Digital innovation and Real estate appraisal:
findings from literature

Agostino Valier

Università degli Studi di Padova

Ezio Micelli

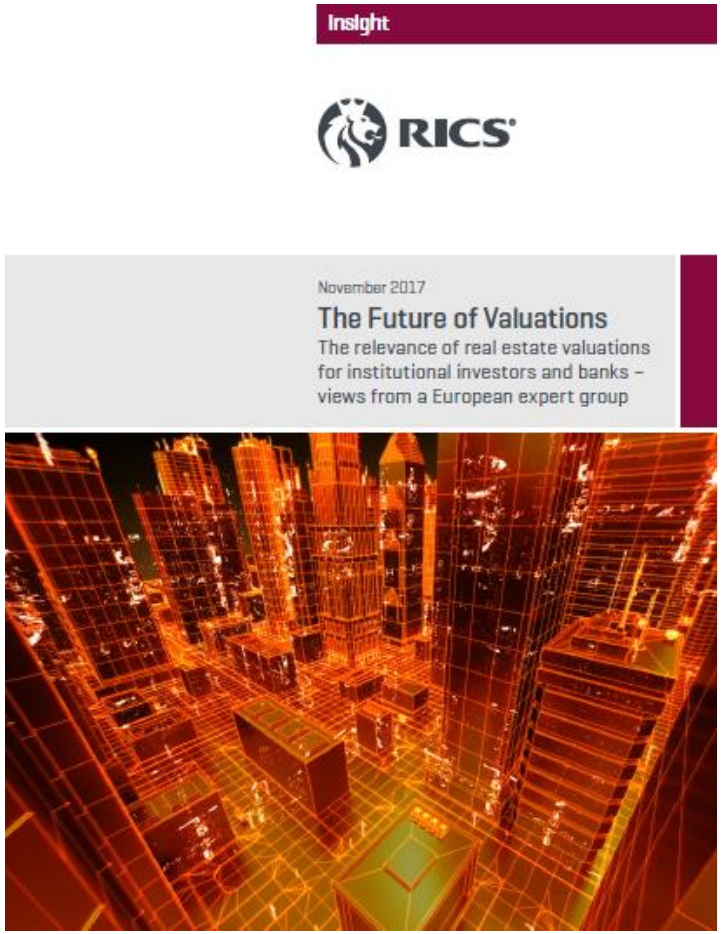
Università Iuav di Venezia

L'Aquila, 17 September 2019

40
1980 - 2019
A.I.S.Re.

The research / background

- Artificial intelligence models learn from the data themselves. These models help - and sometimes replace - human activity in more and more activities.
- The phenomenon of machine learning is intrinsically linked to the increasing availability of information (Big Data).
- Algorithms of self-learning can also be employed in the processes of estimating the value of real estate assets.



The research / purposes

Will artificial intelligence in value estimation be a disruptive or incremental innovation?

- How will the new models relate to manual estimation processes? Will they support or replace the figure of the evaluator?

- Will information theory override economic theory in value determination?

The research / **method**

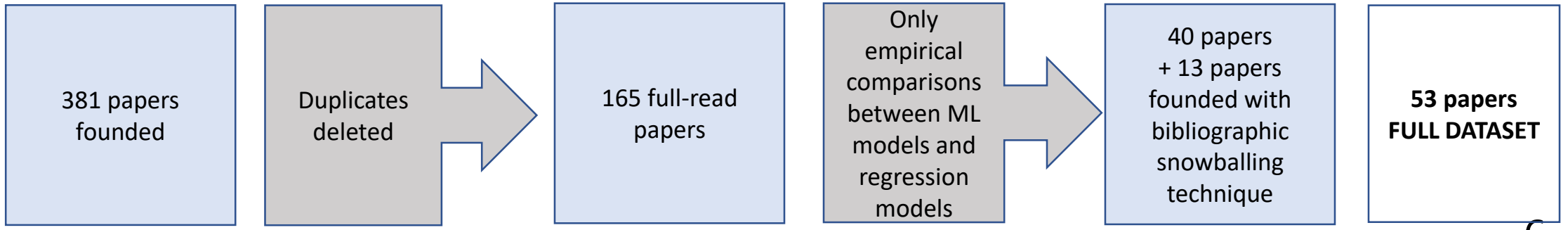
- General Literature review on real estate valuation models using artificial intelligence algorithms.
- The focus is on the evidence emerging from research that tests models on real estate datasets.

The research / **phases**

1	**General literature review and identification of papers to be analysed**
2	**The machine learning models: a proposal for classification.**
3	**Empirical tests: structure and parameters for comparison.**
4	**Results and other evidence from the authors.**

1_General literature review and identification of papers to be analysed

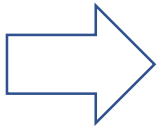
on Scopus database						
• Real estate	+	• Valuation • Appraisal • Automated Valuation Model /AVM • Prices forecasting	+	• Machine learning • Artificial intelligence • Decision/regression trees • Neural networks / ANN • Backpropagation • Genetic algorithm • Bayes • Nearest neighbors / KNN • Vector machine / SVM	36 strings of search	**381 papers**



1_General literature review and identification of papers to be analysed

Research question: will artificial intelligence models replace current estimation procedures?

66 papers test machine learning models on real sales data.

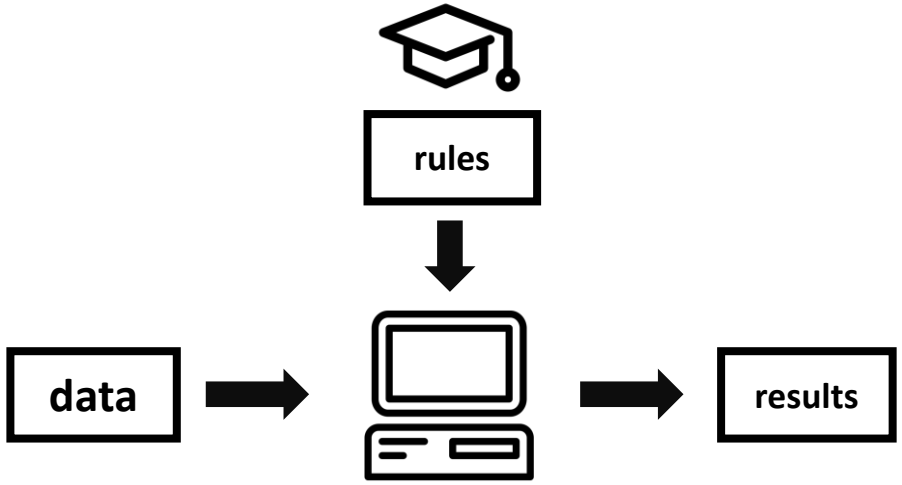


We chose to analyze the 40 papers compares machine learning models with the **regression** technique, identifying it as the mass appraisal model closest to that performed by human activity.

The hedonic price method considers the value of an asset as the sum of the values attributable (hedonic prices) to each of its component characteristics. These attributes are identified and measured with multivariate regression techniques (Rosen, 1974).

Traditional models

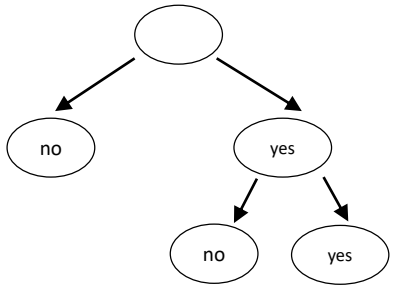
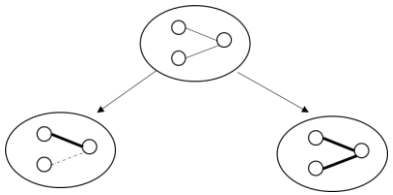
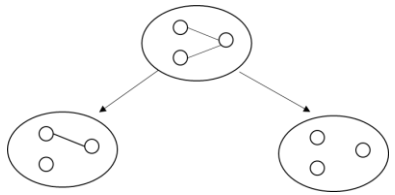
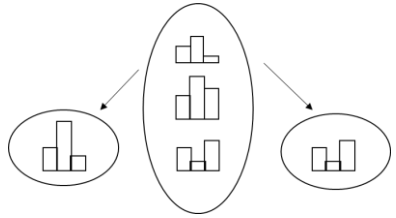
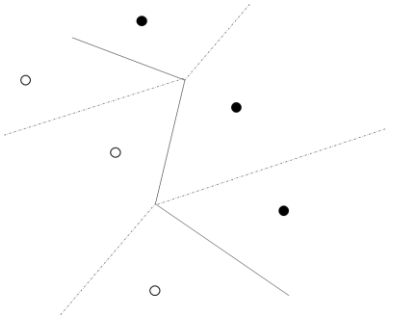
Planning everything in advance



Machine learning models

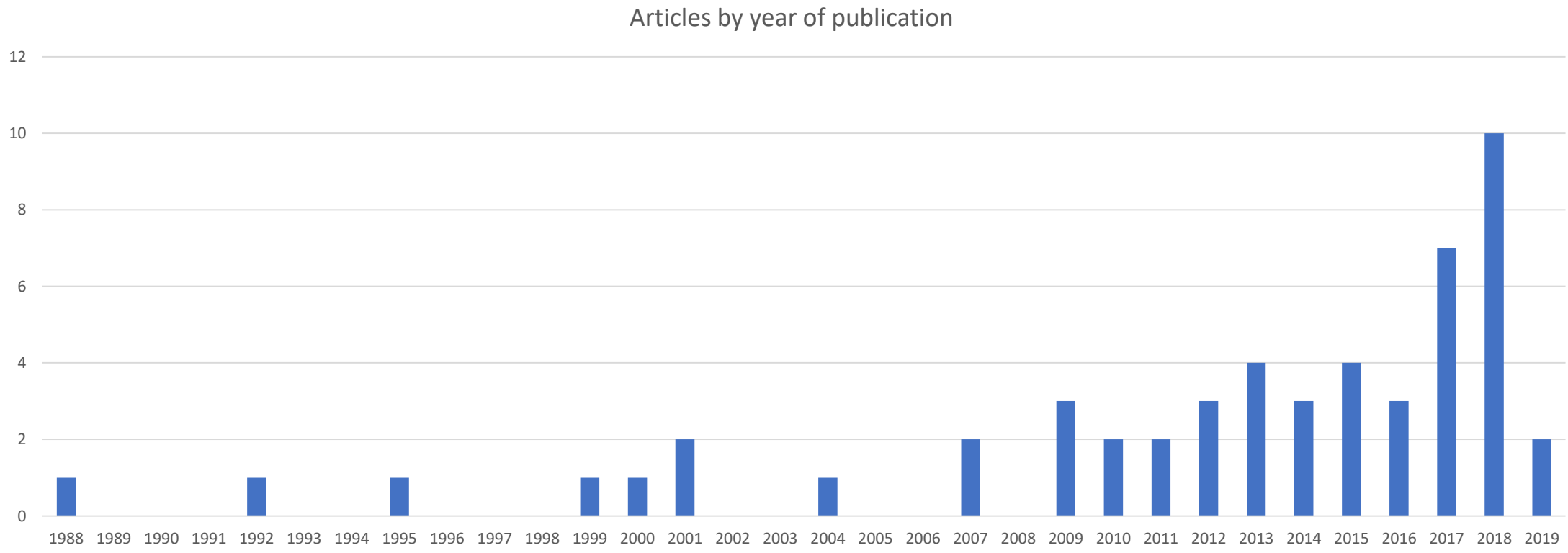
Let the data speak for itself

2_The machine learning models: a proposal for classification.

The symbolists	The connectionists	The evolutionaries	The bayesians	The analogizers
				
The reverse of deduction	*Learning parameteres*	*Evolving structures*	*Weighting evidence*	*Mapping to new solutions*
Decision and regression trees	Artificial neural networks and backpropagation	Genetic algorithms	Bayes' theorem and its derivates	K-Nearest neighbors and Support vector machine

(Pedro Domingos, *The Master algorithm. How the quest for the ultimate learning machine will remake our world*, 2015)

2_The machine learning models: a proposal for classification.



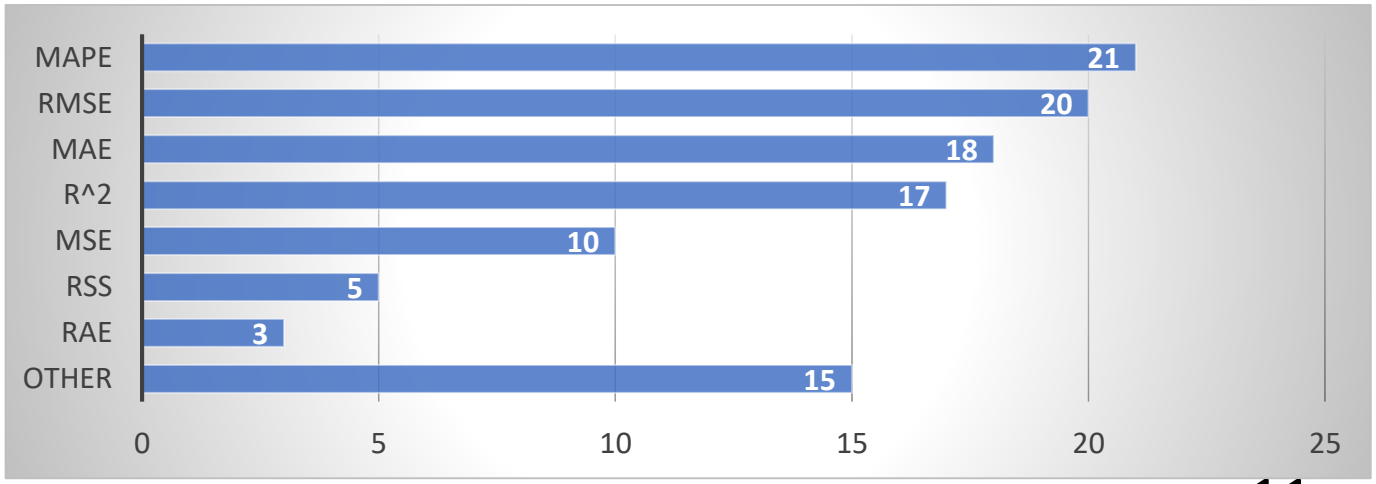
3_Empirical tests: structure and parameters for comparison.

TRAINING SET (~ 70%)

TESTING SET (~30%)

The dataset is divided into two parts. The training set is used for the learning phase of the model. The testing set is used to test the predictive capacity of the model, comparing the results obtained by the model with the output data.

- **MAPE:** Mean absolute percentage error
- **RMSE:** Root mean squared error
- **MAE:** Mean absolute error
- **R^2:** coefficient of determination
- **MSE:** Mean squared error
- **RSS:** Residual sum of squares
- **RAE:** Relative absolute error
- other

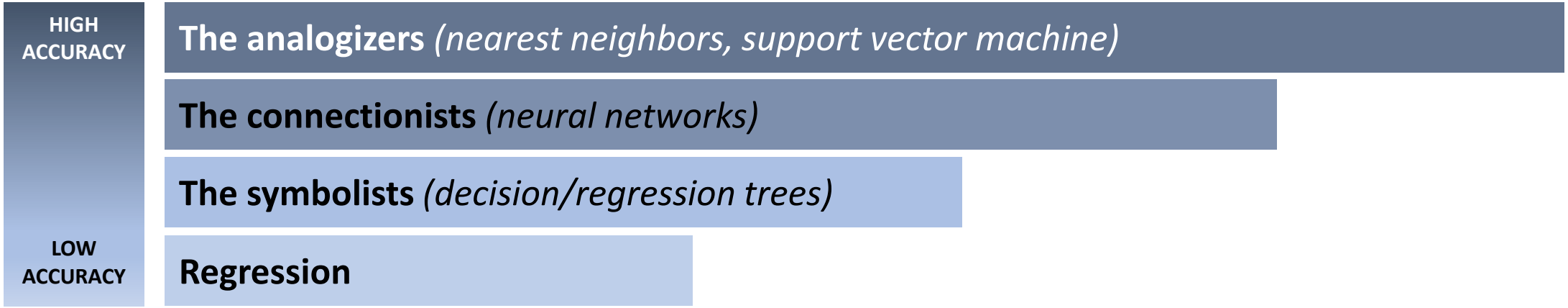


4_Results and other evidence from the authors.

●: higher accuracy ●: lower accuracy

		REGRESSIONE		SIMBOLISTI		CONNESSIONISTI		EVOLUZIONISTI		BAYESIANI		ANALOGISTI		TOT
		>	<	>	<	>	<	>	<	>	<	>	<	
REGRESSIONE	>			4	-	3	-	0	-	0	-	1	-	8
	<			-	5	-	30	-	4	-	1	-	11	51
SIMBOLISTI	>	5	-			1	-	0	-	0	-	1	-	7
	<	-	4			-	5	-	0	-	1	-	3	13
CONNESSIONISTI	>	30	-	5	-			0	-	0	-	3	-	38
	<	-	3	-	1			-	2	-	1	-	7	14
EVOLUZIONISTI	>	4	-	0	-	2	-			0	-	0	-	6
	<	-	0	-	0	-	0			-	0	-	0	0
BAYESIANI	>	1	-	1	-	1	-	0	-			0	-	3
	<	-	0	-	0	-	0	-	0			-	0	0
ANALOGISTI	>	11	-	3	-	7	-	0	-	0	-			21
	<	-	1	-	1	-	3	-	0	-	0			5

4_Results and other evidence from the authors.



Real estate models avm not yet sufficiently investigated

- **The Bayesians**
- **The evolutionaries** (*genetic algorithms*)

4_Results and other evidence from the authors.

	n° of papers
Machine learning models cannot be applied in professional practice because they behave like "black boxes".	●●●●●●●●
In connectionist algorithms, which are stochastics, committes of model perform better.	●●●●●●●
The performance of machine learning models is directly proportional to the amount of data available to them.	●●●●●●
The machine learning models do not have the sensitivity of the evaluator	●●●●●
Model comparisons should not only be assessed in terms of forecasting accuracy, but also time and cost.	●●●
In regression models, artificial intelligence selects variables more effectively than the human brain.	●●

- All machine learning models perform much better in terms of value forecasting ability.
- After an initial enthusiasm in the use of neural networks, several contributions have also analyzed their limits of use.
- In the last years research has focused on algorithms based on the principle of analogy, which demonstrate better accuracy performance.
- The effectiveness of the models is directly proportional to the amount of data.

Open issues

- What role will the sensitivity of the evaluator and the opinion of the experts play in the future?
- Machine learning models often have characteristics of difficult intelligibility, lack of control and irreproducibility of the process. Can these characteristics prevent their use in professional practice?
- The effectiveness of machine learning models is closely linked to the availability of information. So they will only be usable by the big players? What role for the minor players who have more difficulty in accessing the data?

Thank you for your attention

Agostino Valier

University of Padua

agostino.valier@phd.unipd.it

Ezio Micelli

University IUAV of Venice

micelli@iuav.it

1980 - 2019

A.I.S.Re.