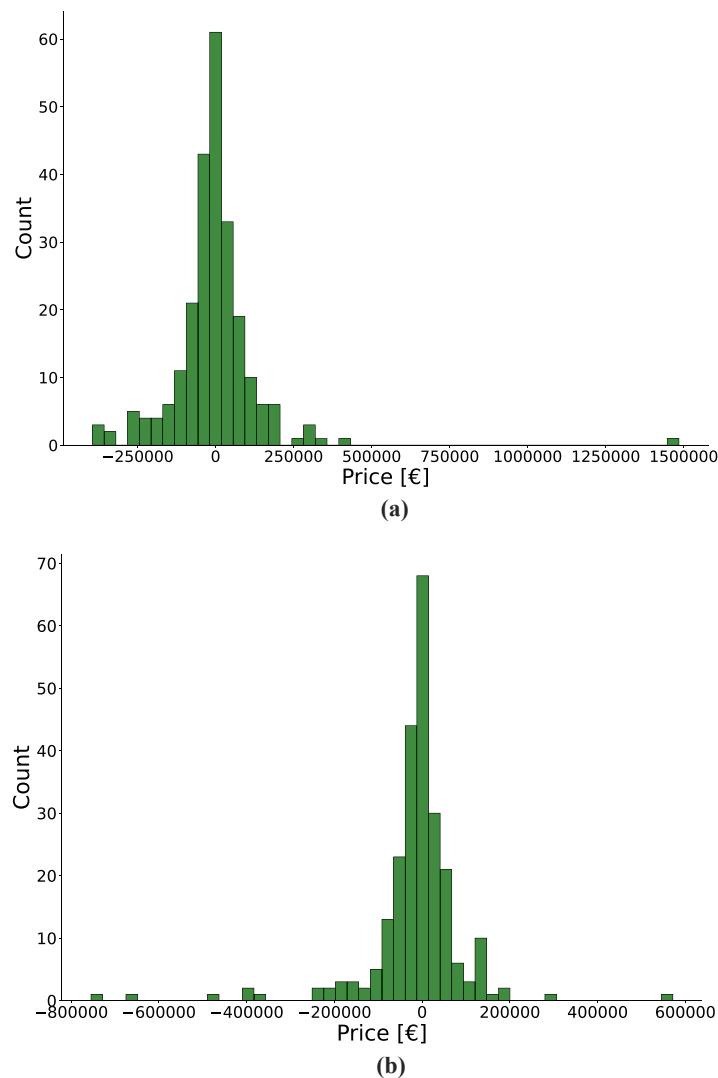


Figure 12.
Distribution of errors
for Brescia (a) and
Varese (b), produced by
XGBoost

are irrelevant to the task in hand (Spiegelhalter, 2019). Therefore, human supervision is always required when using DL techniques since the algorithm cannot comprehend the problem's causality. However, similar to XGBoost results, also the ANN model had a drop in performances for the most expensive properties due to the lack of a sufficient number of high price houses in the training dataset. Many similar studies emphasised the influence on the price of the location profile, gathered by a diverse range of location data sources, such as transportation profile (e.g. distance to nearest train station), education profile (e.g. school zones and ranking), suburb profile based on census data, facility profile (e.g. nearby hospitals, supermarkets) (Gao *et al.*, 2019). A Geographical Information System (GIS) based

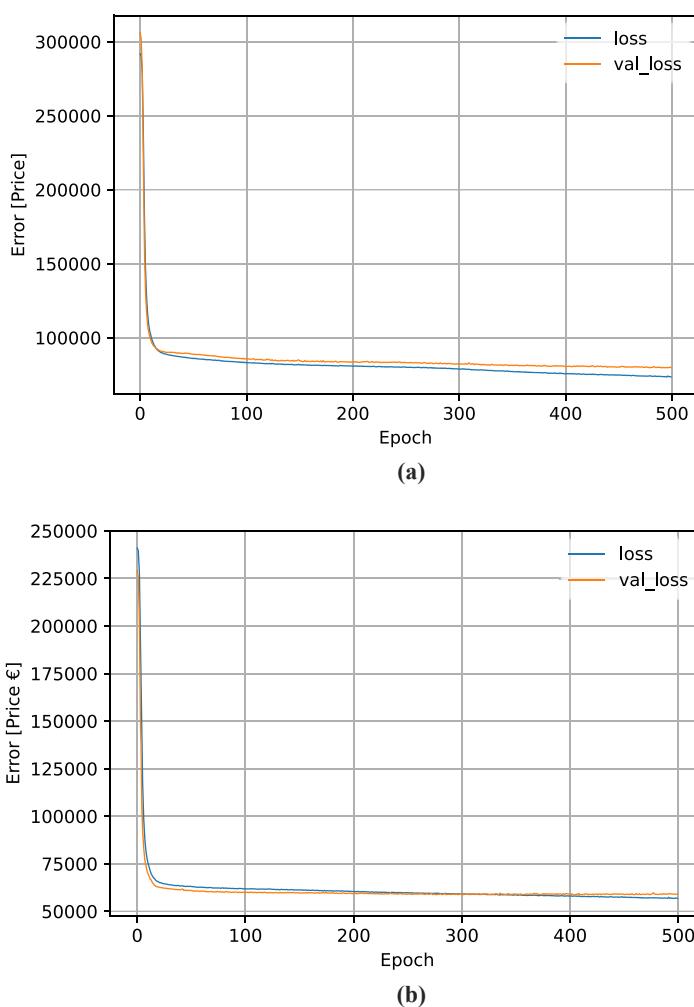


Figure 13.
ANN training and
validation loss history
for the city of (a)
Brescia and (b) Varese

representation of the ANN's error helps prove that the network was able to learn the importance of the location in the training phase. Figure 16 shows the geographical distribution of the error computed as the difference between the actual price in the dataset and the one predicted by the ANN for the city of a) Brescia and b) Varese. Each circle represents a record, the size and colour of the circle are proportional to the error made by the ANN in the prediction. The larger the circle and the darker the colour, the larger the error. In red, the errors where the price is overestimated, i.e. the predicted price is higher than the actual one, in blue those where it is underestimated. Noteworthy, the error is evenly distributed all over the two municipalities proving that the ANN correctly learns the influence of location profile over prices.

In conclusion, this study has demonstrated that modern AI techniques can be applied to predict houses' market price even when the size of datasets used to train the net is small to

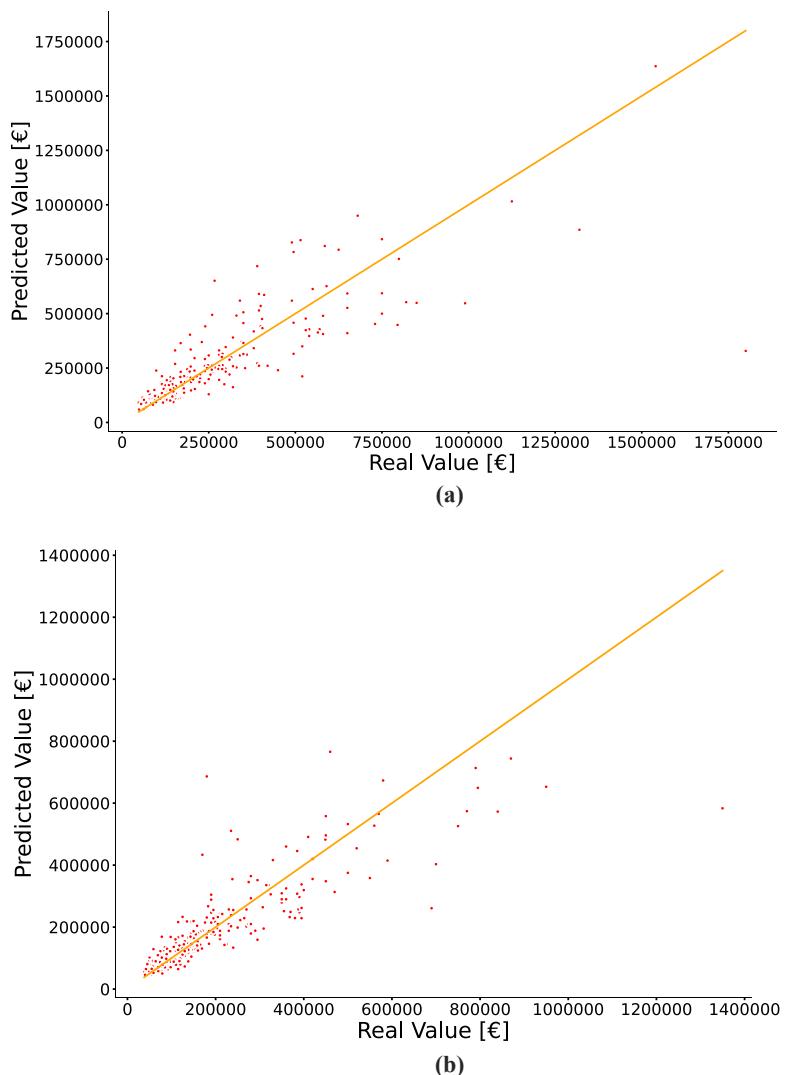


Figure 14.
Regression plot for the ANN algorithm's predictions for Brescia (a) and Varese (b)

medium. Although ML techniques work better with features rich datasets, it is important to stress that only six features were used to train the models in this study. A dataset with more features may lead to higher accuracy, making the prediction model less accessible to non-experts' users and more complex the search for new data to feed the model. All three ML algorithms tested showed weakness in predicting the prices of the most expensive houses, probably due to the dataset, which contained few cases of very expensive flats (see the long right tail in the statistical distribution in Figures 2 and 6). However, the artificial neural network performed better than the other two prediction models by providing accurate predictions for medium to low-priced houses and the lowest error for the most expensive ones.

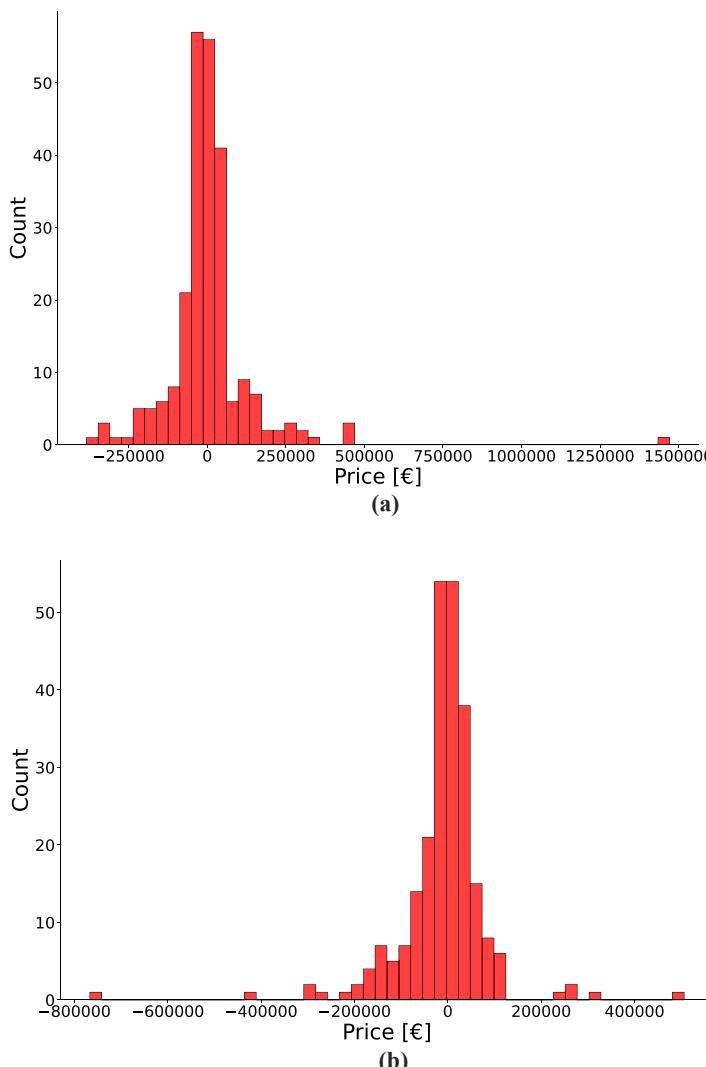


Figure 15.
Distribution of errors
for Brescia (a) and
Varese (b), produced
by ANN

Model	MAE score	Brescia		Varese	
		Relative difference (%)	MAE score	Relative difference (%)	MAE score
ElasticNet	92,988 €	—	62,921 €	—	—
XGBRegressor	81,025 €	13	58,990 €	6	—
ANN	77,015 €	5	56,128 €	5	—

Note(s): The percentage shown in the “Relative difference” column is calculated concerning the model above

Table 4.
MAE Score for the
three prediction
algorithms: the smaller
is the MAE, the better
is the predictor

The debate on the best evaluation method between traditional and AI-driven ones is still open, and the present study has shown that neural networks are the most promising innovative methodology. However, the authors are aware that an adequate accuracy needs well-stocked databases with many features to be exploited commercially. Although such data are available for some cities, data availability is scarce in most cases, such as the one under consideration. Despite this, the AI can still understand much of the dynamics that govern prices. Therefore, the information obtained in this way can be used to support professionals and certainly still needs the supervision of human intervention.

Table 5.
Different distribution of prices between the third and fourth quartile

Quartile	min	Brescia Max	Range	Min	Varese Max	Range
3rd	25,000 €	281,250 €	256,250 €	35,000 €	352,500 €	317,500 €
4th	281,250 €	1,900,000 €	1,618,750 €	352,500	2,040,000 €	1,687,500 €

Table 6.
MAE comparison between the whole test dataset and the test entries considering only prices less than 800,000 €

Model	MAE score	Brescia Relative difference (%)	MAE score	Varese Relative difference (%)
ANN old test data	77,015 €	—	56,128 €	—
ANN new test data	68,365 €	11	49,251 €	12

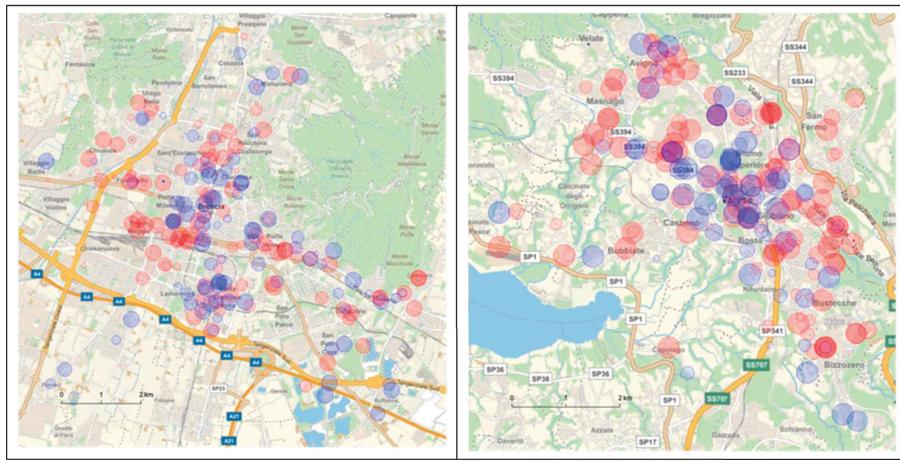


Figure 16.
Geographical distribution of the error for the city of (a) Brescia and (b) Varese

Note(s): Each circle represents a dataset record, the size and colour of the circle are proportional to the error made by the ANN in the prediction. The larger the circle and the darker the colour, the larger the error. In red the errors where the price is overestimated, in blue those where it is underestimated

References

- Abidoye, R.B. and Chan, A.P.C. (2017), "Artificial neural network in property valuation: application framework and research trend", *Property Management*, Vol. 35 No. 5, pp. 554-571, doi: [10.1108/PM-06-2016-0027](https://doi.org/10.1108/PM-06-2016-0027).
- Abidoye, R.B., Chan, A.P.C., Abidoye, F.A. and Oshodi, O.S. (2019), "Predicting property price index using artificial intelligence techniques: evidence from Hong Kong", *International Journal of Housing Markets and Analysis*, Vol. 12 No. 6, pp. 1072-1092, doi: [10.1108/IJHMA-11-2018-0095](https://doi.org/10.1108/IJHMA-11-2018-0095).
- Abiodun, O.I., Jantan, A., Omolara, A.E., Dada, K.V., Mohamed, N.A. and Arshad, H. (2018), "State-of-the-art in artificial neural network applications: a survey", *Heliyon*, Vol. 4 No. 11, p. e00938, doi: [10.1016/j.heliyon.2018.e00938](https://doi.org/10.1016/j.heliyon.2018.e00938).
- Adair, A.S., Berry, J.N. and McGreal, W.S. (1996), "Hedonic modelling, housing submarkets and residential valuation", *Journal of Property Research*, Vol. 13 No. 1, pp. 67-83, doi: [10.1080/095999196368899](https://doi.org/10.1080/095999196368899).
- Arlot, S. and Celisse, A. (2010), "A survey of cross-validation procedures for model selection", *Statistics Surveys*. doi: [10.1214/09-SS054](https://doi.org/10.1214/09-SS054).
- Baldi, P., Brunak, S., Chauvin, Y., Andersen, C.A.F. and Nielsen, H. (2000), "Assessing the accuracy of prediction algorithms for classification: an overview", *Bioinformatics*, Vol. 16 No. 5, pp. 412-424, doi: [10.1093/bioinformatics/16.5.412](https://doi.org/10.1093/bioinformatics/16.5.412).
- Baldini, G., Barboni, M., Bono, F., Delipetrev, B., Duch Brown, N., Fernandez Macias, E., Gkoumas, K., Joossens, E., Kalpaka, A., Nepelski, D., Nunes de Lima, M.V., Pagano, A., Prettico, G., Sanchez, I., Sobolewski, M., Triaille, J.P., Tsakalidis, A. and Urzi Brancati, M.C. (2019), *Digital Transformation in Transport, Construction, Energy, Government and Public Administration*, Publications Office of the European Union, Luxembourg.
- Castelvecchi, D. (2016), "Can we open the black box of AI?", *Nature*, Vol. 538 No. 7623, pp. 20-23, doi: [10.1038/538020a](https://doi.org/10.1038/538020a).
- Chai, T. and Draxler, R.R. (2014), "Root mean square error (RMSE) or mean absolute error (MAE)? -Arguments against avoiding RMSE in the literature", *Geoscientific Model Development*, Vol. 7 No. 3, pp. 1247-1250, doi: [10.5194/gmd-7-1247-2014](https://doi.org/10.5194/gmd-7-1247-2014).
- Chollet, F. (2018), *Deep Learning with Python*, Manning, Shelter Island, New York.
- Dietterich, T. (1995), "Overfitting and undercomputing in machine learning", *ACM Computing Surveys*, Vol. 27 No. 3, pp. 326-327, doi: [10.1145/212094.212114](https://doi.org/10.1145/212094.212114).
- Espey, M. and Lopez, H. (2000), "The impact of airport noise and proximity on residential property values", *Growth and Change*, John Wiley & Sons, Vol. 31 No. 3, pp. 408-419, doi: [10.1111/0017-4815.00135](https://doi.org/10.1111/0017-4815.00135).
- Friedman, J.H. (2001), "Greedy function approximation: a gradient boosting machine", *Annals of Statistics*, Vol. 29 No. 5, pp. 1189-1232, doi: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451).
- Gao, G., Bao, Z., Cao, J., Qin, A.K., Sellis, T., Fellow, IEEE and Wu, Z. (2019), "Location-centered house price prediction: a multi-task learning approach", *ArXiv*, available at: <http://arxiv.org/abs/1901.01774>.
- García-Magariño, I., Medrano, C. and Delgado, J. (2020), "Estimation of missing prices in real-estate market agent-based simulations with machine learning and dimensionality reduction methods", *Neural Computing and Applications*, Vol. 32 No. 7, pp. 2665-2682, doi: [10.1007/s00521-018-3938-7](https://doi.org/10.1007/s00521-018-3938-7).
- Ghojogh, B. and Crowley, M. (2019), "The theory behind overfitting, cross validation, regularization, bagging, and boosting: tutorial", arXiv.
- Grosan, C. and Abraham, A. (2011), "Artificial neural networks", in *Intelligent Systems Reference Library*, Vol. 17, pp. 281-323, doi: [10.1007/978-3-642-21004-4_12](https://doi.org/10.1007/978-3-642-21004-4_12).

- Guresen, E., Kayakutlu, G. and Daim, T.U. (2011), "Using artificial neural network models in stock market index prediction", *Expert Systems with Applications*, Vol. 38 No. 8, pp. 10389-10397, doi: [10.1016/j.eswa.2011.02.068](https://doi.org/10.1016/j.eswa.2011.02.068).
- Ho, W.K.O., Tang, B.-S. and Wong, S.W. (2020), "Predicting property prices with machine learning algorithms", *Journal of Property Research*, Vol. 33 No. 1, pp. 48-70, doi: [10.1080/09599916.2020.1832558](https://doi.org/10.1080/09599916.2020.1832558).
- Huang, H. and Yin, L. (2015), "Creating sustainable urban built environments: an application of hedonic house price models in Wuhan, China", *Journal of Housing and the Built Environment*, Vol. 30 No. 2, pp. 219-235, doi: [10.1007/s10901-014-9403-8](https://doi.org/10.1007/s10901-014-9403-8).
- Jiang, F., Jiang, Y., Zhi, H., Dong, Y., Li, H., Ma, S., Wang, Y., Dong, Q., Shen, H. and Wang, Y. (2017), "Artificial intelligence in healthcare: past, present and future", *Stroke and Vascular Neurology*, Vol. 2 No. 4, pp. 230-243, doi: [10.1136/svn-2017-000101](https://doi.org/10.1136/svn-2017-000101).
- Kaggle (2019), *Kaggle: Your Machine Learning and Data Science Community*, available at: <https://www.kaggle.com/> (accessed 1 August 2021).
- Keskin, B. (2008), "Hedonic analysis of price in the Istanbul housing market", *International Journal of Strategic Property Management*, Vol. 12 No. 2, pp. 125-138, doi: [10.3846/1648-715X.2008.12.125-138](https://doi.org/10.3846/1648-715X.2008.12.125-138).
- Lancaster, K. (1966), "A new approach to consumer theory", *Journal of Political Economy*, Vol. 74, pp. 132-157.
- LeCun, Y., Bengio, Y. and Hinton, G. (2015), "Deep learning", *Nature*, Vol. 521 No. 7553, pp. 436-444, doi: [10.1038/nature14539](https://doi.org/10.1038/nature14539).
- Lee, C.C., Liang, C.M. and Chen, C.Y. (2017), "The impact of urban renewal on neighborhood housing prices in Taipei: an application of the difference-in-difference method", *Journal of Housing and the Built Environment*, Vol. 32 No. 3, pp. 407-428, doi: [10.1007/s10901-016-9518-1](https://doi.org/10.1007/s10901-016-9518-1).
- Lenk, M.M., Worzala, E.M. and Silva, A. (1997), "High-tech valuation: should artificial neural networks bypass the human valuer?", *Journal of Property Valuation and Investment*, Emerald, Vol. 15 No. 1, pp. 8-26, doi: [10.1108/14635789710163775](https://doi.org/10.1108/14635789710163775).
- Limsombunc, V., Gan, C. and Lee, M. (2004), "House price prediction: hedonic price model vs Artificial neural network", *American Journal of Applied Sciences*, Science Publications, Vol. 1 No. 3, pp. 193-201, doi: [10.3844/AJASSP.2004.193.201](https://doi.org/10.3844/AJASSP.2004.193.201).
- Lin, C.C. and Mohan, S.B. (2011), "Effectiveness comparison of the residential property mass appraisal methodologies in the USA", *International Journal of Housing Markets and Analysis*, Vol. 4 No. 3, pp. 224-243, doi: [10.1108/17538271111153013](https://doi.org/10.1108/17538271111153013).
- Lu, Y. (2019), "Artificial intelligence: a survey on evolution, models, applications and future trends", *Journal of Management Analytics*, Vol. 6 No. 1, pp. 1-29, doi: [10.1080/23270012.2019.1570365](https://doi.org/10.1080/23270012.2019.1570365).
- Luttik, J. (2000), "The value of trees, water and open space as reflected by house prices in The Netherlands", *Landscape and Urban Planning*, Vol. 48 Nos 3-4, pp. 161-167, doi: [10.1016/S0169-2046\(00\)00039-6](https://doi.org/10.1016/S0169-2046(00)00039-6).
- Mohd, T., Masrom, S. and Johari, N. (2019), "Machine learning housing price prediction in Petaling Jaya, Selangor, Malaysia", *International Journal of Recent Technology and Engineering*, Vol. 8 No. 2S11, pp. 542-546, doi: [10.35940/ijrte.B1084.0982S119](https://doi.org/10.35940/ijrte.B1084.0982S119).
- Mrsic, L., Jerkovic, H. and Balkovic, M. (2020), "Real estate market price prediction framework based on public data sources with case study from Croatia", *Communications in Computer and Information Science*, doi: [10.1007/978-981-15-3380-8_2](https://doi.org/10.1007/978-981-15-3380-8_2).
- Olson, M., Wyner, A.J. and Berk, R. (2018), "Modern neural networks generalize on small data sets", *Advances in Neural Information Processing Systems*.
- Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T. and French, N. (2003), "Real estate appraisal: a review of valuation methods", *Journal of Property Investment and Finance*, Vol. 21 No. 4, pp. 383-401.

- Park, B. and Kwon Bae, J. (2015), "Using machine learning algorithms for housing price prediction: the case of Fairfax County, Virginia housing data", *Expert Systems with Applications*, Vol. 42 No. 6, pp. 2928-2934, doi: [10.1016/j.eswa.2014.11.040](https://doi.org/10.1016/j.eswa.2014.11.040).
- Rafiei, M.H. and Adeli, H. (2016), "A novel machine learning model for estimation of sale prices of real estate units", *Journal of Construction Engineering and Management*, Vol. 142 No. 2, doi: [10.1061/\(asce\)co.1943-7862.0001047](https://doi.org/10.1061/(asce)co.1943-7862.0001047).
- Rahadi, R.A., Wiryono, S.K., Koesrindartoto, D.P. and Syamwil, I.B. (2015), "Factors influencing the price of housing in Indonesia", *International Journal of Housing Markets and Analysis*, Vol. 8 No. 2, pp. 169-188, doi: [10.1108/IJHMA-04-2014-0008](https://doi.org/10.1108/IJHMA-04-2014-0008).
- Rosen (1974), "Hedonic prices and implicit markets: product differentiation in pure competition sherwin Rosen the", *Tetrahedron Letters*.
- Selim, H. (2009), "Determinants of house prices in Turkey: hedonic regression versus artificial neural network", *Expert Systems with Applications*, Vol. 36 No. 2, pp. 2843-2852, doi: [10.1016/j.eswa.2008.01.044](https://doi.org/10.1016/j.eswa.2008.01.044).
- Spiegelhalter, D. (2019), "The art of statistics: learning from data", *Quantitative Finance*, Vol. 19 No. 8, pp. 1267-1268.
- Taffese, W.Z. (2006), "A survey on application of artificial intelligence in real estate industry", *Proceedings of the Third International Conference on Artificial Intelligence in Engineering and Technology [iCAiET]*.
- Tajani, F., Morano, P. and Ntalianis, K. (2018), "Automated valuation models for real estate portfolios: a method for the value updates of the property assets", *Journal of Property Investment and Finance*, Emerald Group Holdings, Vol. 36 No. 4, pp. 324-347, doi: [10.1108/JPIF-10-2017-0067](https://doi.org/10.1108/JPIF-10-2017-0067).
- Willmott, C.J. and Matsuura, K. (2005), "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance", *Climate Research*, Vol. 30 No. 1, pp. 79-82, doi: [10.3354/cr030079](https://doi.org/10.3354/cr030079).
- Worzala, E., Lenk, M. and Silva, A. (1995), "An exploration of neural networks and its application to real estate valuation", *Journal of Real Estate Research*, American Real Estate Society, Vol. 10 No. 2, pp. 185-201, doi: [10.1080/10835547.1995.12090782](https://doi.org/10.1080/10835547.1995.12090782).
- Wu, H. and Wang, C. (2018), "A new machine learning approach to house price estimation", *New Trends in Mathematical Science*, Vol. 4 No. 6, pp. 165-171, doi: [10.20852/ntmsci.2018.327](https://doi.org/10.20852/ntmsci.2018.327).
- Zhou, X. (2020), "The usage of artificial intelligence in the commodity house price evaluation model", *Journal of Ambient Intelligence and Humanized Computing*, doi: [10.1007/s12652-019-01616-4](https://doi.org/10.1007/s12652-019-01616-4).
- Zou, H. and Hastie, T. (2005), "Regularization and variable selection via the elastic net", *Journal of the Royal Statistical Society. Series B: Statistical Methodology*, Vol. 67 No. 2, pp. 301-320, doi: [10.1111/j.1467-9868.2005.00503.x](https://doi.org/10.1111/j.1467-9868.2005.00503.x).
- Zurada, J.M., Levitan, A.S. and Guan, J. (2006), "Non-conventional approaches to property value assessment", *Journal of Applied Business Research*, Vol. 22 No. 3, doi: [10.19030/jabr.v22i3.1421](https://doi.org/10.19030/jabr.v22i3.1421).

Corresponding author

Luca Rampini can be contacted at: luca.rampini@polimi.it