

Article

Explainable AI for Urban Real-Estate Prediction: A Machine-Learning Framework for Urban Decision Support

Valeria Saiu ^{1,*}  and Matteo Mocci ² ¹ Department of Civil and Environmental Engineering and Architecture, University of Cagliari, Via Santa Croce 67, 09124 Cagliari, Italy² Department of Mathematics and Computer Science, University of Cagliari, Via Ospedale 72, 09124 Cagliari, Italy; matteo.mocci2@unica.it

* Correspondence: v.saiu@unica.it

Abstract

This study introduces RE-VAL (REal-estate VALuation), an explainable framework for urban real-estate analysis that integrates reproducible data acquisition, geographically informed feature processing, predictive benchmarking, and interpretable outputs suitable for decision-support-oriented analysis. Unlike static automated valuation models, the RE-VAL framework is designed to reflect context-dependent market behaviour across heterogeneous urban areas. The comparative evaluation on 1153 residential listings from Cagliari (Italy) showed that MLP achieved the strongest predictive performance, while Random Forest provided the most convincing balance between predictive competitiveness and interpretability. Beyond point estimation, the framework leverages SHAP-based decomposition to translate algorithmic outputs into transparent, monetary-based “Bonus/Malus” adjustment tables. The analysis highlights the presence of potentially non-linear interactions, including a possible premium associated with energy efficiency in prestigious areas, and suggests that the framework can remain informative when incomplete technical data are preserved as potential proxy signals rather than being discarded as noise. Rather than identifying a single predictor, RE-VAL provides a transparent, extensible and decision-oriented workflow for urban real-estate valuation, advancing the integration of explainable artificial intelligence within complex spatial-economic systems.

Keywords: real-estate prediction; urban analytics; explainable AI (XAI); RE-VAL framework; nonlinear value modelling; decision support



Academic Editors: Maria Rosaria Guarini and Kevin H. Zhang

Received: 10 March 2026

Revised: 18 April 2026

Accepted: 2 May 2026

Published: 4 June 2026

Copyright: © 2026 by the authors.

Licensee MDPI, Basel, Switzerland.

This article is an open access article distributed under the terms and conditions of the [Creative Commons Attribution \(CC BY\)](https://creativecommons.org/licenses/by/4.0/) license.

1. Introduction

Property valuation plays a pivotal role in urban engineering, real-estate management, and economic planning, directly influencing taxation policies, infrastructure investment, and sustainable development strategies. In increasingly complex and data-intensive environments, reliable valuation models are essential not only for market actors but also for public administrations operating within mass-appraisal settings. These contexts require standardized, repeatable, and auditable procedures to ensure consistency and transparency, particularly for fiscal and administrative purposes [1,2].

The growing adoption of Automated Valuation Models (AVMs) in the real-estate sector has raised significant questions regarding their alignment with international professional standards, such as the International Valuation Standards (IVS) [3], the European Valuation Standards (EVS) by TEGoVA [4], and the RICS Valuation—Global Standards (Red Book) [5].

These frameworks mandate that valuations be transparent, logical, and based on identifiable market evidence. However, the ‘black-box’ nature of many machine learning models often conflicts with the accountability and auditability required in formal appraisal contexts, such as mortgage lending or legal proceedings.

Traditional hedonic pricing models, while valued for their interpretability and strong theoretical grounding, are fundamentally constrained by their linear structure. They struggle to represent the intricate and non-linear interactions between property characteristics and locational factors that increasingly shape modern urban markets [6]. As documented in the literature, hedonic formulations often fail to accommodate high-dimensional feature spaces, spatial heterogeneity, and complex interaction effects without extensive manual specification or restrictive assumptions [7,8]. Furthermore, urban real-estate systems have become markedly more fragmented due to differentiated neighborhood dynamics and environmental externalities, generating valuation signals that are highly context-dependent and difficult to capture using traditional regression-based approaches alone [9–11].

Recent research highlights that web crawler technologies have become a core infrastructure for real-estate data acquisition, enabling fine-grained and high-frequency observation of market dynamics [12,13]. These developments improve the temporal relevance of valuation outputs, reduce subjective bias in data selection, and support improved generalization across heterogeneous urban contexts. While advances in Artificial Intelligence (AI) and Machine Learning (ML) have significantly improved predictive accuracy—allowing to learn complex, non-linear relationships directly from market evidence—most existing automated valuation systems face three structural challenges.

First, model opacity (the “black-box” problem) limits interpretability and constrains trust in professional and institutional settings. Second, the absence of standardized and reproducible training pipelines hampers benchmarking and model transfer across markets. Third, the widespread assumption of static feature relevance restricts the ability of models to adapt to local market dynamics and context-dependent valuation mechanisms. In this perspective, the value of an advanced valuation framework does not lie solely in maximizing predictive accuracy, but in structuring and interpreting model outputs in ways that remain consistent with valuation theory and professional practice.

To address these challenges, this research introduces RE-VAL (REal-estate VALuation), an explainable methodological framework for urban real-estate analysis that combines reproducible data collection, structured preprocessing, comparative predictive evaluation, and interpretable decision-support outputs. RE-VAL is designed as an end-to-end workflow that supports the full valuation process, from data acquisition and harmonization to model comparison and operational interpretation.

Within this framework, missing and incomplete listing information is not treated only as a technical issue to be corrected, but also as a potential proxy signal whose empirical role requires cautious interpretation. The framework further supports the systematic comparison of multiple predictive models, enabling the identification of robust solutions within a common experimental setting. To enhance transparency and practical usability, the framework integrates SHAP-based analysis to translate model outputs into operational insights, including monetary-based ‘Bonus/Malus’ adjustment tables. This approach enables the exploration of non-linear and context-dependent effects, such as variations in the valuation impact of energy efficiency across different urban areas, while also treating missing technical data as potentially informative proxy signals rather than mere noise.

More specifically, RE-VAL differs from much of the existing machine-learning valuation literature in two respects. First, incomplete listing attributes are not treated only as values to be removed or mechanically corrected, but as potentially informative proxy signals preserved within the modelling workflow under an MNAR-oriented interpre-

tation [14]. Second, explainability is not used only for post hoc feature ranking but is translated into monetary Bonus/Malus adjustments that reflect the comparative logic of professional appraisal practice. In this sense, while recent studies have advanced predictive modelling, web-scraped data pipelines, and SHAP-based transparency [15,16], RE-VAL is positioned as a workflow that explicitly combines missingness-aware preprocessing and monetary interpretability within a common valuation framework.

While the RE-VAL framework integrates established components from machine-learning-based valuation pipelines, its contribution does not lie in proposing a new predictive model. Instead, it advances the role of explainable AI within real-estate valuation by reframing interpretability as an operational component of appraisal logic rather than a purely diagnostic tool.

Specifically, the contribution of this study is twofold. First, it introduces a missingness-aware preprocessing strategy in which incomplete listing attributes are preserved as potential proxy signals under an MNAR-oriented perspective, reflecting the informational structure of real-world digital markets. Second, it translates post hoc explainability outputs into a valuation-oriented framework through monetary Bonus/Malus adjustments, aligning model interpretation with the comparative reasoning required in professional appraisal practice.

By combining these elements within a unified workflow, the paper contributes to the emerging literature on explainable AI by demonstrating how model interpretability can be operationalized in economically meaningful terms, enabling the assessment of valuation logic beyond predictive accuracy alone.

To guide this investigation and address the identified gaps in automated valuation models, this study seeks to answer the following research questions:

- RQ1: To what extent do advanced non-linear machine learning models, e.g., MLP, Random Forest, LightGBM, provide competitive predictive performance relative to simpler linear baselines, and what is gained in explanatory structure and decision-support value?
- RQ2: Can the integration of Explainable AI (XAI) techniques, specifically SHAP values, effectively bridge the gap between ‘black-box’ predictive models and the interpretability requirements of international professional valuation standards (e.g., IVS, RICS)?
- RQ3: Does treating missing data as an informative signal (Missing Not At Random—MNAR) rather than mere noise enhance the model’s capacity to capture context-dependent market behaviours?

By addressing these questions, the study aims to test the hypothesis that a geographically informed and explainable machine-learning framework can provide not only reliable price estimations but also structured and interpretable ‘Bonus/Malus’ insights to support urban decision-making processes.

The remainder of this article is structured as follows. Section 2 reviews the related literature on AI-based valuation, explainability, and the challenges of data ‘missingness’. Section 3 presents the RE-VAL framework and methodological pipeline; Section 4 describes the Cagliari case study and reports the benchmark comparison across the tested predictive models. Section 5 discusses interpretability results and the main methodological insights emerging from the analysis. Section 6 discusses the practical implications of the monetary Bonus/Malus synthesis and the decision-support potential of the framework. Finally, Section 7 provides the conclusions, limitations, and future research directions.

2. Related Work

Research on artificial intelligence (AI) and real-estate valuation has undergone a significant paradigm shift, transitioning from traditional hedonic regressions to sophisticated machine-learning frameworks capable of modelling non-linear and multi-dimensional price dynamics [13]. Early applications of neural networks and ensemble methods consistently outperformed econometric baselines, particularly when combined with systematic hyperparameter optimization and advanced feature engineering [17]. Nevertheless, as noted by Ho et al. [18], the literature consistently reports that no single algorithm performs optimally across different cities, datasets, or valuation purposes, underscoring the critical need for flexible and adaptable modelling strategies.

This functional rigidity is further addressed by Abidoye and Chan [19], who argue that traditional models fail to capture the complex “black-box” of human preferences where the marginal value of a feature is context-dependent. For instance, Kalliola et al. (2021) [20] optimized multilayer perceptrons through Bayesian tuning in the Helsinki market, while Martínez (2023) [21] and Gowda et al. (2025) [22] demonstrated that gradient-boosting and support-vector algorithms drastically surpass linear models when features are standardized and interaction terms are adequately captured. Comparative studies such as Francke and van de Minne (2024) [23] and Zitoune and Arabov (2024) [24] confirmed that hybrid approaches reconcile predictive performance with interpretability, while Çilgin et al. (2023) [15] emphasized that rigorous outlier detection can improve accuracy by over twenty percent.

Parallel to methodological refinement, a second stream of research has focused on data sourcing and automation. The proliferation of web scraping and open-data infrastructures has transformed the empirical foundation of real-estate analytics [12,13]. Berry (2024) [25] conceptualized web scraping as a reproducible research infrastructure, while Üzümcü and Eli Güzel (2023) [26], Santos (2024) [27], and Pineda Montserrat (2024) [28] operationalized automated pipelines for large-scale data collection and cleaning.

Despite this data abundance, a significant limitation persists regarding the treatment of “missingness” in web-scraped data. Following the theoretical framework of Little and Rubin (2019) [14], the proposed RE-VAL framework explicitly considers the possibility that some missing technical information may be informative, and operationalizes this possibility by preserving selected missing or unspecified categories in the modelling workflow. This approach complements recent works like Trindade Neves et al. (2024) [16], who combined open urban data with SHAP-based explainability to enhance transparency in housing forecasts.

Spatial analytics further enrich this infrastructure by identifying how location dictates value gradients. Souza et al. (2021) [29] and Meyberg et al. (2024) [30] used web-scraped datasets to detect spatial autocorrelation and price gradients, demonstrating that location remains the dominant determinant of value even without complete GIS coverage. Furthermore, Helbich et al. (2014) [31] demonstrated that data-driven submarket identification can reveal latent spatial dynamics that traditional administrative boundaries fail to capture. Within this transformation, Cugurullo (2024) [32] and Rey-Blanco et al. (2024) [33] conceptualize “urban AI” as an active component of city governance, co-producing spatial and economic realities rather than merely describing them.

A growing body of research addresses explainability and ethical governance. Recent works recognize that the impact of AI-based valuation extends beyond accuracy to market behavior. Silaghi et al. (2024) [34] and Wheaton and Xu (2024) [35] found that public AI price estimates can influence market transparency only when users understand how valuations are produced. At the methodological frontier, Asimiyu (2024) [36] introduced Explainable Generative AI, while Jaouhari et al. (2024) [36] identified reproducibility and

explainability as critical challenges for next-generation systems. These findings align with simulations by Coletta et al. [37], highlighting the reflexive influence of algorithms on market dynamics.

Despite this progress, a gap remains between algorithmic performance and professional appraisal logic. The proposed RE-VAL framework addresses this gap by structuring urban valuation as an interpretable and operational workflow aligned with the professional comparative approach advocated by D'Amato and Kauko (2017) [38], translating SHAP-based outputs into transparent monetary Bonus/Malus adjustment tables for auditable urban decision-making.

3. Materials and Methods

3.1. RE-VAL Framework Architecture

The proposed RE-VAL (REal-estate VALuation) framework is structured as an integrated analytical pipeline designed to align high-capacity predictive modeling with the formal requirements of urban appraisal. As illustrated in Figure 1, the workflow partitions the valuation process into five functional modules, ensuring reproducibility and methodological transparency across heterogeneous urban contexts.

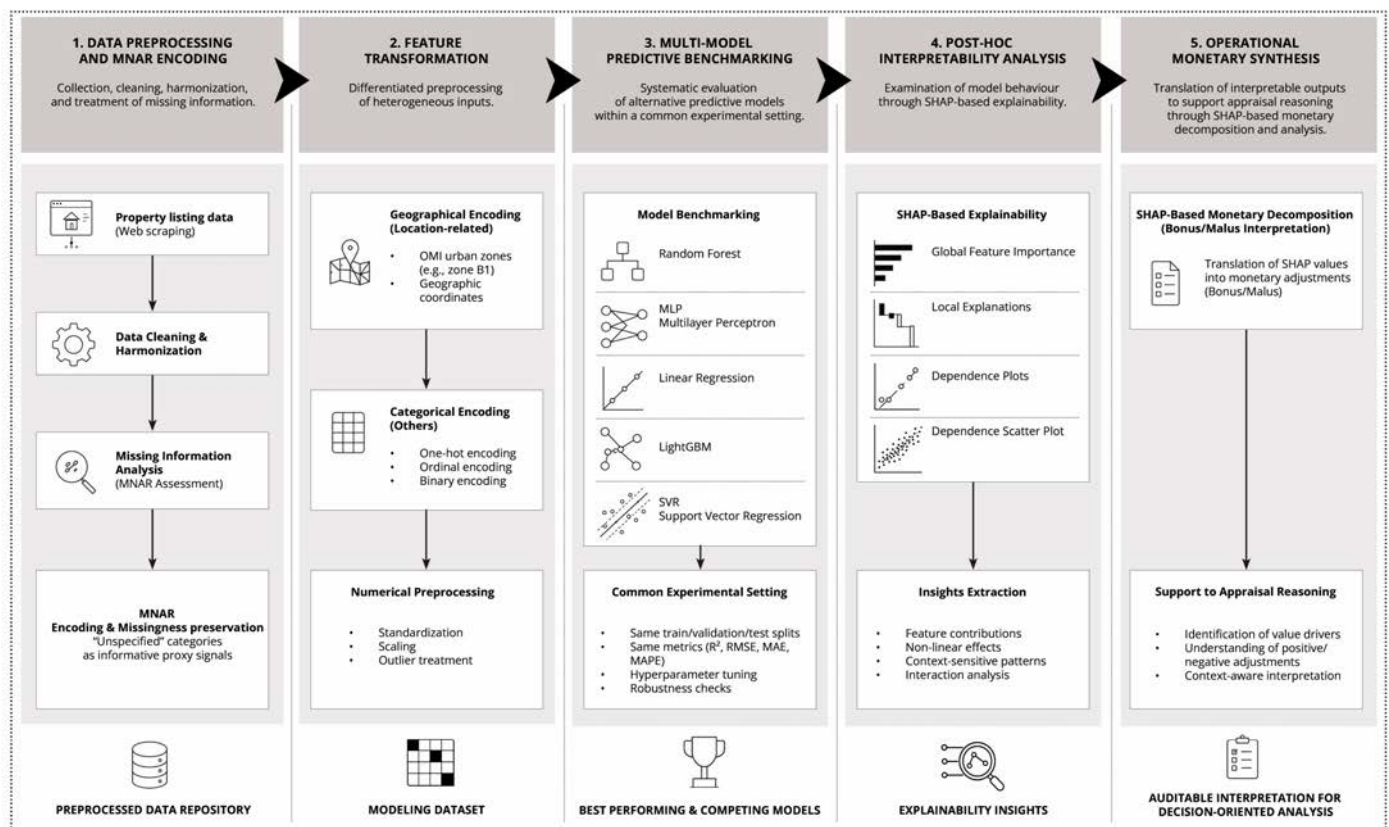


Figure 1. Methodological workflow of the proposed RE-VAL framework for urban real-estate analysis, illustrating the sequence from data preprocessing and feature transformation to predictive benchmarking, interpretability analysis, and decision-support outputs.

In detail:

1. Data Preprocessing and MNAR encoding cleaning, harmonization, and treatment of missing information through informative proxy signals preservation.
2. Feature Transformation: Differentiated preprocessing of heterogeneous tabular inputs, with geographically informed encoding for location-related information, conventional

encoding for the remaining categorical variables, and standard preprocessing for numerical features.

3. Multi-model Predictive Benchmarking: Systematic evaluation of alternative predictive models within a common experimental setting.
4. Post hoc Interpretability Analysis: Examination of model behaviour through SHAP-based explainability to identify feature contributions, non-linear effects, and context-sensitive valuation patterns.
5. Operational Monetary Synthesis: Translation of model outputs into currency-based Bonus/Malus adjustment tables to support appraisal reasoning, scenario analysis, and auditable urban decision-making.

3.2. Data Harvesting and Multidimensional Harmonization

The primary dataset was constructed through automated web-scraping protocols targeting major Italian real-estate listing platforms. This approach captures high-frequency market signals, mitigating the temporal latency inherent in institutional databases. Raw attributes are systematically harmonized into structured taxonomies before model estimation:

- *Property Class*: Grouped into four market segments (economic, middle-range, upscale, and luxury), reflecting increasing levels of prestige and construction quality.
- *Renovation Status*: Coded into four segments (new/under construction, excellent/renovated, good/habitable, and to be renovated).
- *Exposure*: Classified according to the prevailing orientation or openness of the unit (e.g., north, south, east, west, double exposure, internal, external),
- *Window*: Encoded with attention to insulation quality: (single, double, and triple glazing).
- *Heating*: Represented through both heating system type (mainly autonomous vs. centralized) and heating source/delivery system (e.g., air-based, radiator-based, underfloor heating, stove-based), while the corresponding energy supply was retained when specified.
- *Energy Class*: Mapped according to the Italian certification scale (from A to G).

Location attributes are geospatially joined with institutional OMI zones, defined by the Real-Estate Market Observatory of the Italian Revenue Agency, to provide a standardized territorial micro-context [37].

Within this classification, B-zones correspond to central or semi-central areas, typically adjacent to the historic core, and are usually associated with intermediate-to-high property values. C-zones identify peripheral or completion areas, generally characterized by lower density and greater distance from major services and urban amenities. D-zones refer to suburban or extra-urban areas, typically marked by lower density and more mixed settlement patterns. This procedure captures spatial heterogeneity across the urban market and distinguishes between submarkets whose locational effects may not be fully reflected by institutional average values alone.

Following anonymization and fair-use data policies, Python-based tools—specifically *BeautifulSoup* 4.12.3 for HTML parsing and *OpenPyXL* 3.1.5 for data structuring—were employed to automate extraction and integration. This procedure establishes a standardized and workflow designed to support replicability, consistent with recent research emphasizing web scraping as a core infrastructure in real-estate data science [25–27].

3.3. Informative Proxy Signal Encoding (MNAR Management)

A distinctive methodological pillar of the RE-VAL framework is the treatment of incomplete technical specifications as endogenous market signals. Following Little and Rubin's theoretical foundations [14], RE-VAL treats missing data (e.g., window frame

type, exposure, or renovation status) not as noise to be removed, but as informative proxy signals. Based on the Missing Not At Random (MNAR) hypothesis, omissions are explicitly encoded into an ‘Unspecified’ category.

This strategy leverages the correlation between seller reporting behavior and specific market segments, preserving latent information that traditional imputation methods would otherwise sanitize. The final preprocessing stages involve: (1) data anonymization and outlier filtering; (2) categorical encoding and missing signal preservation; (3) median imputation for residual numerical gaps to preserve distributional balance; and (4) feature scaling for continuous variables (Table 1).

Table 1. Preliminary dataset.

Category	Original Data	Variable Type	Transformation for the Model
Target	Sale Price (€)	Continuous (target)	Converted to unit price (€/m ²) as target y
Location	Address/Zone	Categorical	OMI Urban Zones
Size & layout	Net floor area (m ²)	Continuous	Standardized; used to derive unit price
Dwelling Traits	Floor level	Ordinal	0 = ground, 1 = intermediate, 2 = top
	Balcony	Binary	0 = absent, 1 = present
	Garden	Binary	0 = absent, 1 = present
	Parking availability	Binary	0 = absent, 1 = present
Quality and Condition	Property class	Categorical	One-hot (Economy to Premium)
	Renovation status	Categorical	One-hot (including Unspecified as proxy)
	Exposure/Windows	Categorical	One-hot (including Unspecified as proxy)
	Heating source	Categorical	One-hot (including Unspecified as proxy)
	Heating type	Categorical	One-hot (including Unspecified as proxy)
	Window frame type	Categorical	One-hot (including Unspecified as proxy)
	Energy class	Ordinal	Mapped scale (A = 4 ... G = 1; Unspecified = 0)
Accessibility	Elevator	Binary	0 = absent, 1 = present
	Property Availability	Categorical	One-hot encoding (Available, Occupied)
	Wheelchair access	Binary	0 = absent, 1 = present

3.4. Comparative Predictive Benchmarking

The modeling phase is designed as a systematic benchmark to evaluate the trade-off between predictive accuracy and explanatory stability. The suite of algorithms represents divergent mathematical approaches to urban complexity: Linear Regression (additive baseline), SVR (RBF kernel-based), Random Forest (ensemble bagging), MLP (deep non-linear mapping), and LightGBM (tree-based gradient boosting) (Table 2). TabNet was also tested in a preliminary exploratory phase because it is specifically designed for tabular data. However, hyperparameter tuning did not yield competitive results in the present small-sample setting.

Table 2. Regression models evaluated in the study.

Model	Learning Type	Description
Linear Regression	Linear baseline	Simple and interpretable but limited to additive linear relationships.
SVR	Kernel-based	Captures non-linear patterns through Radial Basis Functions (RBF).
Random Forest	Ensemble	Reduces variance through bagging and feature sampling.
MLP Regressor	Deep learning	Multi-Layer Perceptron; learns non-linear mappings via backpropagation.
LightGBM	Gradient boosting	Captures non-linear patterns and feature interactions through tree-based boosting

The predictive component of the RE-VAL framework is designed to support urban real-estate valuation through a standardized and interpretable computational pipeline. Within this framework, processed categorical and numerical features are used to estimate unit property values under a common experimental setting, enabling consistent comparison across alternative regression architectures. This design supports the analysis of predictive behaviour in relation to the heterogeneous and non-linear structure of housing markets, while maintaining coherence with the interpretability and decision-support objectives of the overall framework.

Models are validated through a ten-fold cross-validation protocol, optimizing Mean Absolute Error (MAE) to ensure robustness against market outliers. Table 3 summarizes the key parameters of the optimal configurations of the grid search.

Table 3. Regression models and key parameters.

Model	Key Parameters	Type
SVR (RBF kernel)	$C = 100$, $\epsilon = 1.0$, $\gamma = \text{auto}$	Kernel-based
Linear Regression	Default	Linear baseline
Random Forest	100 trees, max depth = 10, leaf size = 2	Ensemble
MLP	hidden-layers size: 128 and 64 neurons, $\alpha = 0.0001$, LR = 0.005	Deep learning
LightGBM	Leaves = 31, LR = 0.05, N_est = 2000	Gradient Boosting

This design supports the integrated analysis of predictive performance and interpretability within a common valuation workflow. By combining standardized feature processing, comparative model evaluation, and explainability-driven interpretation, RE-VAL promotes methodological transparency and adaptability across diverse urban market contexts.

3.5. XAI-Driven Monetary Synthesis

To align algorithmic outputs with professional appraisal practice, RE-VAL adopts a dual-layer explainability strategy based on SHAP (SHapley Additive Explanations). This allows for both Global Sensitivity Analysis (identifying variables that consistently shape the market) and Local Value Decomposition (decomposing individual predictions into additive feature contributions).

The core innovation lies in the Operational Monetary Synthesis: by denormalizing SHAP values, the framework translates abstract feature contributions into a “Monetary Adjustment Table” (Bonus/Malus). This transforms “black-box” predictions into auditable appraisal evidence, making the framework more directly usable for decision-support-oriented analysis, enabling the interpretation of ‘what-if’ scenarios related to interventions such as renovation or energy retrofitting, consistent with the adjustment logic mandated by IVS and RICS standards.

In addition, interaction and dependence plots are employed to highlight non-linear relationships between features. Combined with SHAP-based feature contribution analysis, these tools enable the examination of how the valuation effect of a characteristic, such as Energy Class, may vary across different urban contexts. This interpretability layer strengthens the connection between conventional appraisal reasoning and contemporary machine-learning analysis, providing both quantitative evidence and qualitative insight into how real-estate value emerges from the interplay of urban context and structural property attributes.

4. Case Study: Residential Market in Cagliari, Italy

To demonstrate the empirical performance and operational utility of RE-VAL, the framework was applied to a dataset of 1153 residential apartments in Cagliari, the capital of Sardinia (Italy) (Figure 2). Cagliari represents a significant case study due to its heterogeneous urban fabric, which combines a historic medieval core, consolidated 20th-century expansion districts, and modern peripheral developments. This spatial fragmentation generates complex price gradients and localized market dynamics, providing an ideal environment to test the model's ability to capture non-linear interactions and neighborhood-specific effects.

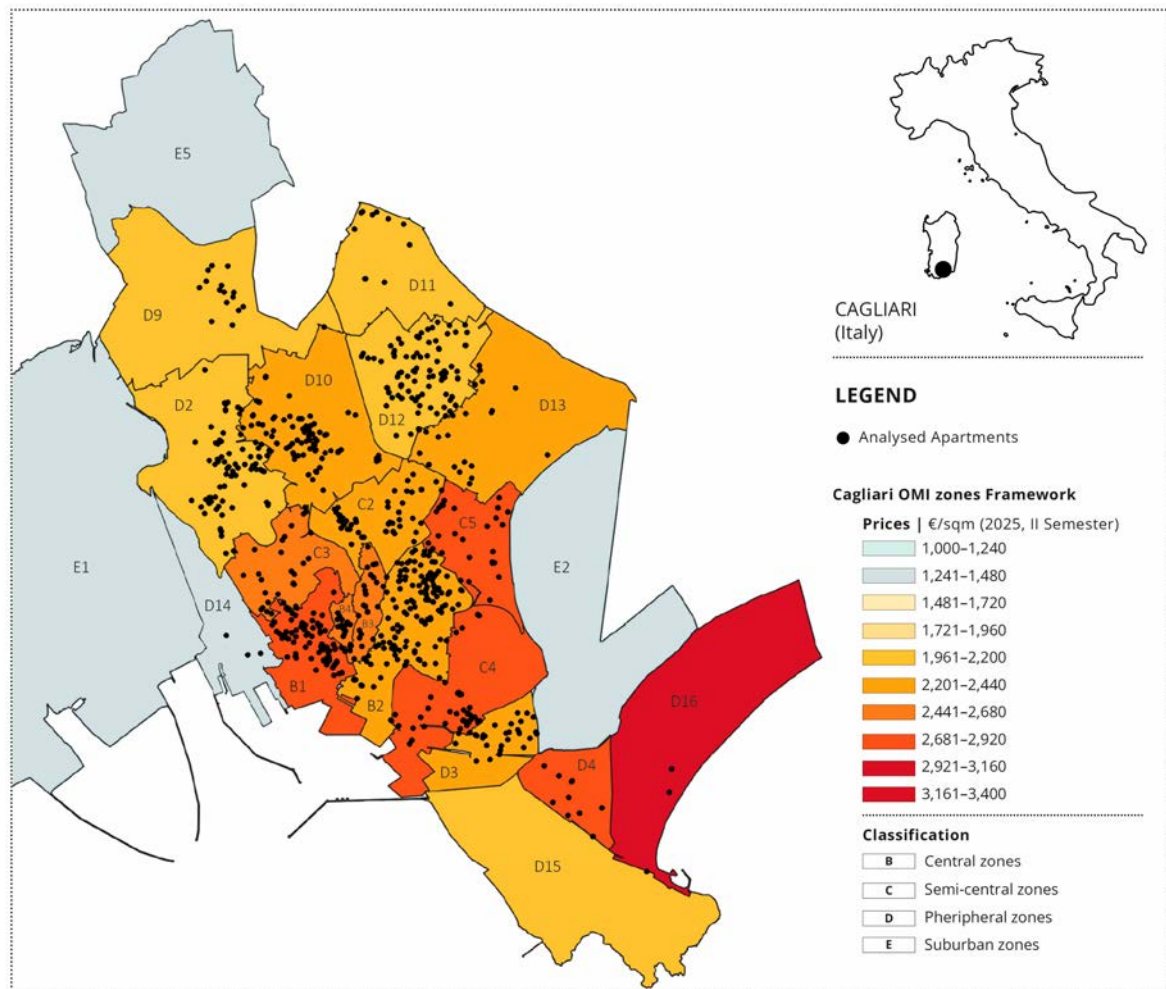


Figure 2. Map of the distribution of the analyzed apartments related to Cagliari's OMI zones. Black dots indicate the analyzed properties, while the color gradient represents the average price per square meter recorded in each OMI zone for the second half of 2025, with increasing values ranging from light blue/yellow to dark red. The alphanumeric labels within each area (e.g., D9, E2, D15) correspond to the official OMI zone codes and are reported for reference purposes only.

4.1. Dataset Composition and Feature Engineering

The dataset was assembled through a targeted web-scraping campaign from major Italian real-estate platforms (*Immobiliare.it*), reflecting real-time market signals. The target variable is the listed unit price (€/m²), derived from the total sale price and net floor area. The temporal coverage includes listings from early 20th-century buildings to newly constructed energy-efficient units (2020–2023). As detailed in Table 1, the features include:

- **Structural Attributes:** Net area, floor level, and presence of balconies/terraces.

- **Qualitative Features:** Renovation status (ranging from “to be refurbished” to “new construction”), heating systems, and window frame types.
- **Energy and Sustainability:** Energy class (A to G), included to capture the possible valuation effect of energy performance in the local market.
- **Location Proxies:** Instead of explicit GPS coordinates, the model utilizes 11 distinct urban zones (categorical proxies) to capture spatial heterogeneity and local prestige effects (e.g., the historic centre Zone B4).

4.2. Missing Data as Informative Signals

A critical aspect of the Cagliari dataset is the presence of incomplete technical specifications in several listings (e.g., missing exposure or window material data). In the RE-VAL implementation, these “missing” entries were not discarded but encoded as a specific category. This choice assumes that, in the Cagliari market, the omission of certain technical details may act as a proxy signal for specific building types or standardized mid-range apartments, allowing the framework to remain informative across heterogeneous data-entry conditions.

4.3. Experimental Setup

The dataset was evaluated using a 10-fold cross-validation protocol to assess the generalizability of results. Within the RE-VAL framework, a set of widely used regression models was benchmarked, including Linear Regression, Support Vector Regression (SVR), Random Forest (RF), Multi-Layer Perceptron (MLP) and LightGBM. Model selection and evaluation were conducted under a common experimental setting, with Mean Absolute Error (MAE) adopted as the primary performance measure because of its robustness to market outliers. Predictive performance was further assessed using Root Mean Squared Error (RMSE), Median Absolute Error (MedAE), and the coefficient of determination (R^2) (Table 4).

Table 4. Ten-fold cross-validation results for tuned models (mean $\pm$ standard deviation).

Model	MAE (€/m ²)	RMSE (€/m ²)	Median AE (€/m ²)	R ²
SVR	706.32 $\pm$ 67.39	948.77 $\pm$ 80.27	529.57 $\pm$ 65.24	0.10
Linear Regression	483.50 $\pm$ 41.69	642.22 $\pm$ 57.70	384.16 $\pm$ 49.40	0.58
Random Forest	475.39 $\pm$ 48.03	647.10 $\pm$ 80.85	357.01 $\pm$ 51.60	0.57
MLP	474.47 $\pm$ 29.00	632.38 $\pm$ 59.44	374.59 $\pm$ 32.70	0.59
LightGBM	508.55 $\pm$ 49.60	688.16 $\pm$ 88.46	385.69 $\pm$ 53.94	0.51

Although MLP achieved the best aggregate metrics, the difference relative to Linear Regression and Random Forest was small. This suggests that, in the present dataset, non-linear models do not provide a large predictive advantage over simpler approaches, even if they may offer greater flexibility in representing heterogeneous valuation patterns.

The comparative evaluation highlighted modest differences across the benchmarked models. Random Forest and MLP achieved the strongest overall predictive performance in the benchmark, although only marginally above Linear Regression with MLP obtaining the best MAE, RMSE, and R^2 , and Random Forest remaining highly competitive while achieving the best Median AE. Among the tested methods, tree-based models proved particularly effective in capturing the heterogeneous structure of the dataset. Overall, these results reinforce the value of the RE-VAL framework as a structured environment for comparative evaluation and interpretable urban real-estate analysis.

4.4. Interpretative Results and Comparative Explainability (SHAP Analysis)

To complement the predictive comparison reported in the previous section, SHAP analysis was used to examine how each model structured its valuation logic, both in terms of global feature importance and local contribution patterns. This step was intended to assess whether the different models converged toward economically plausible decision structures and whether similar aggregate errors concealed substantial differences in interpretability. At the global level, the comparison reveals a clear distinction between models that preserve a balanced multivariate valuation structure and models whose explanatory patterns appear overly concentrated or unstable (Figure 3).

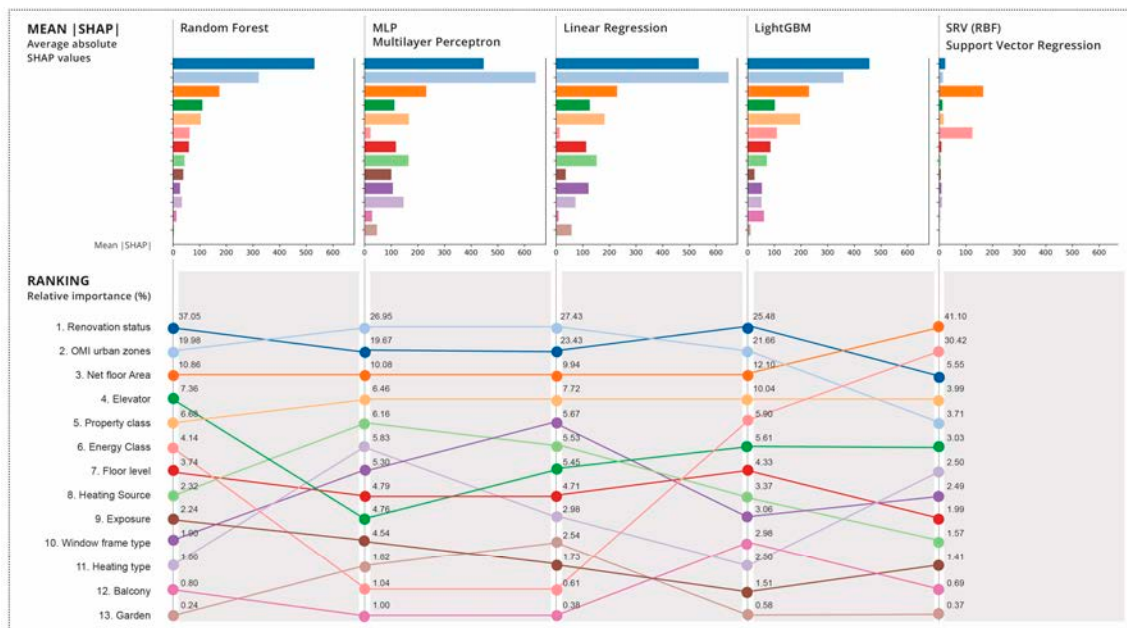


Figure 3. Mean absolute SHAP values and feature rankings across the five benchmarked models. The upper panels report the average absolute SHAP value of each predictor, showing its overall contribution to the predicted unit price within each model. The lower panel compares the ranking of predictors across models, highlighting which variables remain consistently important and which vary in relative importance depending on the modelling approach.

Across LightGBM, Random Forest, Linear Regression, and, more partially, MLP, the most recurrent predictors are OMI Urban Zone, Renovation Status, and Net Floor Area. These are followed, with lower but still relevant contribution, by Property Class, Elevator, Floor Level, and in some cases Energy Class and Heating Source. This hierarchy is economically plausible, since it reflects the combined role of location, dwelling condition, and dimensional characteristics in shaping residential unit prices, while assigning secondary but still meaningful roles to technical and accessibility-related attributes such as floor level, elevator presence, and energy performance.

Although not among the best predictive models in the benchmark, LightGBM provides one of the most balanced explanatory profiles from the perspective of SHAP-based inspection. Its global importance plot distributes relevance across a relatively broad set of variables, with Renovation Status, OMI Urban Zone, and Net Floor Area clearly dominant, followed by Property Class and Energy Class. Random Forest shows a very similar overall hierarchy, again emphasizing renovation, zone, and size as the main value drivers, while retaining non-negligible roles for Elevator, Energy Class, and Floor Level. Linear Regression also captures the main market dimensions, although within a more rigid additive structure in which OMI Urban Zone and Renovation Status dominate more strongly. MLP

partially aligns with these models, since it also gives high importance to OMI Urban Zone, Renovation Status, and Net Floor Area, but its broader ranking appears somewhat less stable. By contrast, SVR shows an excessively concentrated structure dominated by Net Floor Area and especially Energy Class. This profile appears less credible from an appraisal-oriented perspective, because it compresses the valuation process around a narrow subset of predictors. The local SHAP distributions provide a more detailed view of how these variables contribute across observations (Figure 4).

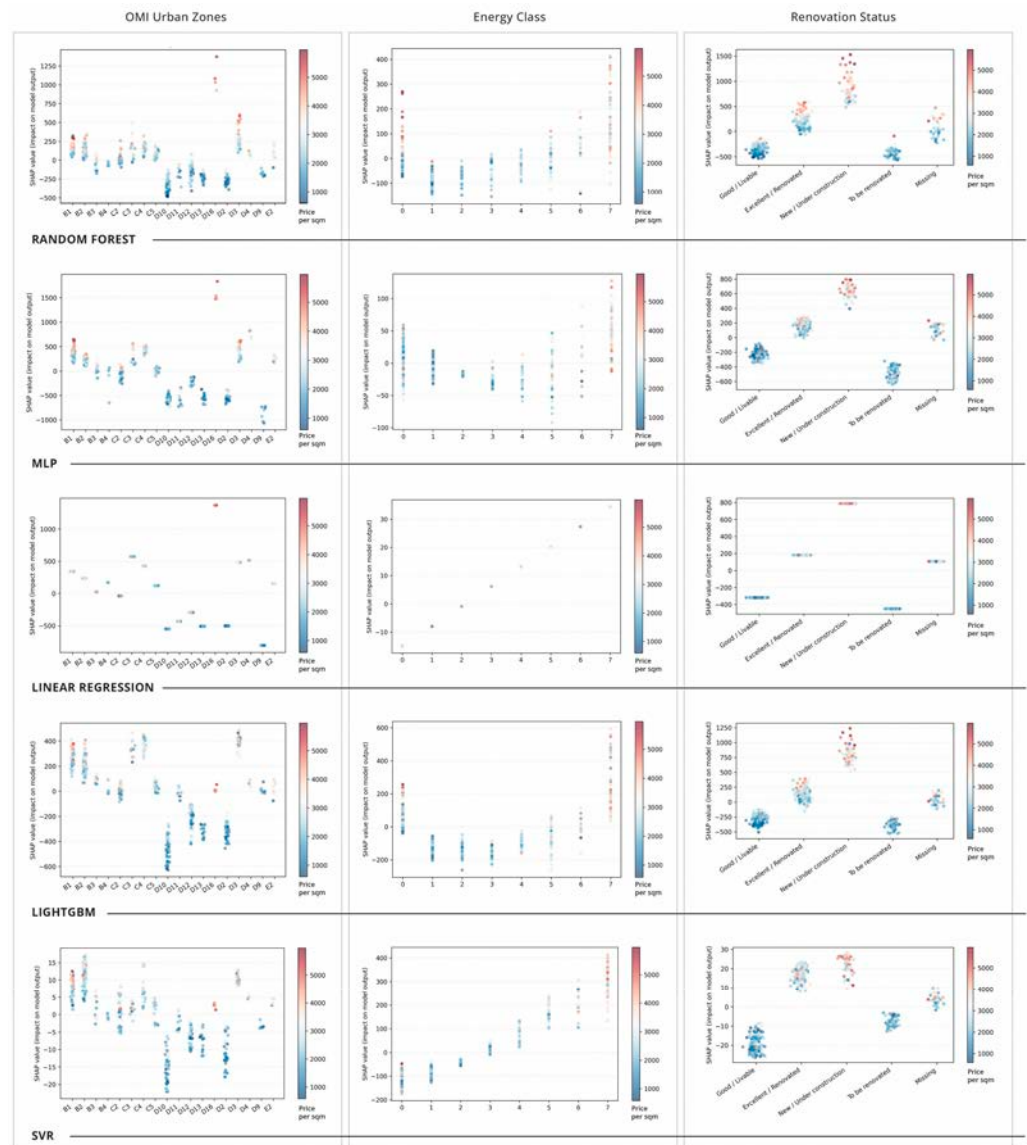


Figure 4. SHAP dependence plots for the five analyzed models and the three selected predictors: OMI urban zone, Energy class, and Renovation status. The x-axis reports the predictor value or category, the y-axis shows the corresponding SHAP value, and point colors indicate the observed price per square meter. The plots show how local feature contributions vary across categories and models.

In LightGBM, the beeswarm shows that positive contributions are systematically associated with renovated or new properties, while lower-quality conditions remain closer to the baseline or exert negative effects. This pattern is economically coherent, since better condition reduces expected refurbishment costs and increases immediate usability. Net Floor Area displays a broad spread of contributions, suggesting that its marginal effect is not fixed but varies according to the broader configuration of the dwelling. This is plausible because, although larger dwellings generally carry greater total value, their

effect on €/m² may depend on layout efficiency, market segment, and location. OMI categories also produce differentiated local impacts, with some zones associated with systematic premiums and others with systematic penalties, which is consistent with the central role of urban location, accessibility, and neighborhood prestige in residential price formation. Taken together, these patterns suggest that LightGBM captures a rich interaction between locational, structural, and qualitative characteristics, even though its predictive performance remains below that of the strongest models.

Random Forest confirms much of the same explanatory structure, but with a more competitive predictive profile. Its beeswarm again shows coherent positive contributions for renovated and new properties, differentiated premiums and discounts across OMI zones, and a flexible effect of Net Floor Area that changes across observations rather than following a fixed linear pattern. From an appraisal perspective, this is a credible result, because the market rarely rewards characteristics such as condition, size, and location through constant additive adjustments. Compared with LightGBM, however, Random Forest appears slightly more segmented, with sharper local jumps and more concentrated extreme contributions in some urban zones. Even so, it remains one of the most convincing models when both predictive competitiveness and explanatory coherence are considered jointly.

Linear Regression provides a different but still useful interpretive profile. The local SHAP patterns are more compact and regular, reflecting the additive nature of the model. Net Floor Area exerts a strong and clearly directional effect, but with less dispersion than in the tree-based models, indicating a more uniform marginal contribution across properties. The same is true for the main categorical predictors, whose effects align along narrower bands. This makes Linear Regression easy to interpret and valuable as a transparent benchmark, but it also confirms its more limited ability to represent context-dependent valuation effects, such as cases in which the premium associated with floor level, energy quality, or size may vary across urban submarkets.

MLP occupies an intermediate position. On the one hand, it is the strongest predictive model in the benchmark, achieving the best MAE, RMSE, and R². On the other hand, its local explanations are less stable than those of the best tree-based models. The beeswarm still shows meaningful directional behavior for Net Floor Area, Renovation Status, and some OMI categories, but it also contains more scattered patterns and isolated high-impact observations that are harder to interpret consistently at the case level. This means that MLP performs very well as a predictor, but its internal valuation logic is less readable than that of Random Forest and, to some extent, LightGBM.

The weakest interpretative profile is observed for SVR. Its beeswarm is heavily driven by Net Floor Area and Energy Class, while many other variables remain tightly compressed around zero. This suggests that the model is failing to recover a sufficiently articulated multivariate logic, instead relying excessively on a narrow subset of predictors. The prominence assigned to Energy Class appears especially difficult to reconcile with the broader cross-model evidence, since the other competitive models treat it as a secondary rather than dominant factor. For this reason, despite being computationally viable, SVR appears less suitable when interpretability is considered a substantive requirement.

From an appraisal and urban-economic perspective, several of the local SHAP patterns are substantively coherent. The differentiated premiums and penalties associated with OMI Urban Zones are consistent with the established role of location in shaping housing values through accessibility, amenity concentration, and neighborhood prestige. The positive contributions associated with renovated or new properties are likewise expected, since better condition reduces anticipated refurbishment costs and increases immediate usability. Floor-related attributes should be interpreted more contextually: higher floors may command a premium because of views, privacy, and natural light, but in buildings

without an elevator the same characteristic may become less attractive, especially for less mobile households.

A similar logic applies to elevator presence, whose value is likely to be stronger in multi-storey buildings and in combination with higher floor levels. Finally, Energy Class appears economically meaningful when it acts as a secondary feature, since better energy performance may plausibly support a premium through lower operating costs, greater comfort, and stronger appeal under sustainability-oriented preferences. These patterns should, however, be interpreted as economically plausible associations recovered by the models rather than as direct causal effects.

A more specific spatial reading also helps contextualize the role of Energy Class and OMI Urban Zone. The differentiated premiums and penalties associated with OMI categories are consistent with the established role of location in shaping housing values through accessibility, amenity concentration, neighborhood prestige, and submarket structure. The spatial distribution of local SHAP contributions further highlights the presence of geographically concentrated explainability patterns, revealing that the influence of locational attributes is not uniformly distributed across the urban system but tends to intensify within central and waterfront areas of Cagliari. These patterns support the interpretation that machine learning models capture localized valuation dynamics and spatial heterogeneity beyond global feature importance rankings (Figure 5).

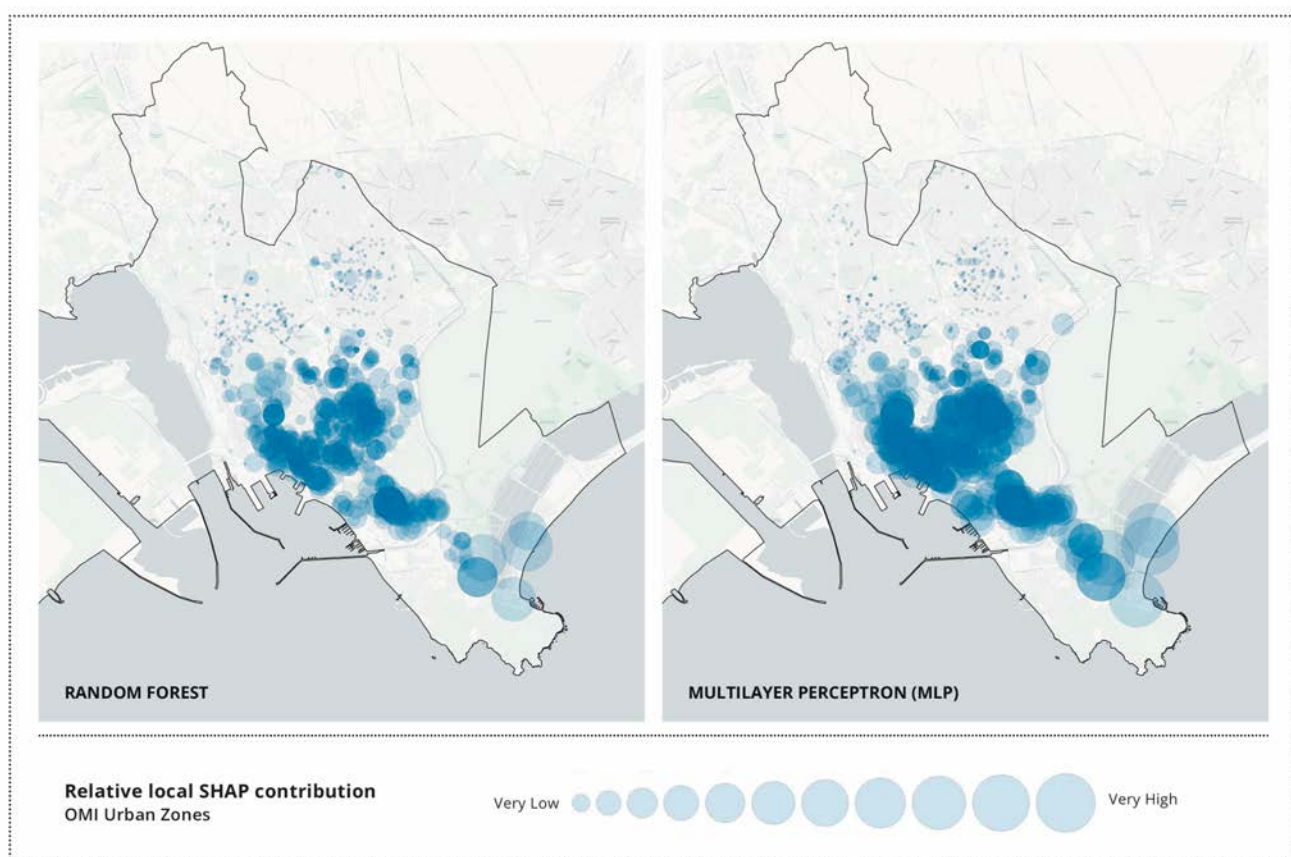


Figure 5. Spatial distribution of local SHAP contributions for OMI urban zones in the Random Forest and Multilayer Perceptron (MLP) models. Bubble size represents the magnitude of local SHAP values, highlighting the spatial concentration and heterogeneity of location effects on housing price predictions across Cagliari.

This interpretation is coherent with previous studies showing that housing prices respond to retail accessibility, spatial gradients, and environmental externalities, while urban

submarkets may exhibit distinct valuation logics even within the same city [9,10,29,30]. Within this framework, the role of Energy Class appears more context-dependent: the present interaction patterns are compatible with the interpretation that better energy performance may command a stronger premium in higher-prestige areas, where buyers may place greater weight on comfort, construction quality, and sustainability-oriented standards. This reading should, however, be understood as an economically plausible interpretation of the observed model behaviour rather than as direct evidence of a causal mechanism. Figure 5 reports SHAP dependence plots for the three selected predictors: OMI urban zone, Energy class, and Renovation status.

A particularly relevant result of the comparative SHAP analysis is that models with similar aggregate metrics can still differ substantially in explanatory quality. MLP, Random Forest, and Linear Regression are the strongest models in predictive terms, but they do not offer the same interpretive properties. MLP is the best pure predictor, yet its local explanatory structure is noisier. Linear Regression remains highly transparent, but its additive form limits its ability to capture context-sensitive variation. Random Forest, by contrast, combines competitive predictive results with a more coherent non-linear explanatory structure, making it the most balanced compromise between predictive competitiveness and interpretability. LightGBM is also interpretively valuable, although its predictive results are weaker than those of the top three models.

In LightGBM and Random Forest Energy Class appears as a secondary but visible predictor, which is compatible with the idea that energy performance may contribute to value under specific market conditions without dominating the appraisal process. This is economically plausible, since more efficient dwellings may offer lower operating costs, improved thermal comfort, and stronger appeal in markets increasingly attentive to sustainability. In Linear Regression and MLP, its role is much weaker, while in SVR it becomes disproportionately important. This divergence suggests that any interpretation of energy-related premiums should remain cautious. The present results do not establish a generalized green premium. Rather, they indicate that energy efficiency may matter, but normally alongside, rather than above, more established drivers such as location, condition, and size.

For missingness-related effects, some models attribute non-negligible contributions to missing categories or missingness indicators, especially for variables such as Renovation Status, Floor Level, or Heating Source. These patterns are compatible with the idea that omitted technical information may retain proxy value, possibly because omission is not random in real-estate listings. At the same time, the present analysis does not isolate missingness experimentally, so these effects should not be interpreted as proof of a causal mechanism. A safer conclusion is that preserving missingness-related signals does not appear to prevent coherent model behavior in the stronger models and may help retain useful information under imperfect listing conditions.

Overall, the SHAP comparison shows that model adequacy cannot be judged on predictive performance alone. MLP stands out as the strongest predictor, Linear Regression remains the clearest interpretive baseline, and Random Forest provides the most convincing balance between predictive competitiveness and explanatory coherence. LightGBM offers an informative non-linear explanatory structure but does not reach the same predictive level, while SVR shows distortions that make them less suitable for operational valuation settings in which interpretability is treated as a substantive requirement rather than a secondary diagnostic.

4.5. Substantive Insights for the RE-VAL Framework

From a substantive perspective, the comparative SHAP analysis supports the idea that the most credible valuation models are those able to jointly represent spatial structure,

property quality, and dimensional heterogeneity. Across the stronger models, the repeated prominence of OMI Urban Zone, Renovation Status, and Net Floor Area is broadly consistent with appraisal logic, while the role of variables such as Floor Level, Elevator, and Energy Class appears more context dependent.

This suggests that the same attribute may not carry a fixed market premium across the city but may instead derive its valuation effect from the interaction between dwelling quality, locational prestige, and the broader urban submarket in which it is embedded. The comparative SHAP analysis further suggests that model adequacy should not be judged on predictive performance alone, but also on the structural credibility of the explanations produced.

Within this benchmark, Random Forest appears to provide the most convincing balance between predictive competitiveness and explanatory coherence, while MLP remains the strongest pure predictor. The results also indicate that preserving missing or unspecified information can remain compatible with coherent model behavior in imperfect listing environments, although no causal claim about missingness can be established from the present design.

5. Discussion: Beyond Predictive Accuracy

The results suggest that model adequacy in urban real-estate valuation cannot be judged on predictive error alone. Although MLP achieved the strongest aggregate predictive performance, Random Forest provided the most balanced combination of predictive competitiveness and explanatory coherence. Linear Regression remained useful as a transparent additive baseline, since the margin between the non-linear models is small. This distinction is important because similar cross-validation metrics may still correspond to very different internal valuation logics. In this context, model evaluation requires not only predictive assessment but also an examination of the economic plausibility of the underlying attribution structure.

In the present case study, the comparative SHAP analysis shows that stronger models are not only those with lower error, but also those that preserve an economically credible balance between locational, structural, and qualitative drivers of value. From an explanatory coherence perspective, this balance represents a key criterion of model adequacy, as it reflects consistency with established valuation principles.

In this perspective, although the predictive performance reflects the inherent noise of listing prices in a fragmented and non-transparent market like Cagliari, the RE-VAL framework's ability to capture the direction and magnitude of feature impacts through SHAP values may be more valuable for public decision-makers than mere point-estimate precision. Rather than focusing solely on the absolute accuracy of a single price, the framework provides a robust understanding of the underlying market dynamics, allowing stakeholders to identify which urban features are consistently associated with price variations and drive or depress property values. From a professional practice perspective, the proposed 'Bonus/Malus' framework operationalizes Explainable AI in a way that resonates with the 'Market Approach' (Comparison Method) defined by International Valuation Standards (IVS) and RICS standards.

In a traditional appraisal, a valuer must apply quantitative or qualitative adjustments to comparable properties to account for differences in features, such as floor level or energy efficiency. By translating SHAP values into monetary feature-level adjustments, RE-VAL provides a data-driven justification for these coefficients, ensuring that the model's reasoning is compliant with professional standards that require the valuer to explain and document the logic behind value adjustments. Compared with existing studies, the present findings are partly convergent and partly more cautious. On the one hand, the recurrent prominence of location-related factors, dwelling condition, and dimensional characteristics

is broadly consistent with previous research showing that accessibility, submarket structure, and environmental context remain central to urban housing values, as highlighted in next sections.

On the other hand, the benchmark reported here does not show a large predictive advantage of advanced non-linear models over the linear baseline, but rather a narrower performance gap than is sometimes suggested in the literature on machine-learning-based valuation [14–21]. In this sense, the present case study supports the view that model superiority is strongly context-dependent and may vary with dataset size, market structure, and feature representation [15,20,21]. Accordingly, the main contribution of RE-VAL lies less in demonstrating a universally dominant predictor than in providing a transparent workflow that combines comparative benchmarking, explanatory inspection, and operational monetary interpretation within a common valuation framework. A methodological comparison between traditional appraisal approaches and the proposed RE-VAL framework is reported in Table 5.

Table 5. Methodological comparison between traditional appraisal and the RE-VAL framework.

Analytical Dimension	Traditional Appraisal (Linear/Static)	RE-VAL Framework (Explainable Workflow)	Operational Impact
Valuation Logic	Additive & Rigid: Assumes constant marginal contributions (e.g., fixed €/m ²).	Non-linear & Interactive: Supports the modelling of variable marginality and cross-feature synergies.	Help identify heterogeneous valuation effects across property characteristics and urban contexts.
Transparency	Expert-driven and partly implicit: relies on professional judgement, with limited formal decomposition of feature-level contributions.	XAI-supported interpretation: SHAP-based decomposition of feature contributions.	Supports alignment with IVS/RICS standards through transparent monetary Bonus/Malus tables.
Resilience	Fragile: Incomplete technical information may reduce consistency and interpretability.	Preserves missingness-related patterns and treats them as potentially informative signals when appropriate.	May improve robustness, or help retain useful information, when working with fragmented and imperfect real-estate listing data.
Final Purpose	Static Point-Estimate: A one-time snapshot of property value.	Decision-support-oriented workflow supports scenario interpretation and what-if reasoning.	Facilitates legal and administrative admissibility (e.g., for taxation or courts) by providing auditable evidence.

A further point of differentiation is that RE-VAL does not treat missingness and explainability as separate technical add-ons: incomplete attributes are preserved as potential proxy signals within the preprocessing logic, and SHAP outputs are carried forward into monetary Bonus/Malus adjustments intended for operational appraisal use.

From an operational standpoint, the framework translates the complexity of Machine Learning into an interpretable ‘Bonus/Malus’ logic, familiar to professional appraisal practice. By interpreting SHAP values, it is possible to quantify how specific features—such as upgrading to Energy Class A or the presence of an elevator—proportionally weight the final price within a specific urban micro-zone. This allows the valuation process to be interpreted as a structured set of additive contributions, consistent with comparative appraisal logic rather than as a purely algorithmic output. This transforms black-box predictions into actionable indicators that can inform incentive policies or urban regeneration strategies.

5.1. Explanatory Balance as a Criterion of Model Adequacy

A key insight of the comparative analysis is that near-similar predictive performance does not imply the same valuation logic. The stronger models differ in how they distribute explanatory relevance across location, property condition, and dimensional characteristics. From an appraisal-oriented perspective, the most credible profiles are those that preserve a balanced role for OMI Urban Zone, Renovation Status, and Net Floor Area, rather than concentrating excessively on a narrow subset of predictors. This balanced structure reflects the expected hierarchy of valuation drivers, where location, condition, and size jointly determine property value.

This is especially relevant for variables such as Energy Class, whose value may be amplified in high-prestige submarkets, and for OMI-zone effects, which plausibly capture differences in accessibility to central functions, service concentration, and environmental attractiveness. Such context-dependent effects suggest that the model captures non-linear valuation mechanisms that remain coherent with urban economic theory. This is one reason why Random Forest appears particularly convincing within the RE-VAL framework: it combines competitive predictive results with a non-linear explanatory structure that remains economically plausible across both global and local inspection.

Regarding the consistency of the findings, the choice of SHAP (Shapley Additive Explanations) over alternative explainability techniques, such as LIME or permutation-based feature importance, is grounded in its solid foundation in coalitional game theory. Unlike local surrogate models that may suffer from instability, SHAP ensures consistency and local accuracy, providing a unique distribution of the ‘payout’ (the price prediction) among the input features. This theoretical robustness is essential for the proposed Bonus/Malus framework, as it ensures that the adjustments derived are not artifacts of the specific explanation method but reflect the actual contribution of property characteristics as learned by the model.

By contrast, weaker models such as SVR show more distorted explanatory hierarchies, suggesting that predictive modelling should be evaluated not only by error metrics, but also by the structural credibility of the valuation logic it produces. From this perspective, explainability becomes a methodological tool for assessing model validity, rather than a purely descriptive output. Furthermore, this approach addresses the need for transparency and accountability in formal valuation processes, including taxation and judicial contexts, where the ‘black-box’ nature of standard ML models has historically been a barrier to legal admissibility. By providing a clear decomposition of how each feature contributes to the final price, the framework supports the valuer’s role as a final responsible auditor of the automated output.

5.2. Robustness to Incomplete Information Through Informative Proxy Signals

Beyond predictive performance and interpretability, an additional methodological insight concerns the treatment of incomplete technical information. In the present study, preserving missing or unspecified categories proved compatible with coherent model behaviour in the stronger models, while weaker models appeared more unstable under the same data conditions. It is crucial to clarify that the identified relationships between missingness and price represent statistical associations within the specific dataset and do not imply a direct causal effect. A clear distinction must be maintained between correlation and causation: the presence of missing data is associated with unobserved property qualities or seller behaviors, but it does not “cause” a change in value.

This does not demonstrate a causal effect of missingness, nor does it establish that all omitted information is substantively meaningful. Moreover, the present analysis does not demonstrate the superiority of MNAR-based encoding over alternative missing-data

treatments. Rather, the results suggest that preserving missingness-related information does not compromise the explanatory coherence of the model and may retain potentially relevant signals in imperfect listing environments.

It does, however, suggest that incomplete listing attributes should not automatically be discarded as pure noise. Within the RE-VAL framework, their preservation can be understood as a pragmatic modelling choice for imperfect digital listing environments, especially when the objective is to retain potentially useful signals without imposing stronger assumptions than the data can support.

6. Practical Implications: Decision-Support Potential of the RE-VAL Framework

A relevant implication of the RE-VAL framework lies in its potential to support more operational and interpretable valuation workflows for appraisers, investors, and public administrations.

6.1. Supporting Context-Sensitive Appraisal Reasoning

A central practical implication of the RE-VAL framework is that it supports a more context-sensitive reading of property value than static linear adjustment logic alone. In operational settings, this means that appraisers, analysts, or public-sector users can inspect whether the contribution of a characteristic such as floor area, renovation condition, or energy performance remains stable across properties or varies depending on urban context. The value of the framework therefore lies not in replacing expert judgement, but in providing a structured analytical basis for understanding why similar properties may position themselves differently within the broader range of local market references. This is particularly relevant in fragmented urban markets, where institutional averages may not fully capture submarket-level variation. This is especially useful when inspecting whether features such as energy performance exhibit stable or context-dependent effects across urban submarkets.

6.2. Operational Use Under Imperfect Listing Conditions

A second practical implication concerns the usability of the framework in the incomplete data environments that characterize many digital real-estate platforms. In applied valuation settings, technical specifications are often unevenly reported, inconsistently entered, or omitted altogether. The RE-VAL workflow addresses this condition by preserving selected missing or unspecified categories rather than requiring complete technical descriptions for all observations. In practical terms, this means that the framework can remain operational even when listing quality is heterogeneous. The present study does not prove that missingness is causally informative, but it does suggest that preserving these signals may help retain useful information and avoid unnecessary data loss in real-world valuation workflows.

6.3. From Point-Estimates to Decision Support: Monetary Bonus/Malus Synthesis

The main decision-support value of the RE-VAL framework lies in its ability to express model behaviour in forms that are more legible to human users. Through the monetary Bonus/Malus synthesis, the estimated value of a property can be broken down into feature-level contributions expressed in the original price scale. This makes automated estimates easier to inspect, compare, and discuss, especially in contexts where a final point estimate alone would be too opaque for professional use.

In practical terms, the framework can support appraisal reasoning, scenario inspection, and communication with stakeholders by showing how specific characteristics contribute positively or negatively to the estimated value. Rather than functioning as a prescriptive

decision tool, it offers an auditable analytical layer that may support more transparent investment, appraisal, and policy-oriented interpretation, although its practical usability still requires validation with relevant stakeholders.

7. Conclusions and Policy Implications

This study introduced RE-VAL, an explainable machine-learning framework designed to support urban real-estate valuation and decision-making. By integrating reproducible data collection with XAI-based interpretation, the framework addresses the “black-box” challenge of advanced predictive models, translating complex non-linear relationships into operational “Bonus/Malus” adjustment tables. Applied to the Cagliari case study, the framework showed how urban valuation can benefit from combining benchmark-based model assessment with post hoc interpretability and operational monetary synthesis. In this sense, the main contribution of the study lies not in the definition of a single predictive architecture, but in the construction of a transparent and transferable workflow for data-informed real-estate analysis in complex urban contexts.

The empirical results highlighted three main contributions. First, the comparative evaluation showed that non-linear models were better suited to the present case study but did not show a large advantage over the linear baseline. Among them, MLP achieved the strongest aggregate predictive performance, while Random Forest emerged as the most balanced model when predictive competitiveness was considered jointly with interpretability. More importantly, the interpretability analysis indicated that aggregate predictive metrics alone do not fully capture the differences in modelling behaviour. While some models achieved comparable scores on standard error measures, the SHAP-based diagnostics revealed substantial differences in their ability to represent heterogeneous, context-sensitive valuation patterns across structural, contextual, and location-related variables. This confirms the value of combining predictive benchmarking with interpretability analysis when studying urban real-estate markets.

Second, the framework suggests the methodological relevance of preserving incomplete listing information as a potentially informative component of the valuation process, while acknowledging that the present study does not isolate missingness experimentally. By preserving and encoding missingness-related patterns within the preprocessing workflow, RE-VAL may support greater robustness in the presence of sparse, inconsistent, or non-standardized web-scraped data. This is particularly relevant for real-world valuation settings, where listing quality is uneven and where incomplete technical specifications may themselves reflect market segmentation, reporting behaviour, or building typologies.

Third, the framework showed how explainability outputs can be translated into forms that are more directly applicable to appraisal reasoning and decision support. Through SHAP-based monetary decomposition, predicted values can be articulated into feature-level Bonus/Malus contributions, making the valuation logic more transparent and operationally interpretable.

Finally, it is crucial to clarify that the insights provided by Explainable AI (XAI) identify recurring market patterns and associations rather than strictly deterministic cause-and-effect relationships. The model describes how the local market perceives and values specific property characteristics at a given time. Therefore, these results should be interpreted as a reliable empirical snapshot of urban market behavior, providing a support tool for decision-making rather than a universal law of causality.

The policy relevance of the framework lies in its capacity to support more transparent and context-sensitive urban analysis. By linking predictive modelling with interpretable monetary synthesis, RE-VAL can help identify how property characteristics, energy performance, and spatial context interact in ways that may be relevant for appraisal practice,

urban regeneration strategies, and sustainability-oriented policy design. Although the present study does not propose a prescriptive policy tool, it provides an operational analytical basis that may support more informed decisions by public administrations, financial actors, and market stakeholders.

Despite these contributions, a significant limitation of this study is the reliance on listing (offer) prices rather than actual transaction prices. In real-estate economics, listing prices represent the initial expectations of sellers and may not reflect the final market value, especially in fluctuating markets where the gap between asking and closing prices can be substantial. This reliance likely influences the model's predictive accuracy, as it captures 'offered' values rather than 'realized' ones. Consequently, the derived Bonus/Malus adjustment coefficients should be interpreted as reflecting the premium or discount perceived by the market during the advertising phase, rather than a definitive measure of transaction value. Despite this limitation, the use of high-frequency listing data remains a vital tool for real-time urban monitoring, provided that valuers and policymakers remain aware of this systematic bias when applying model outputs to formal appraisal contexts.

A further set of limitations concerns components of the framework that remain methodologically promising but not yet fully benchmarked. In particular, the preservation of incomplete listing attributes as potentially informative proxy signals was not compared systematically with alternative treatments such as conventional imputation or exclusion of incomplete observations. Likewise, although spatial context is incorporated through geographically informed variables, the present study does not test residual spatial dependence directly or compare the benchmark with geographically weighted or other spatially explicit models. More generally, the current verification remains limited to cross-validation within a single dataset and reference period, without broader hold-out, temporal, or sensitivity-based assessment.

Future research should therefore pursue these extensions in parallel, so as to clarify when the present workflow captures genuinely informative structure and when some of its effects may instead reflect context-specific or potentially spurious dependencies. Future research could extend the framework in several directions, including the temporal monitoring of market change, validation across additional cities, and deeper integration of scenario-based urban policy analysis. Further work should also explore how shifts in housing demand, regulatory conditions, or socio-economic context alter the explanatory structure of urban valuation over time. Overall, the results suggest that explainable and benchmark-driven workflows such as RE-VAL can provide a useful foundation for more transparent, reproducible, and operationally meaningful urban real-estate analysis.

Author Contributions: Supervision, V.S.; conceptualization, V.S. and M.M.; methodology, V.S. and M.M.; software, M.M.; formal analysis, M.M.; validation, V.S. and M.M.; investigation, V.S. and M.M.; data curation, V.S. and M.M.; writing—original draft preparation, V.S. and M.M.; writing—review and editing, V.S. and M.M.; visualization, V.S. and M.M. All authors have read and agreed to the published version of the manuscript.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The original contributions presented in this study are included in the article. Further inquiries can be directed to the corresponding author.

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Wei, C.; Fu, M.; Wang, L.; Yang, H.; Tang, F.; Xiong, Y. The Research Development of Hedonic Price Model-Based Real Estate Appraisal in the Era of Big Data. *Land* **2022**, *11*, 334. [CrossRef]
2. Simonotti, M. I Procedimenti di Stima su Larga Scala (Mass Appraisal). In *Ce. SET: Quaderni. 8-Mercato Immobiliare, Innovazione e Gestione Dei Catasti Urbani*; Firenze University Press: Florence, Italy, 2002; Volume 1, pp. 446–471. Available online: <https://www.torrossa.com/en/resources/an/2241278> (accessed on 7 April 2026).
3. International Valuation Standards Council (IVSC). Standards. Available online: <https://ivsc.org/standards/> (accessed on 7 April 2026).
4. TEGOVA. European Valuation Standards (EVS). Available online: <http://tegova.org/european-valuation-standards-evs> (accessed on 7 April 2026).
5. RICS Valuation—Global Standards (Red Book). Available online: <https://www.rics.org/profession-standards/rics-standards-and-guidance/sector-standards/valuation-standards/red-book> (accessed on 7 April 2026).
6. Sirmans, G.S.; Macpherson, D.A.; Zietz, E.N. The Composition of Hedonic Pricing Models. *J. Real Estate Lit.* **2005**, *13*, 3–43. [CrossRef]
7. Bárcena, M.J.; Menéndez, P.; Palacios, M.B.; Tusell, F. Alleviating the Effect of Collinearity in Geographically Weighted Regression. *J. Geogr. Syst.* **2014**, *16*, 441–466. [CrossRef]
8. Osland, L. An Application of Spatial Econometrics in Relation to Hedonic House Price Modeling. *J. Real Estate Res.* **2010**, *32*, 289–320. [CrossRef]
9. Jang, M.; Kang, C.-D. Retail Accessibility and Proximity Effects on Housing Prices in Seoul, Korea: A Retail Type and Housing Submarket Approach. *Habitat Int.* **2015**, *49*, 516–528. [CrossRef]
10. Chica-Olmo, J.; Cano-Guervos, R.; Tamaris-Turizo, I. Determination of Buffer Zone for Negative Externalities: Effect on Housing Prices. *Geogr. J.* **2019**, *185*, 222–236. [CrossRef]
11. Ma, J.; Cheng, J.C.P.; Jiang, F.; Chen, W.; Zhang, J. Analyzing Driving Factors of Land Values in Urban Scale Based on Big Data and Non-Linear Machine Learning Techniques. *Land Use Policy* **2020**, *94*, 104537. [CrossRef]
12. Guo, J.; Chiang, S.; Liu, M.; Yang, C.-C.; Guo, K. Can Machine Learning Algorithms Associated with Text Mining from Internet Data Improve Housing Price Prediction Performance? *Int. J. Strateg. Prop. Manag.* **2020**, *24*, 300–312. [CrossRef]
13. Wu, C.; Ye, X.; Ren, F.; Du, Q. Modified Data-Driven Framework for Housing Market Segmentation. *J. Urban Plan. Dev.* **2018**, *144*, 04018036. [CrossRef]
14. Little, R.J.A.; Rubin, D.B. *Statistical Analysis with Missing Data*; John Wiley & Sons: Hoboken, NJ, USA, 2019; ISBN 978-0-470-52679-8.
15. Çılgin, C.; Gökşen, Y.; Gökşen, H. The Effect of Outlier Detection Methods in Real Estate Valuation with Machine Learning [Makine Öğrenimi İle Mülk Değerlemesinde Aykırı Değer Tespit Yöntemlerinin Etkisi]. *İzmir J. Soc. Sci.* **2023**, *5*, 9–20. [CrossRef]
16. Trindade Neves, F.; Aparicio, M.; de Castro Neto, M. The Impacts of Open Data and eXplainable AI on Real Estate Price Predictions in Smart Cities. *Appl. Sci.* **2024**, *14*, 2209. [CrossRef]
17. Hernes, M.; Tutak, P.; Nadolny, M.; Mazurek, A. Real Estate Valuation Using Machine Learning. *Procedia Comput. Sci.* **2024**, *246*, 4592–4599. [CrossRef]
18. Ho, W.K.O.; Tang, B.-S.; Wong, S.W. Predicting Property Prices with Machine Learning Algorithms. *J. Prop. Res.* **2021**, *38*, 48–70. [CrossRef]
19. Abidoye, R.B.; Chan, A.P.C. Artificial Neural Network in Property Valuation: Application Framework and Research Trend. *Prop. Manag.* **2017**, *35*, 554–571. [CrossRef]
20. Kalliola, J.; Kapočiūtė-Dzikiene, J.; Damaševičius, R. Neural Network Hyperparameter Optimization for Prediction of Real Estate Prices in Helsinki. *PeerJ Comput. Sci.* **2021**, *7*, e444. [CrossRef] [PubMed]
21. Martínez Díaz, E.J. *Predicting the Housing Market with Machine Learning*; Polytechnic University of Puerto Rico: San Juan, Puerto Rico, 2023.
22. Mukesh, S.; Gowda, H.M.; Adarsh, H.R.; Darshan, Y.K. Real Estate Price Prediction Using Machine Learning Techniques. *Int. J. Res. Pub. Rev.* **2025**, *6*, 4211–4215.
23. Francke, M.; van de Minne, A. Combining Machine Learning and Econometrics: Application to Commercial Real Estate Prices. *Real Estate Econ.* **2024**, *52*, 1308–1339. [CrossRef]
24. Zitoune, I.; Arabov, M.K. Comparative Analysis of Ensemble and Linear Machine Learning Models in the Task of House Price Prediction. In Proceedings of the 2024 International Russian Automation Conference (RusAutoCon), Sochi, Russia, 8–14 September 2024; pp. 50–55.
25. Berry, N. Modern Web Scraping and Data Analysis Tools to Discover Historic Real Estate Development Opportunities. Ph.D. Thesis, Massachusetts Institute of Technology, Cambridge, MA, USA, 2022.
26. Üzümcü, A.C.; EliGüzel, N. Predictive Analysis Using Web Scraping for the Real Estate Market in Gaziantep. *Bitlis Eren Üniversitesi Fen Bilim. Derg.* **2023**, *12*, 17–24. [CrossRef]

27. Santos, J.M.A. Real Estate Market Data Scraping and Analysis for Financial Investments. Master's Thesis, Universidade do Porto, Porto, Portugal, 2018.
28. Pineda Montserrat, B. Predictive Business Analytics for Real Estate: A Tool for Estimating and Analyzing Housing Prices. Master's Thesis, Universitat Politècnica de Catalunya, Barcelona, Spain, 2024.
29. Souza, T.G.D.; Fonseca, F.D.R.; Fernandes, V.D.O.; Pedrassoli, J.C. Exploratory Spatial Analysis of Housing Prices Obtained from Web Scraping Technique. *Int. Arch. Photogramm. Remote Sens. Spat. Inf. Sci.* **2021**, *XLIII-B4-2021*, 135–140. [[CrossRef](#)]
30. Meyberg, C.; Rendtel, U.; Leerhoff, H. Flat Rent Price Prediction in Berlin with Web Scraping. *ASTA Wirtsch. Sozialstatistisches Arch.* **2024**, *18*, 245–278. [[CrossRef](#)]
31. Helbich, M.; Brunauer, W.; Hagenauer, J.; Leitner, M. Data-Driven Regionalization of Housing Markets. *Ann. Assoc. Am. Geogr.* **2013**, *103*, 871–889. [[CrossRef](#)]
32. Cugurullo, F. New Stories of Urban AI: Exploring the Artificial Intelligence–City Nexus beyond Frankenstein Urbanism. *Urban Geogr.* **2024**, *45*, 1300–1307. [[CrossRef](#)]
33. Rey-Blanco, D.; Arbues, P.; Lopez, F.; Paez, A. A Geo-Referenced Micro-Data Set of Real Estate Listings for Spain's Three Largest Cities. *Environ. Plan. B Urban Anal. City Sci.* **2024**, *51*, 1369–1379. [[CrossRef](#)]
34. Silaghi, V.; Alssadi, Z.; Mathew, B.; Alotaibi, M.; Alqarni, A.; Silaghi, M. Modeling the Feedback of AI Price Estimations on Actual Market Values. *arXiv* **2024**, arXiv:2405.18434.
35. Wheaton, W.C.; Xu, C. Using AI to Improve Price Transparency in Real Estate Valuation. *MIT Cent. Real Estate Res. Pap.* **2024**. [[CrossRef](#)]
36. Jaouhari, A.E.; Samadhiya, A.; Kumar, A.; Šešplaukis, A.; Raslanas, S. Mapping the Landscape: A Systematic Literature Review on Automated Valuation Models and Strategic Applications in Real Estate. *Int. J. Strateg. Prop. Manag.* **2024**, *28*, 286–301. [[CrossRef](#)]
37. Coletta, A.; Prata, M.; Conti, M.; Mercanti, E.; Bartolini, N.; Moulin, A.; Vyetenko, S.; Balch, T. Towards Realistic Market Simulations: A Generative Adversarial Networks Approach. *arXiv* **2021**, arXiv:2110.13287.
38. D'Amato, M.; Kauko, T. *Advances in Automated Valuation Modeling*; Springer: Berlin/Heidelberg, Germany, 2017.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.