Financial Innovation

RESEARCH

Open Access

Toward transparent and accurate housing price appraisal: Hedonic price models versus machine learning algorithms

Sihyun An[1], Yena Song[2], Hanwool Jang[3] and Kwangwon Ahn[4*]

*Correspondence: k.ahn@yonsei.ac.kr

[1] Division of Finance and AI, Hankuk University of Foreign Studies, Yongin, South Korea
[2] Department of Geography, Chonnam National University, Gwangju, South Korea
[3] Department of Finance, Accounting and Risk, Glasgow Caledonian University, Glasgow, UK
[4] Department of Industrial Engineering, Yonsei University, Seoul, South Korea

Abstract

The nonlinearity of hedonic datasets demands flexible automated valuation models to appraise housing prices accurately, and artificial intelligence models have been employed in mass appraisal to this end. However, they have been referred to as "black-box" models owing to difficulties associated with interpretation. In this study, we compared the results of traditional hedonic pricing models with those of machine learning algorithms, e.g., random forest and deep neural network models. Commonly implemented measures, e.g., Gini importance and permutation importance, provide only the magnitude of each explanatory variable's importance, which results in ambiguous interpretability. To address this issue, we employed the SHapley Additive exPlanation (SHAP) method and explored its effectiveness through comparisons with traditionally explainable measures in hedonic pricing models. The results demonstrated that (1) the random forest model with the SHAP method could be a reliable instrument for appraising housing prices with high accuracy and sufficient interpretability, (2) the interpretable results retrieved from the SHAP method can be consolidated by the support of statistical evidence, and (3) housing characteristics and local amenities are primary contributors in property valuation, which is consistent with the findings of previous studies. Thus, our novel methodological framework and robust findings provide informative insights into the use of machine learning methods in property valuation based on the comparative analysis.

Keywords: Hedonic price model, Importance measure, Machine learning, Housing price appraisal

Introduction

Real estate is an essential sector of social and economic systems, and real estate prices fluctuate across time and place, which can affect economic stability (Glaeser et al. 2014; Tchuente and Nyawa 2022). Capturing dynamic patterns in real estate prices requires flexible valuation models that utilize sophisticated algorithms to precisely capture the nonlinearity of housing price datasets. Moreover, housing prices are affected by various attributes (Anderson and West 2006; Czembrowski and Kronenberg 2016; An et al. 2024); therefore, the interpretability of model outcomes has been accentuated to reinforce our understanding of the pattern in property prices. This approach enables us to

Springer

ensure the reliability of decision-making processes based on valuation models (Chuan et al. 2024).

Hedonic price models (HPMs) have been typically used to estimate housing price appraisals. Valuation models have adopted a variety of explanatory variables, such as housing factors, local amenities, demographic characteristics, and spatial or geographic characteristics (McMillan et al. 1980; Chin and Chau 2003; Hong et al. 2020; Wang 2023). However, HPMs assume a linear relationship between hedonic variables and housing prices; thus, these approaches face continuous criticisms regarding whether they can estimate housing prices effectively (Chin and Chau 2003; Hong et al. 2020; Li et al. 2021). These debates focus on the complexity of property prices, which requires more flexible methodological approaches to account for the nonlinear patterns in the data (Potrawa and Tetereva 2022; Lorenz et al. 2023). To this end, artificial intelligence (AI) algorithms have been widely employed recently.

Previous studies have demonstrated that sophisticated algorithms are effective in analyzing real estate data (Shim and Hwang 2018; Yao et al. 2018; Law et al. 2019; Peng et al. 2019; Park and Ryu 2021; Xu and Li 2021). These studies are grounded in scientific and rigorous methodologies; however, they overlook the ex-post analysis and interpretation of the results, lowering the credibility of model outcomes (Dzieciolowski 2019; Ali et al. 2023). Understanding how hedonic variables affect housing prices is crucial for developing urban systems (An et al. 2024) and optimizing the allocation of public resources and services (Hu et al. 2016), thereby facilitating the balanced implementation of housing policies and initiatives (Iban 2022).

Recognizing such concerns in housing price valuation, this study aims to answer the following research questions. (1) Do AI models exert overwhelming predictive power over traditional HPMs in appraising housing prices? (2) Can AI models provide a clear and accurate interpretation of each explanatory variable? (3) Does the SHapley Additive exPlanation (SHAP) method improve interpretability with the support of statistical evidence? Our findings reveal that AI models outperform conventional HPMs in terms of predictive power. Specifically, the random forest (RF) model achieves the robust capability of appraising housing prices regardless of variable combinations. Introducing the SHAP method into AI models enables two-dimensional explainable results, showing the direction and magnitude of each housing factor's influence on price appraisals. This approach helps overcome the challenges of traditional black-box models. Supported by statistical evidence from HPMs, the SHAP method's validity and applicability are reinforced, highlighting its potential as a reliable tool in property valuation.

Accordingly, the novelty and contributions of this study are threefold. First, we provide a comprehensive valuation framework for appraising housing prices, emphasizing its accuracy and transparency. Second, this study contributes to the ongoing discussion regarding the trade-offs of sophisticated AI models, which are at risk of losing their credibility in housing price assessment. Third, our findings advance the existing financial technology field[1] and housing studies by identifying the quantitative and directional impacts of hedonic variables on housing prices.

[1] Previous studies have documented that a technological combination contributes to additional value creation for customers (Tagscherer and Carbon 2023), new business models (To and Chau 2022), and appreciable societal changes (Ali et al. 2023). Li et al. (2023) highlighted that AI and other emerging technologies can promote cost efficiency, increase productivity, and foster innovative processes.

The remainder of this paper is organized as follows. Section "Literature review" reviews the previous literature. Datasets, methodologies, and accuracy metrics are described in section "Materials and methodology". Section "Results and discussion" discusses the results and presents insights into how the effects of hedonic variables on housing prices can be interpreted. Finally, the paper is concluded in section "Conclusion".

Literature review

Nonlinear relationship between hedonic variables and housing prices

HPMs provide intuitive and straightforward interpretability and estimation processes, making them popular for appraising housing prices. Conventionally, the ordinary least squares (OLS) method has been adopted to estimate model parameters. Recently, HPMs have been developed to capture the spatial aspects of real estate markets by applying spatial lag regression (SLR) and spatial autoregressive error models (Pace et al. 1998; Wen et al. 2017; Ahn et al. 2020; An et al. 2023). However, housing prices are often characterized by a nonlinear relationship with hedonic variables (Kim and Bhattacharya 2009) since numerous factors shape their values (Soltani and Lee 2024; Chen et al. 2025). This circumstance indicates that traditional linear models may fall short of precisely capturing the nonlinear interrelationship between hedonic variables and housing prices (Chin and Chau 2003; Hong et al. 2020; Li et al. 2021).

Advanced model specifications, such as generalized linear models (GLMs) (Miles 2008; Wu et al. 2014) and quantile regression approaches (Kim et al. 2015; Kallberg and Shimizu 2025), have been utilized to address this shortcoming. Furthermore, machine learning (ML) and deep learning (DL) models are widely adopted to capture housing price dynamics accurately (Shim and Hwang 2018; Yao et al. 2018; Law et al. 2019; Peng et al. 2019; Park and Ryu 2021; Xu and Li 2021). We note that existing studies have been limited to conducting comparative analyses that identify the predictive power of linear and nonlinear models. In particular, the stability and robustness of property valuation models under various hedonic variables have received little attention. Such investigations are essential for asset pricing (particularly for housing price appraisals) and for ensuring the generalizability of developed models, as a broad spectrum of housing and neighborhood factors can influence property prices.

Interpretability in housing price appraisals

Despite their promising applicability,[2] they suffer from obscurity in explaining their analytical results (Chuan et al. 2024), leading agents to necessitate additional validations in decision-making processes. Recent studies have acknowledged and attempted to address this issue. Table 1 shows that Kim (2023) identified key variables that influence model performance using the regressional ReliefF, and Luo et al. (2023) extracted primary

[2] Recently, AI-based modeling has been widely applied in various domains, such as supply chain management (Lin et al. 2022), air quality prediction (Fu et al. 2023), automobile demand forecasting (Kim 2023), and predicting the influence of innovation (Luo et al. 2023).

Table 1 Relevant literature regarding the interpretation of model results

Authors (year)	Measures	Findings	Interpretation		
			Importance	Direction	Statistical validation
Kim (2023)	Regressional ReliefF	Regressional ReliefF can identify significant input features related to prediction results	O	X	X
Luo et al. (2023)	Variable importance based on feature vectors	When utilizing the RF method, important variables for model outcomes can be identified from feature vectors	O	X	X
Hong et al. (2020)	Gini impurity	Through Gini impurity, the RF model can discern important variables to model outputs	O	X	X
Qiu et al. (2022)	Gini importance	Gini importance can be a helpful method in capturing the importance of urban landscapes concerning housing prices	O	X	X
Iban (2022)	SHAP method	The SHAP method can mitigate the black-box nature of ML algorithms used in mass appraisal systems	O	O	X
Neves et al. (2024)	SHAP method	AI-driven housing price estimations can be advanced by the SHAP method in terms of transparency	O	O	X
Soltani and Lee (2024)	SHAP method	The SHAP method can be employed to better understand how each variable affects housing prices	O	O	X
Dou et al. (2023)	SHAP method	Based on SHAP values, ML algorithms can efficiently reveal the spatial variations of the contribution of each determinant to housing prices	O	O	X

The "Importance" column indicates that the literature investigated the quantitative importance of variables on model outcomes. The "Direction" column indicates that the literature interpreted the model results in terms of directionality. The "Statistical validation" column indicates whether each interpretation result can be supported by statistical evidence in the literature

variables based on feature vectors. Furthermore, some studies employed the Gini importance measure to understand the impacts of variables on housing prices (Hong et al. 2020; Qiu et al. 2022). However, these studies focused on the quantitative importance of hedonic variables, overlooking the directional influence of each hedonic variable on the output of the model.

Table 1 also shows that the SHAP method was introduced to enhance the one-sided interpretability of valuation models. For instance, Iban (2022) employed the SHAP method to identify the local interpretability of explanatory variables, including their directional contributions to housing price appraisals. Similarly, Neves et al. (2024) demonstrated that incorporating the SHAP method improves the transparency of AI-based valuation models (Dou et al. 2023; Soltani and Lee 2024). However, these studies interpreted the results of appraised housing prices and applied the SHAP method without statistical validation. Hence, the applicability of the SHAP method in housing price assessment requires additional review to enhance its reliability and transparency.

Materials and methodology

Data source and data preprocessing

In this study, we used the hedonic dataset provided by Song et al. (2021), which includes transaction prices with housing properties, local characteristics, and the seasonality of each condominium transaction in 2015. This dataset covers the four Korean metropolitan areas of Busan, Daegu, Daejeon, and Gwangju, and Busan was selected as the study area. Among these four cities, Busan is the largest city away from Greater Seoul, and it includes many property transactions that have not been affected by other nearby areas. In addition, Busan boasts a wide variety of local amenities, including the waterfront along the seashore and a well-developed urban transit system (unlike other metropolitan areas in Korea). These characteristics provide an opportunity to explore the impacts of diverse hedonic variables on property prices.

The total number of observations was 61,152, with 17 hedonic variables. As shown in Table 2, the variables can be divided into four categories: housing properties, local amenities, local demographic characteristics, and control of the sales period. Note that some variables were transformed into logarithmic form to bring them close to normal distribution, particularly the explanatory variables (*Greenspace*, *Waterfront*, and *Subway*) and the target variable, i.e., transacted housing prices. The descriptive statistics of each hedonic variable are summarized in Table 3.

Table 4 presents descriptions of the confirmed variable set used to construct the valuation models, clarifying the associated denominations and definitions. Notably, the finalized variable combination conforms to the established housing data structure for property valuation (Ahn et al. 2020; An et al. 2023; Dai et al. 2023). Here, the “Variable” column indicates the nomenclature of each variable listed in Table 2, the “Name” column presents the name of each variable described in the manuscript, and the “Detail” column defines each variable.

Valuation models

HPM based on OLS

HPMs are based on Lancaster’s consumer theory and have served as a representative model for appraising housing prices (Lancaster 1966; Hong et al. 2020). In particular,

Table 2 Explanatory variables

Type		Variables
Housing properties	Unit related	Size of unit, level of the floor
	Complex related	Number of households, parking space per household, construction year, heating type[a]
Local amenities	Transit related	Network distance to nearest subway station[b], number of bus stops,
	Others	Network distance to nearest waterfront[b], network distance to nearest greenspace[b], network distance to CBD, number of entrants to top university
Local demographic characteristics		Population density, ratio of adults with a higher degree
Control of sales period		Spring[c], fall[c], winter[c]

[a] Dummy variable: city gas 1; others 0

[b] This variable is transformed into logarithmic form

[c] Control variables of sales period: spring, fall, and winter

Table 3 Descriptive statistics

Variables	Scale	Min	Max	Mean	Std
Higher degree	Ratio	10.16	59.10	35.40	10.77
Year	Date	1962.00	2015.00	1999.09	9.23
Parking	Ratio	0.00	15.00	1.00	0.50
Area	Ratio	14.74	250.85	79.27	30.27
Households	Ratio	4.00	5239.00	871.75	849.79
CBD	Ratio	244.28	23900.12	9003.08	4917.75
Floor	Interval	−1.00	79.00	10.56	7.90
Greenspace[a]	Ratio	−1.55	8.33	5.21	1.06
Waterfront[a]	Ratio	1.75	7.84	6.08	0.92
Subway[a]	Ratio	0.52	9.52	6.99	1.01
Top school	Ratio	0.00	12.00	1.17	1.85
Spring[b]	Nominal	0.00	1.00	0.31	0.46
Fall[b]	Nominal	0.00	1.00	0.25	0.43
Winter[b]	Nominal	0.00	1.00	0.20	0.40
Pop. density	Ratio	140.41	38720.00	12978.93	7319.40
Heating type[c]	Nominal	0.00	1.00	0.91	0.28
Bus count	Ratio	0.00	46.00	12.20	6.98

[a] This variable is transformed into logarithmic form

[b] Control variables of sales period: spring, fall, and winter

[c] Dummy variable: city gas 1; others 0

the log of real estate price relative to each hedonic variable can be considered a marginal contribution of each hedonic variable to the percentage change in real estate price (McMillan et al. 1980; Chin and Chau 2003; Hong et al. 2020). Consequently, these models are commonly used to examine the relationships between hedonic variables and real estate prices, which can be expressed as follows:

$$\ln p_i = \alpha_0 + \sum_{j=1}^{k} \beta_j x_{ij} + \sum_{l=1}^{3} \gamma_l Q_{il} + e_i, \tag{1}$$

where p_i is the transacted price of each housing i; x_{ij} represents the housing characteristics, local amenities, and local demographic characteristics; Q_{il} is the seasonal dummy variable representing spring, fall, and winter; and β_j and γ_l are the marginal contributions of each hedonic variable and seasonal dummy to the percentage change in real estate price, respectively. Finally, the error term is denoted by e_i.

SLR

SLR was used to address the spatial dependency of the hedonic variables. Note that local housing prices tend to interact with each other, and the strength of these interactions is determined by the distance between the houses. This model assumes that the sale prices of nearby units may influence the prices of other housing units, in addition to their own characteristics (Ahn et al. 2020). In particular, the SLR model contains the spatial lag term Wy in the HPM, and can be expressed as follows (Anselin 2013):

$$y = \rho W y + X\beta + \varepsilon, \tag{2}$$

Table 4 Description of explanatory variables

Variable	Name	Detail
Housing properties		
Size of unit	Area	Unit size aggregated in square meters (m^2)
Floor level	Floor	Housing's floor in a building
Number of households	Households	Number of households in an apartment complex
Parking space per household	Parking	Number of parking spaces divided by the number of households
Heating type	Heating type	The heating type of each building; 1 for city gas and 0 for others
Construction year	Year	The construction year of each apartment
Local environmental amenities		
Network distance to nearest greenspace*	Greenspace*	Network distance to the nearest park, hill, or mountain, in meters
Network distance to nearest waterfront*	Waterfront*	Network distance to the nearest river, stream, pond, or seashore, in meters
Local built environment		
Network distance to nearest subway station*	Subway*	Network distance to the nearest subway station, in meters
Network distance to Central Business District	CBD	Network distance to the city hall, in meters
Number of bus stops	Bus count	Number of bus stops within a 400-m radius of a housing unit
Number of entrants to top university	Top school	Number of Seoul National University entrants from high schools within a 5-km radius of an apartment
Local demographics		
Population density	Pop. density	Number of people per square kilometer (km^2)
Ratio of adults with a higher degree	Higher degree	Proportion of residents with a higher degree divided by the number of people aged 15 or older
Seasonality control		
Spring	Spring	Seasonal dummy indicating that a transaction occurred in March, April, or May
Fall	Fall	Seasonal dummy indicating that a transaction occurred in September, October, or November
Winter	Winter	Seasonal dummy indicating that a transaction occurred in December, January, or February

*Denotes that the variable is transformed into logarithmic form

where y is an $n \times 1$ vector of log-transformed housing prices, W is an $n \times n$ row-standardized spatial weight matrix, ρ is the spatial autoregressive parameter, X is an $n \times (k+3)$ matrix of exogenous hedonic variables and seasonal dummies, β is a $(k+3) \times 1$ vector of regression coefficients of hedonic characteristics and seasonal dummies, and the $n \times 1$ vector ε includes error terms, which are assumed to be normally distributed, homoscedastic, and independent across observations.

With the use of the spatial weight matrix, the SLR model effectively addresses the spatial autocorrelation issue. Each element of the spatial weight matrix is defined as follows:

$$W_{rs} = \begin{cases} \frac{1}{d_{rs}} & \text{for } d_{rs} < D \\ 0 & \text{otherwise} \end{cases},$$

where r and s represent the location of each housing unit by its longitude and latitude, respectively; d_{rs} is the distance between r and s; and D is a distance band.

Equation (2) can be transformed as follows to avoid bias in estimation:

$$(I - \rho W)y = X\beta + \varepsilon. \tag{3}$$

Note that the parameters in Eq. (3) can be estimated in the range of $-1 < \rho < 1$ by minimizing the root mean square error (RMSE) iteratively.

RF

The RF method is an ensemble learning model comprising a set of decision trees. When this model is employed for regression, each tree predictor—expressed as $h_s(X)$ for tree s—provides a continuous prediction value.

Here, each tree selects a random combination of input variables from the training dataset $T = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n)\}$, wherein each input variable can be selected by multiple training subsets. Each training subset (which is not a subset of other training subsets) is used for each individual tree, which is referred to as out-of-bag estimation (Breiman 2001). After numerous trees and their corresponding tree predictors are generated, the RF model aggregates all outputs to yield a robust outcome. Consequently, the RF model attempts to reduce the generalized mean-squared error (GMSE), which is defined as follows:

$$GMSE = E_{X,Y}(Y - h(X))^2, \tag{4}$$

where X is an $n \times (k + 3)$ matrix of exogenous hedonic variables and seasonal dummies, Y is an $n \times 1$ vector of log-transformed housing prices, and $h(X)$ is a tree predictor.

The overfitting issue can be avoided by using a sufficient number of trees, as per the strong law of large numbers. The RF model selects the input randomly (Fig. 1); thus, it is robust against outliers and potential noise and generates analysis results with outstanding accuracy. In addition, the out-of-bag estimation method can achieve unbiased estimation (Breiman 2001).

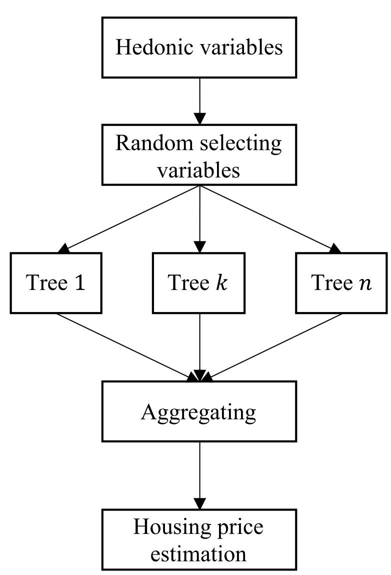

Fig. 1 Structure of the RF model

Deep neural network (DNN)

The DNN has been utilized in numerous studies for various purposes, e.g., disease detection and energy consumption prediction, due to its high prediction accuracy and ability to account for the nonlinear connection between explanatory and dependent variables (Kalogirou and Bojic 2000; Chien et al. 2020). In the architecture of a DNN, an artificial neural network is developed with several layers, including the input layer, hidden layers, and output layer, where each node in every layer is fully connected with all nodes in the next layer.

In this study, a DNN was used to calculate housing prices based on the structure shown in Fig. 2. Here, each input value—passing through hidden layers—is combined with a random initial weight. This input then moves toward the next layer through the activation function. In our model, the rectified linear unit (ReLU) activation function was used for training the model. The ReLU activation function activates only values greater than zero; thus, it deactivates neurons with zero or negative values, preventing the vanishing gradient problem in the learning phase. Finally, the prediction value is calculated in the output layer.

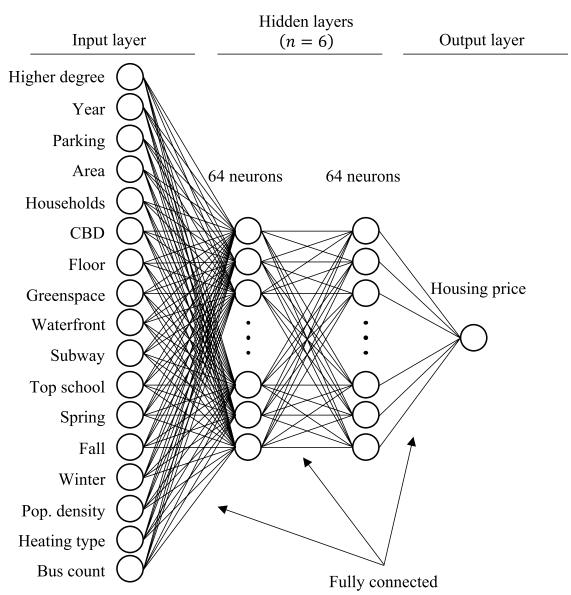

Fig. 2 Structure of the DNN model

Evaluation metrics and importance measures

Evaluation metrics

First, we employ the RMSE and R^2 metrics to assess the accuracy of each model (Taherian et al. 2013; Li et al. 2022; Wang et al. 2022). The RMSE is expressed as follows:

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{N}\left(\widehat{y}_i - y_i\right)^2},$$

where N is the number of observations, i represents each housing unit, and $\widehat{y}$ and y denote the predicted values and the log-transformed real housing prices, respectively. The RMSE value is calculated by the sum of the squared discrepancy between the predicted and observed values; thus, a lower RMSE value indicates better model accuracy.

The R^2 measure indicates the proportion of the variance of a dependent variable that is explained by a set of independent variables in the model (Taherian et al. 2013; Wang et al. 2022). The R^2 measure is calculated as follows:

$$R^2 = \frac{\sum_{i=1}^{N}\left(\widehat{y}_i - \overline{y}\right)^2}{\sum_{i=1}^{N}\left(y_i - \overline{y}\right)^2},$$

where $\overline{y}$ denotes the average of the log-transformed real housing prices. The R^2 measure is frequently considered an indicator of model accuracy, ranging between 0 and 1 (Fang et al. 2022; Wang et al. 2022), where a higher R^2 value indicates higher predictive power.

Importance measures

Permutation importance

The magnitude of multiple regression coefficients cannot guarantee the importance of each explanatory variable in explaining the dependent variable (Karpen 2017; Mizumoto 2023); thus, regression coefficients are considered only to clarify the sign of each hedonic variable's contribution to the response of the model. We rank the quantitative importance of each explanatory variable based on the resultant RMSE and R^2 values (Tran 2022; Kim et al. 2023), which mitigates the bias potentially favoring specific variables in the estimation process (Nembrini et al. 2018). Note that this ranking approach depends on the change in accuracy. In particular, permutation importance can be specified as follows (Tran 2022). When appraising housing prices, we omit a single variable; thus, 16 hedonic variables out of 17 variables are used in the price appraisal process, and this process is repeated for the other hedonic variables, resulting in 17 couples of RMSE and R^2 results. Then, we sort the RMSE (R^2) results in ascending (descending) order. This ranking, based on the RMSE and R^2, can serve as the omitted variable's quantitative importance in appraising housing prices.

Gini importance

Implementing the RF model enables us to identify the importance of each input variable via Gini impurity reduction, which is also referred to as the Gini importance (Nembrini et al. 2018). The RF model splits nodes downward to maximize impurity reduction; thus, the variables utilized for splitting where there is a large decrease in impurity can be progressively considered important (Nembrini et al. 2018). In particular, impurity is determined based on the sample variance within the node (Ishwaran 2015), and the impurity of node k is defined as follows:

$$\widehat{\Delta}(k) = \frac{1}{M}\sum\nolimits_{V_i \in k}\left(S_i - \overline{S}_k\right)^2,$$

where M is the sample size of node k, V_i denotes the input vector, S_i indicates the continuous output based on the input vector V_i, and $\overline{S}_k$ represents the average output value for node k.

Gini importance is simple and can calculate the importance of each input variable with low computational cost; thus, this approach can be employed in a wide range of fields to identify important input variables for output prediction (Tang et al. 2019; Luo et al. 2023). Note that the Gini importance is dependent on the reduction in impurity in the training procedure of tree-based models (Ishwaran 2015; Nembrini et al. 2018), and when the importance of each input variable in the RF model is quantified, the Gini importance ranking can provide robust results (Calle and Urrea 2011). Notably, considerable variability appears when employing permutation importance in an RF model; thus, this study employs the Gini importance to rank the magnitude of each variable's importance. Gini importance is a robust approach to quantify the importance of each explanatory variable regardless of data instability (Calle and Urrea 2011) when an RF model is employed. However, the Gini importance only provides the magnitude of importance. In other words, it does not provide a sign of its association with the predicted values. Thus, we introduce the SHAP method to address this issue, which can yield both the magnitude and sign of the importance of each explanatory variable in the property valuation task.

SHAP

The SHAP value indicates the effect of a given feature on the model prediction process, yielding positive or negative values (Lundberg and Lee 2017; Santos et al. 2022). The mean of the absolute SHAP values can serve as the magnitude of importance (Wang et al. 2022), and the distribution of positive or negative SHAP values reflects the directional importance of each variable (Alonso Robisco and Carbó, 2022; Tan et al. 2023). In this sense, the SHAP method enables us to identify the magnitude and direction of each hedonic variable's contribution to housing price appraisal. The mathematical formula of the SHAP value for any variable i is given as follows:

$$\phi_i(v) = \sum\nolimits_{S \subseteq N-\{i\}} \frac{|S|!(|N|-|S|-1)!}{|N|!}\{v(S \cup \{i\}) - v(S)\},$$

where $\phi_i(v)$ denotes the allocation function; $v(c)$ is the characteristic function, which calculates the contribution of the assigned coalition c; S is the subset of variables excluding variable i; and N is the set of variables in which $|S|$ and $|N|$ are the numbers of elements in S and N, respectively. The first term on the right-hand side, i.e., $\frac{|S|!(|N|-|S|-1)!}{|N|!}$, normalizes the second term, i.e., $\{v(S \cup \{i\}) - v(S)\}$; thus, the output of the allocation function $\phi_i(v)$ is dependent on the weighted average (first term) of the marginal contributions (second term) for variable i.

Experimental design

To compare the performance of the HPMs and ML approaches, we generated 17 cases by removing each hedonic variable from the full model that consists of 17 explanatory variables. Accordingly, each case included 16 explanatory variables to identify the effect of each omitted hedonic variable on the RMSE and R^2 values. The entire dataset was divided into training (70%) and test (30%) sets. Additionally, we considered min–max scaling (Asif et al. 2021) for ML models and identified the scaling effects on the RF and DNN models. Then, we compared the importance rankings based on the RMSE and R^2 values for the HPMs with the Gini importance and SHAP values for the RF model. Finally, we examined the connection between the regression coefficients and the SHAP values.

Results and discussion

Model accuracy

The accuracy of each model is summarized in Fig. 3. We found that the difference between OLS and SLR was moderate in terms of the RMSE and R^2 values. When scaling was not applied to the RF and DNN models, the RF model demonstrated superior accuracy (Asif et al. 2021). In fact, the DNN model without scaling exhibited the lowest accuracy among the four prediction models compared in this study. The DNN model performed better once min–max scaling was applied, showing a 30% increase in the R^2 value and a 50% reduction in the RMSE value. In terms of the effect of each variable on model accuracy, the *Area* variable had the greatest effect on the RF and DNN models. Note that this was observed for the DNN model only after min–max scaling was applied. In the HPM and DNN model without scaling, two variables—*Higher degree* and *Year*—appeared to be important in terms of model accuracy.

This finding corroborates previous results that indicated improvements in DNN model accuracy when min–max scaling was applied (Asif et al. 2021). Note that DNN models accompany complex procedures, e.g., setting parameters, designing hidden layers, and preprocessing data; hence, they involve high calculation durations (Jiang and Shen 2019; Asif et al. 2021). Conversely, the RF model appears more robust than the DNN models, exhibiting consistent performance in terms of the R^2 and RMSE values regardless of the scaling of the hedonic variables.[3]

Importance rank

As shown in Table 5, we measured the importance of each hedonic variable for the two HPMs considering the contribution of each variable to the R^2 and RMSE[4] values. We then calculated the importance rank of each hedonic variable in the RF model via the Gini importance and the SHAP methods. The rationale behind selecting each importance measure for the HPM and RF models is summarized in the Appendix. First, we examined the robustness of the SHAP method in terms of the magnitude of each variable's importance. Note that the DNN model did not perform well compared with the HPM and RF models without min–max scaling. In addition, the RF model exhibited robust fitness and superior accuracy regardless of scaling the hedonic variables and provided Gini importance; thus, the DNN model was excluded from the comparisons.[5]

[3] We perform the model validation with additional statistical indices and the processing time, and the results are summarized in the Appendix.

[4] We also verified the robustness of the permutation importance results, which is summarized in the Appendix.

[5] The DNN model was excluded due to its inconsistent outcomes in terms of accuracy, whereas the HPM and RF models exhibited consistent accuracy regardless of scaling the hedonic variables. The DNN model demonstrated unstable predictive power without scaling. In fact, although DNNs have wide applicability to many disciplines, they have been shown to potentially produce inconsistent outcomes because they primarily depend on data preprocessing and the model architecture (e.g., normalization, the number of layers, and the existence of a regularization kernel) (Renigier-Biłozor et al. 2019).

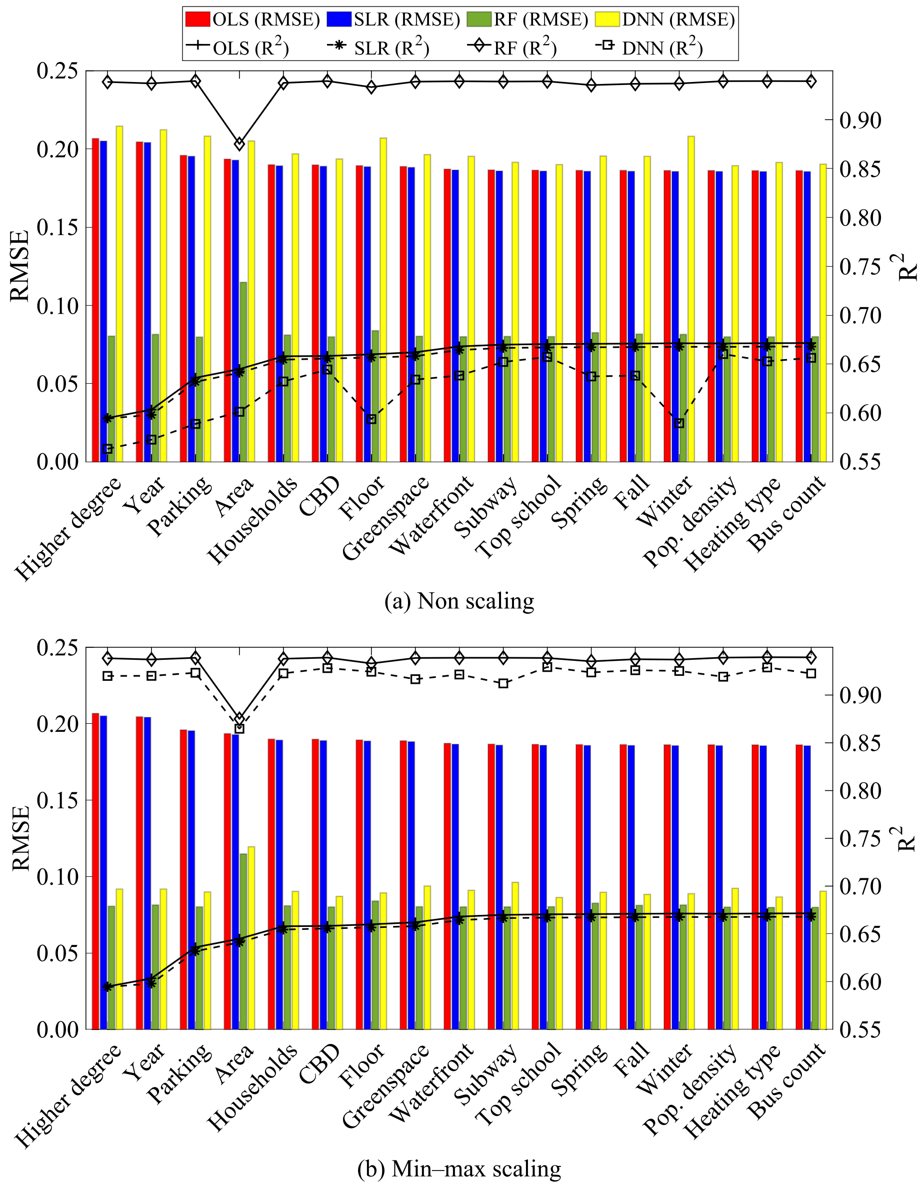

Fig. 3 Performance of the HPMs and ML models

Consequently, we only considered the RF model without min–max scaling for comparisons with the two HPM models. Here, the same input variables[6] were provided for the HPM and RF models.

By examining the top five hedonic variables with importance ranks in the HPM and RF models, we find that the same hedonic variables emerge repeatedly, including the *Higher degree*, *Year*, *Parking*, *Area*, and *Households* variables. In the top 10 ranks, the

[6] The majority of previous studies used HPMs to appraise housing prices with or without the logarithm scale of the hedonic variables (Dai et al. 2023; Fitch et al. 2023; Song et al. 2023). Considering the interpretation of each hedonic variable's behavior (i.e., the regression coefficients and the directions of variable contributions using the SHAP method), we focus on the results derived from the natural and logarithmic scaling approach (rather than normalization).

Table 5 Ranks of hedonic variables by quantitative importance measures and SHAP values

Variables	HPM				ML	
	OLS		SLR		RF	
	R^2	RMSE	R^2	RMSE	Gini	SHAP
Higher degree	1	1	1	1	3	3
Year	2	2	2	2	2	2
Parking	3	3	3	3	1	1
Area	4	4	4	4	5	5
Households	5	5	5	5	4	4
CBD	6	6	6	6	6	6
Floor	7	7	7	7	8	8
Greenspace	8	8	8	8	7	7
Waterfront	9	9	9	9	11	14
Subway	10	10	10	10	9	9
Top school	11	11	11	11	15	16
Spring	12	12	12	12	13	12
Fall	13	13	13	13	14	10
Winter	15	15	15	15	16	15
Pop. density	14	14	14	14	12	13
Heating type	16	16	16	16	17	17
Bus count	17	17	17	17	10	11

The mean of the absolute SHAP values is considered in ranking the magnitude of hedonic variable importance (Wang et al. 2022)

CBD, *Floor*, *Greenspace*, and *Subway* variables further coincide in the three models. Note that most of these variables are related to housing properties or local amenities, which indicates that the local demographics or sales season may not be as significant as other variables in the housing price prediction task. In summary, our findings demonstrate the suitability of the SHAP method for capturing the magnitude of each variable's importance.

The SHAP value and Gini importance of the RF model can directly capture the degree of the effect on the output prediction. In this study, we benchmarked the SHAP method quantitatively against other measures in terms of variable importance, and the results confirmed the substantial applicability of the SHAP method to capture the magnitude of each variable's importance.

SHAP value

The directional contribution of each hedonic variable to explaining housing prices is illustrated in Fig. 4. The higher a hedonic variable value, the higher its SHAP value, which is represented as a bright red value. In contrast, the lower a hedonic variable value, the lower its SHAP value, which is represented as a bright blue value. Note that the hedonic variables shown in Fig. 4 are ordered according to their level of contribution to the housing price appraisal, and the extent and sign of the contribution can be determined via the SHAP distribution. For example, the higher the *Parking* value is, the more the SHAP value tends to be located on the right side (bright red), and the lower the value is, the more the SHAP value tends to be located on the left side (bright blue). This finding indicates that parking spaces positively contribute to housing prices. The variable

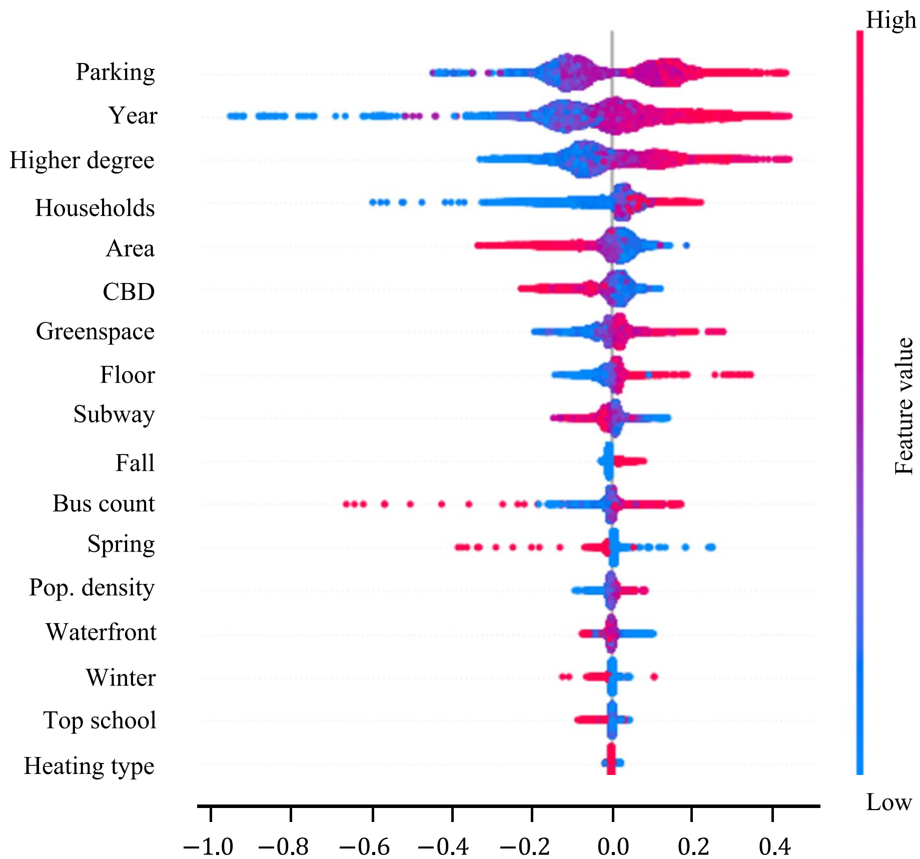

Fig. 4 SHAP value (impact on model output)

related to subway accessibility, i.e., *Subway*, demonstrates the opposite trend and coincides with the findings reported by Ahn et al. (2020). However, not all variables exhibit clear patterns in the SHAP values. For example, we cannot confidently determine the contribution of the *Heating type* variable.

The signs of the regression coefficients in the two HPMs and those of the SHAP values in the RF model are summarized and compared in Table 6.[7] In most cases, the signs of the hedonic variables' contributions in HPMs are the same as those of the SHAP values, with the exception of the *Heating type* variable, which seemingly has the lowest impact on the output prediction. These findings confirm the reliability of our results by comparing the ML model with the HPMs, thereby proving that ML can improve predictability significantly with the help of importance measures, e.g., the Gini importance and SHAP value.

[7] The variables' directional contributions cannot be identified readily by a representative numerical value because positive and negative SHAP values can potentially cancel each other. Alternatively, the directional contribution of each hedonic variable can be identified based on the SHAP values alongside red/blue points distributed in a row (Alonso Robisco and Carbó, 2022; Tan et al. 2023), as shown in Fig. 4.

Table 6 Regression coefficients and signs of the SHAP values

Variables	HPM		ML	Agreement
	OLS	SLR	RF	
	Coef	Coef	Sign	
Higher degree	$0.011^{\ddagger}$	$0.011^{\ddagger}$	+	Y
Year	$0.011^{\ddagger}$	$0.011^{\ddagger}$	+	Y
Parking	$0.168^{\ddagger}$	$0.168^{\ddagger}$	+	Y
Area	$-0.002^{\ddagger}$	$-0.002^{\ddagger}$	−	Y
Households	$5.4 \times 10^{-5\ddagger}$	$5.4 \times 10^{-5\ddagger}$	+	Y
CBD	$-1.0 \times 10^{-5\ddagger}$	$-1.0 \times 10^{-5\ddagger}$	−	Y
Floor	$0.005^{\ddagger}$	$0.005^{\ddagger}$	+	Y
Greenspace	$0.030^{\ddagger}$	$0.031^{\ddagger}$	+	Y
Waterfront	$-0.020^{\ddagger}$	$-0.020^{\ddagger}$	−	Y
Subway	$-0.016^{\ddagger}$	$-0.016^{\ddagger}$	−	Y
Top school	$-0.006^{\ddagger}$	$-0.006^{\ddagger}$	−	Y
Spring	$-0.030^{\ddagger}$	$-0.030^{\ddagger}$	−	Y
Fall	$0.021^{\ddagger}$	$0.022^{\ddagger}$	+	Y
Winter	$-0.026^{\ddagger}$	$-0.026^{\ddagger}$	−	Y
Pop. density	$2.0 \times 10^{-6\ddagger}$	$2.0 \times 10^{-6\ddagger}$	+	Y
Heating type	$-0.009^{\ddagger}$	$-0.007^{\ddagger}$		N
Bus count	$5.8 \times 10^{-4\ddagger}$	$5.7 \times 10^{-4\ddagger}$	+	Y

‡ Indicates significance at the 1% level

Discussion

The results revealed that the RF model effectively accounted for the nonlinearity inherent in hedonic datasets, exhibiting robustness and the highest predictive power among the models used for housing price valuations.[8] This superiority was attributed to the underlying characteristics of the RF model, which leverages numerous trees to avoid overfitting issues (Breiman 2001); this was validated across 17 cases via two evaluation metrics. Despite concerns over the interpretability of sophisticated models noted in previous studies (Dzieciolowski 2019; Ali et al. 2023), the RF model can serve as a transparent valuation model with the SHAP method. The SHAP method offers a two-dimensional interpretation, elucidating the direction and magnitude of hedonic variables' impacts on housing prices. The applicability of the methodological framework proposed in this study in housing price appraisals is validated and supported by the permutation importance measure, Gini importance, and HPM regression coefficients.

The findings of this study suggest that sophisticated models can be enhanced along with traditional approaches to housing price appraisals. ML technologies effectively account for the nonlinearity inherent in datasets; however, the standalone applicability of ML techniques can be limited by insufficient training samples (Vabalas et al. 2019) and/or the demand for transparent decision-making processes (Ali et al. 2023). In this context, traditional instruments complement sophisticated approaches. For example, the

[8] The results of hyperparameter tuning for the RF model are described in the Appendix. Further comparison of the accuracy, derived from RF and GLMs using the test dataset with full hedonic variables, is detailed in the Appendix.

p-values of HPMs provide statistical evidence that supports the SHAP method's reliability for interpretations, as revealed in our results. Thus, the methodological framework employed in this study serves as a precise and transparent valuation instrument, providing interpretable results regarding the magnitude and direction through the integration of these two methods.

Our results regarding the importance rank and directional contribution of hedonic variables are generally consistent with previous findings. For example, Ma et al. (2020) investigated the importance of each hedonic variable using an RF model. They identified the top 15 factors for housing prices, finding that the input variables regarding population and heating type had minimal impacts on housing prices. Similarly, Ogwang and Wang (2003) emphasized the importance of housing characteristics (e.g., property location, types of sliding doors, and the number of rooms) in housing price assessments. The directional interpretations in Fig. 4 corroborate existing empirical evidence—newly constructed housing units on higher floors with substantial parking spaces and the proximity to subway stations increase housing prices (Ahn et al 2020; Soltani et al. 2022; An et al. 2023; Liu and Strobl 2023).

Moreover, our property valuation framework offers comprehensive insights into which housing characteristics and urban resources serve as key determinants for homebuyers. For instance, *Higher degree*, *Parking*, and *Year* emerge as the top three influential hedonic variables contributing to improving the valuation models' predictive power. These findings reflect agents' economic behaviors in housing markets where homebuyers focus on urban education resources (Nguyen-Hoang and Yinger 2011; Wen et al. 2017), the construction year of buildings (Soltani et al. 2022), and the availability of parking spaces (Taylor 2020) when purchasing residential properties. Accordingly, our valuation framework (i.e., the SHAP-integrated RF model) can facilitate the understanding of overarching preference patterns shaped by housing market participants.

Conclusion

This study examined the accuracy of AI models in appraising housing prices compared with traditional approaches and proposed a complementary approach to enhance the interpretability of AI models based on empirical evidence. To address the shortcomings of conventional explainable measures, this study examined the quantitative and directional importance of hedonic variables on housing prices via the SHAP method. Comparing the accuracies of HPMs and AI models, we revealed that the proposed RF-based appraisal model exhibited strong predictive power regardless of the scaling of hedonic variables. Furthermore, the SHAP method applied to the RF model effectively illustrated the directional and quantitative effects of each hedonic variable compared with traditional explainable measures. These findings indicate that the RF model with the SHAP method offers sufficient interpretability with high accuracy and can serve as a simpler alternative to the DNN model. Moreover, our findings demonstrated that the SHAP method can enhance the reliability and transparency of housing price appraisals, and these results are supported by statistical evidence and concrete examples.

The contributions of this study can be generalized in two key ways. First, our research framework is tailored to evaluate real estate prices, particularly residential properties in the form of apartments and condominiums. The appraisal process integrates quantitative methods with aggregated datasets directly linked to housing prices, enabling precise and transparent property price appraisals. Therefore, the valuation framework, supported by relevant information sets, can be broadly adapted to other real estate subsectors, such as commercial, rental, and land development markets. Second, various experiments were conducted to strengthen the robustness of the RF-based housing price appraisal model. These experiments included sensitivity analysis (omitting one hedonic variable per iteration), the use of an external Kaggle housing dataset, and hyperparameter tuning. All the experiments demonstrated the RF model's robust performance in property price evaluation. These findings highlight RF's ability to capture complex and intricate relationships between variables, suggesting its applicability to stock pricing, disease classification, and energy consumption forecasting, where nonlinearity in variable sets is vital.

This study also presents managerial and policy implications for various stakeholders in the market. First, financial institutions (e.g., commercial banks) can utilize our RF-based and SHAP-based property valuation framework to improve the accuracy of collateral valuation for mortgage lending. In particular, SHAP-driven insights into the quantitative and directional importance of housing factors can facilitate decision-making processes and enhance valuation processes with substantial explainability. Furthermore, this transparency in valuation can contribute to the development of balanced urban planning strategies. Regulators and urban designers can leverage these insights to design policies that mitigate local disparities by identifying the unequal distribution of urban amenities (Hu et al. 2016). Furthermore, the SHAP values can guide the prioritization of urban infrastructures and amenities to enhance city environments. Given local governments' budget constraints in housing development and management, our framework offers a cost-effective tool to monitor and capture the dynamics of the real estate market.

This study integrated 17 hedonic variables commonly used in conventional price appraisal models. Future research can explore the incorporation of novel big datasets (e.g., Google street view images, satellite images, and review texts) to enhance property valuation models. Such innovative data sources can provide additional explanatory power and improve model adaptability. The current findings are based on data from a single city; however, the proposed framework can be applied to other regions and generalized to relevant fields. Our results demonstrated that integrating sophisticated technologies with traditional approaches can complement their respective strengths and mitigate their limitations in housing price valuation. Note that this study utilized cross-sectional data; thus, future studies can extend the proposed framework by incorporating the time dimension and/or by integrating the sentiment information of actual residents. This approach can contribute to the further development of more precise, transparent, and practically applicable valuation models.

Appendix

Hyperparameter tuning for the RF model

We conducted hyperparameter tuning to improve the model's predictive power, as measured by the RMSE and R^2. The RF model was used to compare the original and tuned versions because it is free from scaling. Table 7 shows the available hyperparameters

Table 7 Candidates for hyperparameter tuning

Hyperparameter	Control	Tuned value
N of estimators	Number of tree estimators	[100*, 150,200]
Max. features	Fraction of maximum number of variables considered in node splits	[0.5,0.8, 1.0*]
Max. depth	Maximum depth of tree estimator	[50,100, Unlimited*]
Max. leaf nodes	Maximum number of nodes considered in reduction of impurity	[500,750, Unlimited*]

Tuned values with an asterisk are used for our main results

(Lujan-Moreno et al. 2018), their roles, and the corresponding tuned values, resulting in a total of 81 combinations.

Figure 3 shows the results obtained by the RF model using default values (baseline). We then evaluated the top 10 optimal combinations of hyperparameters over the test dataset among the 81 combinations, excluding the baseline case. We compared these with the baseline in terms of the RMSE and R^2 values, as shown in Table 8. The results are rounded to three decimal places, revealing minimal differences between the baseline and tuned models.

After the comparison, we selected the optimal combination shown in Table 8 as the representative of the tuned models. We then compared its predictive power with that of the baseline model under the condition of omitting each hedonic variable. This procedure aligns with our experimental design, as shown in Fig. 3. Figure 5 shows that the accuracy of the tuned model (blue line and purple bars) is similar to that of the baseline model (red line and green bars) in each experiment. Furthermore, the changes in each model's accuracy with the omitted variables reveal identical patterns for both model configurations regarding RMSE and R^2.

Moreover, we investigated the interpretation results in the tuned model via the SHAP method and compared them to those of the baseline model (Fig. 4). The findings reveal that the SHAP values retrieved from the tuned model are similar to those of the baseline model. In particular, the top 10 important variables were identical in both the baseline and tuned models in terms of their quantitative and directional importance. This consistency was also confirmed across other hyperparameter sets, revealing the same conclusions.

In summary, we examined the optimal hyperparameter set to improve model accuracy based on 81 combinations by controlling four hyperparameters. Our findings demonstrated that the baseline's accuracy was compatible with tuned models. Furthermore, the SHAP-based interpretation results were identical to those of the tuned models, confirming the robustness of our findings.

Rationale behind selecting each importance measure for the HPMs and RF

We used different measures for importance ranking to evaluate the robustness of the SHAP value with respect to the magnitude of each variable's importance. While adopting Gini and SHAP values for the RF model, the rationale for ranking variables' importance based on the R^2 and RMSE values for the HPM models is summarized as follows.

Table 8 Comparison of the baseline and tuned models

N of estimators	Max. features	Max. depth	Max. leaf nodes	RMSE	R^2
Baseline					
100	1.0	Unlimited	Unlimited	0.080	0.941
Top 10 optimal combinations					
100	0.8	Unlimited	Unlimited	0.079	0.942
100	0.8	100	Unlimited	0.079	0.942
100	0.8	50	Unlimited	0.079	0.942
150	0.5	50	Unlimited	0.079	0.942
150	0.5	100	Unlimited	0.079	0.942
150	0.5	Unlimited	Unlimited	0.079	0.942
200	0.5	100	Unlimited	0.079	0.942
200	0.5	Unlimited	Unlimited	0.079	0.942
200	0.5	50	Unlimited	0.079	0.942
150	0.8	Unlimited	Unlimited	0.079	0.942

Seventeen hedonic variables are used in model comparisons

First, the Gini importance cannot be inherently applied to HPMs because this measure depends on reducing impurities in the training procedure of a tree-based model (Ishwaran 2015; Nembrini et al. 2018). In addition, the magnitudes of multiple regression coefficients do not necessarily guarantee the importance of each explanatory variable

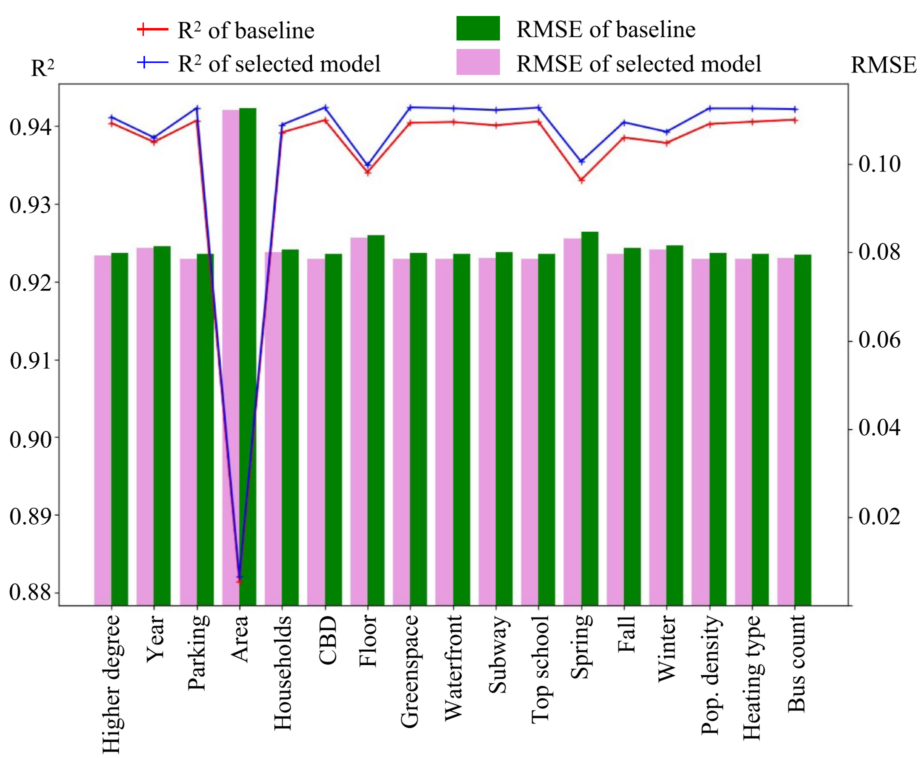

Fig. 5 Accuracy of the baseline and the selected optimal model

(Karpen 2017; Mizumoto 2023), as regression coefficients merely clarify the sign of each variable's contribution to the dependent variable. Accordingly, we quantified the importance of each hedonic variable based on the RMSE and R^2 values (Tran 2022; Kim et al. 2023), which mitigates the bias potentially favoring specific variables in the estimation process (Nembrini et al. 2018). The permutation importance measure depends on the extent of the decrease in accuracy after a single explanatory variable is removed from the full model. After applying this process progressively for each hedonic variable, we determine the extent of the accuracy decrement for each omitted variable, and the ranking results indicate which variable holds more importance in the regression model.

Second, in terms of quantifying variable importance in an RF model, Calle and Urrea (2011) demonstrated that ranking based on Gini impurity for the importance of each input variable could provide robust importance results. Notably, large variability is evident when permutation importance is employed in an RF model. Accordingly, we employed the Gini importance to rank the magnitude of variable importance, which presents a robust approach to quantify variable importance despite instability in the data (Calle and Urrea 2011).

Robustness of permutation importance

We delved deeper into the permutation importance results through the application of the shuffling method, which changes data points to assess variable importance. In this study, we compared the results of the shuffling method with those from our experimental design (omitting method) across various models, including OLS-based regression and SLR models. This evaluation focused on the RMSE metric to evaluate variable importance. As shown in Table 9, the results of permutation importance obtained from the shuffling method largely overlapped with those derived from our experimental design. In particular, both methods reveal similar ranking patterns based on the RMSE metric. Most importantly, nine identical variables consistently ranked within the top 10 ranks across all the studied cases.

Table 9 shows that the results of permutation importance obtained from the shuffling method largely overlapped with those derived from our experimental design. In particular, both methods reveal similar ranking patterns based on the RMSE metric. Most importantly, 9 variables consistently ranked within the top 10 ranks across all the cases. Furthermore, such quantitative importance patterns of hedonic variables are consistent with existing findings, suggesting that housing characteristics and neighborhood amenities are key components shaping housing prices (Ogwang and Wang 2003; Nguyen-Hoang and Yinger 2011; Wen et al. 2017; Soltani et al. 2022; Taylor 2020).

Kaggle housing dataset

To validate the greater predictive power of our flexible models over traditional approaches in the housing price appraisal task, we employed the Kaggle housing dataset (https://kaggle.com/competitions/house-prices-advanced-regression-techniques). Here, the SLR model was not employed in this procedure because this benchmark dataset does not include locational information, e.g., longitudes and latitudes, which are requisite variables to calculate the spatial weight matrix in SLR. Thus, the OLS-based model was

Table 9 Comparisons of two permutation importance measurements

Variables	Shuffling method (10 iterations)				Omitting method	
	OLS		SLR		OLS	SLR
	Rank	Average score	Rank	Average score	Rank	Rank
Higher degree	1	0.0604	1	0.0599	1	1
Year	2	0.0484	2	0.0488	2	2
Parking	3	0.0354	3	0.0354	3	3
Area	4	0.0213	4	0.0212	4	4
Households	6	0.0098	6	0.0097	5	5
CBD	5	0.0132	5	0.0126	6	6
Floor	7	0.0075	7	0.0074	7	7
Greenspace	8	0.0055	8	0.0056	8	8
Waterfront	9	0.0019	9	0.0019	9	9
Subway	11	0.0010	11	0.0009	10	10
Top school	12	0.0008	12	0.0008	11	11
Spring	10	0.0010	10	0.0010	12	12
Fall	15	0.0004	15	0.0004	13	13
Winter	14	0.0006	13	0.0006	15	15
Pop. density	13	0.0006	14	0.0005	14	14
Heating type	16	0.0002	16	0.0002	16	16
Bus count	17	0.0001	17	0.0001	17	17

The average score is the mean value of the importance score, which is obtained by repeating the measurement of permutation importance that shuffles the data points in variables over 10 iterations

used to represent the traditional approach. In addition, the RF model was used for this comparative analysis because of its robustness.

The benchmark dataset includes housing sale prices as the target variable and comprises 80 explanatory variables divided into a training set with 1,460 observations and a test set with 1,459 observations. Prior to developing the property valuation models, we conducted exploratory data analysis and applied feature engineering to refine the dataset and extract informative variables. During this process, we observed that some variables had many missing values. For example, the *PoolQC* variable (indicating the pool quality) had missing values in 99.5% of the total observations, and the *MiscFeature* variable (indicating the miscellaneous feature) was missing in approximately 1,406 of the observations (representing approximately 96.3%). In addition, variables, such as *Utilities* (type of available utilities), exhibited identical values across 1,459 observations, thereby indicating little variability with a minimal impact on housing price dynamics. Thus, we eliminated these variables from the dataset because they were deemed noise rather than meaningful contributors to appraising housing prices.

In addition, we note that there are still missing values in the preprocessed dataset. The benchmark dataset comprises both qualitative and quantitative variables; thus, the missing values were imputed by the missing-indicator method (Zhuchkova and Rotmistrov 2022) for the qualitative variables and the mean imputation (Silva-Ramírez et al. 2011; Khan and Hoque 2020) method for the quantitative variables. The missing-indicator method imputes missing values in a qualitative variable by a numerical value indicating the absence, which has been widely used in previous studies (Trevizo and Lopez 2016;

Weiss et al. 2017). For the quantitative variables, the mean imputation method is a general approach for handling missing values (Silva-Ramírez et al. 2011) with low computational costs (Khan and Hoque 2020) because it replaces the missing values with the average value for the considered variable. These processes preserve most of the observations across the entire dataset for training valuation models. Finally, the training set contained a total of 1,460 observations with 68 variables.

Considering the large number of housing variables in the refined dataset compared with our hedonic dataset, we conducted principal component analysis, which is a general feature engineering technique in the modeling process that effectively reduces the dimensionality of variable sets (Noori et al. 2010; Wang et al. 2016; Zhang et al. 2018). Here, we considered two sets of principal components determined by two points where explainable variances over 90% and 95% of the total variance were achieved. Figure 6 shows the cumulative explainable variance across varying numbers of principal components. In particular, 43 principal components explained 90.23% of the variance, and 51 principal components explained 95.31% of the variance.

Accordingly, we compared the traditional approach (i.e., the OLS regression model) against the sophisticated RF model in terms of the RMSE value. Three cases were investigated: (1) using the full variable set, (2) using the 43 principal components set, and (3) using the 51 principal components set. Table 10 shows the results obtained on the Kaggle housing dataset. The RF model consistently achieved lower RMSE values than those of the OLS-based regression model for all the cases. In our analysis, the case without feature engineering exhibited the highest accuracy in the assessment of housing prices. For each case, the RF model demonstrated an average increase in accuracy of 18% compared with the OLS-based regression model. These results are well aligned with the

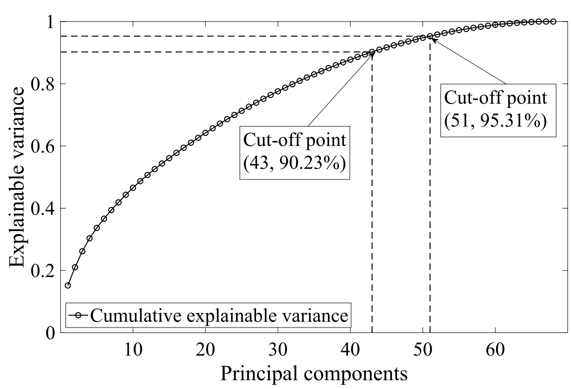

Fig. 6 Cumulative explainable variance of principal components

Table 10 RMSE results obtained on the Kaggle housing dataset

	OLS-based regression			RF		
	(1)	**(2)**	**(3)**	**(1)**	**(2)**	**(3)**
RMSE	0.1749	0.1861	0.1896	0.1451	0.1531	0.1532

The numbers in parentheses indicate each case: (1) Full variable set, (2) 43 principal components, and (3) 51 principal components baseline, demonstrating that the RF model outperforms the HPMs in terms of predictive power regardless of the experimental conditions and scaling.

Housing price appraisal using GLMs

This section investigates the predictive power of GLMs to strengthen our study's findings on housing price appraisals. The standard GLM is expressed as follows (Wood 2011; Bax and Chasomeris 2019):

$$g(u_i) = X_i \mathrm{B},$$

where $g(\cdot)$ denotes the link function. u_i is the conditional expectation of the logarithmic housing price given the row vector of hedonic variables (X_i) for a housing unit i, and B represents the column vector of the regression coefficients.

Referencing the findings of Bailey et al. (2022), we extended the standard GLM by incorporating higher-order terms. The modified GLM forms with the multiplicative and quadratic factor terms are as follows:

$$g(u_i) = \beta_0 + \sum\nolimits_{j=1}^{m} \beta_j x_{ij} + \sum\nolimits_{j=1}^{m-1} \sum\nolimits_{k=j+1}^{m} \beta_{jk} x_{ij} x_{ik}, \tag{5}$$

$$g(u_i) = \beta_0 + \sum\nolimits_{j=1}^{m} \beta_j x_{ij} + \sum\nolimits_{l=1}^{m} \beta_l x_{il}^2, \tag{6}$$

$$g(u_i) = \beta_0 + \sum\nolimits_{j=1}^{m} \beta_j x_{ij} + \sum\nolimits_{j=1}^{m-1} \sum\nolimits_{k=j+1}^{m} \beta_{jk} x_{ij} x_{ik} + \sum\nolimits_{l=1}^{m} \beta_l x_{il}^2, \tag{7}$$

where m is the number of hedonic variables, and β_0 is a constant. β_j and β_l represent the coefficients for the first- and second-order terms of hedonic variables x_i, respectively. β_{jk} represents the coefficient for the interaction terms. We used a normal distribution for the link function of GLMs (James 2002; Srivastava et al. 2013; Bailey et al. 2022) and applied stepwise model selection for optimal fit (Bax and Chasomeris 2019; Bailey et al. 2022). These GLM variants are referred to as GLM1, GLM2, and GLM3, corresponding to Eqs. (5)–(7), respectively.

Table 11 presents the R^2 and RMSE values for the GLMs, OLS-based HPMs, SLR, RF, and DNN when the test dataset with all the hedonic variables is used. The results confirm that GLMs exhibit greater predictive power than those of OLS-based HPMs and

Table 11 Accuracy comparisons between GLMs and other models

	R^2	RMSE
RF	0.9390	0.0813
DNN_s	0.8516	0.1267
GLM1	0.7230	0.1731
GLM2	0.6863	0.1842
GLM3	0.6861	0.1843
OLS	0.6528	0.1938
SLR	0.6510	0.1935
DNN	0.4657	0.2404

The row "DNN_s" indicates the case of the DNN model with min–max scaling

Table 12 Model validation with additional accuracy measures and processing times

	RMSE	R^2	MAE	MAPE	Processing time (sec.)
OLS	0.194	0.653	0.143	0.010	–
SLR	0.194	0.651	0.143	0.010	–
RF	0.081	0.939	0.053	0.004	14.340
DNN	0.186	0.681	0.137	0.009	112.669
DNN_s	0.127	0.850	0.089	0.006	116.708

A lower value of the reported metric indicates that the model has greater predictive power, excluding R^2. The "Processing time (sec.)" column shows the convergence time of the model training process, measured in seconds. The row "DNN_s" indicates the case of the DNN model with min–max scaling

SLRs in terms of both R^2 and RMSE values because of their ability to mitigate biased estimates (López-Laborda et al. 2021). These results are consistent with those of previous studies (Han and Jang 2013; López-Laborda et al. 2021). However, the RF model consistently demonstrates the highest predictive power in all the cases, emphasizing that RF-based valuation models can be considered a desirable approach for precisely appraising housing prices.

Additional accuracy metrics and processing time

We use the additional statistical indices, mean absolute error (MAE) and mean absolute percentage error (MAPE), to further assess the predictive power of the models. These indices are commonly employed metrics in model validation (Ho et al. 2021). We train each model using the complete set of hedonic variables to ensure comparability. We also examine the processing time of the ML and DL models, as convergence speed is a key factor in evaluating the computational efficiency of AI-based systems (Yates and Islam 2021; Abed et al. 2022).

Table 12 presents the predictive power of the HPMs and AI models. Our findings demonstrate that the RF model consistently outperforms the other models across all the evaluation metrics, including the RMSE, R^2, MAE, and MAPE. Furthermore, the RF model requires approximately 13% of the processing time of DNN models, underscoring its computational efficiency. These findings confirm the superiority of the RF model in terms of precision and computational efficiency.

According to Yates and Islam (2021), the computational complexity of the RF model primarily depends on the number of trees. Correspondingly, a smaller ensemble (i.e., a reduced number of tree estimators) consumes fewer computational resources (Yates and Islam 2021). Nevertheless, the fine-tuning experiment for RF in Table 8 shows that the RF configuration with fewer trees can achieve predictive power comparable to those with a larger number of trees. As such, our resultant measures in Table 12 show that the RF model can be a more efficient tool for housing price valuation with fewer trees while maintaining comparable predictive accuracy.

Abbreviations

AI	Artificial intelligence
DL	Deep learning
DNN	Deep neural network
GLM	Generalized linear model

GMSE Generalized mean-squared error
HPM Hedonic price model
MAE Mean absolute error
MAPE Mean absolute percentage error
ML Machine learning
OLS Ordinary least squares
ReLU Rectified linear unit
RF Random forest
RMSE Root mean square error
SHAP SHapley Additive exPlanation
SLR Spatial lag regression

Acknowledgements
The authors would like to thank the editors and the anonymous reviewers for their invaluable comments and suggestions.

Author contributions
All the authors contributed equally.

Funding
This work was supported by (i) the National Research Foundation of Korea grant funded by the Korea government (MSIT) (RS-2025-16067531: Kwangwon Ahn) and (ii) Hankuk University of Foreign Studies Research Fund (0f 2025: Sihyun An).

Availability of data and materials
The datasets used and/or analyzed during the current study are available from the corresponding author on reasonable request.

Declarations

Competing interests
The authors declare that they have no competing interests.

Received: 7 July 2023 Accepted: 6 November 2025
Published online: 31 December 2025

References

Abed M, Imteaz MA, Ahmed AN, Huang YF (2022) Modelling monthly pan evaporation utilising random forest and deep learning algorithms. Sci Rep 12(1):13132
Ahn K, Jang H, Song Y (2020) Economic impacts of being close to subway networks: a case study of Korean metropolitan areas. Res Transp Econ 83:100900
Ali O, Abdelbaki W, Shrestha A, Elbasi E, Alryalat MA, Dwivedi YK (2023) A systematic literature review of artificial intelligence in the healthcare sector: benefits, challenges, methodologies, and functionalities. J Innov Knowl 8(1):100333
Alonso Robisco A, Carbó JM (2022) Measuring the model risk-adjusted performance of machine learning algorithms in credit default prediction. Financ Innov 8(1):70
An S, Ahn K, Bae J, Song Y (2024) Economic impacts of a subway system: exploring local contexts in a metropolitan area. Res Transp Bus Manag 56:101188
An S, Jang H, Kim H, Song Y, Ahn K (2023) Assessment of street-level greenness and its association with housing prices in a metropolitan area. Sci Rep 13(1):22577
Anderson ST, West SE (2006) Open space, residential property values, and spatial context. Reg Sci Urban Econ 36(6):773–789
Anselin L (2013) Spatial econometrics: methods and models, vol 4. Springer, Heidelberg
Asif MR, Bording TS, Barfod AS, Grombacher DJ, Maurya PK, Christiansen AV, Auken E, Larsen JJ (2021) Effect of data preprocessing on the performance of neural networks for 1-D transient electromagnetic forward modeling. IEEE Access 9:34635–34646
Bailey JR, Lauria D, Lindquist WB, Mittnik S, Rachev ST (2022) Hedonic models of real estate prices: GAM models; environmental and sex-offender-proximity factors. J Risk Financ Manag 15(12):601
Bax D, Chasomeris MG (2019) Listing price estimation of apartments: a generalised linear model. J Econ Financ Sci 12(1):1–11
Breiman L (2001) Random forests. Mach Learn 45(1):5–32
Calle ML, Urrea V (2011) Letter to the editor: stability of random forest importance measures. Brief Bioinform 12(1):86–89
Chen Y, Ye Y, Liu X, Yin C, Jones CA (2025) Examining the nonlinear and spatial heterogeneity of housing prices in urban Beijing: an application of GeoShapley. Habitat Int 162:103439
Chien CY, Hsu SW, Lee TL, Sung PS, Lin CC (2020) Using artificial neural network to discriminate Parkinson's disease from other parkinsonisms by focusing on putamen of dopamine transporter SPECT images. Biomedicines 9(1):12
Chin T, Chau K (2003) A critical review of literature on the hedonic price model. Int J Hous Sci Appl 27(2):145–165
Chuan CH, Sun R, Tian S, Tsai WHS (2024) EXplainable artificial intelligence (XAI) for facilitating recognition of algorithmic bias: an experiment from imposed users' perspectives. Telemat Inform 91:102135
Czembrowski P, Kronenberg J (2016) Hedonic pricing and different urban green space types and sizes: insights into the discussion on valuing ecosystem services. Landsc Urban Plan 146:11–19

Dai X, Felsenstein D, Grinberger AY (2023) Viewshed effects and house prices: identifying the visibility value of the natural landscape. Landsc Urban Plan 238:104818
Dou M, Gu Y, Fan H (2023) Incorporating neighborhoods with explainable artificial intelligence for modeling fine-scale housing prices. Appl Geogr 158:103032
Dzieciolowski K (2019) Explaining black-box machine learning models. In: 13th multi conference on computer science and information systems, Porto, Portugal, 16–19 July 2019
Fang F, Ventre C, Basios M, Kanthan L, Martinez-Rego D, Wu F, Li L (2022) Cryptocurrency trading: a comprehensive survey. Financ Innov 8(1):1–59
Fitch RA, Mueller JM, Meldrum J, Huber C (2023) Estimating proximity effects to wildfire fuels treatments on house prices in Cibola National Forest, New Mexico, USA. Landsc Urban Plan 238:104838
Fu L, Li J, Chen Y (2023) An innovative decision making method for air quality monitoring based on big data-assisted artificial intelligence technique. J Innov Knowl 8(2):100294
Glaeser EL, Gyourko J, Morales E, Nathanson CG (2014) Housing dynamics: an urban approach. J Urban Econ 81:45–56
Han C, Jang Y (2013) Effects of debt collection practices on loss given default. J Bank Financ 37(1):21–31
Ho WK, Tang BS, Wong SW (2021) Predicting property prices with machine learning algorithms. J Prop Res 38(1):48–70
Hong J, Choi H, Kim WS (2020) A house price valuation based on the random forest approach: the mass appraisal of residential property in South Korea. Int J Strateg Prop Manag 24(3):140–152
Hu N, Legara EF, Lee KK, Hung GG, Monterola C (2016) Impacts of land use and amenities on public transport use, urban planning and design. Land Use Pol 57:356–367
Iban MC (2022) An explainable model for the mass appraisal of residences: the application of tree-based machine learning algorithms and interpretation of value determinants. Habitat Int 128:102660
Ishwaran H (2015) The effect of splitting on random forests. Mach Learn 99(1):75–118
James GM (2002) Generalized linear models with functional predictors. J R Stat Soc B 64(3):411–432
Jiang Z, Shen G (2019) Prediction of house price based on the back propagation neural network in the keras deep learning framework. In: 6th international conference on systems and informatics, Shanghai, China, 2–4 November 2019
Kallberg JG, Shimizu Y (2025) Crime measures and housing prices: an analysis using quantile regression and spatial autocorrelation. J Real Estate Financ Econ 1–53
Kalogirou S, Bojic M (2000) Artificial neural networks for the prediction of the energy consumption of a passive solar building. Energy 25(5):479–491
Karpen SC (2017) Misuses of regression and ANCOVA in educational research. Am J Pharm Educ 81(8):6501
Khan SI, Hoque ASML (2020) SICE: an improved missing data imputation technique. J Big Data 7(1):37
Kim HG, Hung KC, Park SY (2015) Determinants of housing prices in Hong Kong: a Box-Cox quantile regression approach. J Real Estate Financ Econ 50(2):270–287
Kim JY, Kim D, Li ZJ, Dariva C, Cao Y, Ellis N (2023) Predicting and optimizing syngas production from fluidized bed biomass gasifiers: a machine learning approach. Energy 263:125900
Kim S (2023) Innovating knowledge and information for a firm-level automobile demand forecast system: a machine learning perspective. J Innov Knowl 8(2):100355
Kim SW, Bhattacharya R (2009) Regional housing prices in the USA: an empirical investigation of nonlinearity. J Real Estate Financ Econ 38(4):443–460
Lancaster KJ (1966) A new approach to consumer theory. J Polit Econ 74(2):132–157
Law S, Paige B, Russell C (2019) Take a look around: using street view and satellite images to estimate house prices. ACM Trans Intell Syst Technol 10(5):1–19
Li P, Bastone A, Mohamad TA, Schiavone F (2023) How does artificial intelligence impact human resources performance. Evidence from a healthcare institution in the United Arab Emirates. J Innov Knowl 8(2):100340
Li S, Jiang Y, Ke S, Nie K, Wu C (2021) Understanding the effects of influential factors on housing prices by combining extreme gradient boosting and a hedonic price model (XGBoost-HPM). Land 10(5):533
Li Y, Jiang S, Li X, Wang S (2022) Hybrid data decomposition-based deep learning for Bitcoin prediction and algorithm trading. Financ Innov 8(1):1–24
Lin H, Lin J, Wang F (2022) An innovative machine learning model for supply chain management. J Innov Knowl 7(4):100276
Liu N, Strobl J (2023) Impact of neighborhood features on housing resale prices in Zhuhai (China) based on an (M) GWR model. Big Earth Data 7(1):146–169
López-Laborda J, Marín-González C, Onrubia-Fernández J (2021) Estimating Engel curves: a new way to improve the SILC-HBS matching process using GLM methods. J Appl Stat 48(16):3233–3250
Lorenz F, Willwersch J, Cajias M, Fuerst F (2023) Interpretable machine learning for real estate market analysis. Real Estate Econ 51(5):1178–1208
Lujan-Moreno GA, Howard PR, Rojas OG, Montgomery DC (2018) Design of experiments and response surface methodology to tune machine learning hyperparameters, with a random forest case-study. Expert Syst Appl 109:195–205
Lundberg S, Lee S (2017) A unified approach to interpreting model predictions. In: Adv Neural Info Proc Syst 4765–4774
Luo J, Wang Y, Li G (2023) The innovation effect of administrative hierarchy on intercity connection: the machine learning of twin cities. J Innov Knowl 8(1):100293
Ma J, Cheng JCP, Jiang F, Chen W, Zhang J (2020) Analyzing driving factors of land values in urban scale based on big data and non-linear machine learning techniques. Land Use Policy 94:104537
McMillan ML, Reid BG, Gillen DW (1980) An extension of the hedonic approach for estimating the value of quiet. Land Econ 56(3):315–328
Miles W (2008) Boom–bust cycles and the forecasting performance of linear and non-linear models of house prices. J Real Estate Financ Econ 36:249–264
Mizumoto A (2023) Calculating the relative importance of multiple regression predictor variables using dominance analysis and random forests. Lang Learn 73(1):161–196
Nembrini S, König IR, Wright MN (2018) The revival of the Gini importance? Bioinformatics 34(21):3711–3718

Neves TF, Aparicio M, de Castro Neto M (2024) The impacts of open data and explainable AI on real estate price predictions in smart cities. Appl Sci 14(5):2209

Nguyen-Hoang P, Yinger J (2011) The capitalization of school quality into house values: a review. J Hous Econ 20(1):30–48

Noori R, Khakpour A, Omidvar B, Farokhnia A (2010) Comparison of ANN and principal component analysis-multivariate linear regression models for predicting the river flow based on developed discrepancy ratio statistic. Expert Syst Appl 37(8):5856–5862

Ogwang T, Wang B (2003) A hedonic price function for a northern BC community. Soc Indic Res 61(3):285–296

Pace RK, Barry R, Sirmans CF (1998) Spatial statistics and real estate. J Real Estate Fin Econ 17(1):5–13

Park D, Ryu D (2021) A machine learning–based early warning system for the housing and stock markets. IEEE Access 9:85566–85572

Peng Z, Huang Q, Han Y (2019) Model research on forecast of second-hand house price in Chengdu based on xgboost algorithm. In: 11th international conference on advanced infocomm technology, Jinan, China, 18–20 October 2019

Potrawa T, Tetereva A (2022) How much is the view from the window worth? Machine learning-driven hedonic pricing model of the real estate market. J Bus Res 144:50–65

Qiu W, Zhang Z, Liu X, Li W, Li X, Xu X, Huang X (2022) Subjective or objective measures of street environment, which are more effective in explaining housing prices? Landsc Urban Plan 221:104358

Renigier-Biłozor M, Janowski A, d'Amato M (2019) Automated valuation model based on fuzzy and rough set theory for real estate market with insufficient source data. Land Use Policy 87:104021

Santos OL, Dotta D, Wang M, Chow JH, Decker IC (2022) Performance analysis of a DNN classifier for power system events using an interpretability method. Int J Electron 136:107594

Shim J, Hwang C (2018) Kernel–based geographically and temporally weighted autoregressive model for house price estimation. PLoS ONE 13(10):e0205063

Silva-Ramírez EL, Pino-Mejías R, López-Coello M, Cubiles-de-la-Vega MD (2011) Missing value imputation on missing completely at random data using multilayer perceptrons. Neural Netw 24(1):121–129

Soltani A, Heydari M, Aghaei F, Pettit CJ (2022) Housing price prediction incorporating spatio-temporal dependency into machine learning algorithms. Cities 131:103941

Soltani A, Lee CL (2024) The non-linear dynamics of South Australian regional housing markets: a machine learning approach. Appl Geogr 166:103248

Song J, Nie R, Yuan H, Gao J (2023) Does the renovation of waste-to-energy incineration plants attenuate the housing price gradient? Evidence from a quasi-natural experiment in Shenzhen, China. Energy Econ 120:106633. https://doi.org/10.1016/j.eneco.2023.106633

Song Y, Ahn K, An S, Jang H (2021) Hedonic dataset of the metropolitan housing market—cases in South Korea. Data Brief 35:106877

Srivastava PK, Han D, Ramirez MR, Islam T (2013) Machine learning techniques for downscaling SMOS satellite soil moisture using MODIS land surface temperature for hydrological application. Water Resour Manag 27:3127–3144

Tagscherer F, Carbon CC (2023) Leadership for successful digitalization: a literature review on companies' internal and external aspects of digitalization. Sustain Technol Entrep 2(2):100039

Taherian R, Hadianfard MJ, Golikand AN (2013) A new equation for predicting electrical conductivity of carbon-filled polymer composites used for bipolar plates of fuel cells. J Appl Polym Sci 128(3):1497–1509

Tan B, Gan Z, Wu Y (2023) The measurement and early warning of daily financial stability index based on XGBoost and SHAP: evidence from China. Expert Syst Appl 227:120375

Tang L, Cai F, Ouyang Y (2019) Applying a nonparametric random forest algorithm to assess the credit risk of the energy industry in China. Technol Forecast Soc Change 144:563–572

Taylor EJ (2020) Parking policy: the politics and uneven use of residential parking space in Melbourne. Land Use Policy 91:103706

Tchuente D, Nyawa S (2022) Real estate price estimation in French cities using geocoding and machine learning. Ann Oper Res 308(1–2):571–608

To CKM, Chau KP (2022) Characterizing sustainability materiality: ESG materiality determination in technology venturing. Sustain Technol Entrep 1(3):100024

Tran VQ (2022) Hybrid gradient boosting with meta-heuristic algorithms prediction of unconfined compressive strength of stabilized soil based on initial soil properties, mix design and effective compaction. J Clean Prod 355:13168

Trevizo D, Lopez MJ (2016) Neighborhood segregation and business outcomes: Mexican immigrant entrepreneurs in Los Angeles County. Sociol Perspect 59(3):668–693

Vabalas A, Gowen E, Poliakoff E, Casson AJ (2019) Machine learning algorithm validation with a limited sample size. PLoS ONE 14(11):e0224365

Wang D, Thunéll S, Lindberg U, Jiang L, Trygg J, Tysklind M (2022) Towards better process management in wastewater treatment plants: process analytics based on SHAP values for tree-based machine learning methods. J Environ Manag 301:113941

Wang GY (2023) The effect of environment on housing prices: evidence from the google street view. J Forecast 42(2):288–311

Wang Y, Ma X, Joyce MJ (2016) Reducing sensor complexity for monitoring wind turbine performance using principal component analysis. Renew Energy 97:444–456

Weiss MJ, Bloom HS, Verbitsky-Savitz N, Gupta H, Vigil AE, Cullinan DN (2017) How much do the effects of education and training programs vary across sites? Evidence from past multisite randomized trials. J Res Educ Eff 10(4):843–876

Wen H, Xiao Y, Zhang L (2017) School district, education quality, and housing price: evidence from a natural experiment in Hangzhou, China. Cities 66:72–80

Wood SN (2011) Fast stable restricted maximum likelihood and marginal likelihood estimation of semiparametric generalized linear models. J R Stat Soc Ser B Stat Methodol 73(1):3–36

Wu B, Li R, Huang B (2014) A geographically and temporally weighted autoregressive model with application to housing prices. Int J Geogr Inf Sci 28(5):1186–1204
Xu L, Li Z (2021) A new appraisal model of second-hand housing prices in China's first-tier cities based on machine learning algorithms. Comput Econ 57(2):617–637
Yao Y, Zhang J, Hong Y, Liang H, He J (2018) Mapping fine-scale urban housing prices by fusing remotely sensed imagery and social media data. Trans GIS 22(2):561–581
Yates D, Islam MZ (2021) FastForest: increasing random forest processing speed while maintaining accuracy. Inf Sci 557:130–152
Zhang C, Cao L, Romagnoli A (2018) On the feature engineering of building energy data mining. Sustain Cities Soc 39:508–518
Zhuchkova S, Rotmistrov A (2022) How to choose an approach to handling missing categorical data: (un)expected findings from a simulated statistical experiment. Qual Quant 56(1):1–22

Publisher's Note