Machine Learning with Applications 21 (2025) 100707

Contents lists available at ScienceDirect

Machine Learning with Applications

journal homepage: www.elsevier.com/locate/mlwa

Explaining drivers of housing prices with nonlinear hedonic regressions

Heng Wan [a,*], Pranab K. Roy Chowdhury [a], Jim Yoon [a], Parin Bhaduri [b], Vivek Srikrishnan [b], David Judi [a], Brent Daniel [a]

[a] Pacific Northwest National Laboratory, Richland, WA 99352, USA
[b] Department of Biological and Environmental Engineering, Cornell University, Ithaca, NY 14850, USA

ARTICLE INFO

Keywords:
Baltimore
Hedonic regression
Interpretable machine learning
Sobol sensitivity analysis
Partial dependence plot

ABSTRACT

Housing markets play a critical role in shaping the spatial and demographic evolution of urban areas. Simulating housing price dynamics can enhance projections of future urban development outcomes. However, traditional hedonic regressions for housing prices, which neglect nonlinear interactions among explanatory variables, often exhibit limited predictive performance. While machine learning (ML) methods can provide a more flexible representation of the relationships between predictors, they are often regarded as "black boxes" due to their complexity and lack of transparency. Interpretable ML techniques provide a promising route by combining the flexibility of ML methods with approaches to analyze the relationships between inputs and outputs. In this study, we employ interpretable ML to analyze the patterns driving the housing market in Baltimore, Maryland, USA. We train an Artificial Neural Network (ANN) to predict Baltimore housing prices based on structural characteristics (e.g., home size, number of stories) and locational attributes (e.g., distance to the city center). We then conduct sensitivity and Partial Dependence Plot (PDP) analyses to interpret the fitted ANN model. We find that the ML model achieves higher predictive accuracy and explains 16 % more of housing price variance than a traditional linear regression model. The interpretable ML model also reveals more nuanced and realistic nonlinear relationships between housing sales price and predictors as well as interactive effects underlying Baltimore home price dynamics. For instance, while the linear model indicates a steady housing price increase over time, our interpretable ML model detects a post-2008 decline, with smaller properties experiencing the sharpest drop.

1. Introduction

Housing markets play a critical role in the evolution of urban systems. Housing not only represents the largest share of wealth for most households but also plays a pivotal role in the portfolios of financial intermediaries (Khalifa et al., 2013; Tsatsaronis & Zhu, 2004). In the U. S., housing and its related services contribute to roughly 17–18 % of GDP, highlighting the importance of the housing market for the economy's health (Valadez, 2011). On a more localized level, housing market volatility, such as bubbles, could significantly impact urban resiliency by introducing financial instability and socio-economic distress (Ayub et al., 2020). Therefore, a deeper insight into the housing market is vital for fostering sustainable urban growth and strengthening the local economy (Owusu-Ansah, 2011; Turcu, 2012; Xiao et al., 2017).

A commonly used technique for analyzing the housing market is hedonic regression (Chau & Chin, 2002). Hedonic regression estimates the implicit value of each characteristic of a commodity that contributes to its market price (Herath & Maier, 2010). In real state, hedonic regression estimates property values based on structural attributes (e.g., square footage, age, number of rooms), spatial attributes (e.g., distance to amenities), and neighborhood characteristics (Dubin & Goodman, 1982; Franklin & Waddell, 2003). By estimating and mapping the relationship between property prices and their determinants across locations, hedonic regression offers urban planners and policymakers valuable location-specific insights for improved decision-making (Crespo & Grêt-Regamey, 2013). The general hedonic regression function can be expressed as:

$$P = f(X) + \varepsilon \tag{1}$$

where P is the housing price vector, X is the matrix of housing price predictors, f(X) is a functional form that relates the housing attributes to

* Corresponding author.
E-mail addresses: heng.wan@pnnl.gov (H. Wan), pranab.roychowdhury@pnnl.gov (P.K.R. Chowdhury), jim.yoon@pnnl.gov (J. Yoon), pbb62@cornell.edu (P. Bhaduri), viveks@cornell.edu (V. Srikrishnan), david.judi@pnnl.gov (D. Judi), brent.daniel@pnnl.gov (B. Daniel).

https://doi.org/10.1016/j.mlwa.2025.100707

Available online 4 July 2025

the price, and ε is a vector of random error.

The method used for hedonic regression can strongly influence resulting estimates and insights. Traditionally, hedonic regression is conducted using an Ordinary Least Squares (OLS) linear regression model (Sopranzetti, 2015). The general form for a linear hedonic regression model is:

$$P = \alpha + \beta X + \varepsilon \tag{2}$$

where P is the housing price vector, X is the matrix of housing price predictors, α and β are the vectors of coefficients, and ε is a vector of random error. OLS linear regression is commonly used due to its simplicity and the straightforward interpretability of the relationship between property value and attributes (Owusu-Ansah, 2011). However, current economic theory lacks adequate guidance regarding the appropriate functional relationship between property price and its attributes (Landry et al., 2022). Additionally, linear models are limited in their ability to handle complex nonlinear relationships that may exist between property price and its predictors (Yildirim, 2019). If the functional form of the hedonic equation is assumed to be linear when the “true” data-generating relationships are nonlinear, the resulting estimation of regression coefficients would be inconsistent (Xiao & Xiao, 2017).

Recently, interest in more flexible hedonic regression models, such as those using machine learning (ML) methods, has increased, due to their ability to approximate unknown functional forms and handle potential nonlinear relationships between predictors and response variables (Curry et al., 2002). ML-based hedonic models, especially Artificial Neural Networks (ANNs), have demonstrated a higher degree of property value estimation accuracy than linear hedonic models (Ho et al., 2021; Hong et al., 2020; Selim, 2009).

ANN models are inspired by the structure and functioning of biological neural networks (Almási et al., 2016). An ANN includes three major components: an input layer, hidden layer(s), and an output layer, with each layer containing interconnected neurons (nodes) that communicate with adjacent layers (Wu & Feng, 2018). The input layer receives data from external sources and then transmits it to the nodes in the first hidden layer, where the data undergoes processing through various mathematical operations performed by the neurons. The output from the first hidden layer is then passed to the next hidden layer or the output layer, depending on the network’s structure. Finally, the output layer produces the desired prediction based on the input data. The neurons in different layers are connected by weights, which are adjusted during the training process to optimize the networks’ performance (Jain et al., 1996).

While ANNs offer higher predictive accuracy, they are often criticized for their “black box” nature, which limits clarity regarding the relationships between inputs and outputs (Bonanni, 2019; Kauko, 2003; Lorenz et al., 2022). To deal with this lack of interpretability, several methods have been developed recently to interpret fitted ML models (Murdoch et al., 2019). These methods are designed to address concerns about the “black box” nature of ML models by providing greater transparency and understanding of how these models arrive at their predictions or decisions (Agarwal & Das, 2020).

Sensitivity analysis is a widely used technique for interpreting ML models, as it quantifies and elucidates how input variable changes affect the output (Tunkiel et al., 2020). However, sensitivity analyses do not help to explain the detailed relationships between input variables and output variables. The Partial Dependence Plot (PDP) is a complementary interpretable ML technique to sensitivity analysis that can help to explain the impact of variables on the predictions of a ML model (Agarwal & Das, 2020). The one-way PDP provides a graphical representation of the average marginal effects of one variable on the model output, while holding all other variables constant (Zhao et al., 2022). However, it may not capture the potential interactions between variables (Shi et al., 2023). On the other hand, a multi-way PDP provides a more comprehensive assessment of the relationship between multiple predictors and the model’s predictions, while accounting for potential interactions between the predictors (Shi et al., 2023). For example, Lee (2022) implements a two-way PDP on a trained ANN-based hedonic model to explore how site area and building area interact and influence housing sales prices in South Korea.

In parallel with these developments, advancements in Natural Language Processing (NLP) techniques, particularly the emergence of Large Language Models (LLMs) powered by transformer architectures, provide innovative avenues for housing market analysis (Zhao et al., 2023). Transformers are a neural network architecture capable of processing long sequences of human language (Devlin et al., 2019). By understanding relationships between words and capturing long-range dependencies, transformers have greatly enhanced the feature extraction capabilities of LLMs (De Bellis, 2023). Unstructured data, such as house description texts, which are challenging to integrate into conventional hedonic regression models, can now be processed with NLP techniques to extract relevant information for the prediction of housing prices (Zhang et al., 2024). Recent studies (Zhao et al., 2023; Zhang et al., 2024) have demonstrated the ability of LLMs like GPT and BERT-based models to process textual descriptions and extract nuanced features (e.g., architectural style, subjective quality assessments, neighborhood descriptors) to improve housing price predictions. Tanlamai et al. (2025) further found that LLMs can reduce discrimination-related pricing gaps, such as those rooted in race or income level biases, suggesting that such models could support real estate assessors in fostering a more equitable and inclusive housing market. Furthermore, multi-modal large language models (MM-LLMs), which integrate textual inputs with visual data (e.g., property images or geographic maps), provide additional information for modern housing price prediction tasks (Ma et al., 2024). For instance, Wheaton and Xu (2024) leveraged ChatGPT-4o to analyze property photographs and generate quality scores for each property, which were subsequently integrated with traditionally-used variables (e.g., building area, property age) in housing price prediction models. Their findings demonstrated that incorporating visual data, captured and interpreted through MM-LLMs, substantially enhanced prediction accuracy compared to models relying solely on structured numeric or categorical inputs.

This study combines sensitivity analysis and PDP analyses on a hedonic analysis of Baltimore home prices, introducing an interpretable ML workflow which could reveal complex relationships between housing prices and predictors.

2. Methods

2.1. Study area

The Baltimore Metropolitan Area (BMA) is selected as our study area. With a total land area of 6738 km^2, BMA consists of seven counties including Baltimore County, Anne Arundel County, Baltimore City, Howard County, Harford County, Carroll County, and Queen Anne’s County. According to the U.S. Census Survey, BMA’s population has increased from 2552,994 in 2000 to 2844,510 in 2020 (Manson et al., 2022), which contributes to increasing housing demands (Mulder, 2006). Please see Figure S1 (a) in the supplemental material for the study area map.

2.2. Data collection and preprocessing

We use property transaction and attribute data for the study area from Zillow’s Transaction and Assessment Database (ZTRAX). ZTRAX is a real estate database that covers the entire United States and is exclusively accessible to researchers from academic, non-profit, and government institutions (Zillow, 2020). It provides property attributes data (e.g., square footage, age, lot size acreage, number of bathrooms, etc.) and property transaction data (e.g., housing sales price amount, transaction date, etc.) directly sourced from county recorders' offices (Zillow, 2020).

From the ZTRAX dataset, we have selected a series of structural and locational attributes as predictors for housing sales price, including building area (Area), property age (Age), number of stories (Stories), number of full bathrooms (Baths), lot size area (LotSize), and distance to the city center (CenterDist). While many other predictors are often included in hedonic analysis, we intentionally limit our study to those described above, focusing on a constrained set of indicators and the interpretability of their interactions within an ANN. Given that ZTRAX for our study area encompasses a substantial time period spanning from 1990 to 2019, we have included the number of years elapsed since the beginning year (Years) as a predictor in our analysis, which enables us to account for the impact of time on the property's value. In addition to the aforementioned predictors, the study also considered distance to coastline (CoastDist) as a potential property locational attribute. However, CoastDist was ultimately excluded from the analysis (reasons elaborated in the supplemental material, Section 1).

We extracted all single-family residence property transaction records for the study area from ZTRAX, resulting in 592,513 transactions. Transactions recorded before 1990 were identified as erroneous and subsequently removed, eliminating 35 records. Additionally, transactions lacking key property attributes required for the hedonic analysis, such as building area and lot size, were also excluded. As a result, a total of 544,657 transaction records were retained for analysis. To determine the age of each property at its transaction date, we subtracted the year it was built from the year it was sold.

Some properties had negative age values, which may indicate that they were sold before they were constructed. To account for this, we reclassified all negative age values as 0. We also subtracted 1990 from the transaction year to calculate the number of years elapsed since the beginning year. We then geocoded each property transaction to its corresponding location on the map based on property address on ArcGIS Pro, and only properties with a matching score of 100 were selected. There were 31,737 property transactions that did not match during the geocoding process, resulting in a total of 512,920 property transactions. Based on the geocoded location, we calculated distance to the city center (Downtown Baltimore) for each property transaction. Further refinements to eliminate erroneous data in ZTRAX are detailed in the supplemental material (Section 2), resulting in a final dataset of 482,629 transactions retained after all filtering steps. Table 1 shows the descriptive statistics of all predictors after filtering. Please see Table 1(b) in the supplemental material for the histograms of these predictors.

2.3. Linear regression

We randomly split the property transaction records into two sets: 80 % to a training dataset and 20 % to a testing dataset. We used the training data to fit an OLS linear regression model to predict housing sales price. We then evaluated the performance of the trained OLS linear regression model on the testing datasets by calculating several metrics. Specifically, we computed the coefficient of determination (R^2 values), Root Mean Squared Error (RMSE), and Median Absolute Percentage Error (MnAPE) between the predicted and actual housing sales prices. We preferred MnAPE over the commonly used mean absolute percentage error due to the former's robustness to outliers. As eliminating the outliers can have severe impacts in the hedonic analysis, using MnAPE becomes crucial in accurately assessing the model's performance. We further performed a variance-based sensitivity analysis – Analysis Of Variance (ANOVA) on the fitted OLS linear regression model to calculate the percentage of variance explained by each predictor, which gave the relative importance of each predictor on determining housing sales price (Saltelli et al., 2019).

2.4. Artificial neural networks (ANNs)

As ANN is the most popular machine learning model in real estate price estimation (Čeh et al., 2018), we developed an ANN model with two hidden layers to predict housing sales prices. This two-hidden layer architectu-re is commonly employed in ANN-based hedonic models and is deemed adequate for capturing the nolinearities inherent in hedonic regressions (Selim, 2009). To prepare the data for the ANN model, we standardized the predictors and housing sales price to have a mean of 0 and standard deviation of 1. Although the ANN model is trained using standardized predictors while the linear regression model is trained using original predictors, their accuracies are still comparable as the accuracy of the linear regression model remains the same before and after standardization (Bhalla, 2017). We then split the data into training and testing sets (80 %/20 % splits) as with the OLS linear regression model. More information on ANN model training is included in the supplemental materials (Section 3).

The trained ANN model has two hidden layers, with each hidden layer having 128 neurons. To reduce the risk of converging to a local minimum – a common issue where the optimization algorithm settles on a sub-optimal solution, we trained five separate ANN models, each with a different set of initial weights (Atakulreka & Sutivong, 2007). The final housing price predictions are based on the averaged outputs of this ensemble of five models. This approach not only reduces the likelihood of overfitting but is also supported by prior research, which suggests that an ensemble of moderately trained models often outperforms a single, highly optimized model in predictive accuracy (Bauer & Kohavi, 1999; Dietterich, 2000; Lieske et al., 2018).We used the same accuracy assessment metrics (R^2, RMSE, MnAPE) as those used in OLS linear regression to evaluate the performance of the ANN model on the testing datasets.

2.5. Sobol sensitivity analysis and partial dependence plot analysis

To identify the key variables that explain the variation in the predicted housing sales price, we conducted a Sobol' sensitivity analysis using the SALib Python library (Herman & Usher, 2017). It is worth noting that the first-order Sobol sensitivity index is essentially the same as ANOVA (Archer et al., 1997).

We defined the parameter range for each input variable based on the corresponding value range in the transaction dataset. Then, we generated 2^{14} parameter samples using Saltelli sampling (Please see Section 4 of the supplemental material for the detailed determination of parameter sample number). Saltelli sampling is an efficient sampling technique

Table 1
Descriptive statistics of housing transactions in the Baltimore Metropolitan Area.

Variable name	Interpretation	mean	std	min	max
SalesPrice	Housing sales price in U.S dollars	289,400.70	219,587.90	10,000.00	5000,000
Area	Building area in square feet (Square meters)	2095.8 (194.7)	976.5 (90.7)	93 (8.6)	15,462 (1436.5)
Age	Building age in years	32.2	29.2	0	368
Stories	Number of stories	1.7	0.4	1	4
Baths	Number of full bathrooms	2.1	0.8	0	6
LotSize	Lot size area in square feet (Square meters)	24,777.8 (2301.9)	35,028.8 (3254.3)	435 (40.4)	249,773 (23,204.7)
CenterDist	distance to the city center in meters	24,677.40	13,595.90	253.1	74,587.80
Years	Count of years elapsed since 1990	15.6	6.7	0	29

that generates well-dispersed parameter sets for Sobol' global sensitivity analysis, reducing computational cost (Saltelli et al., 2010). These input sample sets were evaluated by the ANN model, which produced the predicted housing sales price. We then calculated the first-order and total-order Sobol sensitivity indices. The former quantifies the proportion of the total variance directly explained by a single predictor while holding other predictors constant. The latter accounts for the portion of total variance explained by a predictor, considering both its first-order impact and its interactions with other predictors (Fel et al., 2021). Based on the first-order Sobol sensitivity indices, we conducted one-way PDP analysis for predictors with statistically significant contributions to housing price prediction at the 95 % confidence level. We also calculated second-order sensitivity indices for each pair of predictors and their corresponding 95 % confidence intervals based on 1000 bootstrap samples. We focused on pairs of variables that exhibited statistically significant second-order index values (with the 95 % confidence interval excluding 0), suggesting robust interaction effects between the variables. We then conducted two-way PDP analysis on these variable pairs to gain deeper insights into the relationship between the variable interactions and their impact on housing sales price.

We also conducted Shapley Additive Explanations (SHAP) value calculations (supplementary material, Section 5) and permutation feature importance analysis (supplementary material, section 6) for our ANN model to provide complementary insights compared to Sobol analysis.

3. Results

3.1. Model performance in Baltimore's housing market

The ANN model is substantially more accurate than the OLS model (Table 2). With superior predictive performance as evidenced by its higher R^2 value and lower RMSE and MnAPE, the ANN model demonstrates enhanced accuracy over the linear regression model in predicting Baltimore's housing prices. Furthermore, the combined predictors in the ANN model explained a higher portion of the total variance (69 %) compared to the linear regression model (47 %) as indicated in Table 3.

The superior predictive performance of the ANN model can largely be attributed to its ability in capturing complex, non-linear relationships and interactions among predictors—capabilities that linear regression models typically lack. Such modeling enables the ANN to provide more accurate and reliable predictions.

3.2. The direct impacts of predictors on Baltimore housing prices

Both models indicate that building area is the most important predictor of housing price while years elapsed since 1990 ranks the second in Baltimore housing market (Table 3, Fig. 2). The total-order Sobol sensitivity index is much higher than that of the first-order for most of the predictors (Table 3 and Fig. 2), indicating strong interaction effects in the ANN model. Notably, all predictors have a 95 % confidence interval excluding 0 for the first-order Sobol sensitivity index, suggesting a statistically significant contribution of these predictors in directly explaining the variance of the predicted housing sales price in the Baltimore. The low first-order and high total-order Sobol sensitivity index for distance to the city center, number of full bathrooms, building age, and lot size area indicates that these variables mainly contribute to the variance of the predicted housing sales price through their interactions with other variables.

Table 2
Accuracy assessment results for linear regression model and ANN model.

	R^2	RMSE	MnAPE
Linear regression	0.47	159,278	26 %
ANN model	0.63	133,821	19 %

Table 3
Variance-based sensitivity analysis results for the OLS linear regression model and the ANN model.

	ANOVA for OLS linear regression model	Sobol sensitivity analysis for ANN model	
	Percent variance explained	First-order Sobol index	Total-order Sobol index
Area	35.92 % (0.19 %)	0.40 (0.016)	0.58 (0.019)
Years	9.40 % (0.087 %)	0.15 (0.0098)	0.25 (0.0095)
Baths	1.18 % (0.034 %)	0.026 (0.0070)	0.10 (0.0044)
Stories	0.38 % (0.016 %)	0.054 (0.0085)	0.14 (0.0062)
Age	0.17 % (0.012 %)	0.024 (0.0085)	0.15 (0.0073)
LotSize	0.018 % (0.0043 %)	0.0088 (0.0074)	0.11 (0.0055)
CenterDist	0.010 % (0.0021 %)	0.027 (0.0081)	0.14 (0.0074)

Note: ANOVA was conducted for the linear regression model to determine the percentage of variance explained by each predictor. Sobol sensitivity analysis was conducted for ANN to determine the first-order and total-order Sobol sensitivity indices. Values in parenthesis indicate corresponding standard error.

All the predictors in the linear regression model are significantly associated with the housing sales price in Baltimore (Table 4). Specifically, we found that increasing the building area by one square foot (0.093 square meters) is associated with an increase in the housing sales price of $118. Conversely, increasing the property age by one year is associated with a decrease in the housing sales price of $313. The number of stories in the property is also a significant predictor, with an increase of one story leading to a decrease in the housing sales price of $26,921. On the other hand, adding one bathroom to the property is associated with an increase in the housing sales price of $29,818. Additionally, increasing the lot size by one square foot (0.093 square meters) is associated with an increase in the housing sales price of $0.018. Distance to the city center is also positively correlated with the housing sales price, with one meter increase in distance generally increasing the property value by $0.16. It's worth noting that our analysis reveals a positive trend in housing sales price over time, with each year of elapsed time corresponding to an increase in property value of $10,374.

We plotted one-way PDP for each predictor in the ANN model to explore their relationship with predicted housing sales price. We observed that the relationship between each predictor and housing sales price is nonlinear (Fig. 1). The results highlight several significant trends and behaviors:

- Housing sales price remained stable at around $200,000 for the first 7 years (1990 – 1997), followed by a significant increase over the next decade (1997–2007) to over $500,000, and then a sudden drop of $100,000. The decreasing trend of housing sales price lasted for approximately 4 years, then showing an upward trend again (Fig. 1a).
- Housing sales price displays an upward trend with increasing building area, but the rate of increase is not uniform across all ranges of building area (Fig. 1b). It tends to increase faster with the increase

Table 4
OLS linear regression coefficients.

Variable name	Estimate	Std. Error
Intercept	−131,685.97***	1492.70
Area	118.22***	0.43
Age	−313.40***	10.65
Stories	−26,921.06***	696.55
Baths	29,818.66***	424.23
Lotsize	0.018*	0.0085
CenterDist	0.16***	0.021
Years	10,374.30***	39.77

Note: *** significance at the 0.001 level. ** significance at the 0.01 level. * significance at the 0.05 level.

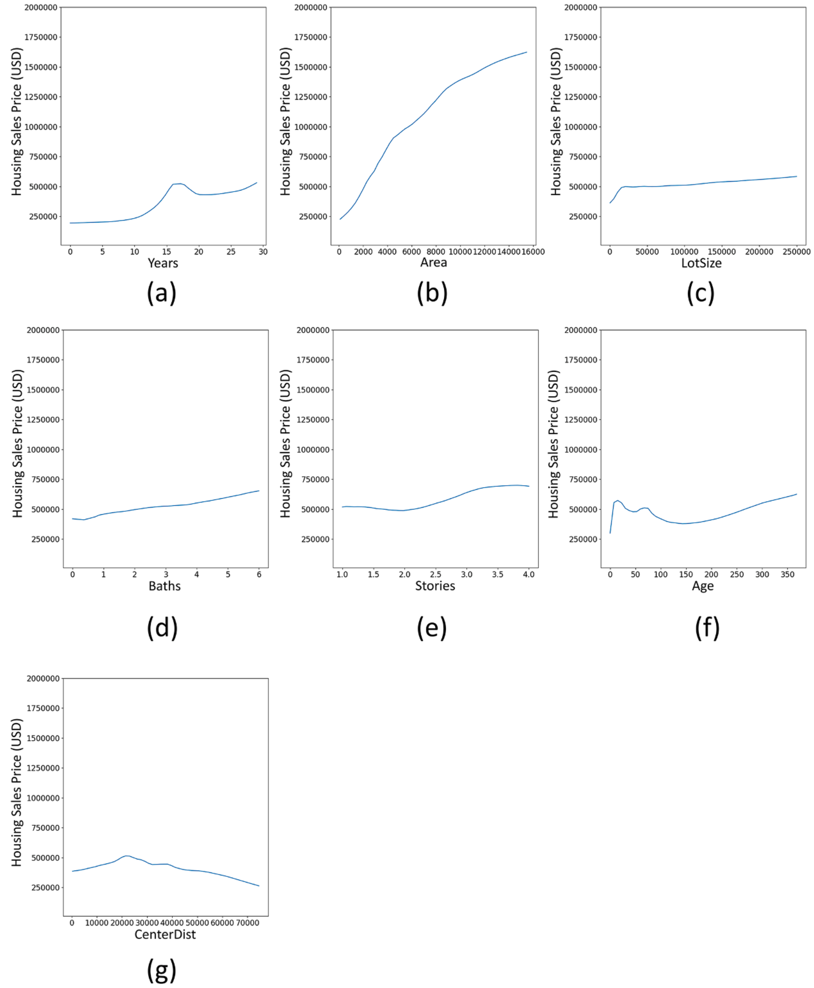

Fig. 1. One-way partial dependence plot for (a) Years (count of years elapsed since 1990); (b) Area (building area in square feet); (c) LotSize (lot size area in square feet) (d) Baths (number of full bathrooms); (e) Stories (number of stories); (f) Age (building age in years); (g) CenterDist (distance to the city center in meters). All plots are showing the same y axis range for consistency.

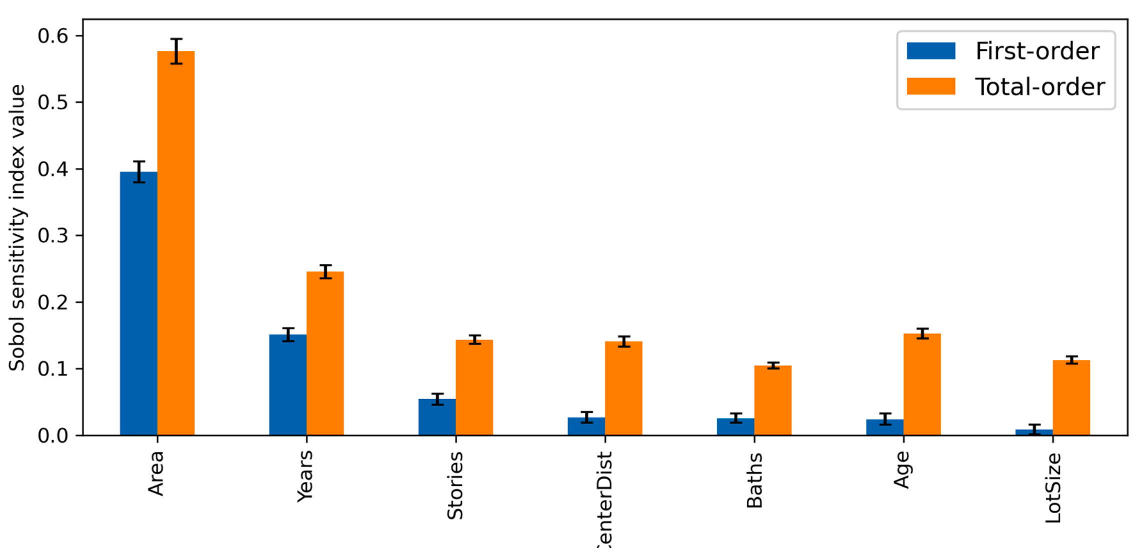

Fig. 2. First-order and total-order Sobol sensitivity index value for the ANN model. Black bars are the 95 % confidence interval based on 1000 bootstrap samples.

of building area for properties with building area <5000 square footage (465 square meters).

- Housing sales price increases with the increase of lot size area, but this increasing trend shows two distinct patterns (Fig. 1c). For properties with lot size area smaller than 20,000 square feet (1858 square meters), one square footage increase in lot size area generally results in $10 increase in sales price while that number is only $0.22 for properties with lot size area greater than 20,000 square feet (1858 square meters).
- The number of full bathrooms has a more uniform relationship with housing sales price as compared to lot size area (Fig. 1d). Adding one full bathroom would increase the housing sales price by about $40,000.
- Properties with three or four stories are typically valued $200,000 higher than those with one or two stories, whereas single-story properties command slightly higher prices than two-story properties (Fig. 1e)
- Housing sales price generally increases around $15,000 per year increase in building age for the first 20 years (Fig. 1f). After this period, the sales price starts to decrease by about $2000 for every building age increase. However, this downward trend reverses for properties aged between 50 and 75 years, where an increase in building age slightly raises the sales price. Additionally, for properties with age exceeding 150 years, their sales price starts to increase again, with one year increase in building age corresponding to approximately $1300 increase in sales price.
- Properties in suburban areas typically command higher sales prices than those in urban or rural areas (Fig. 1 g). As the distance from the city center increases or decreases from roughly 23,000 m, there is a corresponding reduction in the sales price of approximately $4.4 and $6.0 per meter, respectively.

3.3. The interaction effects of predictors on Baltimore housing prices

Three pairs of predictors exhibit a statistically significant interaction effect on the predicted housing sales price (Fig. 3). These pairs are building area and number of stories, lot size area and distance to the city center, and number of full bathrooms and lot size area.

Based on the second-order Sobol sensitivity index plot in Fig. 3, we generated two-way PDPs for three pairs of predictors that exhibited a statistically significant contribution in explaining the output variance. The resulting PDPs were plotted in Fig. 4. Fig. 4 indicates the presence of strong interaction effects between these predictors in determining the housing sales price.

The previous one-way PDP analysis suggested that properties with three or four stories generally command higher prices than those with only one or two stories. However, this trend doesn't consistently hold across all property sizes (Fig. 4a). For properties with a building area under 8000 square feet (743 square meters), the increase in the number of stories correlates only with a marginal rise in housing sales prices. Conversely, for properties exceeding this size threshold, it is the single-story homes that dramatically outpace the others in terms of sales price. In this larger property category, single-story homes achieve much higher sales prices, while two-, three-, and four-story properties tend to have comparable pricing.

Our initial one-way PDP analysis of lot size area indicated that smaller properties, specifically those under 20,000 square feet (1858 square meters), typically experience a more pronounced increase in housing sales prices as lot size increases. However, further insights gained from a two-way PDP between lot size area and distance to the city center indicate that this trend is less pronounced for rural properties located >30,000 m from the city center. In contrast to urban properties, where the price increase significantly slows down once the lot size surpasses 50,000 square feet (4645 square meters), rural properties do not show this marked decline in the rate of price increase until the lot size exceeds 70,000 square feet (6503 square meters). Conversely, the one-way PDP analysis for distance to the city center reveals that housing sales prices generally peak at a distance to the city center of about 23,000 m This peak pricing distance holds true for properties around 25,000 square feet (2323 square meters) (Fig. 4b). For substantially larger properties, those around 250,000 square feet (23,226 square meters), the highest sales price distance increases to about 30,000 m Meanwhile, for much smaller properties with lot sizes under 2000 square feet (186 square meters), their sales prices peak at distances of around 40,000 m from the city center.

Based on the initial one-way PDP analysis for years elapsed since 1990, there's a notable decrease in housing sales prices for a period of approximately four years at year 17. However, Figure 4(c) suggests that this decline is predominantly evident in smaller properties with a building area under 8000 square feet (743 square meters). In contrast, larger properties exceeding 8000 square feet (743 square meters) exhibit a different trend, where the rate of increase in housing sales prices slows down rather than declines as the years progress. Furthermore, Figure 4 (c) also demonstrates that the rate of increase in sales prices with the

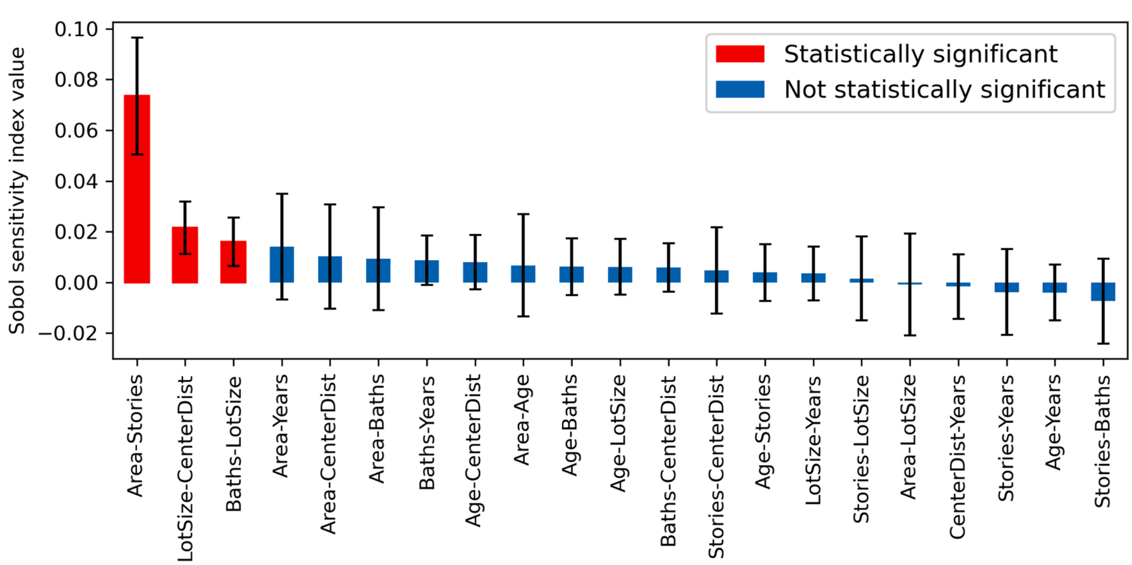

Fig. 3. Second-order Sobol sensitivity index values for each pair of predictors. Red bars represent interactions which are statistically significant in explaining the output variance while blue bars are not statistically significant. Error bars depict the 95 % confidence interval based on 1000 bootstrap samples.

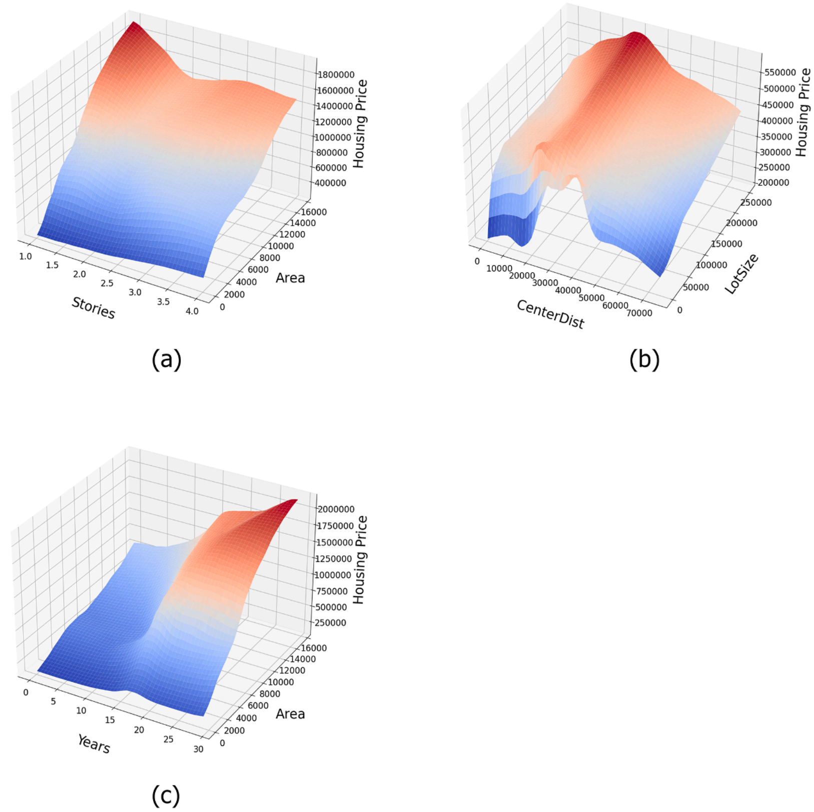

Fig. 4. Two-way partial dependence plot for (a) Stories and Area (number of stories and building area); (b) CenterDist and LotSize (distance to the city center and lot size area); (c) Years and Area (years elapsed since 1990 and building area).

increase of building area intensifies over time.

4. Discussions

While classical models such as random forest often achieve higher accuracy than deep learning for structured data, previous studies (Özdemir, 2023; Wang et al., 2021) have indicated ANN might be particularly suited for housing price regression.

In this study, we did not perform feature selection, as data-driven feature selection could introduce biases and contaminate the comparisons between the linear regression model and the ANN model (Cawley & Talbot, 2010). Instead, we relied on hypotheses about potential contributors to housing price dynamics, carefully selecting predictors that were ideally uncorrelated with one another to ensure a fair comparison between the ANN model and linear regression.

In the variance-based sensitivity analysis, we conducted ANOVA sensitivity analysis on the linear regression model to obtain the percentage of variance explained by each predictor, while ANN model was processed with Sobol sensitivity analysis to calculate first-order, second-order, and total-order Sobol sensitivity indices. Both ANOVA and Sobol sensitivity analysis rank building area and years elapsed since 1990 as the top two predictors explaining the largest variance in the model. This suggests that these two variables are highly important in determining housing sales price in Baltimore. Additionally, when considering the direct contribution from all the predictors combined, the ANN model can explain 69 % of the total variance in housing sales price, while the linear regression model can only explain 47 % of the total variance. This indicates the higher explainability of the ANN model over traditional linear regression models for hedonic regression analysis.

For the ANN model, we adopted PDP analysis to explore the relationships between predictors and housing sales price, as well as to examine the interaction effects among these predictors in Baltimore's housing market. In contrast to linear regression models, which struggle to capture the complex relationships and interaction effects between predictors, PDP analysis applied to the ANN model can reveal more sophisticated and realistic housing price dynamics. While linear regression models can incorporate nonlinear components through feature engineering (e.g., adding interaction terms and applying nonlinear transformations), and nonlinear regression models could enhance the model comparisons, these approaches are inherently constrained by the predefined forms of transformations and functional relationships, for which economic theory offers limited guidance (Landry, 2022). Furthermore, incorporating feature engineering without a clear theoretical guidance can increase the risk of overfitting, as blindly applying transformations or adding interaction terms may lead the model to capture noise or spurious relationships in the data. Additionally, adding interaction terms blindly can introduce multicollinearity, making model coefficients unstable and reduce model interpretability. In contrast, the ANN model automatically learns these complex, nonlinear relationships directly from the data, reducing the need for extensive feature engineering (Rana et al., 2023).

In this study, the linear regression model showed that housing sales price generally decreases as the property gets older. However, one-way PDP analysis for the ANN model indicated that this decreasing trend of housing sales price would be reversed if property is older than approximately 150 years, indicating the historical significance of buildings having a positive impact on the sales price. This trend is particularly reasonable in Baltimore, a city renowned for its numerous historically significant buildings, where the unique architectural character and high-quality materials of older properties often command a market premium (National Trust for Historic Preservation and Urban Land Institute, 2017).

In another example, ANN successfully captured the nuanced and realistic fluctuations of housing sales price over time. Unlike the linear regression model, which simplifies price changes as a linear progression with time, the one-way PDP analysis for the ANN model adeptly captures the market's variability. It specifically identifies the sharp decline in Baltimore housing prices starting in 2008, which lasted for approximately four years before a gradual recovery began. This portrayal is consistent with the trends depicted by Baltimore house price index (U.S. Federal Housing Finance Agency, 2024), which illustrates the profound impact of the 2008 financial crisis and the subsequent gradual recovery that commenced around 2012. Further analysis reveals that while the housing market has been recovering since 2012, it had not returned to its pre-crisis peak by the end of our study period in 2019. This suggests a long-term impact of the crisis on the Baltimore housing market. Our subsequent two-way PDP analysis between building area and years elapsed since 1990 offers deeper insights into these dynamics. While housing sales price in Baltimore generally declined following the 2008 financial crisis, larger properties (building areas greater than 8000 square feet, or 743 square meters) did not show a declining trend but only showed a slower rate of increase over time. The relatively minimal impact of the 2008 financial crisis on large properties can be explained by the fact that most of those large properties belongs to luxury real estate category (Windermere Real Estate, 2023), which typically demonstrate higher resilience during economic downturns due to their desirable locations, exclusive features, and limited supply and steady demand (Atipika Lifestyle Properties, 2024). This example revealed heterogeneous price dynamics not observed in the linear model results.

Our study found that the linear regression model indicated that Baltimore housing sales price generally increases with the distance from the city center. This suggests that property prices in rural areas are generally higher than those in suburban and city-center areas, holding other predictors constant. However, one-way PDP analysis revealed that properties located around 23,000 m from the city center had the highest housing sales price. This distance typically encompasses the outer suburbs of the BMA, as identified by Hanlon and Vicino (2007), suggesting that these areas are more likely to command the highest housing sales prices, rather than rural areas. This trend is consistent with historical patterns of migration, notably the mid-20th century white flight from the city center to suburban areas (Frey, 1979), followed by a shift from inner to outer suburbs after 1980 (Hanlon & Vicino, 2007). Since then, many of the declining inner suburbs have struggled to attract new residents due to the aging housing stock and ongoing white flight, which favored newer, more modern developments in the outer suburbs (Hanlon & Vicino, 2007). These factors collectively contribute to the heightened desirability and valuation of properties in the outer suburbs of Baltimore. It is essential to note, however, that the theoretical relationships between predictors and housing prices could differ dramatically between urban/suburban and rural areas. Therefore, the differences observed in model performance may underscore the need for distinct models tailored to each geographical context rather than reflecting a deficiency in the linear model itself.

The two-way PDP analysis further revealed that distance to the city center could impact how housing sales prices react to changes in lot size area. For urban properties, the initial increase in sales prices with enlargements in lot size is substantially more pronounced compared to rural properties. This heightened sensitivity in urban areas likely stems from the rarity of large properties near city centers, where space is at a premium and larger lots are less common. This scarcity elevates the value of bigger urban properties, reflecting the high demand for more spacious accommodations in densely populated areas.

A major discrepancy of the relationship between housing sales price and predictors revealed by the linear regression and ANN model is the impact of number of stories on housing sales price. While linear regression shows an association between an increasing number of stories and decreasing housing sales price, the one-way PDP analysis of ANN indicates three- or four-story buildings command higher prices than single- or two-story buildings. However, these findings from the two different models are not necessarily contradictory. In the dataset used for this study, most properties are single- or two-story buildings (see Figure S1(b)), with single-story buildings typically having higher sales prices than two-story buildings. Consequently, the linear regression model, observing this predominant trend, infers that additional stories are associated with reduced sales prices. On the other hand, while the ANN model also recognizes the slightly higher prices for single-story buildings over two-story ones, it additionally captures that three- or four-story buildings can attract significantly higher sales prices—a nuance the linear regression model fails to detect due to the relative scarcity of these properties in the dataset. The two-way PDP analysis further reveals the nuanced impact of the number of stories on housing sales price across property sizes. For relatively smaller properties, buildings with multiple stories generally command slightly higher sales prices, possibly due to the efficient use of limited land and the appeal of additional living space without expanding the building's footprint. Conversely, for larger properties, single-story buildings tend to attract much higher prices. This preference could be attributed to the premium placed on accessibility and the expansive, often luxurious layout that single-story homes can offer, especially in larger sizes where land availability is less of a constraint.

The relationships revealed by the PDP analysis look more reasonable and realistic, emphasizing the value of using this method to explore the relationships between predictors and housing sales price and interaction effects. Our analysis combines Sobol sensitivity analysis with PDP analysis to improve the interpretability of ML-based hedonic regression models. The combination of these techniques allows us to reveal more nuanced and realistic relationships between predictors and housing sales price, as well as the interaction effects between predictors. However, we note that PDP analysis has limitations in quantifying high-dimensional interactions. We therefore used Sobol indices to identify influential interactions between parameters.

Our insights about specific housing market dynamics in the BMA may not be directly transferable to other cities, given the differences across urban areas. However, our approach demonstrates a versatile framework that can be applied to other real estate datasets. Caution is necessary when applying this framework to real estate markets characterized by complex, high-dimensional interactions, as the current PDP analysis may have limitations in fully capturing such intricate relationships. Under this circumstance, techniques capable of handling high-dimensional interactions, such as Accumulated Local Effect (ALE) plots, should be employed (Gkolemis et al., 2023). To further enhance the interpretability and predictive accuracy of ML-based hedonic regression models, future research should continue to explore and integrate advanced analytical methods. Other commonly used ML methods, such as Support Vector Machine (SVM), and Random Forest, can be implemented and compared with ANN to see whether the model performance has improved. Additionally, other interpretable ML techniques should be considered for a more comprehensive understanding of housing market dynamics. For example, techniques such as Individual Conditional Expectation (ICE) plots serve as a complementary technique to PDP analysis. While PDP analysis helps visualize the average partial relationship between the predictors and housing sales price, ICE plots can provide detailed insights into how predictors affect the housing sales price for individual properties (Goldstein et al., 2015).

Future research should also explore the potential of multi-modal large language models (MM-LLMs) to integrate diverse data sources, such as unstructured textual data (e.g., property descriptions, agent notes) and visual data (e.g., property images, neighborhood photographs), to improve the accuracy and robustness of housing price prediction models. However, a key challenge lies in the interpretability of predictions generated by LLMs, as their output may not always align with the model's internal reasoning, potentially compromising the interpretability of housing price determinations (Chen et al., 2025). To mitigate these challenges, combining interpretable machine learning models with transformer-based methods could be explored. For example, interpretable machine learning models could be used to quantify and visualize the contribution of specific features extracted by MM-LLMs (e.g., sentiment scores derived from property descriptions, visual quality metrics obtained from property photographs) and highlight how textual and visual inputs influence prediction outcomes.

Future directions should also explore the development of online and adaptive models for housing price prediction to address the dynamic nature of housing markets (Zhang et al., 2025). These models should continuously learn from new data, such as recent housing sales, to improve housing price prediction accuracy. Additionally, integrating local-global modeling approaches can better capture the heterogeneity within housing markets (Zhang et al., 2023), accounting for distinct housing price dynamics in housing submarkets within the study area (Watkins, 2001).

The ZTRAX dataset compiles digitized housing transaction data directly from county recorders' offices, possibly leading to potential overrepresentation of urban areas with robust real estate markets and advanced record-keeping systems. In contrast, rural areas, with fewer transactions and less digitized or incomplete records, are likely underrepresented in this study. As the ZTRAX program has ended, researchers were required to delete all raw ZTRAX data by September 30th, 2023. This might impact the reproducibility of this study and future related research. To address this, we generated a synthetic dataset using the Conditional Tabular Generative Adversarial Network (CTGAN). This synthetic dataset preserves the statistical properties of the original data. However, since CTGAN smooths the data distribution (Liang, 2021), the results based on the synthetic dataset may exhibit slight variations compared to those derived from the original data. Both the synthetic dataset and the code are available at https://github.com/crystalandwan/Hedonic_analysis.git. Alternative property transaction datasets, such as CoreLogic (CoreLogic, 2024), can serve as substitutes for ZTRAX to conduct similar analyses in the future.

5. Conclusions

In this study, we have implemented the traditional OLS linear regression model and the ANN-based regression model for Baltimore housing sales price prediction using a series of property structural and locational attributes as predictors.

We found that the ANN model achieved higher prediction accuracy than the linear regression model, which may be attributed to ANN's capability of dealing with nonlinear relationships between housing sales price and predictors.

While ML-based regression models have the potential to model nonlinear relationships and achieve higher prediction accuracy than traditional linear regression models, they are often not preferred in the field of hedonic regression analysis. This is because ML models, such as ANN, are often considered "black boxes," which means that it is difficult to obtain valuable information about how each predictor impacts the predicted housing sales price. With the development of the interpretable ML techniques (e.g., Sobol sensitivity analysis, PDP analysis), it is possible to implement ML-based regression models in the hedonic regression analysis for achieving high-accuracy prediction while also providing valuable information about the relationships between housing sales price and predictors.

Variance-based sensitivity analyses, such as Sobol sensitivity analysis, are one possible solution to obtain the relative importance ranking of each predictor in determining housing sales price in ML-based regression models. To further explore the detailed relationships between housing sales price and predictors, we conducted one-way PDP analysis for each predictor in the ANN model. Moreover, we observed that the total-order Sobol sensitivity index was much higher than the first-order index for each predictor, indicating potential interaction effects between predictors on determining housing sales price. As a result, we calculated the second-order Sobol sensitivity indices for each pair of predictors. Our analysis revealed that three pairs of predictors showed statistically significant contributions to explaining housing sales price variation. We further analyzed these three pairs using two-way PDP analysis, which helped us to better understand the interaction effects between these predictors.

Compared to linear regression, integrating Sobol sensitivity analysis with PDP analysis for the ANN model allows for a more nuanced and realistic understanding of the relationships between predictors and housing sales prices. This approach provides valuable insights into housing market behaviors across a range of factors while accounting for intricate interaction effects, thereby offering a deeper and more comprehensive view of the dynamics influencing housing prices.

Declaration of generative AI and AI-assisted technologies in the writing process

During the preparation of this work the author(s) used ChatGPT-4.0 (OpenAI) in order to revise sentence structure, and check grammar and spelling. After using this tool/service, the author(s) reviewed and edited the content as needed and take(s) full responsibility for the content of the published article.

Funding

This work was supported by the Multisector Dynamics program area of the U.S. Department of Energy, Office of Science, Office of Biological and Environmental Research as part of the multi-program, collaborative Integrated Coastal Modeling (ICoM) project.

Declaration of competing interest

The authors declare the following financial interests/personal relationships which may be considered as potential competing interests:

David Judi reports financial support was provided by US Department of Energy. If there are other authors, they declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgement

We acknowledge the property transaction data used in our study, which is provided by Zillow through the Zillow Transaction and Assessment Dataset (ZTRAX). More information on accessing the data can be found at http://www.zillow.com/ztrax. The results and opinions are those of the author(s) and do not reflect the position of Zillow Group. Additionally, we would like to acknowledge the assistance of ChatGPT, an AI language model developed by OpenAI, for its help in checking grammar and spelling, as well as in revising sentence structure. After using this tool, we reviewed and edited the content as needed and take full responsibility for the content of this publication.

Supplementary materials

Supplementary material associated with this article can be found, in the online version, at doi:10.1016/j.mlwa.2025.100707.

Data availability

The authors do not have permission to share data.

References

Agarwal, N., & Das, S. (2020). Interpretable machine learning tools: A survey. In *2020 IEEE symposium series on computational intelligence* (pp. 1528–1534). IEEE. https://doi.org/10.1109/SSCI47803.2020.9308260.

Almási, A. D., Woźniak, S., Cristea, V., Leblebici, Y., & Engbersen, T. (2016). Review of advances in neural networks: Neural design technology stack. *Neurocomputing, 174*, 31–41. https://doi.org/10.1016/j.neucom.2015.02.092

Archer, G. E. B., Saltelli, A., & Sobol, I. M. (1997). Sensitivity measures, anova-like techniques and the use of bootstrap. *Journal of Statistical Computation and Simulation, 58*(2), 99–120. https://doi.org/10.1080/00949659708811825

Atakulreka, A., & Sutivong, D. (2007). Avoiding local minima in feedforward neural networks by simultaneous learning. In *20th Australian joint conference on artificial intelligence* (pp. 100–109). Springer. https://doi.org/10.1007/978-3-540-76928-6_12
Atipika Lifestyle Properties. (2024). The growing market for luxury housing: A comprehensive analysis. Retrieved from https://atipika.com/en/the-growing-market-for-luxury-housing-a-comprehensive-analysis/. Accessed May 28, 2024.
Ayub, B., Naderpajouh, N., Boukamp, F., & McGough, T. (2020). Housing market bubbles and urban resilience: Applying systems theory. *Cities (London, England), 106*, Article 102925. https://doi.org/10.1016/j.cities.2020.102925
Bauer, E., & Kohavi, R. (1999). An empirical comparison of voting classification algorithms: Bagging, boosting, and variants. *Machine Learning, 36*, 105–139. https://doi.org/10.1023/A:1007515423169
Bhalla, D. (2017). When and why to standardize a variable. Retrieved from https://www.listendata.com/2017/04/how-to-standardize-variable-in-regression.html. Accessed May 23, 2024.
Bonanni, L. (2019). The democratic aspect of machine learning: Limitations and opportunities for Parkinson's disease. *Mov Disord, 34*(2), 164–166. https://doi.org/10.1002/mds.27600
Cawley, G. C., & Talbot, N. L. (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *The Journal of Machine Learning Research, 11*, 2079–2107.
Čeh, M., Kilibarda, M., Lisec, A., & Bajat, B. (2018). Estimating the performance of random forest versus multiple regression for predicting prices of the apartments. *ISPRS International Journal of Geo-Information, 7*, 168. https://doi.org/10.3390/ijgi7050168
CoreLogic (2024). CoreLogic property data . https://www.corelogic.com/360-property-data/.</Dataset>.
Chau, K. W., & Chin, T. L. (2002). A critical review of literature on the hedonic price model. *International Journal for Housing Science and Its Applications, 27*(2), 145–165.
Chen, Y., Benton, J., Radhakrishnan, A., Uesato, J., Denison, C., Schulman, J., Somani, A., Hase, P., Wagner, M., Roger, F., Mikulik, V., Bowman, S., Leike, J., Kplan, J., & Perez, E. (2025). *Reasoning models don't always say what they think*. Anthropic Research. https://assets.anthropic.com/m/71876fabef0f0ed4/original/reasoning_models_paper.pdf.
Crespo, R., & Grêt-Regamey, A. (2013). Local hedonic house-price modelling for urban planners: Advantages of using local regression techniques. *Environment and Planning B: Planning and Design, 40*(4), 664–682. https://doi.org/10.1068/b38093
Curry, B., Morgan, P., & Silver, M. (2002). Neural networks and non-linear statistical methods: An application to the modelling of price–quality relationships. *Computers & Operations Research, 29*(8), 951–969. https://doi.org/10.1016/S0305-0548(00)00096-4
De Bellis, A. (2023). Structuring the unstructured: An LLM-guided transition. In *Proceeding of the Doctoral Consortium at ISWC 2023 Co-Located with 22ndInternational Semantic Web Conference*. https://ceur-ws.org/Vol-3678/paper12.pdf.
Devlin, J., Chang, M. W., Lee, K., & Toutanova, K. (2019). Bert: Pre-training of deep bidirectional transformers for language understanding. In *Proceedings of the 2019 conference of the North American chapter of the association for computational linguistics: Human language technologies* (pp. 4171–4186). https://doi.org/10.18653/v1/N19-1423
Dietterich, T. G. (2000). Ensemble methods in machine learning. In *Multiple classifier systems, 1857* pp. 1–15). Berlin, Heidelberg: Springer. https://doi.org/10.1007/3-540-45014-9_1
Dubin, R. A., & Goodman, A. C. (1982). Valuation of education and crime neighborhood characteristics through hedonic housing prices. *Population and Environment, 5*(3), 166–181. https://doi.org/10.1007/BF01257055
Fel, T., Cadène, R., Chalvidal, M., Cord, M., Vigouroux, D., & Serre, T. (2021). Look at the variance! Efficient black-box explanations with sobol-based sensitivity analysis. *Advances in Neural Information Processing Systems (NeurIPS)*.
Franklin, J. P., & Waddell, P. (2003). A hedonic regression of home prices in King County, Washington, using activity-specific accessibility measures. In *Proceedings of theTransportation Research Board82nd Annual Meeting*. Washington DC, US: Transportation Research Board.
Frey, W. H. (1979). Central city white flight: Racial and nonracial causes. *American Sociological Review, 44*, 425–448. https://doi.org/10.2307/2094885
Gkolemis, V., Dalamagas, T., & Diou, C. (2023). DALE: Differential Accumulated Local Effects for efficient and accurate global explanations. In *Asian Conference on Machine Learning* (pp. 375–390). PMLR.
Goldstein, A., Kapelner, A., Bleich, J., & Pitkin, E. (2015). Peeking inside the black box: Visualizing statistical learning with plots of individual conditional expectation. *Journal of Computational and Graphical Statistics, 24*, 44–65. https://doi.org/10.1080/10618600.2014.907095
Hanlon, B., & Vicino, T. J. (2007). The fate of inner suburbs: Evidence from metropolitan Baltimore. *Urban Geography, 28*, 249–275. https://doi.org/10.2747/0272-3638.28.3.249
Herath, S., & Maier, G. (2010). *The hedonic price method in real estate and housing market research: A review of the literature. SRE-Discussion papers at.* Vienna, Austria: WU Vienna University of Economics and Business.
Herman, J., & Usher, W. (2017). SALib: An open-source Python library for sensitivity analysis. *Journal of Open Source Software, 2*(9), 97. https://doi.org/10.21105/joss.00097
Ho, W. K., Tang, B. S., & Wong, S. W. (2021). Predicting property prices with machine learning algorithms. *Journal of Property Research, 38*(1), 48–70. https://doi.org/10.1080/09599916.2020.1832558
Hong, J., Choi, H., & Kim, W. (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
Jain, A. K., Mao, J., & Mohiuddin, K. M. (1996). Artificial neural networks: A tutorial. *Computer, 29*(3), 31–44. https://doi.org/10.1109/2.485891
Khalifa, S., Seck, O., & Tobing, E. (2013). Housing wealth effect: Evidence from threshold estimation. *Journal of Housing Economics, 22*, 25–35. https://doi.org/10.1016/j.jhe.2012.08.004
Kauko, T. (2003). On current neural network applications involving spatial modelling of property prices. *Journal of Housing and the Built Environment, 18*(2), 159–181. https://doi.org/10.1023/A:1023977111302
Landry, C. E., Turner, D., & Allen, T. (2022). Hedonic property prices and coastal beach width. *Applied Economic Perspectives and Policy, 44*, 1373–1392. https://doi.org/10.1002/aepp.13197
Lee, C. (2022). Training and interpreting machine learning models: Application in property tax assessment. *Real Estate Management and Valuation, 30*(1), 13–22. https://doi.org/10.2478/remav-2022-0002
Liang, T. (2021). How well generative adversarial networks learn distributions. *Journal of Machine Learning Research, 22*, 1–41.
Lieske, D. J., Schmid, M. S., & Mahoney, M. (2018). Ensembles of ensembles: Combining the predictions from multiple machine learning methods. In G. Humphries, D. Magness, & F. Huettmann (Eds.), *Machine learning for ecology and sustainable natural resource management: 109–121. Machine learning for ecology and sustainable natural resource management*. Cham: Springer. https://doi.org/10.1007/978-3-319-96978-7_5.
Lorenz, F., Willwersch, J., Cajias, M., & Fuerst, F. (2022). Interpretable machine learning for real estate market analysis. *Real Estate Economics, 51*, 1178–1208. https://doi.org/10.1111/1540-6229.12397
Mulder, C. H. (2006). Population and housing: A two-sided relationship. *Demographic Research, 15*, 401–412. https://doi.org/10.4054/DemRes.2006.15.13
Murdoch, W. J., Singh, C., Kumbier, K., Abbasi-Asl, R., & Yu, B. (2019). Definitions, methods, and applications in interpretable machine learning. *Proceedings of the National Academy of Sciences, 116*(44), 22071–22080. https://doi.org/10.1073/pnas.1900654116
National Trust for Historic Preservation and Urban Land Institute. (2017). *Building on baltimore's history: The partnership for building reuse*. November. Report for National Trust for Historic Preservation and Urban Land Institute.
Owusu-Ansah, A. (2011). A review of hedonic pricing models in housing research. *Journal of International Real Estate and Construction Studies, 1*(1), 19.
Özdemir, M., Yıldız, K., & Büyüktanır, B. (2023). Housing price estimation with deep learning: A case study of sakarya Turkey. *BSEU Journal of Science, 9*, 138–151. https://doi.org/10.35193/bseufbd.998331
Rana, P. S., Modi, S. K., Yadav, A. L., & Singla, S. (2023). Comparative analysis of tree-based models and deep learning architectures for tabular data: Performance disparities and underlying factors. In *2023 International Conference on Advanced Computing & Communication Technologies* (pp. 224–231). IEEE. https://doi.org/10.1109/ICACCTech61146.2023.00044.
Saltelli, A., Aleksankina, K., Becker, W., Fennell, P., Ferretti, F., Holst, N., Li, S., & Wu, Q. (2019). Why so many published sensitivity analyses are false: A systematic review of sensitivity analysis practices. *Environmental Modelling & Software, 114*, 29–39. https://doi.org/10.1016/j.envsoft.2019.01.012
Saltelli, A., Annoni, P., Azzini, I., Campolongo, F., Ratto, M., & Tarantola, S. (2010). Variance based sensitivity analysis of model output. Design and estimator for the total sensitivity index. *Computer physics communications, 181*, 259–270. https://doi.org/10.1016/j.cpc.2009.09.018
Selim, H. (2009). Determinants of house prices in Turkey: Hedonic regression versus artificial neural network. *Expert Systems with Applications, 36*(2), 2843–2852. https://doi.org/10.1016/j.eswa.2008.01.044
Shi, H., Yang, N., Yang, X., & Tang, H. (2023). Clarifying relationship between PM2. 5 concentrations and spatiotemporal predictors using multi-way partial dependence plots. *Remote Sensing, 15*(2), 358. https://doi.org/10.3390/rs15020358
Sopranzetti, B. J. (2015). Hedonic regression models. In C. Lee, & & J. Lee (Eds.), *Handbook of financial econometrics and statistics* (pp. 2119–2134). New York: Springer.
Ma, Y., Zhang, T., & Zhan, G. (2024). An LLM-based intelligent system for the evaluation of property geographical environment. In *2024 International Symposium on Intelligent Robotics and Systems* (p. 258). IEEE. https://doi.org/10.1109/ISoIRS63136.2024.00057.
Manson, S., Schroeder, J., Riper, D. V., Kugler, T., & Ruggles, S. (2022). IPUMS National Historical Geographic Information System. *IPUMS Data*. https://doi.org/10.18128/D050.V17.0.</Dataset>. v17.0.
Tanlamai, J., Khern-am-nuai, W., & Cohen, M. C. (2025). Generative AI and Price Discrimination in the Housing Market. *Preprint*. Available at SSRN https://ssrn.com/abstract=4764418.
Tsatsaronis, K., & Zhu, H. (2004). What drives housing price dynamics: Cross-country evidence. *BIS Quarterly Review*.
Tunkiel, A. T., Sui, D., & Wiktorski, T. (2020). Data-driven sensitivity analysis of complex machine learning models: A case study of directional drilling. *Journal of Petroleum Science and Engineering, 195*, Article 107630. https://doi.org/10.1016/j.petrol.2020.107630
Turcu, C. (2012). Local experiences of urban sustainability: Researching housing market renewal interventions in three English neighbourhoods. *Progress in planning, 78*, 101–150. https://doi.org/10.1016/j.progress.2012.04.002
U.S. Federal Housing Finance Agency (2024). All-transactions house price index for Baltimore-Columbia-Towson, MD (MSA). Retrieved from https://fred.stlouisfed.org/series/ATNHPIUS12580Q. Accessed May 25, 2024.

Valadez, R. M. (2011). The housing bubble and the GDP: A correlation perspective. *Journal of Case Research in Business and Economics, 3*, 1–18.

Wang, P. Y., Chen, C. T., Su, J. W., Wang, T. Y., & Huang, S. H. (2021). Deep learning model for house price prediction using heterogeneous data analysis along with joint self-attention mechanism. *IEEE access : Practical innovations, open solutions, 9*, 55244–55259. https://doi.org/10.1109/ACCESS.2021.3071306

Watkins, C. A. (2001). The definition and identification of housing submarkets. *Environment & planning A, 33*, 2235–2253. https://doi.org/10.1068/a34162

Windermere Real Estate. (2023). Mansion vs estate: Features of luxury properties. Retrieved from https://www.windermere.com/blog/mansion-vs-estate-features-of-luxury-properties. Accessed May 28, 2024.

Wheaton, W. C., & Xu, C. (2024). Using AI to improve price transparency in real estate valuation. *MIT Center for Real Estate Research*. https://doi.org/10.2139/ssrn.5056315. *Paper* No. 24/16. Available at SSRN: H t t p s ://ssrn.com/abstract=5056315 or.

Wu, Y., & Feng, J. (2018). Development and application of artificial neural network. *Wireless Personal Communications, 102*, 1645–1656. https://doi.org/10.1007/s11277-017-5224-x

Xiao, Y., Chen, X., Li, Q., Yu, X., Chen, J., & Guo, J. (2017). Exploring determinants of housing prices in Beijing: An enhanced hedonic regression with open access POI data. *ISPRS International Journal of Geo-Information, 6*, 358. https://doi.org/10.3390/ijgi6110358

Xiao, Y., & Xiao, Y. (2017). Hedonic housing price theory review. *Urban morphology and housing market* (pp. 11–40). Singapore: Springer. https://doi.org/10.1007/978-981-10-2762-8_2

Yildirim, H. (2019). Property value assessment using artificial neural networks, hedonic regression and nearest neighbors regression methods. *Selcuk University Journal of Engineering, Science & Technology, 7*(2), 387–404. https://doi.org/10.15317/Scitech.2019.207

Zhang, H., Li, Y., & Branco, P. (2024). Describe the house and I will tell you the price: House price prediction with textual description data. *Natural Language Engineering, 30*, 661–695. https://doi.org/10.1017/S1351324923000360

Zhang, J., Li, Y., Li, Q., & Xiao, W. (2023). Variance-constrained local–global modeling for device-free localization under uncertainties. *IEEE Transactions on Industrial Informatics, 20*, 5229–5240.

Zhang, J., Xue, J., Li, Y., & Cotton, S. L. (2025). Leveraging online learning for domain-adaptation in Wi-Fi-based device-free localization. *IEEE Transactions on Mobile Computing*. https://doi.org/10.1109/TMC.2025.3552538

Zhao, X., Yang, H., Yao, Y., Qi, H., Guo, M., & Su, Y. (2022). Factors affecting traffic risks on bridge sections of freeways based on partial dependence plots. *Physica A: Statistical Mechanics and Its Applications, 598*, Article 127343. https://doi.org/10.1016/j.physa.2022.127343

Zhao, Y., Chetty, G., & Tran, D. (2023). Real estate price prediction on GenerativeLanguage models. In *2023 IEEE Asia-Pacific Conference on Computer Science and Data Engineering* (pp. 1–7). IEEE. https://doi.org/10.1109/CSDE59766.2023.10487658.

Zillow. (2020). *Housing Data*. https://www.zillow.com/research/data/.</Dataset>.