



UNIVERSITÀ  
DEGLI STUDI  
DI PADOVA

*Università degli Studi di Padova*

*Padua Research Archive - Institutional Repository*

Who performs better? AVMs vs hedonic models

*Original Citation:*

*Availability:*

This version is available at: 11577/3339932 since: 2020-05-12T09:52:05Z

*Publisher:*

Emerald Group Publishing Ltd.

*Published version:*

DOI: 10.1108/JPIF-12-2019-0157

*Terms of use:*

Open Access

This article is made available under terms and conditions applicable to Open Access Guidelines, as described at <http://www.unipd.it/download/file/fid/55401> (Italian only)

(Article begins on next page)

# AUTHOR QUERIES

**Note:** *It is crucial that you NOT make direct edits to the PDF using the editing tools as doing so could lead us to overlook your desired changes. Edits should be made via the 'Comments' feature.*

## AUTHOR PLEASE ANSWER ALL QUERIES

- AQ1—Please confirm that the **given name(s)** and **surname(s)** have been identified correctly and are presented in the desired order, and please carefully verify the spelling of all authors' names.
- AQ2—Please check the accuracy of the affiliation(s) of each author and make changes as appropriate. Please ensure to include City and Country names while revising the affiliation as these are mandatory information per style.
- AQ3—Please confirm that the provided email “valier.agostino@gmail.com” is the correct address for official communication.
- AQ4—Please check the short title created for this article (seen on the right side top of the first page).
- AQ5—Reference McCluskey and Anand, 1999 is cited in the text but not provided in the reference list. Please provide it in the reference list or alternatively delete the citations from the text.
- AQ6—There is currently no acknowledgements and funding included. Please confirm if this is correct or provide the acknowledgements and funding section
- AQ7—Please provide the page number(s) for the reference “Mu *et al.*, 2014; Mangialardo *et al.*, 2019; Huang, 2019.

AQ: 4

# Who performs better? AVMs vs hedonic models

AVMs vs  
hedonic models

AQ: 1

Agostino Valier

*Department of Civil Building and Environmental Engineering,  
Universita degli Studi di Padova, Padova, Italy*

AQ: 2

## Abstract

**Purpose** – In the literature there are numerous tests that compare the accuracy of automated valuation models (AVMs). These models first train themselves with price data and property characteristics, then they are tested by measuring their ability to predict prices. Most of them compare the effectiveness of traditional econometric models against the use of machine learning algorithms. Although the latter seem to offer better performance, there is not yet a complete survey of the literature to confirm the hypothesis.

**Design/methodology/approach** – All tests comparing regression analysis and AVMs machine learning on the same data set have been identified. The scores obtained in terms of accuracy were then compared with each other.

**Findings** – Machine learning models are more accurate than traditional regression analysis in their ability to predict value. Nevertheless, many authors point out as their limit their black box nature and their poor inferential abilities.

**Practical implications** – AVMs machine learning offers a huge advantage for all real estate operators who know and can use them. Their use in public policy or litigation can be critical.

**Originality/value** – According to the author, this is the first systematic review that collects all the articles produced on the subject done comparing the results obtained.

**Keywords** Real estate, Mass appraisal, Valuation, AVM, Automated valuation models, Machine learning, Econometric model

**Paper type** Literature review

Received 18 December 2019

Revised 20 January 2020

Accepted 24 January 2020

## Introduction

Artificial intelligence is bringing about a radical change in many activities traditionally carried out by human work: among them, real estate valuation. Innovation affects the nature of evaluations, operational procedures and the skills required of the professional sector (Rics, 2017).

Frey and Osborne (2017) have carried out an extensive survey that assigns to each profession the degree of possible computerization, that is, the possibility that the work currently done by man can be entirely replaced by the work of a machine. In this survey, the profession of real estate valuers is estimated to be susceptible to computerization at 90%. The whole field of evaluation therefore wonders what the future of estimates will be, what impact automatic value prediction models will have on professional evaluation practice (Cook, 2015).

Automated value prediction models are gradually replacing the evaluator's work. In the past, these models only used regression analysis. Now these models are improved by self-learning algorithms. The new learning techniques are able to provide predictions with a very high degree of accuracy.

Many stakeholders are interested in the use of machine learning models in mass appraisal: among them, real estate companies, public authorities, banks and so on. The use of these new techniques requires the inclusion, within the traditional evaluation groups, of new professional figures such as data analysts.

Some real estate companies already successfully use machine learning models in estimates. The best known case is the home valuation model Zestimate© by the American agency Zillow. It is not yet known, at least to the author's knowledge, if machine learning models are used in public policies. For banks the possibility of using self-learning algorithms was introduced by the Basel II Accord in 2004. It allows the use of statistical methods to



---

monitor the value of real estate and identify those that need verification. However, it does not specify the nature of these statistical methods, the size and characteristics of the data sets to be used.

On the one hand, machine learning models provide rapid, reliable and low-cost estimates. On the other hand, these models are often black boxes that are difficult to be controlled. The scientific literature has also dealt extensively with the use of machine learning algorithms within automatic value prediction models, inaugurating a critical debate on the potential for use and limits of these models.

### *Econometric models and machine learning models*

Machine learning models in real estate valuation can be used in mass appraisal techniques. Mass appraisal techniques are defined by the IAAO (International Association of Assessing Officers) as the process of valuing a group of properties as of a given date and using common data, standardized methods and statistical testing (IAAO, 2013).

The analysis profiles with which scientific production deals with self-learning models in mass appraisal evaluations can be traced back to three: theoretical, methodological and empirical. From a theoretical point of view, the economic theories on which the automatic value estimation procedures are based are investigated (Mooya, 2009, 2017). Methodological research proposes new evaluation models or proposes a classification of existing ones (D'Amato and Kauko, 2017; Glumac and Des Roisiers, 2018). Finally, empirical research tests evaluation models on real estate data data sets – sale prices or asking prices – quantifying the forecasting capacity of the models. Accuracy measurement is never an end in itself, but is the starting point from which to reflect on estimation models and their possibilities of use.

The comparative testing trend is by far the most widespread in the literature on automated valuation model (AVM) machine learning. The measure of predictive capacity always adopts the same protocol. The data set available to authors is divided into two parts: the training set and the testing set. The first set includes 70–80% of the total data and is used for the training phase of the model, in which the computer works with the input data  $x$  (the property features) and the output data  $y$  (the final prices). The computer identifies the function that best explains the value dependent variable. The remaining part of the data set (the testing set) is used to test the obtained model. The input data ( $x$ ) of the testing set are processed by the algorithm formed on the training set, then the output values provided by the model ( $\hat{y}$ ) are compared with the output values of the testing set ( $y$ ). The smaller the difference between  $\hat{y}$  and  $y$ , the more the model can be declared effective in its ability to predict the value.

There are numerous statistical indicators measuring the difference between the predicted values and the actual values. The choice of this indicator is not a marginal choice, in many researches (e.g. Lasota *et al.*, 2009) the order in which models are distributed according to their forecasting capacity varies according to the indicator considered.

Although there are also articles that only compare machine learning models with each other, most comparison articles compare machine learning models with the traditional econometric model of hedonic prices. Multiparametric regression analysis is in fact the most widely used for mass appraisal techniques. It determines the extent to which each variable contributes to the variation of the final price, assigning each one a numerical coefficient.

The comparison of machine learning models with the regression analysis does not only represent the comparison between different techniques. It can be seen as the comparison between artificial intelligence and human intelligence, between the estimate carried out by machine and the manual estimate. Many authors assume that the regression analysis reflects the traditional process that the evaluator performs when estimating assets.

Kauko and D'Amato (2008) in their scientific production use an effective terminology to name the two classes just described. On the one hand, the “orthodox” models, which use a

---

hedonic approach, quantify the relationship between the price of the property and its characteristics. On the other hand, “heretical” models, which instead adopt a statistical approach, read the patterns emerging from the distribution of data. For the terminology used in this article, I will use the subdivision between traditional models and machine learning models.

### *General literature review*

There are different types of machine learning models. The debate has always focused on artificial neural networks (ANNs), devoting much more attention to them than to other algorithms. [Borst \(1991\)](#) is the first to quantify the capacity of ANNs to provide reliable estimates. [Do and Grudnitski \(1992\)](#) are instead the first to inaugurate the comparative researches of more models on the same data set. They test the superior effectiveness of neural networks on multiple regression in recognizing the price of 136 homes in San Diego. In many other subsequent papers the performance of ANNs exceeds, sometimes in significant terms, that of traditional models ([Amri and Tularam, 2012](#); [Kutasi and Badics, 2016](#)).

The common enthusiasm for the predictive capacity of neural networks comes to a halt with the research of [Worzala et al. \(1995\)](#). It is the first research to criticize the works of [Borst \(1991\)](#) and of [Do and Grudnitski \(1992\)](#). The authors work on a set of 288 houses in Fort Collins, therefore a larger sample of data than the two previous researches. They compare neural networks/multiple regression on three samples: the complete sample (case 1), the sample consisting of cases that fell within the price range analyzed by Do and Grudnitski (case 2) and finally – to compare with the case of Borst, who had used very similar goods between them – a sample consisting of houses belonging to the same postal code (case 3). Case 1 found values of accuracy almost identical for both models. In case 2, the performance classification varied according to the type of software used. Only in case 3 – therefore with a very homogeneous data sample – the neural networks performed better than the regression analysis. The authors therefore question the absolute superiority of the neural networks over the traditional models. They make such superiority correspond to specific conditions of the data set or of the software employed.

Similar conclusions have been reached by [Lenk et al. \(1997\)](#), [McGreal et al. \(1998\)](#), [McCluskey et al. \(2013\)](#). [Nguyen and Cripps \(2001\)](#) effectively study the relationship that binds neural networks to the amount of data available to them. Using a data set of 3,906 observations and 108 times the comparison with data sets of different sizes, they show that ANNs exceed the predictive capacity of multiple regression only when the sample is of medium-large size.

Algorithms by analogy work by researching the behavior of cases similar to the case under investigation, in order to predict the behavior of the latter. They are defined as lazy forms of learning. The most known are the  $k$ -nearest neighbors, where the investigated variable – the price – is the average of the values that the variable assumes in the number  $k$  of the closest cases. [Isakson \(1988\)](#) uses the technique to predict the value of 143 real estate properties in Dallas divided between apartment, industrial, office and retail. In all four types the nearest neighbors achieve better performance than the OLS (Ordinary Least Squares) method. In the same contribution, however, the author notes that the technique is effective in cases where the value to be predicted has characteristics close to the average of the available data, while it proves more inadequate in cases where the object to be evaluated is a statistical outlier. In other research, the  $k$ -nearest neighbors record values of accuracy lower than those obtained from other models ([Borde et al., 2017](#)).

These models use complex geometries, more able to work on multidimensionality than the traditional Euclidean geometry. Nevertheless, they present strong criticalities when the number of data – and therefore the number of sizes – increases ([Cover and Hart, 1967](#)). This

---

has led to limiting their use in mass appraisal models, although they prove effective in some phases of the evaluation: McCluskey and Anand (1999) use them to identify the most significant comparable within a hybrid model, then entrusting the determination of the price to neural networks and genetic algorithms.

The criticalities found in the nearest neighbors led to a gradual abandonment of the techniques of work on hyperspaces, until the introduction of the support vector machines (SVM) introduced by Vapnik. Their applications in real estate valuation models are many and generally show great effectiveness in predicting value. Kontrimas and Verikas (2011) identify the Support Vector Regression (SVR) as the most effective predictor of value. Similar conclusions are reached by numerous other researches, mostly coming from the Asian area (Zhang, 2012; Yeh *et al.*, 2013; Mu *et al.*, 2014; Wang *et al.*, 2014; Huang, 2019). However, it is not possible to assign to the SVM the role of best algorithm for the absolute evaluation. It always depends on the nature of the data available. For example, in their comparison with ANNs, some authors identify them as more effective (Lam *et al.*, 2009), others as less effective (Abidoye *et al.*, 2019).

Finally, the genetic algorithms are based on the same principle as the fitness function of Darwin. The Italian academic world is giving an important contribution in testing genetic algorithms in real estate evaluation. These models have in fact proved effective in predicting the value of real estate in Naples (Del Giudice *et al.*, 2017), Potenza (Manganelli *et al.*, 2015) and Bari, Naples and Rome (Morano *et al.*, 2018).

All the authors point out that all the models described so far present, against an undeniable predictive effectiveness, an element of criticality that lies in their character of black box. It is difficult to observe the role that the single parameters play in the variation of the value, defining in numerical coefficients the causal relationships between the prices and the characteristics of the assets (Yacim and Boshoff, 2018). This is not true in the case of regression trees, for which it is possible to know the value assumed by the data in each step of the self-learning path. For this reason, they are also defined as white box models.

These models are very often at risk of being overfitted. To overcome this problem, training periods are limited or specific techniques such as pruning are used. Random forest models, which are models of ensemble learning resulting from the aggregation of several regression trees, are more successful. The models of ensemble learning combine several individual models within a single metamodel. In this case, many regression trees are combined in a random forest model, which offers better performance than that offered by each model considered individually (Graczyk *et al.*, 2010; Antipov and Pokryshevskaya, 2012).

In the research by Mullainathan and Spiess (2017), decision trees show less predictive capacity than regression analysis, which is in turn outdated by random forest models. In the estimate of 7,400 residential transactions in the city of Ljubljana, the coefficient of determination ( $R^2$ ) recorded by random forests is 34 percentage points higher than that obtained by the method of least squares (Ceh *et al.*, 2018). Kok *et al.* (2017) use data from 36,000 single houses in California, Florida and Texas to test random forest models. They apply the model in three different cases. In the first two cases the model predicts the market value. In the third case it predicts the values of NOI (Net Operating Income). In the second case, moreover, the NOI values of the assets have been inserted within the real estate data necessary to predict the market value. Random forest models have proved more effective than the minimum square method in the first and third cases. Therefore, when NOI is included in the input data, regression analysis is even more effective.

Reading the results from time to time reported would seem to confirm a greater ability to predict self-learning models than the traditional econometric approach. However, there is no research that can empirically confirm this hypothesis. There is a need of a survey on a data set of articles as broad and representative as possible, quantifying the results emerging from the literature produced.

**Methodology**

This research aims to quantify whether and how machine learning models in the literature have been more accurate than traditional models. In order to answer this question, only one type of article was analyzed: those testing on the same data set regression analysis and machine learning models. The results of the comparisons were then reported in a table that counts the cases of higher accuracy and those of lower accuracy.

The identification of the set of articles was made on the Scopus online database. This was chosen, because it represents one of the most complete and reliable database in the field of estimation disciplines. The preliminary study of the literature on the subject of self-learning algorithms allowed the authors to identify the words that recur most frequently in abstracts and article titles. These terms were then used in research on Scopus. The total search strings used are 36. They're the result of a combination of a term delimiting the operational scope of real estate, four terms indicating evaluation practice and nine terms describing the models of self-learning. These words have been combined with each other through the appropriate

T1 Boolean logical connectors as can be seen in the table (Table I).

The Scopus research, which was updated to July 2019, identified a total of 381 articles. The elimination of the numerous duplicates then reduced the sample to 165 articles.

Subsequently, the data set was limited to 40 articles containing the application of one or more machine learning evaluation models and a regression analysis model (lin, log-lin or log log) tested on the same data set of real estate data. The data set of 40 articles was subsequently expanded with the technique of bibliographic snowballing, that is, the inclusion of new contributions that were found within the lists of bibliographic references of other articles. In fact, many of the articles read referred to the content of other papers not present in the texts initially identified. The most significant bibliographic references were therefore added, enriching the sample with 13 units, for a total of 53 articles used.

For each article, the results obtained by the models were reported in terms of accuracy. They were also read in full, with particular attention to the conclusions reached by the authors of each paper. The reflections contained in each paper were used as the basis on which to base the conclusions of this research.

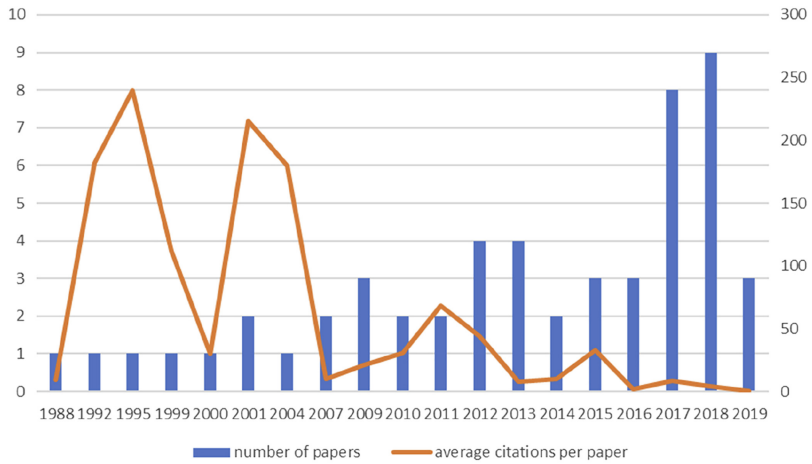
*Data description*

F1 **Figure 1** shows the distribution of 53 articles per year (columns in the graph) and the average citations per article (bars). The average citations per article have been calculated by dividing the sum of all citations obtained from articles published during the year by the number of articles of the year. (source: Scopus)

The graph shows that, as the number of articles per year increases, the average number of citations received by articles decreases. This shows that in literature, experimentation with machine learning models is increasing year after year. On the other hand, however, the critical

|               |     |  |     |  |
|---------------|-----|--|-----|--|
| "Real estate" | AND | (1) "Valuation*"<br>(2) "Appraisal*"<br>(3) "Automated valuation model*"<br>OR "avm"<br>(4) "Price* forecast*" | AND | (1) "Machine learning"<br>(2) "Artificial intelligence" OR "ai"<br>(3) "Decision tree*" OR "regression tree*"<br>(4) "Neural network*" OR "ann"<br>(5) "Backpropagation"<br>(6) "Genetic algorithm*"<br>(7) "Bayes*"<br>(8) "Nearest neighbour*" OR "knn"<br>(9) "Vector machine" OR "svm" |
|---------------|-----|--|-----|--|

**Table I.**  
Strings of search used in Scopus



**Figure 1.** Paper per year and average citations per papers

Source(s): Scopus

debate on models is decreasing. Recently published articles are not mentioned in other articles, nor is their content a subject for reflection.

In the content of the article, [Table II](#) shows statistics on the size of the data sets used. The number and characteristics of the variables used could not also be reported. This is because many articles did not report the characteristics of the variables or were so inconsistent that it was difficult to produce summary statistics. The data sets are around a few hundred or a few thousand cases.

[Table III](#) shows which statistical indicators the articles use to measure the distance between the predicted value (y) and the actual value (y). Many articles use more than one statistical indicator to measure the same comparison.

### Results

All the results obtained have been reported in a table. The table, for each type of algorithm, provides the number of articles in which they have more or less predictive capacity compared to any other type of model.

Each unit of the table represents an examined article. The unit is placed in the column > when that class of algorithms placed in the row has recorded greater predictive capacity with respect to the class of algorithms of the corresponding column. Vice versa, the unit is marked in the column <. If the comparison varies according to the statistical indicator used, the results obtained by using the indicator most frequently used in the literature, as in [Table III](#), have been reported.

**Table II.** Dimensions of the data sets

|                |         |
|----------------|---------|
| Count          | 53      |
| Mean           | 11.088  |
| SD             | 20.957  |
| minimum        | 40      |
| Lower quartile | 300     |
| Median         | 2.266   |
| Upper quartile | 9.795   |
| Maximum        | 90.275  |
| Sum            | 587.679 |



T4 The machine learning algorithms have been synthesized in five categories: tree, neural networks, genetic algorithm, nearest neighbors, support vector machines (Table IV).

The results obtained allow an objective and quantitatively verifiable identification of which models have been indicated in the literature as effective and reliable for the mass evaluation of real estate assets.

T5 Table V shows the superiority of machine learning models over traditional models. The number of comparisons in which the values obtained by machine learning were more accurate than those found with regression analysis are four times the number of comparisons in which they showed less accuracy.

Within the machine learning models, it is not possible to draw up a ranking in order of accuracy. Limiting to the papers reported in this research, SVMs “win” on regression 8 to 1, neural networks 29 to 6, trees 12 to 3, while *k*-nearest neighbors “equalize” 3 to 3.

Reading the table offers a wide range of information. The most frequent comparison is between ANNs and regression analysis. This is also determined by the appearance of neural

|                       | Description                    |   | Min value | Max value | Desirable outcome | <i>n</i> ° of papers |
|-----------------------|--------------------------------|---|-----------|-----------|-------------------|----------------------|
| MAPE                  | Mean Absolute Percentage Error | $\frac{1}{N} \sum_{i=1}^N \frac{ y_i - \hat{y}_i }{y_i} \cdot 100\%$  | 0         | 100%      | minute            | 21                   |
| RMSE                  | Root Mean Squared Error        | $\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}$   | 0         | $\infty$  | minute            | 20                   |
| MAE                   | Mean Absolute Error            | $\frac{1}{N} \sum_{i=1}^N  y_i - \hat{y}_i $  | 0         | $\infty$  | minute            | 18                   |
| <i>R</i> <sup>2</sup> | Coefficient of determination   | $\frac{\sum_{i=1}^N (\hat{y}_i - avg(y))^2}{\sum_{i=1}^N (y_i - avg(y))^2}$   | 0         | 100%      | Closer to 100%    | 17                   |
| MSE                   | Mean Squared error             | $\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2$  | 0         | $\infty$  | minute            | 10                   |
| RSS                   | Residual Sum of Squares        | $\sum_{i=1}^N (y_i - \hat{y}_i)^2$  | 0         | $\infty$  | minute            | 5                    |
| RAE                   | Relative Absolute Error        | $\frac{\sum_{i=1}^N  y_i - \hat{y}_i }{\sum_{i=1}^N  y_i - avg(y) }$  | 0         | $\infty$  | minute            | 3                    |
| r                     | Linear correlation coefficient | $\frac{\sum_{i=1}^N (y_i - avg(y))(\hat{y}_i - avg(\hat{y}))}{\sqrt{\sum_{i=1}^N (y_i - avg(y))^2} \sqrt{\sum_{i=1}^N (\hat{y}_i - avg(\hat{y}))^2}}$ | -1        | 1         | Closer to 1       | 3                    |
| COD                   | Coefficient of Dispersion      | $\frac{100}{avg(y)} \cdot \frac{\sum_{i=1}^N  \hat{y}_i - avg(y) }{N}$  | 0         | 100%      | minute            | 3                    |
| RSE                   | Relative Squared Error         | $\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - avg(y))^2}$  | 0         | $\infty$  | minute            | 2                    |
| RRSE                  | Root Relative Squared Error    | $\sqrt{\frac{\sum_{i=1}^N (y_i - \hat{y}_i)^2}{\sum_{i=1}^N (y_i - avg(y))^2}}$   | 0         | $\infty$  | minute            | 2                    |
| NDEI                  | Nondimensional Error Index     | $\frac{\sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - \hat{y}_i)^2}}{std(y)}$  | 0         | $\infty$  | minute            | 2                    |

**Table III.**  
Accuracy scores

| Manual appraisal                     | Artificial intelligence   |
|--------------------------------------|---|
| HEDONIC MODELS (regression analysis) | -TREES (regression trees and random forest)<br>-ANN (artificial neural networks)<br>-NN (nearest neighbors)<br>-SVM (support vector machine)<br>-GENETIC (genetic algorithms) |

**Table IV.**  
The types of algorithms

Table V.  
Results

|                     | Regression analysis |   | Trees |   | Neural networks |   | N-neighbors |   | SVM |   | Genetic algorithms |    |
|---------------------|---------------------|---|-------|---|-----------------|---|-------------|---|-----|---|--------------------|----|
| Regression analysis | >                   | < | >     | < | >               | < | >           | < | >   | < | >                  | <  |
| Trees               | 12                  | 3 | 12    | 6 | 29              | 3 | 3           | 1 | 8   | 0 | 5                  | 13 |
| Neural networks     | 3                   | 3 | 6     | 6 | 3               | 4 | 0           | 0 | 4   | 0 | 0                  | 57 |
| N-neighbors         | 6                   | 6 | 6     | 1 | -               | 4 | 1           | 1 | 8   | 0 | 0                  | 24 |
| SVM                 | 3                   | 4 | 4     | 8 | 4               | 1 | 1           | 1 | 1   | 0 | 0                  | 10 |
| Genetic algorithms  | 8                   | 0 | 2     | 0 | 1               | 0 | 0           | 0 | 1   | 0 | 0                  | 37 |
|                     | 5                   | 0 | 0     | 0 | 1               | 0 | 1           | 0 | 0   | 0 | 0                  | 21 |
|                     | 5                   | 0 | 0     | 0 | 0               | 0 | 0           | 0 | 0   | 0 | 0                  | 5  |
|                     | 0                   | 0 | 0     | 0 | 0               | 0 | 0           | 0 | 0   | 0 | 0                  | 12 |
|                     | 0                   | 0 | 0     | 0 | 0               | 0 | 0           | 0 | 0   | 0 | 0                  | 21 |
|                     | 0                   | 0 | 0     | 0 | 0               | 0 | 0           | 0 | 0   | 0 | 0                  | 5  |
|                     | 0                   | 0 | 0     | 0 | 0               | 0 | 0           | 0 | 0   | 0 | 0                  | 5  |
|                     | 0                   | 0 | 0     | 0 | 0               | 0 | 0           | 0 | 0   | 0 | 0                  | 0  |

networks, within the scientific debate, since the 1940s. Regression trees almost never prove effective. In cases where the “tree” category has greater accuracy, they are random forest algorithms. SVR has only recently appeared on the scene. Nevertheless, it shows very good results so far. It is still too early to say that SVM will replace neural networks, although the results obtained from this research seem to demonstrate this. Genetic algorithms are proving always more effective than regression, although five articles are too few to demonstrate absolute superiority.

It cannot be said with certainty that a machine learning algorithm is always more effective than another machine learning algorithm. The effectiveness of each model is related to the characteristics of the data. Only in the case of the  $k$ -nearest neighbors the self-learning model does not prove to be an effective way.

Reading the table confirms that machine learning models are more effective than the traditional hedonic approach. However, considering these results as the only results of this research would be limiting. The results in the table are partial results, since they only concern the accuracy characteristic.

Models can be evaluated according to two capacities: inferential capacity and predictive capacity. The first consists in the model's ability to identify cause–effect relationships between explained variables and independent variables. The second lies in the model's ability to process output results corresponding to the value of real data. Almost all the conclusions obtained by the authors confirm that the econometric approach has good inferential capacities and poor predictive capacities. Machine learning models, vice versa (Baldominos *et al.*, 2018; Pérez-Rave *et al.*, 2019).

The traditional models of regression – linear or logarithmic – are inferential procedures: they explain the cause–effect relationship that the independent variables have on the dependent variable. Their inferential power makes them valid tools on which to base inductive processes to describe generally valid behaviors on a given statistical population (Mangialardo *et al.*, 2019).

The main objective of machine learning models is to predict the value of  $y$ . The identification of the line (straight or curved) that explains the relationship between  $y$  and each variable is of less importance. It is not possible to derive inferential hypotheses for lack of numerical coefficients  $\beta$ . Taking up the expression of Mullainathan and Spiess (2017): the first estimate  $\beta$ , the second estimate  $y$ .

The main limit of machine learning lies in overfitting. Its powerful predictive capacity often runs the risk of being ineffective when confronted with new data, different from those with which it has trained. To overcome this critical situation, techniques of regularization and refinement of parameters are used. At the same time, each model is only valid for the data with which it was designed. The addition of new data or the modification of existing data will not necessarily result in the model retaining or increasing its predictive capacity. It may, on the contrary, see it decrease.

## Conclusion

The research wondered in what proportion in literature machine learning models had proved more effective than the traditional hedonic approach. Until now, literature cited this superiority but there was no research that systematically quantified this superiority, at least to the knowledge of the author. It found 57 cases in which artificial intelligence models were more accurate in predicting value, compared to 13 cases in which regression performed better. Within the machine learning models, it is not possible to draw up a classification in order of accuracy. If for predictive characteristics the machine learning AVMs are clearly more effective, on the front of the inferential capacity they are less so. This has significant repercussions in the field of operational use.

---

Each valuer, when defining a mass appraisal model, is at a crossroads: on the one hand, traditional econometric models and, on the other, machine learning models. Although both share the same objective – estimating market value – their methods of use and their characteristics are very different.

Machine learning models are data-driven models: the form they take and their effectiveness depend entirely on the data available to them. This makes them difficult to use for public policies, where the evaluation process must guarantee fairness of treatment for all the cases concerned and maintain the same efficiency over time. The self-learning models are not able to guarantee the same requirements of accuracy in the face of the arrival of new data to be estimated. This could lead to complaints from individuals who feel damaged by their assessments.

On the other hand, the high performance achieved in forecasting real estate prices makes machine learning models attractive to all operators who evaluate, manage or trade real estate assets. Investors can use them to evaluate possible investments or transactions of which they are a party. Similarly, valuation service providers can use self-learning algorithms to offer reliable estimates to their clients. The creation of machine learning models will be possible only to those who have access to the information with which to train and optimize learning. Small independent evaluators are unlikely to have enough data and skills to create their own models. They will be able to take advantage of the services sold to them by the largest players in the industry. Technological innovation will therefore bring radical changes to the current structure of the professional sector of evaluators (Abidoye and Chan, 2017).

Finally, a conclusive element in favor of modern learning techniques. They are able to work with Big Data, not only in their vastness but also in their variety (Choi and Varian, 2012). In this research only variables expressed through numerical values or categories were considered. But machine learning models can also work with very different types of data: for example, photographic images. Numerous researches study evaluation processes using real estate photos (Poursaeed *et al.*, 2018). The use of new information sources such as images (including satellite images), the movements traced by the devices we use daily, the rating assigned to businesses – just to name a few – may prove to be effective predictors of value. They may also partially make up for the lack of information sources traditionally complained of by many in the real estate sector.

AQ: 6

## References

- Abidoye, R.B. and Chan, A.P.C. (2017), “Valuers’ receptiveness to the application of artificial intelligence in property valuation”, *Pacific Rim Property Research Journal*, Vol. 23 No. 2, pp. 175-193.
- Abidoye, R.B., Chan, A.P.C., Abidoye, F.A. and Oshodi, O.S. (2019), “Predicting property price index using artificial intelligence techniques”, *International Journal of Housing Markets and Analysis*, Vol. 12 No. 6, pp. 1072-1092.
- Amri, S. and Tularam, G.A. (2012), “Performance of multiple linear regression and nonlinear neural networks and fuzzy logic techniques in modelling house prices”, *Journal of Mathematics and Statistics*, Vol. 8 No. 4, pp. 419-434.
- Antipov, E.A. and Pokryshevskaya, E.B. (2012), “Mass appraisal of residential apartments: an application of random forest for valuation and a CART-based approach for model diagnostics”, *Expert Systems with Applications*, Vol. 39 No. 2, pp. 1772-1778.
- Baldominos, A., Blanco, I., Moreno, A., Iturrarte, R., Bernardez, O. and Afonso, C. (2018), “Identifying real estate opportunities using machine learning”, *Applied Sciences*, Vol. 8 No. 11, p. 2321.

- Borde, S., Rane, A., Shende, G. and Shetty, S. (2017), "Real estate investment advising using machine learning", *International Research Journal of Engineering and Technology*, Vol. 4, No. 3, pp. 1821-1825.
- Borst, R.A. (1991), "Artificial neural networks: the next modelling/calibration technology for the assessment community", *Property Tax Journal*, Vol. 10 No. 1, pp. 69-94.
- Ceh, M., Kilibarda, M., Lisec, A. and Bajat, B. (2018), "Estimating the performance of random forest versus multiple regression for predicting prices of the apartments", *ISPRS International Journal of Geo-Information*, Vol. 7 No. 5, p. 168.
- Choi, H. and Varian, H. (2012), "Predicting the present with google trends", *The Economic Record*, Vol. 88, pp. 2-9.
- Cook, D. (2015), "RICS futures: turning disruption from technology to opportunity", *Journal of Property Investment & Finance*, Vol. 33 No. 5, pp. 456-464.
- Cover, T.M. and Hart, P.E. (1967), "Nearest neighbor pattern classification", *IEEE Transactions on Information Theory*, Vol. 13 No. 1, pp. 21-27.
- D'Amato, M. and Kauko, T. (2017), *Advances in Automated Valuation Modeling*, Springer International Publishing, New York.
- Del Giudice, V., De Paola, P. and Forte, F. (2017), "Using genetic algorithms for real estate appraisals", *Buildings*, Vol. 7 No. 4, p. 31.
- Do, A.Q. and Grudnitski, G. (1992), "A neural network approach to residential property appraisal", *Real Estate Appraiser*, Vol. 58 No. 3, pp. 38-45.
- Frey, C.B. and Osborne, M.A. (2017), "The future of employment: how susceptible are jobs to computerisation?", *Technological Forecasting and Social Change*, Vol. 114, pp. 254-280.
- Glumac, B. and Des Roisiers, F. (2018), "Real estate and land property automated valuation systems: a taxonomy and conceptual model", working paper 09, Luxembourg Institute of socio-economic research, Luxembourg, April 2018.
- Graczyk, M., Lasota, T., Trawiński, B. and Trawiński, K. (2010), "Comparison of bagging, boosting and stacking ensembles applied to real estate appraisal", in Nguyen, N.T., Thanh, M. and Świątek, L. (Eds), *Intelligent Information and Database Systems*, Springer, New York, pp. 340-350.
- Huang, Y. (2019), "Predicting home value in California, United States via machine learning modeling", *Statistics, optimization and information computing*, Vol. 7 No. 1.
- IAAO (2013), "Standard on mass appraisal of real property", available at: <https://www.iaao.org/media/standards/StandardOnMassAppraisal.pdf> (accessed 4 December 2019).
- Isakson, H.R. (1988), "Valuation analysis of commercial real estate using the nearest neighbors appraisal technique", *Growth and Change*, Vol. 19 No. 2, pp. 11-24.
- Kauko, T. and D'Amato, M. (2008), *Mass Appraisal Methods: An International Perspective for Property Valuers*, Wiley-Blackwell, Hoboken.
- Kok, N., Koponen, E.L. and Martínez-Barbosa, C.A. (2017), "Big data in real estate? From manual appraisal to automated valuation", *Journal of Portfolio Management*, Vol. 43 No. 6, pp. 202-211.
- Kutasi, D. and Badics, M.C. (2016), "Valuation methods for the housing market: evidence from Budapest", *Acta Oeconomica*, Vol. 66 No. 3, pp. 527-546.
- Kontrimas, V. and Verikas, A. (2011), "The mass appraisal of the real estate by computational intelligence", *Applied Soft Computing*, Vol. 11 No. 1, pp. 443-448.
- Lam, K.C., Yu, C.Y. and Lam, C.K. (2009), "Support vector machine and entropy based decision support system for property valuation", *Journal of Property Research*, Vol. 26 No. 3, pp. 213-233.
- Lasota, T., Makos, M. and Trawinski, B. (2009), "Comparative analysis of Neural Network models for premises valuation using SAS enterprise miner", in Nguyen, N T., Katarzyniak, R.P. and Janiak, A. (Eds), *New Challenges in Computational Collective Intelligence*, Springer, Berlin Heidelberg, pp. 337-348.

AQ: 7

- 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*, Vol. 15 No. 1, pp. 8-26.
- Manganelli, B., De Mare, G. and Nesticò, A. (2015), "Using genetic algorithms in the housing market analysis", in Gervasi, O., Murgante, B., Misra, S., Gavrilova, M.L., Rocha, A.M.A.C., Torre, C.M., Taniar, D. and Apduhan, B.O. (Eds), *Computational Science and its Applications -ICCSA 2015*, Springer International Publishing, New York, pp. 36-45.
- Mangialardo, A., Micelli, E. and Sacconi, F. (2019) "Does sustainability affect real estate market values? Empirical evidence from the office buildings market in Milan (Italy)", *Sustainability*, Vol. 11 No. 1.
- McCluskey, W.J., McCord, M., Davis, P.T., Haran, M. and McIlhatton, D. (2013), "Prediction accuracy in mass appraisal: a comparison of modern approaches", *Journal of Property Research*, Vol. 30 No. 4, pp. 239-265.
- McGreal, S., Adair, A., McBurney, D. and Patterson, D. (1998), "Neural networks: the prediction of residential values", *Journal of Property Valuation and Investment*, Vol. 16 No. 1, pp. 57-70.
- Mooya, M.M. (2009), "Market value without a market: perspectives from transaction cost theory", *Urban Studies*, Vol. 46 No. 3, pp. 687-701.
- Mooya, M.M. (2017), "Automated valuation models and economic theory", in D'Amato, M. and Kauko, T. (Eds), *Advances in Automated Valuation Modeling*, Springer International Publishing, New York, pp. 33-57.
- Morano, P., Tajani, F. and Locurcio, M. (2018), "Multicriteria analysis and genetic algorithms for mass appraisals in the Italian property market", *International Journal of Housing Markets and Analysis*, Vol. 11 No. 2, pp. 229-262.
- Mu, J., Wu, F. and Zhang, A. (2014), "Housing value forecasting based on machine learning methods", *Abstract and Applied Analysis*, Vol. 4.
- Mullainathan, S. and Spiess, J. (2017), "Machine learning: an applied econometric approach", *The Journal of Economic Perspectives*, Vol. 31 No. 2, pp. 87-106.
- Nguyen, N. and Cripps, A. (2001), "Predicting housing value: a comparison of multiple regression analysis and artificial neural networks", *Journal of Real Estate Research*, Vol. 22 No. 3, pp. 313-336.
- Pérez-Rave, J.I., Correa-Morales, J.C. and González-Echavarría, F. (2019), "A machine learning approach to big data regression analysis of real estate prices for inferential and predictive purposes of real estate prices for inferential and predictive purposes", *Journal of Property Research*, Vol. 36 No. 1, pp. 59-96.
- Poursaeed, O., Matera, T. and Belongie, S. (2018), "Vision-based real estate price estimation", *Machine Vision and Applications*, Vol. 29 No. 4, pp. 667-676.
- Rics (2017), "The future of valuations", available at: <https://www.rics.org/globalassets/rics-website/media/knowledge/research/insights/future-of-valuations-insights-paper-rics.pdf> (accessed 4 December 2019).
- Wang, X., Wen, J., Zhang, Y. and Wang, Y. (2014), "Real estate price forecasting based on SVM optimized by PSO", *Optik*, Vol. 125 No. 3, pp. 1439-1443.
- 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*, Vol. 10 No. 2, pp. 185-201.
- Yacim, J.A. and Boshoff, D.G.B. (2018), "Impact of Artificial Neural Networks training algorithms on accurate prediction of property values", *Journal of Real Estate Research*, Vol. 40 No. 3, pp. 375-418.
- Yeh, H.C., Hsieh, T.K. and Wang, T.S. (2013), "Study on the relationship between city and district average price by GAOT in taipei", *Applied Mechanics and Materials*, Vol. 370, pp. 2043-2049.

---

Zhang, S.H. (2012), "Application of Support Vector Machine in determination of real estate price",  
*Advanced Materials Research*, Vol. 461, pp. 818-821.

AVMs vs  
hedonic models

**Corresponding author**

AQ: 3 Agostino Valier can be contacted at: [valier.agostino@gmail.com](mailto:valier.agostino@gmail.com)

---

For instructions on how to order reprints of this article, please visit our website:

[www.emeraldgroupublishing.com/licensing/reprints.htm](http://www.emeraldgroupublishing.com/licensing/reprints.htm)

Or contact us for further details: [permissions@emeraldinsight.com](mailto:permissions@emeraldinsight.com)