ISSN: 1648-715X / eISSN: 1648-9179

VILNIUS TECH
Vilnius Gediminas Technical University

INTERNATIONAL JOURNAL OF STRATEGIC PROPERTY MANAGEMENT

2025
Volume 29
Issue 5
Pages 350–376
https://doi.org/10.3846/ijspm.2025.25137

EXPLAINABLE AI-BASED MASS APPRAISAL: INSIGHTS FROM MACHINE LEARNING APPLICATIONS IN KOREA'S RESIDENTIAL PROPERTY MARKET

Woo-sung KIM [1], Insu CHOI [2], Myounggu LEE [1*]

[1] *Department of Business Administration, Konkuk University, Seoul, South Korea*
[2] *Division of Business Administration, Dongduk Women's University, Seoul, South Korea*

Article History:
- received 2 February 2025
- accepted 28 August 2025

Abstract. Mass appraisal plays a pivotal role in real estate management, facilitating property tax assessments, mortgage evaluations, and urban planning across large geographical areas. In regions like Korea, where real estate markets are rapidly evolving, valuation models based on multiple linear regression are valued for their simplicity and interpretability but often fall short in capturing complex market dynamics. In contrast, machine learning (ML) models, while addressing non-linear relationships between property characteristics and market values and offering superior predictive performance, are often criticized for their "black-box" nature, which raises concerns over interpretability in transparency-critical domains like property tax assessments and policy planning. To address these concerns, this study investigates the application of Explainable AI (XAI) techniques in the mass appraisal of residential properties in Korea, integrating XAI methods with both multiple linear regression and random forest models. Using SHAP (SHapley Additive exPlanations) and PFI (Permutation Feature Importance) values, the study analyzes feature importance and predictive contributions, offering insights into the factors driving property valuations. Additionally, a temporal analysis was conducted by segmenting the data into time intervals to examine how feature importance and predictive contributions evolve over time. By combining high predictive performance with transparent and interpretable insights, the findings demonstrate that XAI can enhance the usability of both traditional and advanced automated valuation models (AVMs) for real-world decision-making in the Korean real estate sector.

Keywords: automatic valuation model, mass appraisal, explainable AI (XAI), SHAP (SHapley Additive exPlanations), Permutation Feature Importance (PFI), temporal analysis.

* Corresponding author. E-mail: *myounggulee@konkuk.ac.kr*

1. Introduction

Mass appraisal, as defined by the International Association of Assessing Officers (IAAO), refers to "the process of valuing a group of properties as of a given date using common data, standardized methods, and statistical testing". This method is extensively employed by local governments for property taxation, as it allows for the efficient and timely valuation of a large number of properties (IAAO, 2017). Given that property tax assessments must be conducted regularly—particularly in markets with rapid changes in economic conditions—mass appraisal provides a pragmatic solution for achieving both cost-effectiveness and timeliness in valuation processes. International regulatory frameworks, such as the Basel II Accord established by the Basel Committee on Banking Supervision (BCBS) in 2008, further underscore the importance of regular property valuation. These regulations require financial institutions to monitor the value of collateral frequently, mandating at least one valuation annually, with more frequent assessments recommended in highly volatile markets (Hong et al., 2020). The growing emphasis on accurate and timely property valuations reflects their critical role in maintaining financial stability. As a result, the demand for robust and scalable mass appraisal models has significantly increased.

Nevertheless, valuing real estate presents distinctive challenges due to the heterogeneous nature of properties. Unlike homogeneous commodities, residential properties are characterized by immobility, durability, and variability in both physical features and market conditions (McCluskey et al., 2000). Furthermore, a multitude of factors—including market segmentation and government intervention—can influence property prices, adding complexity to the valuation process. The fundamental objective of valuation models is to produce a reliable, credible, and cost-effective estimate of a property's market value at a given point in time. Achieving this goal is essential not only for property tax assessments but also for broader applications such as portfolio risk management, insurance valuations, and urban planning (Glumac & Des Rosiers, 2021).

To address the challenges inherent in valuing heterogeneous real estate properties, traditional valuation methods often fall short, necessitating the use of more advanced approaches such as automated valuation models (AVMs). AVMs leverage advanced statistical techniques and machine learning algorithms to analyze large datasets, offering enhanced accuracy and consistency over traditional methods (Kok et al., 2017). Recent research emphasizes the growing importance of AVMs in the real estate sector, primarily due to their advantages in terms of speed, scalability, and cost-efficiency (Wang & Li, 2019). These models not only meet increasing regulatory demands for frequent and precise property valuations but also provide a robust framework for managing the intricacies of modern real estate markets. A survey by the International Association of Assessing Officers (IAAO) highlights that AVMs offer notable advantages in property assessment, including improved accuracy, consistency, fairness, cost reductions, and increased operational efficiency (Bidanset & Rakow, 2022). Countries such as Australia, Sweden, Northern Ireland, New Zealand, Singapore, Malaysia, and the United States have successfully adopted computer-assisted mass appraisal (CAMA) systems that rely on automated appraisal models (Dimopoulos & Moulas, 2016).

As AVMs continue to advance, machine learning techniques are being increasingly integrated, offering significant improvements over traditional models. Some machine learning methods are particularly effective at capturing the non-linear relationships between property characteristics and market values, which extend beyond the linear assumptions of linear regression-based models. According to Kok et al. (2017), machine learning-based automated valuation models are being used in cases such as House Canary's automated valuation model and Zillow's much-discussed "Zestimate". While these machine learning models have demonstrated superior predictive accuracy, they also introduce significant challenges in terms of interpretability—an essential factor in fields such as real estate, where transparency and accountability are paramount. As Wang and Li (2019) note, "because the target of mass appraisal is a large number of properties, and the valuation results need to be explained to the public, the basic needs are convenient operation and simple understanding". This underscores the critical importance of interpretability in mass appraisal, where clear explanations of valuation outcomes are necessary to maintain public trust and ensure regulatory compliance. Furthermore, the interpretability of real estate price determinants can serve as a key differentiating factor for real estate platform services, such as personalized property recommendation systems and targeted advertising strategies.

To address these concerns, the development of Explainable AI (XAI) techniques has emerged as a promising solution. XAI seeks to bridge the gap between the high predictive power of machine learning models and the need for transparency by clarifying how these models generate their predictions. In real estate, where valuation decisions directly affect stakeholders such as property owners, financial institutions, and regulatory bodies, the ability to provide understandable and accessible justifications for property values is crucial. Several studies have begun utilizing XAI techniques to analyze the non-linear effects of real estate characteristics and economic variables on property prices. These approaches have provided valuable insights into how factors such as location, property size, and macroeconomic conditions contribute to price fluctuations in ways that traditional models often fail to capture. However, existing studies often focus on individual machine learning models without comparing the explanatory capabilities of linear and non-linear models. This is a significant limitation because the SHAP and PFI values, which represent the importance and contribution of each feature, can vary depending on the algorithm used. Linear models, such as multiple linear regression, assume a linear relationship between independent variables and the dependent variable, which simplifies the interpretation of feature importance. However, random forests, based on decision tree algorithms, partition the feature space into multiple regions through recursive splitting. This fundamental difference in algorithmic principles means that even when using the same features, the importance and contribution values derived from SHAP or PFI can vary significantly between the two models. Comparing these variations across algorithms is crucial for understanding the trade-offs between model interpretability and predictive accuracy as well as for providing robust and consistent insights into the factors influencing property valuations (cf. While traditional linear regression models offer greater interpretability owing to their inherent linear structure, they frequently fall short in predictive performance. On the other hand, Random Forest models tend to outperform to the linear regression model in terms of accuracy yet provide limited interpretability without the aid of supplementary methods). Moreover, existing research seldom incorporates temporal analyses to evaluate how feature importance and predictive contributions evolve over time. This is particularly important in the context of real estate, where property values are influenced by dynamic factors such as shifting consumer preferences, economic conditions, and urban development trends. For instance, the demand for certain property characteristics, such as proximity to public transportation or availability of green spaces, may change over time due to societal or economic shifts. By analyzing these temporal changes, we can better understand how the drivers of property valuation adapt to evolving market conditions, providing more accurate and context-sensitive insights for stakeholders.

In this paper, we evaluate and compare the interpretability and predictive capabilities of linear and non-linear models using XAI techniques. For the linear model, we employed multiple linear regression (MLR), which assumes a direct and straightforward relationship between the independent variables and the target variable. For the non-linear model, we utilized random forest (RF), a tree-based ensemble method known for its ability to capture complex interactions among features. These two models were selected as they are widely recognized and commonly applied in various mass appraisal systems due to their effectiveness and adaptability in property valuation tasks. To evaluate the importance and contribution of features across these models, we employed SHAP (SHapley Additive exPlanations) and PFI (Permutation Feature Importance) methods. While previous studies often emphasize the superior predictive performance of some machine learning models, such as random forests, in capturing the complex and non-linear characteristics of the housing market, this study extends the discussion by applying XAI to identify the specific features and interactions where these effects play a critical role. Through this approach, the research provides a deeper understanding of how the models address these complexities, which are often challenging to uncover using traditional linear methods. Furthermore, this study incorporates a temporal analysis by segmenting the data into discrete time intervals to investigate the dynamic evolution of feature importance over time. Changes in demographic structures and shifts in consumer preferences play a critical role in shaping housing selection behaviours, and these changes are subsequently reflected in property values. By examining these temporal trends, the study provides useful insights into how societal and economic transformation influence property valuations, highlighting the dynamic and adaptive nature of real estate markets. We expect that these findings will contribute to a more comprehensive understanding of the balance between model interpretability and predictive performance, while offering enhanced transparency into the complex interactions that shape housing values.

The rest of the paper is organized as follows. Section 2 provides a review of the relevant literature, focusing on prior research related to mass appraisal models and associated methodologies. In Section 3, we outline the techniques and data analysis processes employed in this study, including an introduction to the individual machine learning algorithms utilized as well as the XAI. Section 4 describes the dataset used in the analysis, along with key summary statistics. The results of the analysis are presented in Section 5. Finally, the conclusions and implications of the study are discussed in the concluding section.

2. Literature review

In this section, we present a review of the literature relevant to this study, focusing on hedonic models and machine learning techniques commonly employed in automated valuation models (AVMs) for mass appraisal. While traditional methods such as the comparable method, investment method, profit method, and residual method are well-established for real estate market value estimation through extensive research, they present challenges for large-scale property assessments due to significant labor demands and methodological limitations (Pagourtzi et al., 2003). For a comprehensive review of these valuation methods, refer to studies by Pagourtzi et al. (2003), Gabrielli and French (2021), and Binoy et al. (2022).

Over the years, the hedonic pricing model, which is primarily based on multiple linear regression, has emerged as one of the most widely employed approach for estimating real estate prices in the context of mass appraisal. Rooted in Lancaster's consumer theory (Lancaster, 1966) and further developed by Rosen (1974), the hedonic model conceptualizes real estate as a heterogeneous good composed of a bundle of features that provide utility to buyers. These features—ranging from structural features, such as the number of rooms and square footage, to locational features, like proximity to transportation, school, and commercial centers—are key determinants of property values. The model assumes that a property's price can be decomposed into the individual contributions of these features, which can be estimated as regression coefficients (Rosen, 1974). This decomposition enables the identification of the implicit prices of each feature, offering a granular perspective on how different property features influence overall value.

A key advantage of the hedonic pricing model, particularly when based on multiple linear regression, is its high degree of interpretability (Wang & Li, 2019). The interpretability stems from the model's ability to assign distinct and constant coefficients to each property features, quantifying their individual impact on property value. For instance, an increase in square footage or proximity to essential services can be directly linked to a proportional change in property value, as indicated by their respective coefficients. This transparency is critical in real estate markets, where stakeholders—ranging from property owners and investors to policymakers and assessors—require clear and easily interpretable insights into the drivers of property values. The linear structure of the hedonic model allows for straightforward hypothesis testing and comparison of features, making it a robust tool for explaining market trends and informing decision-making processes.

Building upon the theoretical framework of hedonic pricing models, these approaches have been widely adopted in both academic research and practical applications to investigate how property values are influenced by their underlying features. Studies employing these models have analyzed diverse factors, including environmental influences such as air quality and green space availability, as well as socioeconomic elements like neighborhood safety and income levels. The most frequently used features in hedonic pricing models pertain to structural characteristics, which are closely associated with the physical and functional features of a property, including property type, age, heating systems, number of bedrooms, other rooms, and available amenities. Numerous studies have demonstrated that factors such as the number of bedrooms (Li & Brown, 1980; Fletcher et al., 2000), the number of bathrooms (Garrod & Willis, 1992), and the overall floor area are positively influence property prices (Rodriguez & Sirmans, 1994; Carroll et al., 1996). For example, Garrod and Willis observed that having a single garage increases the house price by 6.9%, while a double garage contributes nearly three times that amount. Additionally, the inclusion of central heating was found to raise the property's value by approximately 6.5%. In Forrest et al. (1996), the number of garages and the type of heating system were considered as feature variables in the analysis. Chau and Chin (2003) applied a hedonic model to quantify the influence of structural features such as building services (including lifts and air conditioning systems), floor level (in multi-storey buildings), available facilities (such as swimming pools, gymnasiums, and tennis courts), and the overall structural quality, including design, materials, and fixtures, on property prices. In addition to structural characteristics, another crucial category of variables frequently used in hedonic pricing models involves locational features. These features capture the impact of a property's location on its value, which is often quantified through proximity to amenities, transportation networks, and the central business district (McMillan et al., 1992; Adair et al., 2000). Several studies emphasize the importance of accessibility, showing that properties near employment centers, schools, and public transport generally command higher prices (Follain & Jimenez, 1985). Some studies have shown that buyers are willing to pay a premium for properties with desirable views, such as those overlooking lakes or golf courses (Mok et al., 1995; Rodriguez & Sirmans, 1994). For example, properties with ocean frontage tend to command significantly higher prices compared to those with partial or no views (Benson et al., 1998). Additionally, higher floor levels in multi-storey buildings, which often provide superior views, are associated with increased property values. In contrast, Tse and Love (2000) found that properties in Hong Kong with a cemetery view tend to experience a decline in value as such views are often regarded as inauspicious in Chinese culture, symbolizing death and associated with negative feng shui (geomancy). Neighbourhood features further play a significant role in shaping property values. They reflect the broader social and environmental characteristics surrounding a property. These features encompass socioeconomic variables, such as the income levels and educational attainment of residents, as well as access to local services like schools, hospitals, and public transportation (Chau & Chin, 2003).

While the traditional hedonic pricing model is grounded in the assumption that property values are determined by intrinsic, property-specific characteristics, such as structural attributes, location, and neighborhood amenities, some studies have acknowledged that broader economic and contextual factors may also influence real estate prices. Externalities such as crime rates, pollution levels, and traffic noise often diminish property values, with higher crime rates and noise pollution leading to further decreases (Li & Brown, 1980; Clark & Herrin, 2000; Espey & Lopez, 2000). The neighborhood features, therefore, capture the quality of life offered by the surrounding area, making them a key consideration in real estate valuation models.

Accordingly, certain research has incorporated macroeconomic indicators or external market variables into extended multiple linear regression models to better capture the multifaceted determinants of housing values. For instance, Duan et al. (2021) employed a semi-logarithmic OLS model and a vector autoregression (VAR) framework to examine the dynamic effects of macro-level indicators on housing prices in Beijing, using macroeconomic data (specifically GDP). Their findings revealed that macroeconomic shocks produce varying impacts on housing prices over time, with money supply having a significant long-term positive effect. Similarly, Sayin et al. (2022) analyzed the impact of macroeconomic variables—including the dollar exchange rate, consumer price index, industrial production index, and housing loan interest rate—on housing prices using a linear regression approach, highlighting the role of economic conditions in shaping real estate values.

Although the inclusion of macroeconomic variables departs from the foundational assumptions of the classical hedonic framework, such extensions reflect an ongoing effort to adapt valuation models to the realities of dynamic and interconnected housing markets. These regression-based models aim to enhance explanatory power by accounting for external influences that, while not intrinsic to individual properties, affect price formation at the aggregate level.

While the primary strength of hedonic models lies in their simplicity and ease of interpreting regression coefficients, they have been critiqued for the strong assumptions they impose, particularly regarding linearity (Chau & Chin, 2003; Malpezzi, 2003). The conventional functional form of the hedonic pricing model simplifies the complexities of household preferences and housing markets by assuming that the effects of each feature are constant and separable. This implies several restrictive conditions, such as perfect competition, market equilibrium, and uniform preferences across markets. Consequently, the accuracy of ordinary least squares (OLS)-based models can be compromised, especially when real-world housing markets exhibit complexities or non-linear relationships. For instance, in markets segmented by housing size or income groups, or when household preferences for certain features are non-linear, the model fails to account for these intricacies. The rigidity of the traditional hedonic pricing model limits its ability to capture such complexities, as it is unable to reflect the dynamic and interrelated nature of market characteristics. Zurada et al. (2011) highlight these limitations, noting that issues such as functional form misspecification, variable interactions, multicollinearity, and non-linearity contribute to imprecise or unstable coefficients.

To address these limitations, machine learning-based models have been developed as an alternative to traditional OLS-based hedonic pricing models. These models offer greater flexibility by accommodating non-linear relationships and interactions between variables, thereby better reflecting the complexities of real-world housing markets. Many machine learning techniques can capture intricate patterns in data without relying on rigid pre-specified assumptions about the model's functional form (Antipov & Pokryshevskaya, 2012). Consequently, several studies have highlighted the effectiveness of decision trees and their ensemble models in property valuation.

For example, in Reyes-Bueno et al. (2018), decision tree models were applied to a dataset of land plot transactions in the rural sector of Vilcabamba parish in southern Ecuador. Similarly, Fan et al. (2006) employed the decision tree method to explore the relationship between housing prices and characteristics, identifying key determinants of property values in Singapore's resale public housing market. Antipov and Pokryshevskaya (2012) applied the random forest algorithm, an ensemble of decision trees, to a residential apartment dataset from Saint Petersburg, Russia, to assess its performance in property valuation. Hong et al. (2020) compared the performance of the random forest algorithm and linear regression using apartment transaction data from the Gangnam district in Seoul, South Korea. In addition to random forest models, boosted tree algorithms such as XGBoost, LightGBM, and CatBoost have been increasingly applied in property valuation studies due to their ability to minimize error and handle complex nonlinear relationships between variables. Boosted trees are particularly effective as they iteratively refine predictions by addressing the errors of previous iterations, thereby enhancing overall predictive accuracy. Examples of studies applying boosted tree algorithms to mass appraisal include Hong and Kim (2022), Ostrikova and Selyutin (2024), and Wang et al. (2020). Various machine learning techniques, such as artificial neural networks (ANN) and support vector machines, have also been applied. The backpropagation neural network approach was proposed to address the mass appraisal of real estate in Weihai city, China, by Zheng et al. (2022). McCluskey et al. (2012) examined the performance of an ANN compared to various multiple regression techniques, using data from the Lisburn District Council area in Northern Ireland. Studies by Pi-ying (2011), Yasnitsky et al. (2021), Torres-Pruñonosa et al. (2021), and Chen et al. (2024) have also applied artificial neural networks to mass appraisal problems. In studies by Kontrimas and Verikas (2011), and Bilgilioğlu and Yılmaz (2023), the support vector regression method was employed to estimate real estate prices.

Despite the predictive superiority of machine learning models, their lack of interpretability poses a significant challenge. Often described as "black boxes", these models obscure the mechanisms behind their predictions, making validation and transparency difficult. This limitation is particularly problematic in real estate valuation, where explainability is essential for fostering stakeholder trust and ensuring the adoption of predictive models (Worzala et al., 1995). Transparent and defensible models are critical, particularly when valuation outcomes have significant financial or regulatory implications. The trade-off between predictive accuracy and interpretability remains a fundamental barrier to the adoption of machine learning methods in property valuation. As McCluskey et al. (2012) note, "Although multiple regression does have its weaknesses, it is an accepted and standard method for predictive modeling. From an industry perspective, having a transparent and ultimately defensible model is a prerequisite. The black box approach of ANNs is a major impediment to undertaking price modeling for mass appraisal". Without interpretability, addressing stakeholder concerns about fairness, accountability, and potential biases becomes increasingly challenging.

In recent years, XAI techniques have been developed to enhance the interpretability of machine learning models. These approaches mitigate the limitations of traditional machine learning approaches, which often function as "black boxes", by providing insights into how models generate their predictions (Lundberg et al., 2019; Xu et al., 2019; Chen et al., 2020; Lenaers et al., 2024; Trindade Neves et al., 2024; Teoh et al., 2023; Su et al., 2021). Methods such as SHAP and LIME (Local Interpretable Model-Agnostic Explanations) increase transparency by identifying the contribution of individual features to the final output. By applying XAI, researchers and practitioners can partially address the interpretability challenges of machine learning models, making them more accessible and understandable while preserving a degree of their high predictive accuracy. Several studies have utilized explainable AI techniques in mass appraisal contexts to assess the significance of key factors and their respective contributions to price determination. For example, Iban (2022) and Teoh et al. (2023) combined tree-based algorithms with XAI techniques, using SHAP to provide local explanations for model predictions. Similarly, Tchuente (2024) conducted an experiment on the French real estate market, applying machine learning models alongside Shapley values to improve the interpretability of predictions. Krämer et al. (2023) utilized the XGBoost algorithm in conjunction with Accumulated Local Effects (ALE) plots to analyze value-determining effects of structural, locational, and socio-economic features using a dataset of 81,166 residential properties from seven major German cities. Lenaers et al. (2024) collected data on Belgian residential rental properties and developed rent prediction models using random forest and XGBoost algorithms, comparing their performance to linear regression; SHAP feature importance and summary plots were used to interpret key predictors. Chen et al. (2020) investigated the effects of urban environmental factors on residential housing prices in Shanghai using multisource data and employed SHAP to interpret the influence of these factors. Trindade Neves et al. (2024) demonstrated that integrating proprietary and open data significantly improves real estate price prediction using XGBoost in smart cities, with SHAP providing transparency into key predictors such as property size, location, accessibility to amenities, and socio-economic indicators.

Building on previous applications of explainable AI (XAI) in real estate analysis, this study evaluates and compares the interpretability and predictive performance of linear and nonlinear models using SHAP and Permutation Feature Importance (PFI). To identify the key factors influencing housing prices, a balanced feature set encompassing structural, neighborhood, locational, and macroeconomic variables was carefully constructed, representing an advancement over prior studies that often focused on a narrower range of predictors. Multiple linear regression (MLR) and random forest (RF) were selected as representative models commonly employed in mass appraisal systems, reflecting their differing capacities to capture relationships between variables. Although previous studies have incorporated XAI techniques in real estate research, few have systematically examined how linear and nonlinear models differ in their representation of feature importance and the mechanisms underlying their predictions. Moreover, limited attention has been given to understanding how feature importance evolves over time, despite the dynamic nature of housing markets driven by demographic changes, infrastructure developments, and shifting consumer preferences.

The main contributions of this study are as follows. First, a comparative analysis of feature importance values obtained through PFI and SHAP was conducted for both the linear regression and random forest models to assess model interpretability. This comparison enables an evaluation of the consistency of feature importance across models with differing functional assumptions; features exhibiting high importance in both models suggest a stable and robust influence on apartment prices, irrespective of model form. Discrepancies in feature importance rankings between the models provide insights into potential nonlinearities and interaction effects among predictors, where features identified as important exclusively in the random forest model imply complex or conditional relationships not captured by linear regression. Furthermore, the cross-model comparison facilitates an assessment of the explanatory adequacy of XAI techniques. While linear regression inherently offers interpretability through model coefficients, SHAP and PFI provide a more granular and comprehensive understanding of feature contributions, particularly under conditions of structural complexity. For random forest models, where internal decision structures are inherently opaque, the role of XAI methods becomes indispensable. Second, temporal analyses of PFI and SHAP values were conducted to investigate how the importance of key features evolves over time. By tracking changes in feature contributions across different periods, this study reveals dynamic shifts in market drivers and offers new insights into the adaptability and sensitivity of both linear and nonlinear models to changing real estate environments. This longitudinal perspective highlights the necessity of incorporating temporal dynamics into mass appraisal modeling and demonstrates the added value of XAI methods in capturing complex, evolving patterns within the housing market.

3. Methodology

Our aim is to investigate how different predictive factors contribute to estimating house prices by integrating XAI methodologies, including SHAP and PFI, with standard machine learning regression models. While advanced models with superior performance are available, we focus on fundamental machine learning models commonly used in practice. For this study, we selected two representative predictive models: a linear hedonic model based on multiple regression analysis and a non-linear random forest algorithm. These models were chosen for their widespread applicability in house price prediction and their balance between predictive performance and interpretability. This section provides an overview of the algorithms employed in this study and their underlying methodologies. Additionally, it focuses on the XAI methodologies utilized, specifically SHAP and PFI. These approaches are employed to systematically evaluate the contributions of individual predictive factors, enhancing the interpretability of the models. Detailed explanations of SHAP and PFI are provided to demonstrate their application in analyzing and interpreting the predictive models used in this study.

All analyses were conducted on a workstation equipped with an AMD Ryzen 5 7500F 6-Core Processor (3.70 GHz) and 64 GB of RAM, operating on a 64-bit Windows system. The computational environment was based on Python 3.11.9. Random Forest and Multiple Linear Regression models were implemented using the scikit-learn library. To interpret model outputs, SHapley Additive exPlanations (SHAP) were computed using the shap package, and Permutation Feature Importance (PFI) was derived using the permutation_importance function from scikit-learn. Data transformations were performed using pandas and numpy.

3.1. Predictive models

3.1.1. Multiple linear regression model

The hedonic pricing model, widely used in the valuation of real estate and other goods, is theoretically rooted in Lancaster's characteristics demand theory (Lancaster, 1966) and Rosen's extension of this theory (Rosen, 1974). Lancaster posited that consumers derive utility not directly from goods themselves but from the composite characteristics or features these goods possess. For example, in the context of real estate, consumers value a house based on its specific features, such as size, location, proximity to amenities, and environmental factors, rather than the house as a singular entity. Rosen expanded upon Lancaster's theory by proposing the hedonic pricing model, which argues that the price of a good is the aggregate value of its features. In an equilibrium market, each feature is assumed to have a unique implicit price, which collectively determines the overall price of the good. This theoretical framework implies that the price of a product, such as a house, can be modeled as a function of its characteristics, allowing us to statistically estimate the contribution of each feature to the price through regression analysis.

In this study, we utilize a pricing model based on the multiple linear regression method, which assumes a linear relationship between the price of a house and its explanatory variables. The model can be expressed mathematically as follows:

$$p_i = \beta_0 + \sum_{j=1}^{k} \beta_j x_{i,j} + \varepsilon_i. \quad (1)$$

where: p_i represents the natural logarithm of the price of the i-th house. Taking the natural logarithm of house prices is a standard approach in the linear pricing models as it helps to address skewness in the distribution of housing prices and ensures that the model better satisfies the assumptions of linear regression, such as homoscedasticity (Hong et al., 2020). Additionally, the logarithmic transformation allows for a more interpretable interpretation of the coefficients, where each coefficient (β_j) can be understood as the percentage change in price associated with a one-unit change in the corresponding explanatory variable. The model assumes that the relationship between the dependent variable (p_i) and the explanatory variables ($x_{i,j}$) is additive and linear. Here, β_0 is the intercept term, which represents the predicted value of p_i when all explanatory variables are zero. β_j denotes the regression coefficient for the j-th explanatory variable, capturing the magnitude and direction of the relationship between $x_{i,j}$ and p_i. ε_i represents the error term for the i-th observation, accounting for variations in the dependent variable that cannot be explained by the explanatory variables. The error term is assumed to follow a normal distribution with a mean of zero and constant variance, satisfying the Gauss-Markov assumptions required for unbiased and efficient estimation.

The regression coefficients in the linear pricing model are estimated using the Ordinary Least Squares (OLS) method, which minimizes the sum of squared residuals between the observed and predicted prices. The estimated coefficients ($\hat{\beta}_j$) can be interpreted as the marginal contribution of each feature to the price of a house, assuming all other factors remain constant. The fitted model can be expressed as:

$$\hat{p}_i = \hat{\beta}_0 + \sum_{j=1}^{k} \hat{\beta}_j x_{i,j} + \varepsilon_i. \quad (2)$$

While the OLS approach ensures unbiased and efficient estimates under its assumptions, deviations from these assumptions can affect the validity of the results. For example, the presence of multicollinearity among explanatory variables can inflate the standard errors of the estimated coefficients, reducing their reliability. Similarly, omitted variable bias may arise if relevant factors influencing house prices are not included in the model, potentially distorting the results. Therefore, the proper specification of the functional form and the careful selection of explanatory variables are essential to maintain the robustness of the model. By focusing on the linear relationship between house prices and their features, the linear pricing model decomposes the total price of a property into the values of its individual characteristics. This decomposition provides actionable insights into the relative importance of different factors, such as proximity to public amenities, neighborhood quality, or structural features. Despite its simplicity, this framework remains one of the most widely used tools in empirical real estate research, offering a balance between interpretability and analytical rigor.

3.1.2. Random forest model

Another algorithm employed in this study is the Random Forest algorithm, which is based on decision tree principles. This algorithm demonstrated the best performance compared to the other machine learning prediction algorithms we evaluated (see Appendix Table A1). Decision Trees (DTs) are a foundational decision support tool in machine learning that utilize a tree-like structure to model decision-making processes. Each node in a decision tree represents a decision based on a specific feature and its corresponding threshold, while the branches signify the outcome of the decision. For example, given a node split based on feature A with threshold T, a sample with $A < T$ will follow the left branch, while a sample with $A \geq T$ will proceed along the right branch. Decision trees are versatile and can be applied to both classification and regression problems. In classification, each terminal node (leaf) represents a class, and predictions are made by traversing the tree from the root to a leaf node. In regression, the process involves predicting continuous values by averaging the target variable within each terminal node, which defines a local approximation of the data. The construction of a decision tree involves recursively selecting features and thresholds to split the data into increasingly homogeneous subsets. This process is guided by specific metrics, such as reduction in variance (for regression) or Gini impurity and information gain (for classification). The algorithm grows the tree by iteratively adding nodes, splitting data at each step to create subgroups that maximize the chosen splitting criterion. While decision trees are highly interpretable and straightforward to construct, they are prone to overfitting, especially when grown to full depth without pruning. This limitation often results in a model that performs well on the training data but poorly generalizes to unseen data.

Random Forest (RF) addresses the overfitting problem inherent in single decision trees by employing an ensemble learning technique. It combines the predictions of multiple uncorrelated decision trees to produce a robust and accurate model. Each tree in a random forest is independently trained on a bootstrap sample of the original dataset, with a random subset of predictors considered at each node split. This randomness in both the data and feature selection enhances model diversity and reduces the correlation among trees, leading to improved generalization performance. The predictions of the random forest model are aggregated by averaging the outputs of all trees for regression tasks. For example, if a random forest is trained on a housing price dataset, each tree independently predicts a price based on its training subset, and the final prediction is obtained by averaging the outputs of all trees. This ensemble approach not only reduces variance but also mitigates the risk of overfitting, a key advantage over individual decision trees.

A unique strength of the Random Forest algorithm lies in its robustness when dealing with high-dimensional datasets and a mixture of variable types. The research used one-hot encoding to transform categorical data into binary indicators which served as input for model training. The tree construction process of the algorithm chooses random feature subsets for each node split to create diverse models which help prevent overfitting. This reduces the dimensionality burden and mitigates the risk of overfitting, which is often a challenge in traditional regression models, such as OLS or neural networks. Furthermore, RF is well-suited for capturing nonlinear relationships and interactions among variables, making it particularly advantageous in complex domains like real estate mass appraisal, where property features such as location, brand, and heating system exhibit nonlinear and heterogeneous effects on housing prices.

One of the practical advantages of RF is its relative simplicity in training and interpretation. It requires only two primary hyperparameters: the number of trees in the forest and the maximum depth of each tree. Increasing the number of trees generally improves stability and predictive accuracy without a significant computational burden, while adjusting tree depth allows the model to balance precision and overfitting. We explored the optimal RF parameters using the grid search method and confirmed that the best performance was achieved when the number of trees was set to 10 among {5, 10, 15, 20} hyperparameters, and the maximum tree depth was set to 5 among {5, 10, 15, 20} hyperparameters. Additionally, RF retains a degree of interpretability by enabling feature importance analysis, where the contribution of each variable to the model's predictive performance is quantified. This feature is especially valuable in understanding the relative importance of different housing features, facilitating actionable insights for stakeholders in real estate markets. While RF models are computationally efficient and exhibit strong predictive performance, they can sometimes lack the interpretability of simpler models like OLS. However, the incorporation of explainable AI techniques, such as SHAP, can bridge this gap by providing insights into individual predictions.

3.1.3. Evaluating model performances

To evaluate and compare the predictive performance of the multiple linear regression model and the Random Forest algorithm, we employed four widely used performance metrics: the coefficient of determination (R^2), Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE) and Mean Absolute Error (MAE) (see Appendix Table A1). These metrics provide complementary insights into the accuracy and reliability of the models' predictions. The multiple linear regression model was trained using the natural logarithm of house prices ($p_i = \ln y_i$), which stabilizes variance and addresses skewness in housing price distributions, thus improving the model's adherence to the assumptions of linear regression. In contrast, the Random Forest model was trained directly on the original price values (y_i). To ensure consistency in performance evaluation, predictions from the regression model ($\hat{p}_i$) were exponentiated ($\hat{y}_i = \exp(\hat{p}_i)$) to transform them back to the original price scale before calculating the performance metrics. Predictions from the Random Forest model ($\hat{y}_i$) were already on the original price scale. The coefficient of determination (R^2) measures the proportion of variance in the actual prices (y_i) that is explained by the model. It is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2}, \tag{3}$$

where: y_i represents the actual price of the i-th house; $\hat{y}_i$ is the predicted price (either directly from the Random Forest model or exponentiated from the regression model's log-transformed predictions); $\bar{y}$ is the mean of the actual prices calculated as $\bar{y} = \left(1/n\right)\sum_{i=1}^{n} y_i$; n is the total number of observations in the dataset. In this formula, the numerator $\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2$ represents the residual sum of squares (RSS), which captures the variance unexplained by the model, while the denominator $\sum_{i=1}^{n}\left(y_i - \bar{y}\right)^2$ represents the total variance in the actual prices. Higher R^2 values indicate better model performance, with 1 representing a perfect fit.

The Mean Absolute Percentage Error (MAPE) evaluates the average percentage deviation between actual and predicted prices, offering an intuitive, scale-independent measure of prediction accuracy. *MAPE* is defined as:

$$MAPE = \frac{1}{n}\sum_{i=1}^{n}\left|\frac{y_i - \hat{y}_i}{y_i}\right| \times 100. \tag{4}$$

Lower MAPE values indicate higher prediction accuracy and reflect how closely the predicted prices align with the actual prices as a percentage of the actual values. By combining R^2 and *MAPE*, the evaluation provides a comprehensive view of the models' predictive capabilities. The R^2 metric highlights the proportion of variance explained by the model, while *MAPE* captures the relative prediction error in practical terms.

The Root Mean Square Error (RMSE) constitutes a performance metric that calculates the square root of the mean squared residuals between predicted values and observed outcomes. This evaluation criterion applies disproportionate weighting to substantial errors compared to minor deviations, rendering it particularly responsive to outlying observations. It is calculated as:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}. \tag{5}$$

The Mean Absolute Error (MAE) quantifies the arithmetic mean of the absolute differentials between model predictions and empirical observations, irrespective of directional orientation. In contrast to RMSE, the MAE applies uniform weighting across the error distribution, thereby providing a more equitable assessment of predictive accuracy in contexts where anomalous observations are not of significance. It is defined as:

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right|. \quad (6)$$

3.2. Explainable artificial intelligence (XAI)

As machine learning models become increasingly complex, their decision-making processes have often been criticized as "black boxes", making it challenging for practitioners to interpret and trust their predictions. While these models, such as Random Forests, achieve high predictive accuracy, their lack of transparency limits their applicability in domains where interpretability is crucial, such as finance, healthcare, and real estate. To address this challenge, XAI methods have been developed, providing insights into how features influence model predictions and enabling stakeholders to better understand, validate, and trust the outputs of machine learning models.

Among the prominent XAI methods, SHAP has emerged as a robust tool for interpreting complex model predictions. Introduced by Lundberg and Lee (2017), SHAP leverages Shapley values from cooperative game theory to quantify the contribution of each feature to an individual prediction (Shapley, 1953). This method calculates the marginal contribution of a feature by evaluating the changes in the model's prediction when the feature is included or excluded from different subsets of input variables. The Shapley value for a feature i is computed as:

$$\phi_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\left(|N| - |S| - 1\right)!}{|N|!}\left[f_x\left(S \cup \{i\}\right) - f_x\left(S\right)\right], \quad (7)$$

where: N represents the set of all features; S is a subset of N excluding feature i; $f_x\left(S\right)$ is the model's prediction considering only the features in S; $f_x\left(S \cup \{i\}\right)$ is the prediction when feature i is added to S. The Shapley value, ϕ_i, thus quantifies the contribution of feature i by averaging its marginal contributions across all possible subsets of features.

SHAP provides a mathematically sound and interpretable framework for understanding how features interact to produce a specific prediction. Each SHAP value ϕ_i captures the impact of a feature not in isolation but in the context of other features, offering a holistic view of the prediction process. For example, in a real estate valuation model, a positive SHAP value for "proximity to schools" indicates that this feature positively influences the predicted housing price, while a negative SHAP value for "age of the property" suggests a reduction in price. These values are particularly useful in applications requiring transparency and accountability, as they allow users to trace back the prediction to its contributing factors. Furthermore, SHAP values satisfy the efficiency property, ensuring that the sum of all feature contributions equals the difference between the model's prediction and a baseline output, reinforcing their interpretability.

To improve computational efficiency for tree-based models, Lundberg et al. (2018) proposed TreeSHAP, a specialized algorithm designed to calculate SHAP values for decision trees and Random Forests. TreeSHAP maintains the theoretical properties of traditional SHAP while significantly reducing computational costs. In this study, TreeSHAP was employed to compute SHAP values for the Random Forest model, with the global importance of each feature quantified as the mean absolute value of its SHAP values across all instances:

$$I_j = \frac{1}{n}\sum_{i=1}^{n}\left|\phi_j^{(i)}\right|, \quad (8)$$

where: $\phi_j^{(i)}$ is the SHAP value of feature j for instance i; I_j represents the global importance score of feature j. This global importance measure aggregates the localized effects of a feature across the entire dataset, providing insights into its average influence on model predictions.

In contrast to SHAP's instance-specific explanations, Permutation Feature Importance (PFI) focuses on the global relevance of features by evaluating their impact on model performance. Originally introduced by Breiman (2001) as part of the Random Forest algorithm, PFI measures the change in performance when the values of a specific feature are randomly shuffled. By disrupting the relationship between the feature and the target variable, PFI quantifies how much the model depends on the feature to make accurate predictions. The importance of feature j is calculated as:

$$I_j^{PFI} = \frac{1}{M}\sum_{i=1}^{M}\left(\text{Metric}_{\text{baseline}} - \text{Metric}_{\text{permuted}(j)}\right), \quad (9)$$

where: M is the number of permutations; $\text{Metric}_{\text{baseline}}$ is the model's performance on the original dataset; $\text{Metric}_{\text{permuted}(j)}$ is the performance after permuting the values of feature j. In this study, the Mean Absolute Percentage Error (MAPE) was used as the performance metric. A significant drop in performance after shuffling a feature indicates its critical role in predictions, while minimal changes suggest that the feature contributes little to the model's accuracy.

PFI provides a straightforward and intuitive measure of feature importance, making it appealing for a wide range of machine learning applications. However, PFI assumes independence among features, which can lead to biased importance estimates in datasets with correlated features. For example, in real estate data, highly correlated features such as "square footage" and "number of bedrooms" may share predictive information. Shuffling one feature could inadvertently affect the importance score of the other, complicating interpretation. Despite this limitation, PFI remains a valuable tool for understanding the global structure of a model's dependencies.

Together, SHAP and PFI offer complementary insights into feature importance. SHAP excels in providing granular, instance-level explanations, enabling users to dissect individual predictions into their contributing factors. On the other hand, PFI emphasizes a feature's overall impact on model performance, highlighting its global relevance across the dataset. By combining these methods, this study achieves a balanced approach to interpretability, addressing both local and global perspectives. This dual approach ensures that the machine learning models used for housing price prediction are not only accurate but also transparent, offering actionable insights into the factors driving predictions and enhancing trust in the model's outcomes.

4. Data

Gangnam District in Seoul, South Korea, was selected as the focal area for data collection and analysis in this study. Known for its high real estate values and modern urban infrastructure, Gangnam provides an ideal context for investigating the factors that influence property markets in a highly competitive and developed urban environment. Moreover, its high population density and the large volume of real estate transactions concentrated within a compact area make it particularly well-suited for applying XAI methods. This approach enables a detailed analysis of the key determinants of housing prices and their temporal dynamics, allowing for a comprehensive exploration of how various factors interact and evolve over time.

Situated south of the Han River, Gangnam is one of Seoul's most prominent and affluent areas. Spanning 39.5 square kilometers with a population exceeding 560,000, it ranks as the third-largest administrative district in the city. Gangnam is divided into 22 administrative subdivisions, or "dongs", which facilitate efficient governance and community management. Figure 1 presents two maps: the map on the left shows the location of Gangnam District within Seoul, and the map on the right illustrates the administrative subdivisions ("dongs") of Gangnam District. Often regarded as a benchmark for urban development and residential desirability, the district has become a focal point for both domestic and international attention. The real estate market in Gangnam is exceptional in South Korea, consistently ranking among the most expensive nationwide. As of October 2024, the average apartment price in Gangnam was approximately 93.95 million KRW per pyeong (3.3 square meters), equivalent to about 28.47 million KRW or approximately $20,300 per square meter, assuming an exchange rate of 1,400 KRW/USD (Maeil Business Newpaper, 2024). This premium pricing far exceeds the average apartment price in Seoul and is significantly higher than the national average. The elevated property values reflect the district's combination of desirable features, including its exceptional educational resources, advanced infrastructure, and high standard of living.

A key factor contributing to the desirability of Gangnam District is its well-established educational environment. The district hosts a concentration of prestigious schools and elite private academies, commonly known as "hagwons", which attract families seeking superior educational opportunities for their children. This strong emphasis on education has positioned it as a central driver of housing demand, particularly in neighborhoods located in close proximity to these institutions. Consequently, this demand has significantly influenced the rise in property prices within the area.

Moreover, Gangnam's appeal is further enhanced by its extensive and meticulously planned urban infrastructure. The district is seamlessly integrated into the broader Seoul metropolitan area through a comprehensive network of subway lines, major roads, and bus routes, ensuring efficient access to key business hubs, cultural landmarks, and recreational facilities. Residential developments in Gangnam, including high-rise apartment complexes and mixed-use buildings, are designed to cater to the needs of affluent residents, offering premium features such as private parking facilities, advanced security systems, and landscaped green spaces. In addition to its robust infrastructure, Gangnam provides an array of high-quality amenities that contribute to an elevated standard of living. These include world-class medical facilities, upscale shopping

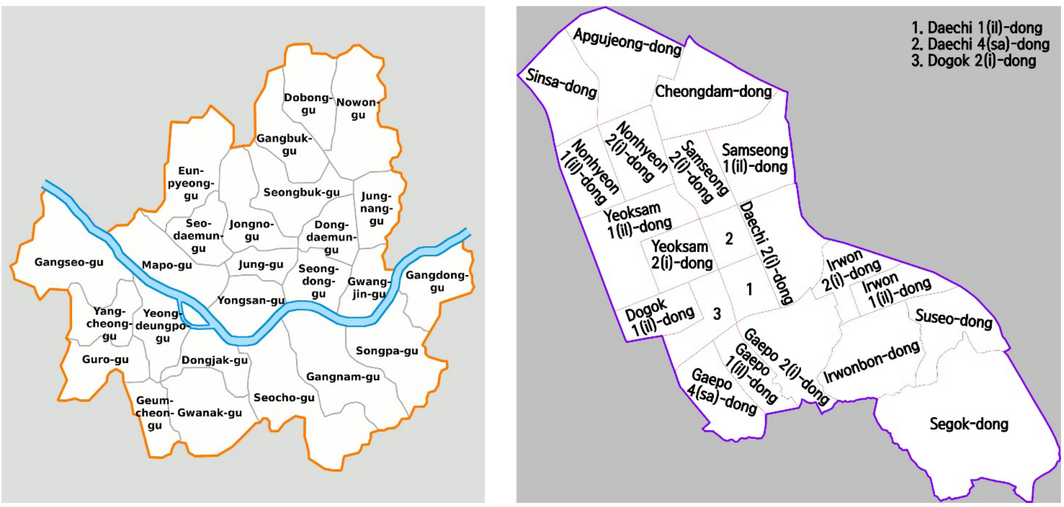


Figure 1. Location of Gangnam district in Seoul and its administrative division (source: Wikipedia, 2025)

centers, and diverse recreational spaces, which collectively enhance the attractiveness of the district for both domestic and international buyers. As a result, Gangnam has firmly established itself as one of the most sought-after residential areas in Seoul, with steady transaction activity that reflects its continued desirability.

This steady level of transactions provides a valuable dataset for analyzing market dynamics using advanced data-driven models, as demonstrated in prior studies (Hong et al., 2020). The availability of such stable data enables the identification and examination of temporal variations in key determinants of property prices, offering insights into the interactions between socioeconomic and infrastructural factors.

Although the study focuses exclusively on Gangnam District and utilizes data from the period 2006–2017, this scope was intentionally defined to serve the study's methodological aims. Gangnam has long functioned as a price-leading and policy-sensitive submarket within Seoul, characterized by consistently high property values, active transaction volumes, and strong influence on price trends across other regions, especially in response to government policy changes (Al-Yahyaee et al., 2021; Bae & Joo, 2020). These characteristics make Gangnam a strategic setting for developing and validating explainable AI (XAI) models aimed at capturing the multifaceted and evolving drivers of housing prices. Furthermore, the chosen timeframe encompasses a number of significant macroeconomic shifts and real estate policy changes. Notably, loan-to-value (LTV) and debt-to-income (DTI) regulations were repeatedly tightened or relaxed during this period, often in direct response to speculative pressures in high-demand areas like Gangnam. For example, DTI limits were first introduced in Gangnam in 2005 and subsequently adjusted multiple times throughout the following decade. The study period also includes major policy shifts such as redevelopment restrictions and housing supply initiatives, providing a diverse and dynamic policy context for evaluating model performance.

While the dataset does not reflect the most recent market conditions, its extended temporal span enables the identification of long-term structural patterns—such as evolving preferences for large-scale apartment complexes, the influence of building age and redevelopment expectations, and the increasing premium associated with branded housing. These phenomena remain analytically valuable, as the core determinants of apartment prices—such as structural, locational, and neighborhood attributes—retain relevance regardless of market cycle. Although geographically limited, the analytical framework used in this study is generalizable and can be applied to other urban areas with comparable data availability. Rather than aiming to predict contemporary prices, the study's primary contribution lies in improving the interpretability of machine learning-based mass appraisal models using XAI, while tracing how the importance of key features shifts over time under varying market and policy conditions.

A total of 15,162 apartment transaction records from Gangnam District, spanning the period from 2006 to 2017, were obtained from the Ministry of Land, Infrastructure, and Transport (MOLIT) in South Korea. These records represent approximately 40% of all apartment transactions that occurred in Gangnam during the specified timeframe. This proportion reflects the subset of transactions retained after data cleaning procedures, during which entries with missing or incomplete values in key variables—such as geospatial coordinates, building characteristics, or other essential property attributes—were excluded to ensure data quality and analytical consistency. The variables used in the analysis are presented in Table 1. Note that the property price is the target variable analyzed in this study.

The structural features represent inherent characteristics of the property. In this study, we include construction year, area, and floor level of a property as key structural variables. While construction year is often associated with the entire apartment complex, it has been included as a variable in this study because buildings within the same complex may have been constructed at different times, leading to variations in property characteristics. Area refers to the total floor area of the property, while Floor level captures the vertical position of the unit within the building. These variables are considered essential for understanding the structural differences that influence property prices.

Neighborhood features represent the shared characteristics of the apartment complex and its surrounding environment. These features include apartment brand, the number of available units within the building, the total number of buildings in the complex, parking availability, heating system, floor area ratio (FAR), building coverage ratio (BCR), and the tallest and shortest building heights within the apartment complex. The apartment brand variable is calculated using a ranking system based on data from the Korea Institute of Corporate Reputation. Brands ranked within the top 15 are assigned scores according to their rank, with the highest-ranked brand receiving a score of 15, the second-ranked brand a score of 14, and so on. Brands ranked below 15th are assigned a score of 0. This scoring system reflects the reputation and desirability of the brand in influencing property prices. The parking availability variable represents the average number of parking spaces per household within the apartment complex. The floor area ratio (FAR) is calculated as the ratio of the total floor area (gross floor area) to the total land area, while the building coverage ratio (BCR) represents the ratio of the building's footprint to the total land area. The inclusion of the tallest and shortest building heights as variables captures the vertical variation in the complex's design, which can influence property values. These neighborhood features collectively provide insights into the shared physical and environmental factors that impact the value of properties within an apartment complex. The locational features used in this study, which also influence property prices, include the dong (administrative division) and accessibility to nearby facilities. The facilities considered in the analysis are national parks, high schools, redevelopment areas, universities, general hospitals, museums, and subway stations. The information on the administrative division (dong) of the property was obtained from the dataset provided by the Ministry of Land, Infrastructure, and Transport (MOLIT), while the distances to the nearest facilities were calculated using data retrieved through the MAP open Application Programming Interface (API). To ensure spatial precision, all distance-based locational variables were calculated using the geospatial coordinates (latitude and longitude) of each apartment complex as reported in the transaction dataset. Euclidean distances to the nearest facilities were computed based on these coordinates, using data retrieved via the Naver Map Open API. This approach provides a high spatial resolution at the apartment-complex level, offering greater granularity than methods based on aggregated administrative units.

Previous studies have shown that macroeconomic factors can significantly impact the housing market (Miller et al., 2011). In this study, the relevant macroeconomic variables include the transaction period (year and quarter), the size of the economy (gross domestic product, GDP), economic growth rate (percentage growth in GDP), the land price fluctuation rate in Seoul, and the mortgage interest rate. These variables are measured annually. Descriptive statistics for the numerical variables are provided in Table 2.

Table 1. Variables

Category	Variables	Unit
Target variable	Price	Korean won (KRW)
Structural features	Construction year Area Floor level of a property	Year m^2 Floor level
Neighborhood features	Apartment brand Number of units in the apartment complex Number of buildings in the apartment complex Parking lot Heating system Floor area ratio (FAR) Building coverage ratio (BCR) The tallest building height The shortest building height	Ranking of prominent apartment brands (scores are assigned based on rankings up to 15th place, while a score of 0 is given for brands ranked below 15th, otherwise 1) Number of units Number of buildings Number of parking spot/number of units 0 if an apartment has a central heating system Otherwise, the value is set to 1 Ratio Ratio Floor level Floor level
Locational features	Dong (administrative subdivisions) Distance to the nearest national park Distance to the nearest high school Distance to the nearest redevelopment area Distance to the nearest university Distance to the nearest general hospital Distance to the nearest museum Distance to the nearest subway station	The name of the “dong” (categorical variable) Meter Meter Meter Meter Meter Meter Meter
Macro variable	Transaction Year and quarter Gross domestic product (GDP) Economic growth rate Land price fluctuation rate in Seoul Mortgage loan interest rate	Year and quarter Billion won % % %

Table 2. Descriptive statistics

Variables	Mean	Median	Standard deviation	Min	Max
Price (ten-thousands won)	84137.4	77000	0.569857	1000	570000
Construction year	1993.32	1993	0.005186	1978	2014
Area	72.6455	59.98	0.4933	16.78	273.83
Floor level	7.66	6	0.74	–1	45
Apartment brand	1.25234	0	2.70865	0	15

End of Table 2

Variables	Mean	Median	Standard deviation	Min	Max
Number of units in the complex	1666.96	805	1.09	7	5040
Number of buildings in the complex	28.42	8	1.44	1	124
Parking lot	1.0292	1.00	0.5823	0.27	4.53
Floor area ratio (FAR)	258.52	237	0.81	72	2435
Building coverage ratio (BCR)	24.91	19	0.60	12	204
The tallest building height	14.08	14	0.54	4	46
The shortest building height	11.26	12	0.50	3	26
Distance to national park	1060.18	1053.21	0.377388	86.1079	2142.47
Distance to high school	535.593	522.609	0.443323	31.8829	1531.52
Distance to redevelopment area	639.331	571.583	0.65956	0.00	3758.56
Distance to university	3369.99	3527.72	0.379691	24.5866	7136.5
Distance to general hospital	1037.13	965.389	0.493903	41.6327	3470.83
Distance to museum	986.687	1038.93	0.380867	87.4901	3323.86
Distance to subway station	660.919	557.849	0.581247	47.487	2559.07
GDP (billion won)	349295.82	353743	0.16	240439	446835
Economic growth rate	3.362	3.2	0.529	–1.9	7.4
Land price fluctuation rate	0.00770145	0.0158096	39.504	–2.64275	0.351297
Mortgage interest rate	6.14895	5.90853	0.0972837	5.2633	7.41544

5. Results and discussions

This study employed a 5-fold cross-validation approach, which divides the dataset into five subsets, using four subsets for training and one for validation in each iteration. This ensures a robust evaluation of the model's predictive performance by minimizing the impact of data partitioning on the results. Using the approach, for the linear regression model, the results showed an R^2 value of 0.892 and a MAPE of 0.12227. In comparison, the random forest model achieved an R^2 value of 0.970 and a MAPE of 0.05928. Numerous studies have highlighted the superior predictive performance of random forest models compared to linear regression models, making this result somewhat expected (Hong et al., 2020). Since multicollinearity caused by correlations among variables can pose a problem in both multiple linear regression models and Permutation Feature Importance (PFI) analysis, we examined the Variance Inflation Factor (VIF) scores for all variables. The analysis revealed that most VIF values were below 10, suggesting that multicollinearity does not significantly impact our results (see Appendix Table A2). Also, as mentioned earlier, XAI techniques were incorporated to enhance the interpretability of the model. The mean absolute SHAP values and Permutation Feature Importance scores (PFI) were applied to analyze the importance of features in predicting property prices and to provide insights into the relationships between features and the target variable. The insights derived from SHAP and PFI analyses were further utilized to identify key drivers of property prices, offering actionable information to support decision-making in the real estate domain. The analysis was conducted using Python and relevant libraries, including SHAP and scikit-learn, ensuring the reproducibility and transparency of the results.

5.1. Comparison of random forest and linear regression methods

The Mean Absolute SHAP values and Permutation Feature Importance scores were calculated for both a linear regression model and a random forest model, and the results are presented in Tables 3 and 4. It is noteworthy that, in the linear regression model, the target variable was transformed by applying a logarithmic function to apartment prices in order to improve predictive performance. This transformation stabilizes the variance of the target variable and enables the model to better capture relationships between features and prices. However, as a consequence, the SHAP and PFI values derived from the linear regression model reflect the importance of features on the logarithmic scale rather than the original price scale. Therefore, the numerical values of SHAP and PFI from the linear regression model cannot be directly compared to those from the random forest model, which evaluates feature importance on the original price scale. Nonetheless, within each model, these metrics provide valuable insights into the relative importance of features and their contributions to the target variable. Additionally, by examining the rankings and patterns of feature importance across both models, it is possible to identify which variables are consistently significant and how their influence may differ depending on the modeling approach. This analysis enables a nuanced interpretation of the factors driving apartment prices.

Table 3. Explanatory indicators for apartment prices in the linear regression model

Rank	Linear regression model			
	Mean absolute SHAP		Permutation feature importance	
1	Area	0.2879490	Area	0.0270373
2	Number of units in the complex	0.1652130	Number of units in the complex	0.0124664
3	Dong	0.1650627	Dong	0.0106721
4	Construction year	0.1363777	Construction year	0.0071438
5	GDP	0.0979463	The tallest building height	0.0053472
6	The tallest building height	0.0932182	GDP	0.0050100
7	Transaction year and quarter	0.0804335	Apartment brand	0.0032981
8	The shortest building height	0.0658740	Transaction year and quarter	0.0030130
9	Number of buildings	0.0492220	The shortest building height	0.0029705
10	Apartment brand	0.0454338	Number of buildings	0.0020738
11	Heating system	0.0446536	Heating system	0.0011156
12	Mortgage interest rate	0.0354506	Mortgage interest rate	0.0009495
13	Building coverage ratio (BCR)	0.0213643	Floor area ratio (FAR)	0.0006287
14	Floor area ratio (FAR)	0.0213598	Building coverage ratio (BCR)	0.0004180
15	Floor level	0.0109712	Floor level	0.0001145
16	Economic growth rates	0.0064508	Parking lot	0.0000675
17	Parking lot	0.0063428	Economic growth rates	0.0000547
18	Land price fluctuation rate	0.0028206	Land price fluctuation rate	0.0000272
19	Distance to the nearest national park	0.0025191	Distance to the nearest national park	0.0000147
20	Distance to the redevelopment area	0.0019281	Distance to the nearest subway station	0.0000056
21	Distance to the nearest general hospital	0.0013635	Distance to the nearest national museum	0.0000013
22	Distance to the nearest subway station	0.0011362	Distance to the nearest university	0.0000001
23	Distance to the nearest national museum	0.0009954	Distance to the nearest high school	–0.0000006
24	Distance to the nearest high school	0.0005801	Distance to the redevelopment area	–0.0000031
25	Distance to the nearest university	0.0001085	Distance to the nearest general hospital	(0.0000044)

Table 4. Explanatory indicators for apartment prices in the random forest model

Rank	Random forest model			
	Mean absolute SHAP		Permutation feature importance	
1	Area	22098.10	Area	0.45935900
2	Number of buildings	10862.90	Number of buildings	0.25295000
3	GDP	6397.02	GDP	0.12923800
4	Dong	3643.78	Dong	0.05789090
5	The shortest building height	2576.39	Building coverage ratio (BCR)	0.04592850
6	Number of units in the complex	2439.17	Number of units in the complex	0.03644470
7	Mortgage interest rate	1378.81	Mortgage interest rate	0.03068410
8	Floor area ratio (FAR)	1344.20	Floor area ratio (FAR)	0.02470100
9	The tallest building height	1339.18	The shortest building height	0.02383690
10	Parking lot	1270.30	Land price fluctuation rate	0.02169290
11	Building coverage ratio (BCR)	1152.89	The tallest building height	0.01770560
12	Land price fluctuation rate	1011.33	Parking lot	0.01360550
13	Construction year	944.63	Economic growth rates	0.01288790

End of Table 4

	Random forest model			
Rank	Mean absolute SHAP		Permutation feature importance	
14	Apartment brand	833.71	Construction year	0.01247240
15	Floor level	605.82	Floor level	0.01231590
16	Transaction year and quarter	586.77	Transaction year and quarter	0.00925989
17	Economic growth rates	569.85	Apartment brand	0.00768583
18	Distance to the redevelopment area	261.18	Distance to the nearest national park	0.00480895
19	Distance to the nearest national park	205.44	Distance to the redevelopment area	0.00429915
20	Heating system	180.99	Distance to the nearest university	0.00326974
21	Distance to the nearest university	113.81	Distance to the nearest high school	0.00306591
22	Distance to the nearest subway station	104.85	Distance to the nearest subway station	0.00303488
23	Distance to the nearest high school	92.14	Distance to the nearest national museum	0.00302023
24	Distance to the nearest national museum	87.55	Distance to the nearest general hospital	0.00265319
25	Distance to the nearest general hospital	84.96	Heating system	0.00173744

The comparison between the linear regression and random forest models reveals both similarities and differences in evaluating feature importance for predicting apartment prices. Both models identified key features, such as "Area", "Number of units in the complex", "Dong", "GDP", and "Construction year", as highly significant, suggesting that these variables have strong intrinsic relationships with apartment prices regardless of the modeling approach. However, notable differences were observed in the rankings of certain variables, reflecting the distinct characteristics of each model.

For example, in the linear regression model, "Number of units in the complex" was ranked as a more important feature compared to "Number of buildings". In contrast, the random forest model assigned higher importance to "Number of buildings" and "Building coverage ratio (BCR)". Both variables represent aspects of the scale of an apartment complex, but their differing rankings reflect the fundamental characteristics of the two modeling approaches. "Number of units in the complex" has a more linear relationship with apartment prices, which aligns with the assumptions of the linear regression model. As a result, it is prioritized in the feature ranking within this model. On the other hand, "Number of buildings", while less linearly correlated with apartment prices, captures more complex interactions and nonlinear relationships that the random forest model, based on decision-tree algorithms, is well-equipped to detect. Additionally, "Building coverage ratio (BCR)" complements "Number of buildings" by providing a related but distinct measure of the physical scale and layout of apartment complexes, further emphasizing the random forest model's ability to account for these features. These differences suggest that the random forest model interprets apartment scale through a combination of interacting and nonlinear factors, whereas the linear regression model prioritizes simpler, more direct relationships.

In the linear regression model, "Construction year" was ranked higher in importance compared to the random forest model. This difference can be attributed to the distinct characteristics of the two modeling approaches. "Construction year" often has a linear or near-linear relationship with apartment prices, as newer buildings are generally associated with higher market values due to better infrastructure, design, and amenities. The linear regression model captures this straightforward relationship effectively, leading to a higher ranking for this feature. In contrast, the random forest model relies on decision-tree-based partitions to capture nonlinear relationships and interactions among features. While "Construction year" may still influence apartment prices in this framework, its impact might be partially distributed across interactions with other variables, such as "Building coverage ratio (BCR)", "Floor area ratio (FAR)" or "Number of units in the complex". As a result, its individual importance is diluted, leading to a lower ranking in the random forest model. Moreover, the random forest model may prioritize features that have clear thresholds or significant nonlinearity, which "Construction year" does not exhibit as strongly.

Interestingly, in the linear regression model, "Tallest building height" was ranked higher in importance compared to "Shortest building height". Conversely, in the random forest model, the ranking was reversed, with "Shortest building height" being more important. The linear regression model prioritizes "Tallest building height" due to its more direct and linear relationship with apartment prices, as taller buildings often signify premium developments. On the other hand, the random forest model might treat “Shortest building height” as a threshold-based feature that helps partition the data into distinct groups (e.g., complexes with low-rise vs. high-rise buildings). This thresholding capability allows the random forest model to assign greater importance to features that aid in making splits, even if their direct correlation with the target variable is weaker. Ultimately, this suggests that the distinction between high-priced and low-priced apartment complexes is more strongly influenced by the floor level of the lowest-height buildings within a complex, rather than by the floor level of the tallest building.

The feature “Heating system” exhibits a notable difference in importance rankings between the linear regression and random forest models. In the linear regression model, “Heating system” is ranked 11th, whereas it is ranked much lower in the random forest model—20th in SHAP and 25th (the lowest rank) in PFI. In the linear regression model, “Heating system” is treated as a categorical variable, represented through dummy variables. These dummy variables allow the model to capture distinct effects associated with each heating type. Since the linear regression model explicitly evaluates each feature’s contribution independently, “Heating system” retains a relatively higher ranking as a direct explanatory variable. In contrast, the random forest model assigns significantly lower importance to “Heating system”. This can be explained by the model’s ability to capture complex interactions and nonlinear relationships among features. Several factors contribute to this outcome. First, features such as “Number of buildings” or “Number of units in the complex” may already encapsulate aspects of the apartment complex’s overall structure, reducing the additional explanatory power of “Heating system”. Since the random forest model can identify and prioritize features that capture broader patterns, it is likely that the influence of “Heating system” is effectively absorbed by these more comprehensive variables, particularly those representing the scale and characteristics of the apartment complex. Second, as a categorical variable, “Heating system” might contain specific subcategories that are indirectly linked to apartment price variations. The random forest model, with its decision-tree-based structure, is adept at leveraging detailed interactions between variables and identifying when a feature is redundant or when its effect is sufficiently captured by related features. If certain subcategories of “Heating system” are associated with characteristics already well-represented by other features, the model may assign it a lower importance. Third, “Heating system” may influence apartment prices indirectly, through interactions with variables that reflect the broader features of the complex or its location. For instance, the presence of a particular heating system could be indicative of specific construction standards or regional preferences, which are captured through more directly relevant features in the random forest model. This indirect relationship could lead to “Heating system” having a reduced standalone impact, as the model relies on other features to represent its underlying effect on apartment prices. While both models identify GDP as a key macroeconomic variable reflecting the overall economic conditions at the time of the transaction, they diverge in how they prioritize other macroeconomic features. In the linear regression model, “Transaction year and quarter” ranks higher, likely because it directly captures temporal trends, serving as a proxy for time-sensitive economic fluctuations. In contrast, the random forest model assigns greater importance to “Mortgage interest rate”, which can significantly impact housing affordability and market dynamics. The random forest’s ability to capture nonlinear relationships and threshold effects likely accounts for this difference, as changes in mortgage interest rates interact with other variables in ways that influence apartment prices beyond what a simple temporal trend can capture.

Variables related to proximity to social infrastructure, such as “Distance to the nearest subway station” and “Distance to the nearest general hospital”, were consistently ranked among the least important features in both models. This suggests that when reflecting the locational characteristics of apartments, both models consider variables like “Dong” to be sufficient for capturing the influence of location. The relatively low importance of proximity variables may also be attributed to the fact that areas such as Gangnam are generally well-equipped with convenient transportation and essential living facilities. As a result, direct distance measures fail to provide significant differentiation in value, further diminishing their importance in explaining apartment prices.

These findings underscore the fundamental differences between the two modeling approaches. Linear regression provides a straightforward and interpretable framework, well-suited for capturing strong linear correlations but limited in its ability to model complex, nonlinear interactions. In contrast, the random forest model excels at identifying nonlinear relationships and interactions, leveraging its decision-tree-based structure to evaluate feature importance more comprehensively. This distinction highlights the importance of selecting a modeling approach that aligns with the data’s underlying structure and the specific objectives of the analysis. In Figure 2, we show the interpretation of two test set samples’ predictions: we use the samples with the best accuracy for the property price. The red and blue arrows illustrate the impact of individual variables on the predicted outcome, with red indicating positive contributions and blue indicating negative contributions. The intersection point of these arrows represents the final predicted value for a given observation. The y-axis also includes a grey marker denoting the overall mean of the dependent variable across all observations, which serves as the baseline prediction. The sequential shifts from this baseline to the final predicted value reflect the cumulative effects of the individual variables, either increasing or decreasing the prediction accordingly. This result indicates that the influence of each factor on individual apartment prices aligns with the theoretical frameworks.

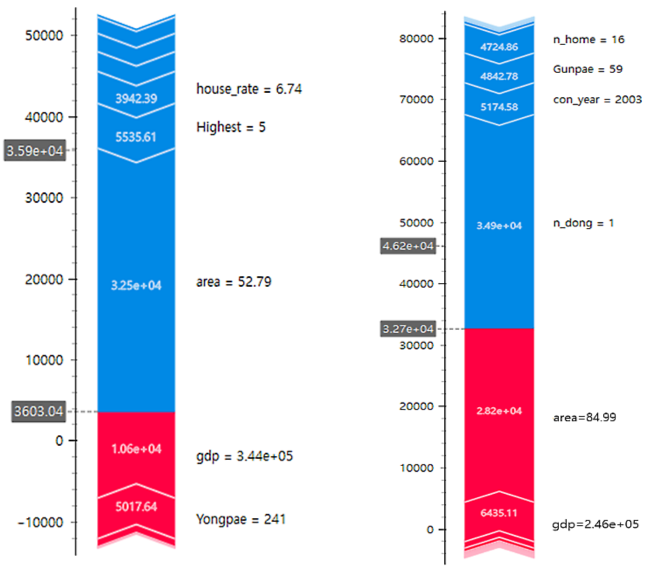


Figure 2. Two predictions interpretations (Left: Linear regression, Right: Random forest)

5.2. Temporal analysis of feature importance

In this section, we analyze the SHAP and PFI values of selected features over time to examine their temporal variations in predicting apartment prices. By applying both linear regression and random forest models, we track how feature importance evolves over time, allowing us to identify trends in key determinants of apartment prices. This temporal analysis helps uncover whether certain factors consistently influence prices or if their impact shifts over different periods. Additionally, by comparing SHAP and PFI values across time windows, we can assess whether model interpretability remains stable or if different features become more dominant in different market conditions. We selected variables that exhibited relatively clear trends in importance across different time periods, including area, number of parking spaces per unit, building age, number of housing units, number of buildings, prestigious apartment brand, distance to the nearest general hospital, distance to the nearest national park, distance to the nearest university, and distance to the nearest subway station. These features were categorized according to the classifications presented in Table 1.

To account for temporal changes, we applied a rolling time window approach, similar to a moving average, where feature importance was evaluated using overlapping three-year periods. Specifically, we computed SHAP and PFI values for 2007–2009, then shifted the window forward by one year to compute values for 2008–2010, and so on. This method ensures a smoother assessment of temporal trends by capturing gradual changes in feature importance while maintaining sufficient data for model training and evaluation in each period. Since the data was segmented into distinct time windows, macroeconomic variables such as GDP, mortgage interest rates, and economic growth rates lost their relevance as they remained fixed within each interval. Consequently, these variables were excluded from the model to ensure a more meaningful analysis of time-dependent feature importance.

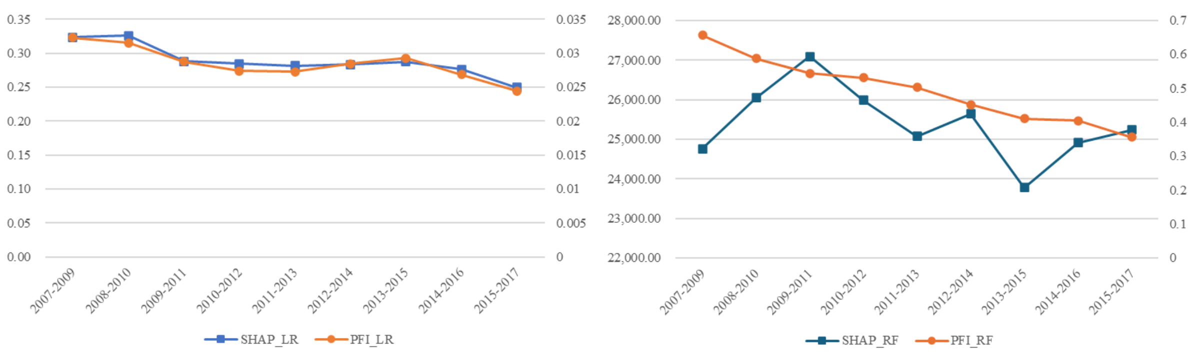
5.2.1. Structural features

Structural characteristics, particularly area and construction year, exhibited notable temporal variations in their importance for predicting apartment prices, as indicated by SHAP and PFI values from both the linear regression and random forest models. These trends provide insights into shifting housing preferences and broader demographic changes over time.

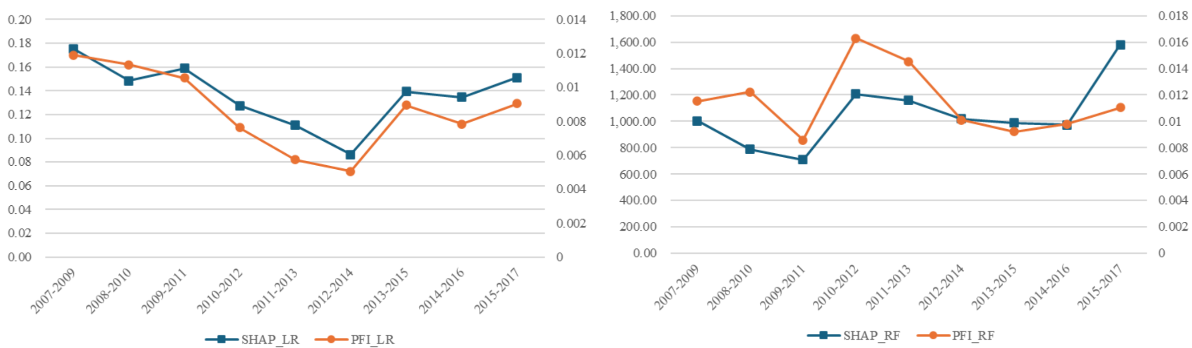
First, the importance of area in explaining apartment prices has shown a gradual decline over time, as indicated by both SHAP and PFI values (see Figure 3a). The decreasing PFI values suggest that, over time, variations in area have had a diminishing impact on predictive accuracy, meaning that other factors have become more influential in determining apartment prices. This implies that while area was once a dominant factor in price formation, its role has weakened as buyers increasingly consider other characteristics such as location, amenities, and neighborhood features. Interestingly, the SHAP values exhibit different patterns depending on the model. In the linear regression model, SHAP values show a steady decline over time, indicating a consistent reduction in the contribution of area to price predictions. This suggests that the role of area in explaining apartment prices has gradually diminished in a linear fashion. In contrast, the random forest model shows slight fluctuations in SHAP values, reflecting the model's ability to capture nonlinear interactions between area and other features. These fluctuations may be attributed to changes in market conditions, shifting buyer preferences, or interactions with emerging influential factors in different time periods. Several demographic and socioeconomic factors may underlie this trend. The ongoing decline in birth rates, the aging population, and the transition of the baby boomer generation (born approximately between 1946 and 1964) into retirement have led to a reduction in average household size. As family structures evolve, the demand for large family-oriented apartments decreases, while smaller, more manageable housing units become increasingly preferred. Additionally, lifestyle changes, such as the rising number of single-person and dual-income households, contribute to a preference for compact, efficient living spaces rather than expansive homes that require greater maintenance. The decreasing importance of area in price determination may reflect a broader redefinition of housing value, where structural size is no longer the primary determinant of apartment prices, and other factors such as location, accessibility, and housing quality take precedence.

Second, while construction year (building age) might intuitively suggest a negative impact on apartment prices due to physical depreciation, its relationship with price appears to be more complex, particularly in markets where redevelopment potential plays a role in property valuation. The results indicate that the importance of construction year has varied over time, potentially reflecting shifts in market conditions and redevelopment expectations (see Figure 3b left). In the linear regression model, SHAP and PFI values exhibit a gradual decline until 2012–2014, followed by an increase. This trend may suggest that construction year became relatively less influential in explaining apartment prices during this period, possibly due to weakened redevelopment expectations. Additionally, the declining importance of construction year up to 2012–2014 could partly reflect the natural depreciation of older buildings, as property values tend to decrease over time due to physical aging. This suggests that during this period, the effect of building age may have been driven more by depreciation rather than redevelopment potential. However, after 2014, SHAP and PFI values began to rise again, which could indicate a shift in market conditions where redevelopment expectations became more prominent. This pattern implies that while the declining importance of construction year in 2012–2014 may have been partially influenced by the aging effect, the subsequent increase suggests that broader market factors, such as renewed investor interest in redevelopment, played a key role in shaping property values. In contrast, the random forest model exhibits greater fluctuation in SHAP values over time, suggesting that the influence of construction year may be more context-dependent and influenced by nonlinear interactions with other variables. Unlike the smoother trend observed in the linear regression model, SHAP and PFI values in the random forest model fluctuate over time without a clear long-term trajectory. This variability may indicate that construction year's importance is affected by interactions with factors such as zoning regulations, land values, and market-driven redevelopment incentives, which may change dynamically. Rather than reflecting a gradual increase or decrease, the random forest model appears to capture short-term variations, where construction year becomes particularly relevant in specific periods, possibly due to policy adjustments, investment patterns, or shifts in redevelopment activity. However, the extent to which these fluctuations correspond to external economic or policy changes require further investigation.

For instance, during strong housing markets, older apartments—especially those approaching eligibility for redevelopment—tend to increase in value, as redevelopment opportunities introduce speculative price premiums. Conversely, during market downturns, the importance of construction year declines, as the probability of redevelopment projects decreases. Between 2007 and 2009, the importance of construction year in explaining apartment prices was relatively high, reflecting strong redevelopment expectations. However, following the global financial crisis, this importance diminished, corresponding with a period of stagnation in the housing market. After 2013, the importance of construction year began to rise again, likely due to the market recovery and renewed speculation regarding redevelopment potential.



a) Area



b) Construction year

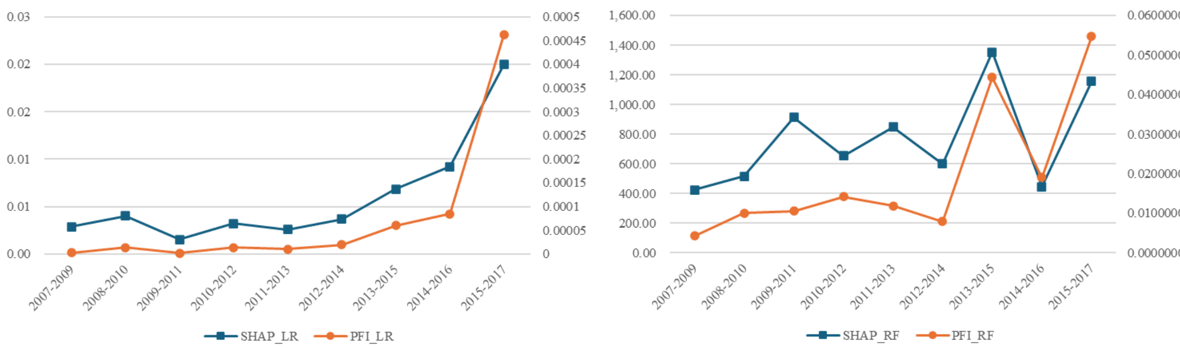
Figure 3. Evolution in structural characteristic indicators over time (Left axis: SHAP, Right axis: PFI)

This distinction underscores the complementary nature of the two modeling approaches. Linear regression provides insight into long-term structural trends, such as gradual shifts in redevelopment expectations over extended market cycles. In contrast, random forest captures short-term, localized variations that may reflect the impact of policy changes, investment cycles, and regional redevelopment initiatives. While the smoother trend in the linear regression model suggests that redevelopment expectations may be linked to macroeconomic cycles, the fluctuations observed in the random forest model imply that construction year's importance may be more sensitive to market-specific conditions and policy interventions. Further research would be needed to more precisely determine the drivers of these variations over time.

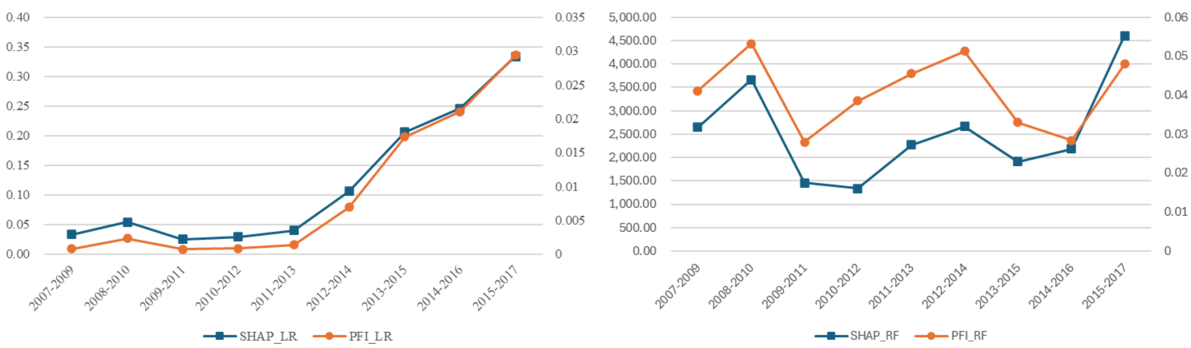
5.2.2. Neighbourhood features

Neighbourhood features represent shared characteristics of apartment complexes, distinguishing them from structural features, which describe an individual unit's inherent features. In other words, neighbourhood features include factors such as the scale of the apartment complex, shared facilities, and overall living environment, all of which influence apartment prices collectively rather than on a unit-by-unit basis. The temporal variations in the importance of these features provide insights into how housing demand has evolved in response to broader urban development trends and changes in buyer preferences.

The importance of parking spaces in explaining apartment prices has exhibited a gradual upward trend over time (see Figure 4a). This reflects the growing demand for parking availability in residential complexes, a trend that aligns with rising vehicle ownership rates. According to the Seoul Vehicle Registration Statistics (Seoul Metropolitan Government, 2025), the vehicle registration rate in Gangnam-gu, Seoul, was 42.02% relative to the population in 2007, with an average of 0.92 vehicles per household. By 2017, the vehicle registration rate had risen to 42.63%, with an average of 1.15 vehicles per household. This increase in vehicle ownership suggests a corresponding rise in demand for parking infrastructure, making parking availability an increasingly important determinant of apartment prices. Beyond the increase in vehicle ownership, several regulatory and market-driven factors may have contributed to the rising importance of parking spaces. In recent years, regulations on parking space allocation in apartment complexes have been strengthened. According to the 2012 revision of the "Regulation on Housing Construction Standards", the required number of parking spaces per household in newly built apartments was increased. In metropolitan areas, the minimum parking requirement per household was adjusted from approximately 0.7–0.8 spaces to 1.0–1.5 spaces, ensuring that new apartment complexes provide more parking facilities than older ones. Furthermore, major construction firms, such as Raemian (Samsung C&T) and Xi (GS Construction), have introduced larger parking areas and advanced parking facilities in



a) Parking lots



b) Number of units in the complex

Figure 4. *To be continued*

c) Number of buildings

d) Apartment brand

Figure 4. Evolution in neighborhood characteristic indicators over time (Left axis: SHAP, Right axis: PFI)

newly developed complexes. These additions reflect growing consumer demand for better parking infrastructure, which is not only influenced by vehicle ownership but also by the integration of modern amenities such as electric vehicle (EV) charging stations, enhanced security features, and smart parking management systems**.** The increasing presence of such features suggests that parking availability is evolving beyond a basic necessity to become a premium residential feature that contributes to property valuation

The number of units and the number of buildings in an apartment complex both represent the overall scale of the development. However, their importance in determining apartment prices has shown contrasting trends over time. While the number of housing units has exhibited a gradual upward trend in importance (see Figure 4b left), the significance of the number of buildings has steadily declined (see Figure 4c). This divergence likely reflects the increasing prevalence of high-rise apartment complexes in urban areas, particularly in Gangnam-gu, Seoul, where vertical expansion has become the dominant development strategy. Larger apartment complexes tend to command higher prices, as they typically offer extensive shared facilities, well-developed infrastructure, and a more comprehensive range of amenities, making them more attractive to buyers. Additionally, these large-scale developments are often strategically located in prime areas, further reinforcing their desirability. The growing importance of the number of housing units suggests that buyers are placing greater emphasis on the overall residential environment, prioritizing large, well-equipped complexes over smaller developments. In contrast, the declining significance of the number of buildings in explaining apartment prices may be attributed to changes in apartment design and construction practices. In recent years, there has been a shift toward the development of taller apartment buildings, accommodating a larger number of housing units within fewer structures. As a result, newer high-rise apartment complexes tend to have fewer buildings compared to older, low-rise developments with a similar number of units. This shift implies that the traditional role of the number of buildings as an indicator of apartment complex size has diminished, as the total number of housing units has become a more relevant factor in assessing apartment prices. This reflects an evolving preference for efficient land use and high-density residential developments in urban areas.

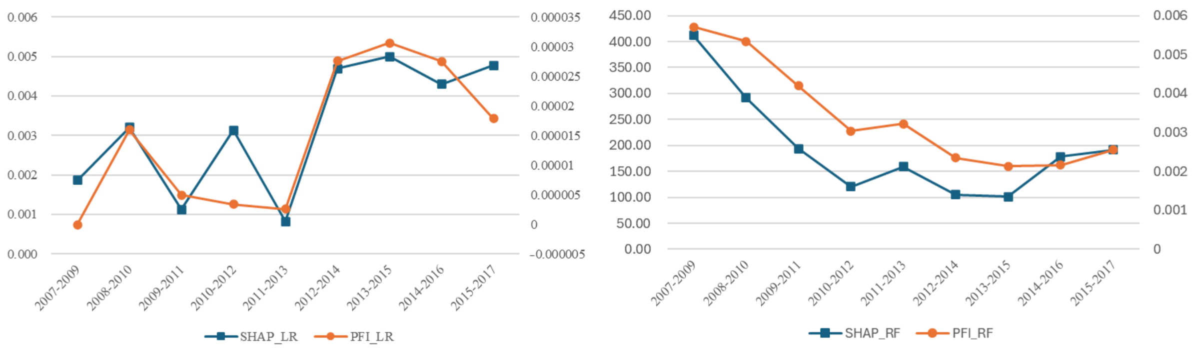
The importance of well-known apartment brands in explaining apartment prices has also exhibited an upward trend in the linear regression model (see Figure 4d left). This can be attributed to the premium pricing associated with large-scale, high-rise apartment complexes developed by major construction firms. The increasing significance of brand reputation suggests that buyers perceive branded apartments as higher in quality, reliability, and long-term investment value. Furthermore, if the prices of branded apartments were similar to those of existing non-branded apartments, the brand's impact on pricing would have been negligible. However, the observed upward trend in the linear regression model indicates that the premium associated with branded apartments has steadily increased over time, reinforcing the growing role of brand perception in price formation. In contrast, the random forest model exhibits fluctuations in the importance of brand features over time, suggesting that the influence of brand reputation may be more context-dependent. Unlike the linear regression model, which captures long-term trends, the random forest model reflects short-term variations, where brand importance may fluctuate based on market conditions, supply dynamics, and interactions with other property features. Notably, the random forest model shows a peak in PFI and SHAP values around 2012–2014, which aligns with the period when several large-scale apartment complexes developed by major construction firms were newly occupied. The increase in brand importance during this period suggests that the introduction of high-profile, branded apartment complexes had a substantial impact on price formation, temporarily amplifying the influence of brand reputation. This variability implies that while brand reputation remains a key factor in apartment pricing, its impact may be amplified in certain market conditions—such as during periods of high brand-new apartment supply or strong market sentiment—while being relatively less significant in other periods. The difference in results between the two models highlights the need to consider both long-term brand value appreciation and short-term market-driven fluctuations when analyzing the role of branding in apartment price formation.

Also, the evolution of neighbourhood feature importance from 2006 to 2017, as illustrated in Figures 4, closely mirrors the dynamic shifts in South Korea's real estate market during this period. As the market moved from heavy regulation in the mid-2000s to aggressive deregulation post-financial crisis, and back to tighter control after 2017, buyers' valuation criteria for apartment complexes also evolved. Physical characteristics of apartment complexes—such as Parking lots, Number of units in the complex, and Apartment brand—became increasingly influential, particularly after 2013, reflecting a policy environment that favoured large-scale, branded, and redevelopment-driven housing. These results underscore how policy direction, urban redevelopment initiatives, and macroeconomic conditions (e.g., prolonged low-interest rate situation; low interest rates can facilitate the inflow of investment capital into the apartment market) reshaped the salience of neighborhood attributes in apartment price formation.

5.2.3. Locational features

As mentioned in the previous section, features representing the distance to major facilities generally do not have a significant impact on apartment prices. However, for some proximity variables, trends in SHAP and PFI values emerge over time, indicating that their influence is not entirely negligible.

Notably, in almost all explanatory variables except for proximity variables, the variables based on linear regression models exhibited clearer trends compared to those based on random forest models. This suggests that most housing and neighborhood characteristic variables, excluding proximity, can explain apartment prices in a linear fashion. In other words, since these variables have a linear

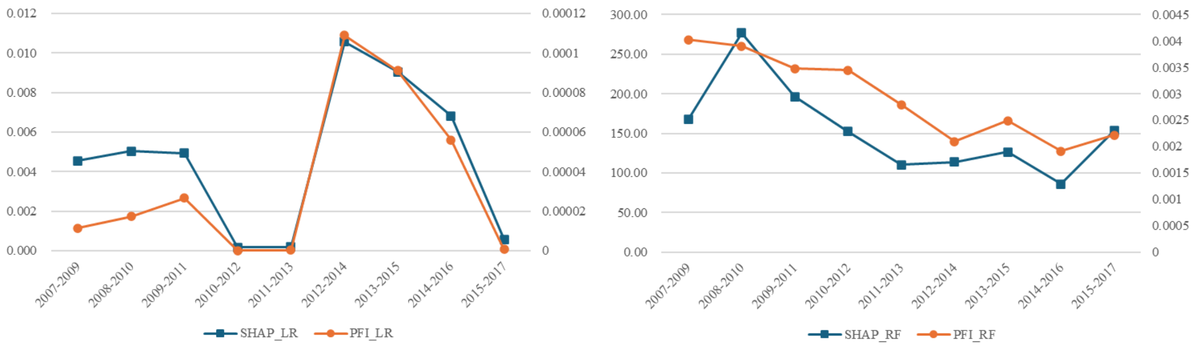


a) Distance to the nearest general hospital

b) Distance to the nearest national park

Figure 5. *To be continued*

c) Distance to the nearest university



d) Distance to the nearest subway station

Figure 5. Evolution in locational characteristic indicators over time (Left axis: SHAP, Right axis: PFI)

relationship with apartment prices at different time points, using a linear model to calculate the importance of the variables results in more distinct patterns. In fact, factors such as apartment area, parking spaces per unit, building age, and, among neighborhood characteristics, prestigious apartment brand and the number of housing units, are known to have a linear relationship with apartment prices (Moreira de Aguiar et al., 2014; Wittowsky et al., 2020; Hong et al., 2020; Hong & Kim, 2022).

In contrast, as shown in Figure 5, proximity variables did not show a clear trend in the linear regression model-based explanatory indicators but demonstrated a relatively clearer trend in the random forest-based explanatory indicators. This suggests that using a linear model does not effectively predict housing prices based on locational features. From the perspective of apartment suppliers, there is an incentive to set higher prices for properties with better access to surrounding social infrastructure. However, from the perspective of buyers, the process of increasing willingness to pay based on distances is still not well understood compared to structural and neighborhood characteristics. The fluctuating importance of locational variables over time suggests that their impact on apartment prices may be influenced by external factors such as urban development, transportation improvements, and changing buyer preferences, which are better captured by non-linear models like random forest.

In addition, the declining importance of locational proximity factors, such as the distance to the nearest general hospital, national park, university, and subway station suggests a shift in consumer preference from traditional accessibility metrics to internal quality and infrastructure within the housing complex. This suggests that shifts in the apartment market and buyer preferences have a significant impact on apartment prices.

6. Conclusions

This study evaluates and compares the interpretability and predictive performance of linear and non-linear models using XAI techniques in the context of real estate mass appraisal. By employing multiple linear regression (MLR) and random forest (RF), we systematically analyzed how these models represent feature importance and how their predictive mechanisms differ. Utilizing SHAP and PFI, we identified key features that influence apartment prices and examined how the ranking of these features varies depending on the modeling approach.

The comparison between the linear regression and random forest models reveals both similarities and differences in evaluating feature importance for predicting apartment prices. Both models identified key features, such as area, number of units in the complex, dong (administrative division), GDP, and construction year, as highly significant, indicating their strong intrinsic relationships with apartment prices. However, notable differences were observed in feature rankings due to the distinct characteristics of each model. The linear regression model assigned higher importance to the number of units in the complex, while the random forest model prioritized number of buildings and building coverage ratio (BCR). This suggests that the linear model captures straightforward relationships, whereas the random forest model accounts for complex interactions. Similarly, construction year was ranked higher in the linear model, likely due to its direct correlation with apartment prices, while its importance was lower in the random forest model, where it interacted with other variables. Additionally, the two models differed in their treatment of tallest and shortest building height, with the linear model emphasizing the tallest building height and the random forest model assigning greater importance to shortest building height, suggesting its role as a threshold-based feature in distinguishing high- and low-priced apartment complexes. Lastly, proximity variables, such as distances to subway stations or hospitals, ranked among the least important features in both models, indicating that location-related features were sufficiently captured by other variables, such as administrative division (dong).

Beyond the general feature importance rankings, this study also incorporated a temporal analysis by segmenting the data into discrete time intervals to investigate the dynamic evolution of feature importance over time. Among structural features, area and construction year exhibited notable variations, as indicated by SHAP and PFI values in both linear regression and random forest models. The declining importance of area suggests a shift in housing preferences, where buyers increasingly prioritize factors beyond size, such as amenities and location. While the linear model showed a steady decline in SHAP values, the random forest model captured fluctuations, likely reflecting nonlinear interactions with evolving market conditions. Similarly, construction year demonstrated a decline in importance until 2012–2014, possibly due to aging effects reducing redevelopment expectations, followed by a resurgence as market recovery renewed investor interest in redevelopment. The random forest model showed greater variability in construction year's importance, indicating that its influence is more sensitive to policy changes, investment cycles, and localized redevelopment dynamics. Neighborhood features, such as the number of housing units, gained importance over time, aligning with the increasing preference for large-scale apartment complexes, while the significance of the number of buildings declined as high-rise developments became more common. Additionally, the rising importance of branded apartments in the linear model reflects the growing market premium associated with well-known developers, while fluctuations in the random forest model suggest that branding effects are more context-dependent, peaking during periods of high-profile apartment completions. These findings highlight the distinct advantages of linear and nonlinear models, with the former capturing broad market trends and the latter identifying short-term variations shaped by external factors.

From a practical standpoint, these findings offer meaningful implications for real estate professionals, policymakers, and urban planners. The integration of XAI techniques enhances the transparency of mass appraisal systems, allowing stakeholders to better understand the drivers of property valuations. By distinguishing between variables that exhibit stable importance and those that fluctuate based on market conditions, this study provides valuable insights for improving model selection and appraisal methodologies. Additionally, the results suggest that while linear models offer clearer interpretability, non-linear models like random forest can capture hidden interactions and temporal variations that traditional approaches may overlook. SHAP values provide individualized, case-specific explanations for each prediction by quantifying the marginal contribution of every feature. In practice, a policymaker reviewing property tax assessments can use SHAP explanations to trace why two similar units may have significantly different valuations—e.g., due to differences in brand reputation, building scale, or available parking lots—which fosters procedural transparency and supports citizen accountability. Furthermore, PFI and SHAP-based global feature importance analysis highlights market trends that are often invisible in traditional regression models. The temporal segmentation in this study revealed a consistent increase in the influence of physical complex characteristics (e.g., number of units, brand, parking space) and a declining influence of proximity-based features (e.g., distance to subway or university). This finding has significant implications: real estate professionals can leverage such patterns to recalibrate investment strategies or development priorities based on evolving buyer preferences, while urban planners can assess which infrastructural factors are becoming influential in shaping housing demand. Within mortgage evaluations frameworks, Explainable Artificial Intelligence (XAI) methodologies, particularly SHAP, facilitate comprehension of the determinant factors underlying property valuation mechanisms that inform loan-to-value (LTV) ratio calculations. For example, when algorithmic models attribute elevated valuations predominantly to brand reputation coefficients and parking infrastructure availability, lending officers can critically evaluate whether such variables constitute reliable indicators of sustained collateral stability. Moreover, this interpretability enables financial institutions to articulate the rationale supporting automated lending determinations to both regulatory authorities and clientele, thereby enhancing confidence and regulatory adherence in automated appraisal protocols. In taxation, XAI tools help ensure fair and consistent property assessments by revealing which factors drive valuation differences. SHAP explanations enhance transparency, allowing taxpayers to understand and challenge their assessments, thereby strengthening trust and equity in tax systems.

Despite these contributions, the study has certain limitations. The analysis was conducted within the Korean real estate market, and further research is needed to assess the generalizability of the findings to other housing markets with different structural and economic conditions. Especially, the study may be subject to potential sampling bias, as the data is confined to the Gangnam District, a high-value urban area. In future research, we aim to enhance the generalizability of the study by incorporating data from a more diverse range of regions. Also, while SHAP and PFI provide valuable insights into feature importance, they do not fully capture causal relationships between explanatory variables and price changes. Future studies could expand on this work by integrating causal inference techniques or applying XAI-driven methods to broader datasets, including commercial and mixed-use properties. Lastly, the static nature of SHAP explanations, which rely on a fixed model structure, may limit their effectiveness in capturing temporal dynamics in housing price determinants. Future research could address this limitation by employing dynamic modeling approaches (e.g., Recurrent Neural Networks, Bayesian Dynamic Models) or time-aware XAI techniques to better reflect evolving market conditions.

Author contributions

W. Kim and M. Lee conceived the study and were responsible for the design and development of the data analysis. W. Kim was responsible for data collection and M. Lee and W. Kim were responsible for data analysis and interpretation. I. Choi was responsible for setting the direction of data analysis and conducting theoretical research.

Disclosure statement

There are no conflicts of interest.

References

Adair, A., McGreal, S., Smyth, A., Cooper, J., & Ryley, T. (2000). House prices and accessibility: The testing of relationships within the Belfast urban area. *Housing Studies*, *15*(5), 699–716. https://doi.org/10.1080/02673030050134565

Al-Yahyaee, K. H., Mensi, W., Ko, H. U., Caporin, M., & Kang, S. H. (2021). Is the Korean housing market following Gangnam style? *Empirical Economics*, *61*, 2041–2072. https://doi.org/10.1007/s00181-020-01931-2

Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of random forest for valuation and a CART-based approach for model diagnostics. *Expert Systems with Applications*, *39*(2), 1772–1778. https://doi.org/10.1016/j.eswa.2011.08.077

Bae, Y., & Joo, Y. M. (2020). The making of Gangnam: Social construction and identity of urban place in South Korea. *Urban Affairs Review*, *56*(3), 726–757. https://doi.org/10.1177/1078087419827645

Benson, E. D., Hansen, J. L., Schwartz, A. L., & Smersh, G. T. (1998). Pricing residential amenities: The value of a view. *The Journal of Real Estate Finance and Economics*, *16*, 55–73. https://doi.org/10.1023/A:1007785315925

Bidanset, P. E., & Rakow, R. (2022). Survey on the use of automated valuation models (AVMs) in government assessment offices: An analysis of AVM use, acceptance, and barriers to more widespread implementation. *Journal of Property Tax Assessment & Administration*, *19*(2). https://doi.org/10.63642/1357-1419.1250

Bilgilioğlu, S. S., & Yılmaz, H. M. (2023). Comparison of different machine learning models for mass appraisal of real estate. *Survey Review*, *55*(388), 32–43. https://doi.org/10.1080/00396265.2021.1996799

Binoy, B. V., Naseer, M. A., Kumar, P. A., & Lazar, N. (2022). A bibliometric analysis of property valuation research. *International Journal of Housing Markets and Analysis*, *15*(1), 35–54. https://doi.org/10.1108/IJHMA-09-2020-0115

Breiman, L. (2001). Random forests. *Machine Learning*, *45*, 5–32. https://doi.org/10.1023/A:1010933404324

Carroll, T. M., Clauretie, T. M., & Jensen, J. (1996). Living next to godliness: Residential property values and churches. *The Journal of Real Estate Finance and Economics*, *12*, 319–330. https://doi.org/10.1007/BF00127540

Chau, K. W., & Chin, T. L. (2003). A critical review of literature on the hedonic price model. *International Journal for Housing Science and its Applications*, *27*(2), 145–165.

Chen, C., Ma, X., & Zhang, X. (2024). Empirical study on real estate mass appraisal based on dynamic neural networks. *Buildings*, *14*(7), Article 2199. https://doi.org/10.3390/buildings14072199

Chen, L., Yao, X., Liu, Y., Zhu, Y., Chen, W., Zhao, X., & Chi, T. (2020). Measuring impacts of urban environmental elements on housing prices based on multisource data—a case study of Shanghai, China. *ISPRS International Journal of Geo-Information*, *9*(2), Article 106. https://doi.org/10.3390/ijgi9020106

Clark, D. E., & Herrin, W. E. (2000). The impact of public school attributes on home sale prices in California. *Growth and Change*, *31*(3), 385–407. https://doi.org/10.1111/0017-4815.00134

Dimopoulos, T., & Moulas, A. (2016). A proposal of a mass appraisal system in Greece with CAMA system: Evaluating GWR and MRA techniques in Thessaloniki Municipality. *Open Geosciences*, *8*(1), 675–693. https://doi.org/10.1515/geo-2016-0064

Duan, J., Tian, G., Yang, L., & Zhou, T. (2021). Addressing the macroeconomic and hedonic determinants of housing prices in Beijing metropolitan area, China. *Habitat International*, *113*, Article 102374. https://doi.org/10.1016/j.habitatint.2021.102374

Espey, M., & Lopez, H. (2000). The impact of airport noise and proximity on residential property values. *Growth and Change*, *31*(3), 408–419. https://doi.org/10.1111/0017-4815.00135

Fan, G. Z., Ong, S. E., & Koh, H. C. (2006). Determinants of house price: A decision tree approach. *Urban Studies*, *43*(12), 2301–2315. https://doi.org/10.1080/00420980600990928

Fletcher, M., Gallimore, P., & Mangan, J. (2000). Heteroscedasticity in hedonic house price model. *Journal of Property Research*, *17*(2), 93–108. https://doi.org/10.1080/095999100367930

Follain, J. R., & Jimenez, E. (1985). Estimating the demand for housing characteristics: A survey and critique. *Regional Science and Urban Economics*, *15*(1), 77–107. https://doi.org/10.1016/0166-0462(85)90033-X

Forrest, D., Glen, J., & Ward, R. (1996). The impact of a light rail system on the structure of house prices: A hedonic longitudinal study. *Journal of Transport Economics and Policy*, *30*, 15–29.

Gabrielli, L., & French, N. (2021). Pricing to market: Property valuation methods – a practical review. *Journal of Property Investment & Finance*, *39*(5), 464–480. https://doi.org/10.1108/JPIF-09-2020-0101

Garrod, G. D., & Willis, K. G. (1992). Valuing goods' characteristics: An application of the hedonic price method to environmental features. *Journal of Environmental Management*, *34*(1), 59–76. https://doi.org/10.1016/S0301-4797(05)80110-0

Glumac, B., & Des Rosiers, F. (2021). Practice briefing – Automated valuation models (AVMs): Their role, their advantages and their limitations. *Journal of Property Investment and Finance*, *39*(5), 481–491. https://doi.org/10.1108/JPIF-07-2020-0086

Hong, J., & Kim, W.-s. (2022). Combination of machine learning-based automatic valuation models for residential properties in South Korea. *International Journal of Strategic Property Management*, *26*(5), 362–384. https://doi.org/10.3846/ijspm.2022.17909

Hong, J., Choi, H., & Kim, W.-s. (2020). A house price valuation based on the random forest approach: The mass appraisal of residential property in South Korea. *International Journal of Strategic Property Management*, *24*(3), 140–152. https://doi.org/10.3846/ijspm.2020.11544

Iban, M. C. (2022). An explainable model for the mass appraisal of residences: The application of tree-based machine learning algorithms and interpretation of value determinants. *Habitat International*, *128*, Article 102660. https://doi.org/10.1016/j.habitatint.2022.102660

International Association of Assessing Officers. (2017). *Standard on mass appraisal of real property*. Kansas City, Missouri, USA.

Kok, N., Koponen, E. L., & Martínez-Barbosa, C. A. (2017). Big data in real estate? From manual appraisal to automated valuation. *Journal of Portfolio Management*, *43*(6), 202–211. https://doi.org/10.3905/jpm.2017.43.6.202

Kontrimas, V., & Verikas, A. (2011). The mass appraisal of the real estate by computational intelligence. *Applied Soft Computing*, *11*(1), 443–448. https://doi.org/10.1016/j.asoc.2009.12.003

Krämer, B., Nagl, C., Stang, M., & Schäfers, W. (2023). Explainable AI in a real estate context–exploring the determinants of residential real estate values. *Journal of Housing Research*, *32*(2), 204–245. https://doi.org/10.1080/10527001.2023.2170769

Lancaster, K. J. (1966). A new approach to consumer theory. *Journal of Political Economy*, *74*, 132–157. https://doi.org/10.1086/259131

Lenaers, I., Boudt, K., & De Moor, L. (2024). Predictability of Belgian residential real estate rents using tree-based ML models and IML techniques. *International Journal of Housing Markets and Analysis*, *17*(1), 96–113. https://doi.org/10.1108/ijhma-11-2022-0172

Li, M. M., & Brown, H. J. (1980). Micro-neighborhood externalities and hedonic housing prices. *Land Economics*, *56*(2), 125–141. https://doi.org/10.2307/3145857

Lundberg, S. M., Erion, G. G., & Lee, S. I. (2018). *Consistent individualized feature attribution for tree ensembles*. arXiv. https://doi.org/10.48550/arXiv.1802.03888

Lundberg, S. M., Erion, G., Chen, H., DeGrave, A., Prutkin, J. M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., & Lee, S. I. (2019). *Explainable AI for trees: From local explanations to global understanding*. arXiv. https://doi.org/10.48550/arXiv.1905.04610

Lundberg, S., & Lee, S. I. (2017). *A unified approach to interpreting model predictions*. arXiv. https://doi.org/10.48550/arXiv.1705.07874

Maeil Business Newpaper. (2024, November 18). *KB real estate average sale price survey Seocho-gu ranked 2nd with 86.75 million won. Lowest place, Dobong-gu, KRW 26.69 million.* https://www.mk.co.kr/en/realestate/11170899

Malpezzi, S. (2003). Hedonic pricing models: A selective and applied review. *Housing Economics and Public Policy*, *1*, 67–89. https://doi.org/10.1002/9780470690680.ch5

McCluskey, W. J., Deddis, W. G., Lamont, I. G., & Borst, R. A. (2000). The application of surface generated interpolation models for the prediction of residential property values. *Journal of Property Investment and Finance*, *18*(2), 162–176. https://doi.org/10.1108/14635780010324321

McCluskey, W., Davis, P., Haran, M., McCord, M., & McIlhatton, D. (2012). The potential of artificial neural networks in mass appraisal: The case revisited. *Journal of Financial Management of Property and Construction*, *17*(3), 274–292. https://doi.org/10.1108/13664381211274371

McMillan, D., Jarmin, R., & Thorsnes, P. (1992). Selection bias and land development in the monocentric model. *Journal of Urban Economics*, *31*, 273–284. https://doi.org/10.1016/0094-1190(92)90056-Q

Miller, N., Peng, L., & Sklarz, M. (2011). House prices and economic growth. *The Journal of Real Estate Finance and Economics*, *42*(4), 522–541. https://doi.org/10.1007/s11146-009-9197-8

Mok, H. M., Chan, P. P., & Cho, Y. S. (1995). A hedonic price model for private properties in Hong Kong. *The Journal of Real Estate Finance and Economics*, *10*, 37–48. https://doi.org/10.1007/BF01099610

Moreira de Aguiar, M., Simões, R., & Braz Golgher, A. (2014). Housing market analysis using a hierarchical–spatial approach: The case of Belo Horizonte, Minas Gerais, Brazil. *Regional Studies*, *Regional Science*, *1*(1), 116–137. https://doi.org/10.1080/21681376.2014.934391

Ostrikova, A., & Selyutin, V. (2024, March). Machine learning for mass valuation of residential real estate. In *Future of Information and Communication Conference* (pp. 570–578). Springer Nature Switzerland. https://doi.org/10.1007/978-3-031-53960-2_37

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: A review of valuation methods. *Journal of Property Investment & Finance*, *21*(4), 383–401. https://doi.org/10.1108/14635780310483656

Pi-ying, L. (2011). Analysis of the mass appraisal model by using artificial neural network in Kaohsiung city. *Journal of Modern Accounting and Auditing*, *7*(10), Article 1080.

Reyes-Bueno, F., García-Samaniego, J. M., & Sánchez-Rodríguez, A. (2018). Large-scale simultaneous market segment definition and mass appraisal using decision tree learning for fiscal purposes. *Land Use Policy*, *79*, 116–122. https://doi.org/10.1016/j.landusepol.2018.08.012

Rodriguez, M., & Sirmans, C. F. (1994). Quantifying the value of a view in single-family housing markets. *Appraisal Journal*, *62*, 600–603.

Rosen, S. (1974). Hedonic prices and implicit markets: Product differentiation in pure competition. *Journal of Political Economy*, *82*(1), 34–55. https://doi.org/10.1086/260169

Sayın, Z. M., Elburz, Z., & Duran, H. E. (2022). Analyzing housing price determinants in Izmir using spatial models. *Habitat International*, *130*, Article 102712. https://doi.org/10.1016/j.habitatint.2022.102712

Seoul Metropolitan Government. (2025). *Automobile registration status* [Web page]. https://news.seoul.go.kr/traffic/archives/341

Shapley, L. S. (1953). Stochastic games. *Proceedings of the National Academy of Sciences*, *39*(10), 1095–1100. https://doi.org/10.1073/pnas.39.10.1095

Su, T., Li, H., & An, Y. (2021). A BIM and machine learning integration framework for automated property valuation. *Journal of Building Engineering*, *44*, Article 102636. https://doi.org/10.1016/j.jobe.2021.10263

Tchuente, D. (2024). Real estate automated valuation model with explainable artificial intelligence based on Shapley values. *The Journal of Real Estate Finance and Economics*, 1–39. https://doi.org/10.1007/s11146-024-09998-9

Teoh, E. Z., Yau, W. C., Ong, T. S., & Connie, T. (2023). Explainable housing price prediction with determinant analysis. *International Journal of Housing Markets and Analysis*, *16*(5), 1021–1045. https://doi.org/10.1108/ijhma-02-2022-0025

Torres-Pruñonosa, J., Garcia-Estevez, P., & Prado-Roman, C. (2021). Artificial neural network, quantile and semi-log regression modelling of mass appraisal in housing. *Mathematics*, *9*(7), Article 783. https://doi.org/10.3390/math9070783

Trindade Neves, F., Aparicio, M., & de Castro Neto, M. (2024). The impacts of open data and eXplainable AI on real estate price predictions in smart cities. *Applied Sciences*, *14*(5), Article 2209. https://doi.org/10.3390/app14052209

Tse, R. Y., & Love, P. E. (2000). Measuring residential property values in Hong Kong. *Property Management*, *18*(5), 366–374. https://doi.org/10.1108/02637470010360669

Wang, D., & Li, V. J. (2019). Mass appraisal models of real estate in the 21st century: A systematic literature review. *Sustainability*, *11*(24), Article 7006. https://doi.org/10.3390/su11247006

Wang, J., Sui, J., Zhang, Z., Qi, J., Liu, N., & Lv, J. (2020, May). Empirical analysis of I-GBDT to improve the accuracy of mass appraisal method. *Journal of Physics: Conference Series*, *1550*, Article 032074. https://doi.org/10.1088/1742-6596/1550/3/032074

Wikipedia. (2025). Gangnam-gu. In *Korean Wikipedia*. https://ko.wikipedia.org/wiki/%EA%B0%95%EB%82%A8%EA%B5%AC

Wittowsky, D., Hoekveld, J., Welsch, J., & Steier, M. (2020). Residential housing prices: Impact of housing characteristics, accessibility and neighbouring apartments – a case study of Dortmund, Germany. *Urban*, *Planning and Transport Research*, *8*(1), 44–70. https://doi.org/10.1080/21650020.2019.1704429

Worzala, E., Lenk, M., & Silva, A. (1995). An exploration of neural networks and its application to real estate valuation. *Journal of Real Estate Research*, *10*(2), 185–201. https://doi.org/10.1080/10835547.1995.12090782

Xu, F., Uszkoreit, H., Du, Y., Fan, W., Zhao, D., & Zhu, J. (2019). Explainable AI: A brief survey on history, research areas, approaches and challenges. In J. Tang, M.-Y. Kan, D. Zhao, S. Li, & H. Zan (Eds.), *Natural language processing and Chinese computing* (pp. 563–574). Springer International Publishing. https://doi.org/10.1007/978-3-030-32236-6_51

Yasnitsky, L. N., Yasnitsky, V. L., & Alekseev, A. O. (2021). The complex neural network model for mass appraisal and scenario forecasting of the urban real estate market value that adapts itself to space and time. *Complexity*, *2021*(1), Article 5392170. https://doi.org/10.1155/2021/5392170

Zheng, Y., Yang, B., Zhang, R., Bai, Z., & Sun, Y. (2022, December). Mass appraisal of real estate prices using improved BP neural network with policy evaluation. In *2022 IEEE Conference on Telecommunications*, *Optics and Computer Science* (*TOCS*) (pp. 1036–1041). IEEE. https://doi.org/10.1109/TOCS56154.2022.10015915

Zurada, J., Levitan, A., & Guan, J. (2011). A comparison of regression and artificial intelligence methods in a mass appraisal context. *Journal of Real Estate Research*, *33*(3), 349–388. https://doi.org/10.1080/10835547.2011.12091311

APPENDIX

Table A1. Predictive performance comparison of mass appraisal models (Top: 5-fold cross validation, Bottom: 10-fold cross validation)

5-fold cross validation	RMSE	MAE	MAPE	R^2
Multiple linear regression	20852.602	14353.423	0.122	0.892
Decision tree	9560.756	4830.354	0.072	0.958
Random forest	8114.391	4044.365	0.059	0.970
XGBoost (eXtreme gradient boosting)	11189.555	7312.372	0.113	0.942
CatBoost (Categorical boosting)	8489.308	4834.046	0.077	0.967
kNN	35159.391	20326.893	0.323	0.427
10-fold cross validation	RMSE	MAE	MAPE	R^2
Multiple linear regression	20843.526	14346.761	0.233	0.799
Decision tree	9220.677	4572.057	0.069	0.961
Random forest	7840.665	3910.092	0.060	0.972
XGBoost (eXtreme gradient boosting)	11173.802	7300.551	0.112	0.942
CatBoost (Categorical boosting)	8373.855	4719.512	0.075	0.967
kNN	34884.449	19991.033	0.316	0.436

Table A2. Variance Inflation Factor (VIF) of independent variables

	Variance Inflation Factor (VIF)
Construction year	2.301
Area	4.904
Floor level	1.681
GDP	1.362
Economic growth rate	1.227
Land price fluctuation rate	1.310
Mortgage interest rate	1.136
Distance to national park	1.371
Distance to high school	1.241
Distance to redevelopment area	1.384
Distance to university	6.568
Distance to general hospital	2.226
Distance to museum	1.238
Distance to subway station	1.868
Apartment brand	2.371
Number of units in the complex	2.229
Number of buildings in the complex	7.335
Parking lot	2.023
Floor area ratio (FAR)	9.708
Building coverage ratio (BCR)	2.164
The tallest building height	4.612
The shortest building height	2.884