



Comparison of Feature Importance Measures as Explanations for House Price Estimation in Vadodara City, India

Muditkumar D. Mankad

Assistant Professor

Department of Geography

Faculty of Science, The Maharaja Sayajirao University of Baroda, Vadodara, India

Abstract: Understanding the features influencing house prices across different residential areas is crucial for spatial economic analysis and accurate appraisal. This study applies eXplainable Artificial Intelligence (XAI) to identify key determinants of house prices in Vadodara, Gujarat, India. The eXtreme Gradient Boosting (XGBoost) model was used for price estimation, and various feature importance measures were compared, including XGBoost feature importance (global, model-specific), Permutation Feature Importance (PFI) (global), and SHapley Additive exPlanations (SHAP) (local), both model-agnostic. XGBoost's performance was evaluated against the Hedonic Price Model (HPM) using Ordinary Least Squares (OLS) with R^2 , RMSE, and MAPE metrics. Results indicate that the most influential features vary across importance measures, and combining multiple explanation techniques enhances reliability. The study highlights the value of integrating XAI approaches into large-scale house price estimation and urban housing research to improve model transparency and stakeholder understanding.

Keywords: House Price; XAI; XGBoost; PFI; SHAP.

1. Introduction

Rapid urbanization and commercialization in India have intensified the influence of house prices on land value, shaping urban planning and development. Since spatial characteristics strongly affect property prices (Straszheim, 1975), analyzing their impact is crucial. Traditionally, the Hedonic Price Model (HPM) (Rosen, 1974) has been widely used to evaluate influential features such as location, neighborhood, accessibility, and environmental factors. However, HPM suffers from limitations including collinearity, linear assumptions, and weak robustness (Ju et al., 2016; Li et al., 2019), reducing its modeling accuracy.

To address these issues, machine learning (ML) techniques offer powerful alternatives, especially for complex, high-dimensional data. Yet, limited real estate data in Indian cities like Vadodara demand the use of feature importance methods to reduce model complexity and improve interpretability (Chanasit et al., 2021). Since many ML models act as “black boxes,” eXplainable Artificial Intelligence (XAI) helps uncover how features influence model predictions (Das & Rad, 2020).

XAI approaches include global methods—explaining overall model behavior—and local methods—explaining individual predictions. The eXtreme Gradient Boosting (XGBoost) algorithm (Chen & Guestrin, 2016) is a robust tree-based ML model with built-in global feature importance. Complementary techniques include Permutation Feature Importance (PFI), a global, model-agnostic approach, and SHapley Additive exPlanations (SHAP), a local, model-agnostic method for instance-level interpretation.

This study applies XGBoost for house price estimation in Vadodara, Gujarat, using XGBoost feature importance, PFI, and SHAP to identify key determinants of price variations. The model's performance is also compared with the HPM using Ordinary Least Squares (OLS).

2. Study Area

Over the years, Vadodra has come up as one of the fastest growing industrial centres of India (Figure 1). This has resulted into more demand for residential houses in the city. Also, since bigger cities now have limited possibility for further property development due to scarcity of available land, the focus of real estate developers have shifted to mid-level and smaller Indian cities that are having more potential. Vadodra is one of the mid-level cities that have gained the attention of developers and investors from every quarter of the Gujarat State and the country.

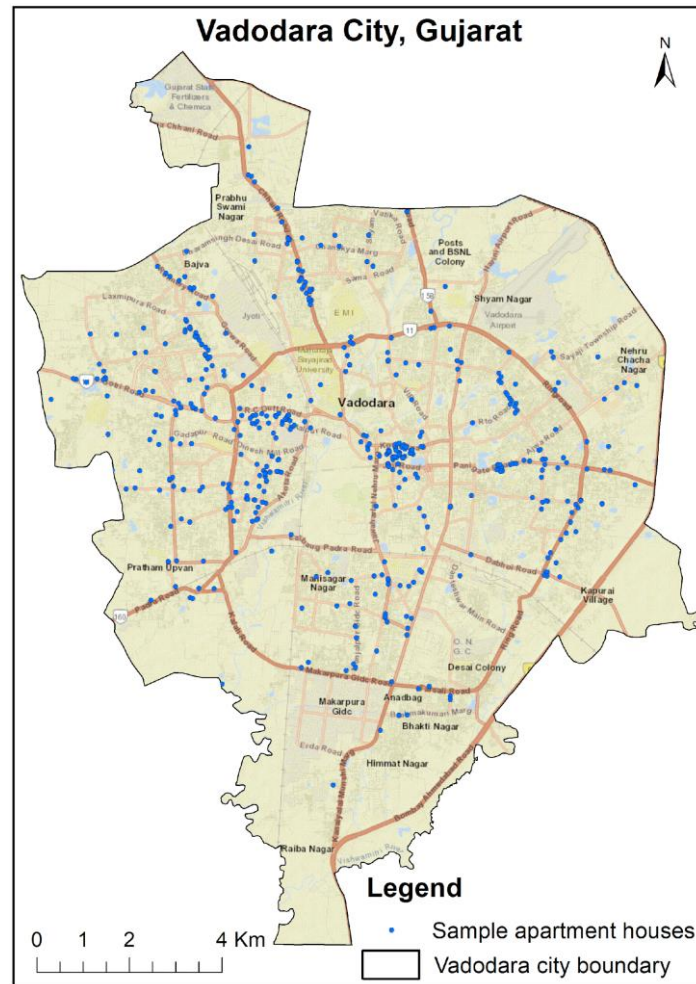


Figure 1: Study Area Map - Vadodra City, Gujarat (India)

Source: Adapted and modified from India Street Map, ESRI.

3. Analytic Framework

The first step was the collection of multi-source geodatasets, including house prices, structural attributes of the houses, neighbourhood features, facilities and amenities data and Point of Interest (POI) data. Second, the quantification of the dependent variable (saleable price of houses) and independent variables/features. Third, the models, HPM and XGBoost, and features importance techniques, XGBoost feature importance (model specific) and PFI and SHAP (model-agnostic), were implemented to identify the varying spatial effects of influential features on house prices. Finally, the results were visualized and interpreted focusing on the importance scores of influential features in estimating house prices.

3.1 Data Collection

House price and selected structural data (apartment type, carpet area, saleable area, age of house, number of bedrooms etc.) for 602 apartment houses sold during last quarter of 2024 spreading across the entire Vadodra city were collected from Liasas Foras, a real estate research company. Legal status and exterior conditions of the apartment houses were verified for all sample apartment houses to ensure authenticity of the data. For these new and sold apartment houses, primary survey was conducted to verify the selected structural data and facilities/amenities data were collected. The descriptive statistics of the influential features of house price for sample apartment houses is represented in Table 1.

Table 1: Descriptive statistics of the influential features of house price

SR. No.		Mean	1st Quartile	3rd Quartile	Standard Deviation
Dependent Feature					
1	House Price in INR (Hprice) (per sq.ft.)	2802	2380	3119	641
Independent - Structural Features					
2	No. of Storey	5	4	5	2
3	No. of Bedrooms	3	2	3	1
4	Age (years)	6	2	9	6
5	Carpet Area (sq.ft.)	965	735	1120	368
6	Saleable Area (sq.ft.)	1429	1050	1735	544
Independent - Location Features (in m)					
7	Distance to nearest major road (DisNMR)	511	174	718	447
8	Distance to nearest local market (DisNLM)	529	295	685	317
9	Distance to nearest school (DisNS)	538	257	698	365
10	Distance to nearest public transport facility (DisNPTF)	349	147	506	260
11	Distance to nearest recreational facility (DisNRecF)	1036	668	1356	524
12	Distance to nearest health service (DisNHS)	419	185	549	332
13	Distance to central transport facility (DisCTF)	4738	3863	5515	1198
14	Distance to city center/central market (DisCCCM)	7188	5839	8556	1613
15	Distance to nearest noise producing facility (DisNNoPF)	501	319	639	260
16	Distance to nearest exit from city (DisNEFC)	863	412	1206	617
17	Distance to highway (DisH)	5780	3478	7618	2566

Source: Author

Spearman's rank correlation coefficient method which is non-parametric was used to understand the degree of association between different variables. The result of correlation matrix is represented in the form of heat map (Figure 2).

3.2 Factor Quantification

To ensure that the house price data matched to the assumptions of the regression model, this study used the HPM in logarithmic form. Also, XGBoost feature importance, PFI and SHAP used logarithmic form of house price and all influential features. According to previous studies, house prices are affected by many features, and different influential features result in spatial variations in house prices. Therefore, a total of 18 influential features for house prices were included from the perspectives of structure, location, facilities, and amenities (Kang et al., 2020; Yuan et al., 2020; Du et al., 2020).

3.2.1 Structural Features

Based on the data procured from a real estate information and research company in India, five structural features such as the no. of storey, no. of bedrooms, house age (years), carpet area (sq.ft.), and saleable area (sq.ft.) (Table 1) were used to represent the type, age, shape, and size of the residence. House area and the its floor number, both have a direct impact on the lives of residents, affecting features such as comfort, views, and light exposure. Generally, house price decreases/remains constant with increasing floor number keeping other features constant in the study area. Practically, carpet area is the actual usable area enclosed within the walls for the given residential unit. Whereas, saleable area is the total built-up area of the house which is over and above the carpet area including the walls with their thickness and area under common spaces such as outer passage/lobby, staircases, lift, common facilities etc. Most property developers in the study area now offer the

houses based on the carpet area and not on the saleable area particularly after the Real Estate (Regulation and Development) Act, 2016 (RERA), implemented by the Government of India.

3.2.2 Locational features

The location's impact was evaluated considering latitude and longitude along with specific location features, which were further categorised into geographical (proximity to city centre/central market, local market and city exit points) infrastructural facilities (proximity to major road, school, public transport facility, public health service, central transport facility and highway) and neighbourhood (proximity to recreational facility and noise producing facility) (Table 1). The descriptive statistics of sample apartment houses is represented in Table 1. Like many cities of India, Vadodara's city centre is occupied by central market therefore, city centre and central market both are considered as a single feature. In general, higher the distance from/to facility, lower is the locational advantage. Except the nearness to noise producing facility, which is considered as a negative attribute i.e. higher the distance from/to facility, higher is the locational advantage. The actual road distance measured through existing road network from/to the sample apartment houses is incorporated in the present study which is implemented using ArcGIS network analyst.

3.3 Modelling Methods

3.3.1 Hedonic Price Model (HPM)

The HPM is one of the commonly used method for studying the market values of the influential features of house prices. The HPM is generally implemented through ordinary linear regression, which can reveal the marginal price of house property. The HPM can take a variety of transformations, such as logarithmic, semi-logarithmic or exponential. The present study has modelled both the house prices and their influential features in logarithmic form to solve/reduce the heteroscedasticity problem.

The study has used OLS linear regression model to implement HPM. To measure model performances, some of the most common accuracy metrics i.e. R^2 , Root Mean Squared Error (RMSE) and Mean Absolute Percentage Error (MAPE) are used. Coefficient of determination (R^2) shows how dependent and independent variables are correlated in the regression. A higher R^2 indicates better fit. It is crucial to consider its value when interpreting feature importance. RMSE is a metric that defines the difference between the actual value and predicted value by calculating the square root of the average of squared errors. Lower RMSE means a better fit for the estimation. Whereas, MAPE is the sum of prediction errors divided by the actual values and shows the mean absolute deviation as a percentage.

3.3.2 XGBoost Model

XGBoost is a boosting integration model that combines the gradient lift algorithm and decision trees, specifically using several preferred weak learners (i.e., decision trees) to complete the learning task (Chen & Guestrin, 2016; Ren et al., 2017; Gumus & Kiran, 2017). It prioritizes simpler models through regularization, addresses overfitting with a regularization term in each iteration, and optimizes computation by sorting and storing features in blocks for reuse.

The present study implemented hyperparameter automated search to determine the optimal tuned parameters of XGBoost and return the best fit model. The hyperparameters are tuned using a sophisticated bayesian optimization technique that trains the model with various possible range of parameters until a best fit model is obtained. The present study has incorporated Hyperopt python library in google colaboratory environment to implement bayesian optimization hyperparameter tuning.

3.3.3 XGBoost Feature Importance

XGBoost feature importance scores indicate how useful each feature was in building the decision trees of the model. Higher scores mean a feature was more frequently used in making key decisions within the trees, and this is calculated for each feature, allowing for ranking and comparison. Importance is determined by how much a feature improves performance (e.g. weight, gain based on Gini index or a custom error function) at each split point, weighted by the number of data points the node affects. The present work has considered 'weight' as performance measure as the goal is to compare feature importance across multiple models. The feature importances across all trees in the model are then averaged.

3.3.4 Permutation Feature Importance (PFI)

Permutation Feature Importance (PFI) (Breiman, 2001) is a widely used model-agnostic global method applicable to both linear and complex models like XGBoost. It provides a global understanding of feature importance by measuring how model performance changes when a feature's values are randomly shuffled. After training the model and establishing a baseline performance score, each feature is permuted, and the drop in performance indicates its importance. This process is repeated to obtain averaged importance scores with error estimates. PFI can be applied to both training and test sets: the train-set reveals model reliance on features, while the test-set reflects their contribution to generalization (Molnar, 2020; Adler & Painsky, 2022). In this study, PFI was implemented on the train-set to assess the model's dependence on each feature.

3.3.5 SHAP Feature Importance

SHAP (SHapley Additive exPlanations) (Lundberg & Lee, 2017) is a local, model-agnostic feature importance method that applies concepts from cooperative game theory to interpret machine learning models. It quantifies each feature's contribution to a specific prediction by comparing model outputs with and without that feature, thus revealing both individual relevance and inter-feature relationships (Kumar et al., 2020; Iban & Sekertekin, 2022). SHAP values indicate how each feature influences the final prediction and how they interact within the model (Wojtuch et al., 2021; Yamaguchi, 2020). In this study, SHAP was employed to identify and interpret the most influential structural and locational features affecting house prices estimated by the XGBoost model. All analyses, including HPM, XGBoost, and feature importance measures (XGBoost, PFI, SHAP), were implemented in the Google Colaboratory environment.

4. Results

To understand the degree of association between a dependent feature house price and independent features including 18 structural and locational, a spearman's rank correlation matrix was calculated (Figure 2).

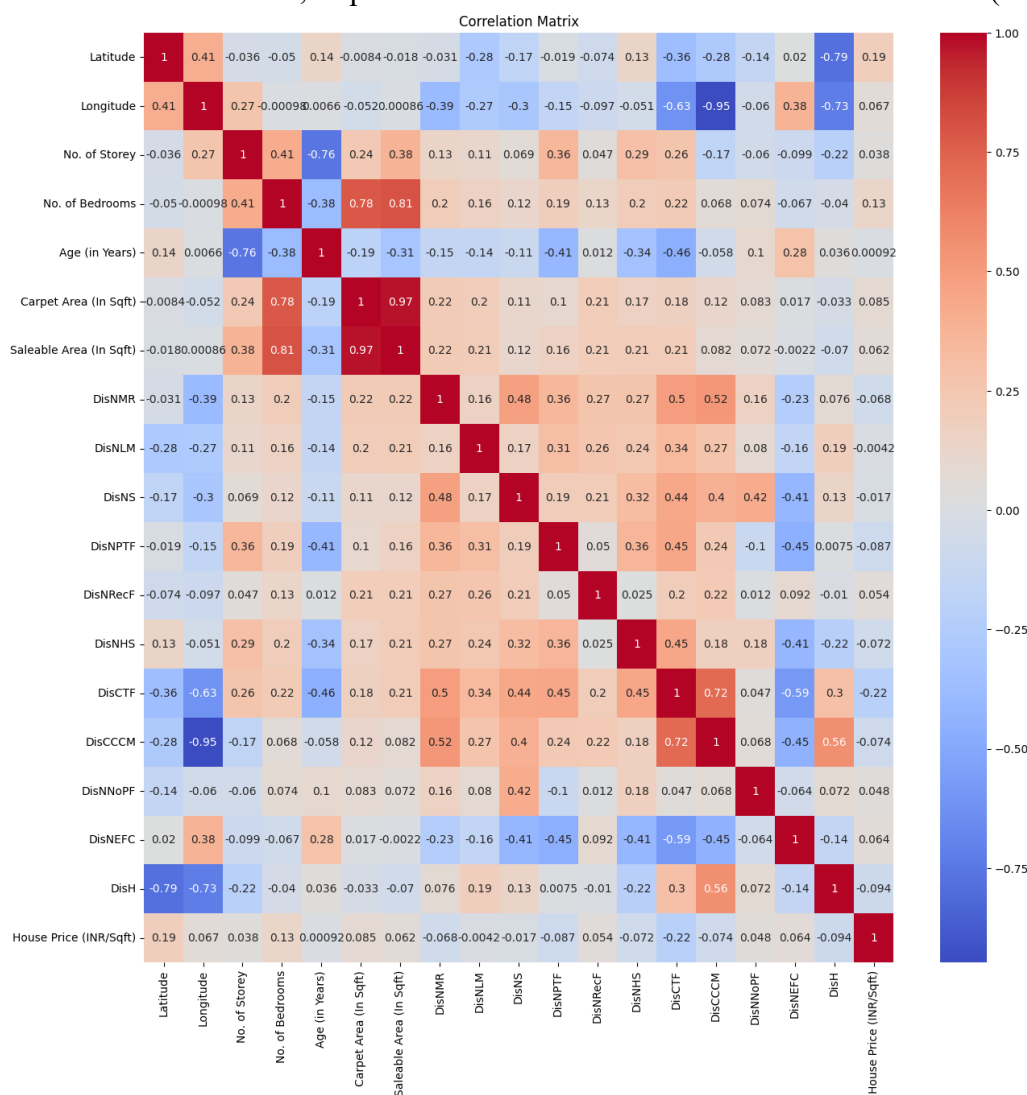


Figure 2: Spearman's Rank Correlation Matrix Heat Map

Source: Author

Two independent structural features namely carpet area and saleable area show very high correlation as expected. This is because higher carpet area leads to increase in saleable area in proportion (Figure 2). Also, "latitude" and "longitude" depict weak positive correlation with "no. of storey" and "no. of bedrooms". This might suggest that houses with more stories and bedrooms tend to be located in specific geographical areas. Also, strong positive correlation between number of bedrooms and both the carpet area and saleable area suggests that houses with a greater number of bedrooms are generally larger in size. There is moderate negative correlation observed between age of the house and carpet/saleable area and house price which could imply that older houses might be smaller in size and have lower house price. In case of locational features, all depict somewhat weak to moderate correlations among locational and between locational and structural features except two instances. One, strong negative correlation between longitude and distance to city center/central market and distance to highway and second, moderate to strong correlation between distance to central market and distance to highway/distance to central transport facility suggesting positive locational dependence.

4.1 Results of the Hedonic Price Model (HPM)

To examine the factors influencing house price variations in Vadodara, the Hedonic Price Model (HPM) using Ordinary Least Squares (OLS) was applied. House price was considered the dependent variable, with eighteen structural (e.g., number of storeys, bedrooms, age, carpet area, saleable area) and locational features (e.g., latitude, longitude, and distances to key amenities and facilities) as independent variables. The model achieved an R^2 of 0.34, indicating that 34% of price variability was explained by these features. Nine variables—five structural and four locational (latitude, and distances to recreational facility, central transport facility, and highway)—were statistically significant ($\alpha = 0.05$) (Table 2). The model yielded RMSE = 0.18 and MAPE = 1.80%, indicating relatively high estimation error (Table 3). These results highlight the limitations of OLS in capturing spatial variability in house prices, justifying the use of advanced methods like XGBoost.

Table 2: OLS result

	Coefficients	Standard Error	t Stat	P-value	Significance
Intercept	7.915500	0.009	900.519	0.000	*
Latitude	0.061900	0.014	4.346	0.000	*
Longitude	0.016500	0.055	0.300	0.764	
No. of Storey	0.047800	0.015	3.098	0.002	*
No. of Bedrooms	0.080200	0.016	5.084	0.000	*
Age (years)	-0.047500	0.017	-2.745	0.006	*
Carpet Area (sq.ft.)	0.343400	0.079	4.331	0.000	*
Saleable Area (sq.ft.)	-0.406300	0.082	-4.978	0.000	*
DisNMR	-0.009300	0.011	-0.828	0.408	
DisNLM	0.012000	0.010	1.159	0.247	
DisNS	0.013300	0.011	1.189	0.235	
DisNPTF	-0.014400	0.011	-1.259	0.209	
DisNRecF	0.038300	0.011	3.333	0.001	*
DisNHS	0.003800	0.011	0.337	0.736	
DisCTF	-0.064700	0.020	-3.293	0.001	*
DisCCCM	-0.000800	0.042	-0.019	0.985	
DisNNoPF	0.000076	0.010	0.008	0.994	
DisNEFC	-0.005500	0.012	-0.464	0.643	
DisH	0.066200	0.033	2.008	0.045	*

*significant at 0.05 level

for variable description, refer Table1

Source: Author

4.2 Results of XGBoost model

To analyze the variation in house prices across Vadodara, the XGBoost model was employed to capture non-linear relationships among eighteen structural and locational features. The model was trained using optimized hyperparameters, including max-depth, learning rate, n-estimators, subsample, colsample-bytree, gamma, min-child-weight, reg-alpha, and reg-lambda. XGBoost achieved strong performance with an R^2 of 0.84, explaining 84% of the price variation, and low error values (RMSE = 0.15, MAPE = 1.57) (Table 3). Compared to the Hedonic Price Model (HPM), XGBoost exhibited significantly higher accuracy and lower estimation error. The tuned model was thus selected as the optimal approach for further analysis using XGBoost feature importance (model-specific) and model-agnostic methods—Permutation Feature Importance (PFI) and SHapley Additive exPlanations (SHAP).

Table 3: Accuracy Metrics of the HPM and XGBoost Models

Model	R^2	RMSE	MAPE
HPM	0.34	0.18	1.80
XGBoost	0.84	0.15	1.57

Source: Author

4.3 Results of XGBoost Feature Importance

XGBoost estimates feature importance based on feature weights (F-scores). As shown in Figure 3, the top five influential features are distances to the highway, nearest city exit, noise-producing facility, city center/central market, and the distances to central transport facility and health service (both with equal scores). Among these, distance to the highway has the highest F-score, highlighting its dominant role in predicting house prices. Proximity to highways, city exits, and central markets enhances accessibility but may also introduce disamenities such as noise or congestion, influencing property values. Distance to noise-producing facilities shows a negative relationship with price, while proximity to amenities like recreational areas and schools also contributes moderately. Structural attributes (e.g., number of bedrooms, number of storeys, house age, and carpet area) have lesser influence. Overall, XGBoost results indicate that distance-related locational features are the strongest determinants of house price variation in Vadodara.

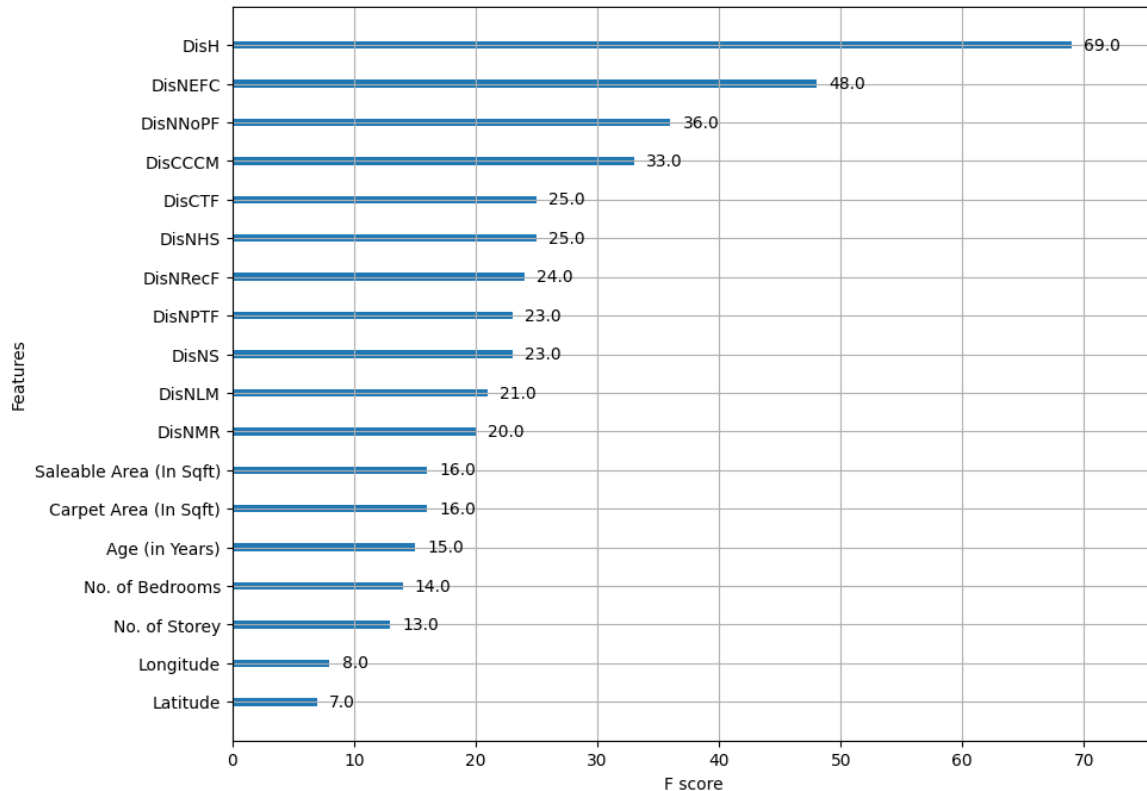


Figure 3: XGBoost Feature Importance - F Score

Source: Author

4.4 Results of Permutation Feature Importance (PFI)

Figure 4 presents the Permutation Feature Importance (PFI) results, showing the relative weights of features influencing house prices. Feature importance was measured by the change in R^2 when each variable was shuffled, with higher weights indicating greater influence. Distance to the central transport facility ranked highest, confirming that proximity to major transport hubs significantly increases house prices. Carpet area and saleable area followed, suggesting that larger homes command higher prices. The age of the house ranked third, with newer properties generally valued higher. Latitude also showed a stronger spatial effect than longitude. Distance to the highway ranked fifth, reflecting the benefit of accessibility balanced against potential noise and congestion. Overall, structural features like carpet area and house age, along with key locational factors—particularly transport accessibility—are the most influential. Other variables, such as distance to schools, markets, and public facilities, had relatively minor effects on house price estimation.

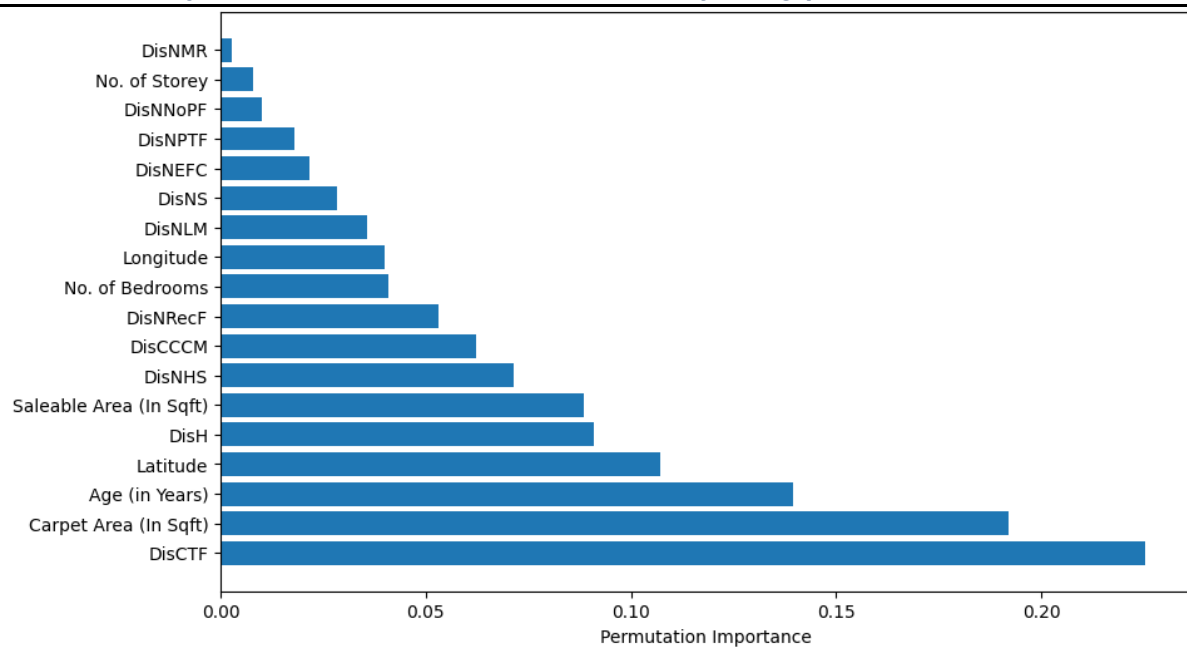


Figure 4: Permutation Feature Importance Weights

Source: Author

4.5 Results of SHAP Feature Importance

Figure 5 presents the SHAP feature importance results, illustrating each feature's impact and direction on house price predictions. The vertical order indicates overall importance, while the horizontal spread (SHAP value) shows whether a feature increases or decreases the predicted price. The top five influential features are number of bedrooms, distance to central transport facility (DisCTF), age of the house, saleable area, and latitude.

The number of bedrooms shows a strong positive effect, indicating that larger houses with more rooms tend to be higher priced. Shorter distances to central transport facilities also increase prices, reflecting Vadodara's centrally located bus and railway stations. Newer houses (lower age) and larger saleable areas positively influence prices. Latitude shows mixed effects, suggesting spatial variability.

Other features, such as carpet area and number of storeys, also contribute positively, while proximity to highways or noise sources negatively affects prices. Overall, SHAP highlights structural and locational factors jointly shaping house price variation.



Figure 5: Feature Importance using SHAP – Feature and SHAP Value

Source: Author

5. Conclusion

Machine learning has become vital for modeling spatial phenomena, though interpreting results remains a challenge, particularly when uncovering underlying spatial processes. Advances in interpretable machine learning now enable better understanding of model behavior. This study analyzed the factors influencing house prices in Vadodara, Gujarat, using the XGBoost model to capture non-linear relationships among eighteen structural and locational features. Bayesian optimization was used for hyperparameter tuning, yielding strong performance ($R^2 = 0.84$, RMSE = 0.15, MAPE = 1.57) compared to the Hedonic Price Model ($R^2 = 0.34$, RMSE = 0.18, MAPE = 1.80).

To interpret the model, three feature importance methods—XGBoost (global, model-specific), PFI (global, model-agnostic), and SHAP (local, model-agnostic)—were applied. Results showed differing influential features, with distance to the central transport facility identified as significant across all methods. SHAP emphasized structural features more strongly, highlighting bedrooms and house age as key factors. Overall, combining multiple interpretability techniques provides a comprehensive understanding of house price variations and supports more transparent spatial economic analysis.

Funding

This work was supported by Research and Consultancy Cell, The Maharaja Sayajirao University of Baroda, Vadodara, India. [Grant Number GCU/RCC/2021-22/15-32/503].

References

- Adler, A. I., & Painsky, A. (2022). Feature importance in gradient boosting trees with cross-validation feature selection. *Entropy*, 24(5), 687. <https://doi.org/10.3390/e24050687>.
- Breiman, L. (2001). Random forests. *Machine learning*, 45, 5-32. <https://doi.org/10.1023/A:1010933404324>.
- Chanasit, K., Chuangsuwanich, E., Suchato, A., & Punyabukkana, P. (2021). A real estate valuation model using boosted feature selection. *IEEE Access*, 9, 86938-86953. <https://doi.org/10.1109/ACCESS.2021.3089198>.
- Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794). <https://doi.org/10.1145/2939672.2939785>.
- Das, A., & Rad, P. (2020). Opportunities and challenges in explainable artificial intelligence (xai): A survey. *arXiv preprint arXiv:2006.11371*. <https://doi.org/10.48550/arxiv.2006.11371>.
- Du, Z., Wang, Z., Wu, S., Zhang, F., & Liu, R. (2020). Geographically neural network weighted regression for the accurate estimation of spatial non-stationarity. *International Journal of Geographical Information Science*, 34(7), 1353-1377. <https://doi.org/10.1080/13658816.2019.1707834>.
- Gumus, M., & Kiran, M. S. (2017, October). Crude oil price forecasting using XGBoost. In *2017 International conference on computer science and engineering (UBMK)* (pp. 1100-1103). IEEE. <https://doi.org/10.1109/UBMK.2017.8093500>.
- Iban, M. C., & Sekertekin, A. (2022). Machine learning based wildfire susceptibility mapping using remotely sensed fire data and GIS: A case study of Adana and Mersin provinces, Turkey. *Ecological Informatics*, 69, 101647. <https://doi.org/10.1016/j.ecoinf.2022.101647>.
- Ju, H., Zhang, Z., Zuo, L., Wang, J., Zhang, S., Wang, X., & Zhao, X. (2016). Driving forces and their interactions of built-up land expansion based on the geographical detector—a case study of Beijing, China. *International Journal of Geographical Information Science*, 30(11), 2188-2207. <https://doi.org/10.1080/13658816.2016.1165228>.
- Kang, J., Lee, H. J., Jeong, S. H., Lee, H. S., & Oh, K. J. (2020). Developing a forecasting model for real estate auction prices using artificial intelligence. *Sustainability*, 12(7), 2899. <https://doi.org/10.3390/su12072899>.
- Kumar, C. S., Choudary, M. N. S., Bommineni, V. B., Tarun, G., & Anjali, T. (2020, July). Dimensionality reduction based on shap analysis: a simple and trustworthy approach. In *2020 international conference on communication and signal processing (ICCSP)* (pp. 558-560). IEEE. <https://doi.org/10.1109/ICCSP48568.2020.9182109>.
- Li, H., Wei, Y. D., Wu, Y., & Tian, G. (2019). Analyzing housing prices in Shanghai with open data: Amenity, accessibility and urban structure. *Cities*, 91, 165-179. <https://doi.org/10.1016/j.cities.2018.11.016>.
- Lundberg, S. M., & Lee, S. I. (2017). A unified approach to interpreting model predictions. *Advances in neural information processing systems*, 30. <https://doi.org/10.5555/3295222.3295230>.
- Molnar, C. (2020). *Interpretable machine learning*. Lulu.com.
- Ren, X., Guo, H., Li, S., Wang, S., & Li, J. (2017). A novel image classification method with CNN-XGBoost model. In *Digital Forensics and Watermarking: 16th International Workshop, IWDW 2017, Magdeburg, Germany, August 23-25, 2017, Proceedings 16* (pp. 378-390). Springer International Publishing. https://doi.org/10.1007/978-3-319-64185-0_28.
- 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>.
- Straszheim, M. R. (1975). Front matter," An Econometric Analysis of the Urban Housing Market". In *An Econometric Analysis of the Urban Housing Market* (pp. 16-0). NBER.
- Wojtuch, A., Jankowski, R., & Podlowska, S. (2021). How can SHAP values help to shape metabolic stability of chemical compounds?. *Journal of cheminformatics*, 13, 1-20. <https://doi.org/10.1186/s13321-021-00542-y>.
- Yamaguchi, K. (2020, December). Intrinsic meaning of shapley values in regression. In *2020 11th international conference on awareness science and technology (ICAST)* (pp. 1-6). IEEE. <https://doi.org/10.1109/ICAST51195.2020.9319492>.
- Yuan, F., Wei, Y. D., & Wu, J. (2020). Amenity effects of urban facilities on housing prices in China: Accessibility, scarcity, and urban spaces. *Cities*, 96, 102433. <https://doi.org/10.1016/J.CITIES.2019.102433>.