International Journal of Advanced and Applied Sciences, 12(12) 2025, Pages: 237-243

Contents lists available at Science-Gate

International Journal of Advanced and Applied Sciences

Journal homepage: http://www.science-gate.com/IJAAS.html

Application of machine learning algorithms in real estate valuation

Tsolmon Sodnomdavaa [1], Erdenetsogt Gurbazar [1,*], Tegshjargal Sodnomdavaa [2]

[1]*Department of Economics, Mandakh University, Ulaanbaatar, Mongolia*
[2]*Management School, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia*

ARTICLE INFO

Article history:
Received 24 July 2025
Received in revised form
23 November 2025
Accepted 1 December 2025

Keywords:
Housing valuation
Machine learning
Deep neural network
Price prediction
Transitional economies

ABSTRACT

Accurately valuing residential properties is difficult in transitional economies because market data are fragmented, many transactions occur informally, and price trends are unstable. These conditions reduce the reliability of traditional appraisal methods. This study systematically compares nine machine learning models, covering regression, kernel-based, ensemble, and deep learning approaches, using a dataset of 9,326 housing listings from Ulaanbaatar, Mongolia. The methodology includes extensive hyperparameter tuning, five-fold cross-validation, and district-level validation to ensure robust findings. Model performance was assessed on a separate test set using R^2, Mean Squared Error (MSE), and Mean Absolute Error (MAE). The deep neural network (DNN) achieved the highest accuracy (R^2 = 0.918; MSE = 0.051), outperforming both XGBoost and Random Forest, while Support Vector Regression (SVR) showed the weakest results. The most influential price factors were district, total area, floor level, garage availability, and number of windows. Some interior characteristics, such as parquet or tile flooring, were linked to lower prices, suggesting a buyer preference for more modern designs. The study also presents a Docker-based web application for real-time price prediction, demonstrating the practical value of these models in settings with limited data. By examining Mongolia's secondary housing market, this research offers new evidence on the potential of machine learning to improve transparency and support decision-making in real estate valuation.

1. Introduction

Accurate real estate valuation is crucial for buyers, sellers, investors, and policymakers, as systematic mispricing can lead to distorted investment signals, increased credit risk, and flawed fiscal decisions (Pagourtzi et al., 2003; Ghysels et al., 2013). Conventional approaches, such as hedonic pricing and comparative market analysis, depend on restrictive assumptions and limited feature sets, which limit their ability to capture nonlinear effects, regime shifts, and heterogeneous behaviors in practice (Pagourtzi et al., 2003). These constraints are amplified in transitional and post-socialist markets where data fragmentation, informal transactions, and rapid urban change reduce the reliability and transferability of standard appraisal frameworks.

Machine learning offers a flexible alternative that models complex, nonlinear interactions and accommodates high-dimensional signals without requiring the perspectivization of functional forms (Abidoye and Chan, 2018; Gao et al., 2022). Across multiple contexts, ensemble and deep models have outperformed classical benchmarks in price prediction and mass appraisal while enabling integration of spatial and multi-source features and delivering interpretable diagnostics through permutation importance or SHAP-style analyses (Ho et al., 2021; Potrawa and Tetereva, 2022; Iban, 2022; Soltani et al., 2022; Baur et al., 2023). Parallel advances in valuation governance and professional guidance underscore the need to align automated valuation models with transparent, sustainable market development and explainability expectations (Renigier-Biłozor et al., 2022). At the same time, emerging work on index construction and appraisal accuracy in commercial settings demonstrates that gains from machine learning depend on rigorous validation design and error measurement that guard against overfitting and selection bias (Calainho et al.,

* Corresponding Author.
Email Address: erdenetsogt@mandakh.edu.mn (E. Gurbazar)
https://doi.org/10.21833/ijaas.2025.12.021
Corresponding author's ORCID profile:
https://orcid.org/0009-0009-6713-3215

2024; Deppner et al., 2025; Shcherbakov et al., 2013).

Despite this progress, the empirical literature remains concentrated in data-rich environments, leaving limited evidence on how competing algorithms perform in markets characterized by incomplete information, volatile demand, and aging secondary stock. Studies that address spatial and temporal dependency emphasize that model design needs to reflect local data-generating processes, including neighborhood effects and transaction frictions prevalent in emerging systems (Soltani et al., 2022; Zhang et al., 2021). Consequently, there is a practical and scientific need for comparative evaluations tailored to transitional contexts that document not only accuracy but also stability and transparency under realistic data constraints.

This study addresses these challenges by evaluating nine algorithms, ranging from regularized linear models to deep neural networks, on 9,326 second-hand apartment listings from Ulaanbaatar, Mongolia. The research makes three main contributions: it provides evidence tailored to a post-socialist housing market with unique data risks, employs systematic hyperparameter tuning and cross-validation for reliable model comparison, and develops a Docker-based tool for real-time price estimation. Rather than asserting universal applicability, the study offers a replicable framework for similar data-limited urban markets, emphasizing accuracy, clarity, and practical relevance.

2. Literature review

The rapid advancement of artificial intelligence (AI), particularly machine learning (ML), has transformed various commercial sectors by leveraging enhanced computational capabilities and utilizing large-scale data. Within real estate, ML has emerged as a powerful tool for analyzing complex datasets to forecast property values, enhancing decision-making, mitigating risks, and improving the efficiency of appraisals and investments (Gao et al., 2022; Ho et al., 2021). For instance, an analysis of office property price indices across ten major Chinese cities from 2005 to 2021 demonstrated that Gaussian Process Regression outperformed traditional econometric models, such as autoregressive and nonlinear neural network models, in predictive accuracy (Jin and Xu, 2025). Similarly, studies in developed markets, such as Australia and Hong Kong, have shown that ML algorithms like Random Forest and Gradient Boosting achieve superior performance by capturing non-linear relationships in property data (Choy and Ho, 2023; Gao et al., 2022). These studies typically rely on standardized datasets, including sales records, demographic information, and property attributes such as location and amenities, to deliver precise valuations (Baldominos et al., 2018).

Beyond valuation, recent methodological contributions stress the importance of handling data imperfections in ML applications. Atoum (2025) emphasized that the robustness of predictive outcomes strongly depends on the choice of evaluation metrics and strategies for addressing missing data, a concern highly relevant for transitional housing markets, where incomplete information is prevalent. Likewise, Shah et al. (2025) demonstrated through a comprehensive comparative study that algorithm performance can vary significantly depending on the domain and data type, underscoring the importance of holistic model assessment when deploying ML techniques. These insights suggest that real estate valuation research should integrate both advanced algorithms and rigorous evaluation practices to achieve sustainable improvements in predictive accuracy.

Despite these advances, most existing work focuses on newly constructed properties or data-rich markets in North America, Europe, and developed Asian economies. For example, Gao et al. (2022) utilized ML to model property values in Greater Sydney, while Choy and Ho (2023) demonstrated that ML models outperformed hedonic regressions by 12.9% in R^2 scores in Hong Kong's policy-driven market. Such studies benefit from robust data ecosystems and formal appraisal mechanisms, which enable the integration of advanced techniques, such as hybrid AI-BIM models, for near-perfect valuation accuracy. However, these approaches are less applicable in transitional economies, where data scarcity, informal transactions, and volatile market dynamics pose significant challenges.

The valuation of aging housing stock in post-socialist urban contexts, such as Ulaanbaatar, Mongolia, remains underexplored. Transitional economies face fragmented data ecosystems, high market volatility driven by rapid urbanization, and a reliance on informal transactions (Tegshjargal et al., 2025). By integrating methodological advances from broader ML research with context-specific data challenges, this study contributes to filling this gap. Specifically, it applies nine ML algorithms to Mongolia's post-socialist secondary housing market, conducts rigorous validation procedures to ensure robustness, and introduces a Docker-based web application for real-time valuation, thereby extending theoretical modeling into practical deployment for data-scarce environments.

3. Methodology

This study evaluates nine machine learning algorithms to forecast second-hand apartment prices in Ulaanbaatar, Mongolia, using a dataset of 9,326 listings collected from www.unegui.mn as of October 6, 2024. The dataset included variables commonly used in real estate valuation, such as floor area, number of rooms, floor level, construction year, building material, and locational attributes. Categorical features (e.g., building type, district) were transformed into dummy variables, while continuous variables (e.g., size, age) were standardized to ensure comparability across models. Outliers, defined as observations that are more than three standard deviations from the mean, were minimized to reduce distortion effects (Abidoye and Chan, 2018). Missing values, particularly prevalent in secondary housing transactions, were imputed using k-nearest neighbors (KNN) methods to retain sample size without introducing bias (Potrawa and Tetereva, 2022).

Nine algorithms (linear regression, ridge regression, Lasso regression, etc.) were selected to balance traditional statistical approaches with modern machine learning methods. The inclusion of both linear and nonlinear methods allowed for systematic comparison between parsimonious econometric models and flexible ML approaches (Gao et al., 2022; Ho et al., 2021). Prior research demonstrates that ensemble models such as Random Forest and XGBoost achieve higher predictive accuracy in property valuation tasks, particularly when capturing nonlinear effects and feature interactions (Baur et al., 2023; Soltani et al., 2022). To address concerns of overfitting and ensuring external validity, the dataset was split into training (70%), validation (15%), and test sets (15%). Hyperparameter tuning was conducted using grid search combined with five-fold cross-validation. This procedure enabled the systematic exploration of model configurations, including the number of estimators and the depth of trees for ensemble models, penalty coefficients for regularized regressions, and learning rates for boosting algorithms (Shcherbakov et al., 2013).

Model performance was assessed using multiple forecast error measures to capture both average error and robustness across observations. The primary metrics included the coefficient of determination (R^2), root mean squared error (RMSE), mean absolute error (MAE), and mean absolute percentage error (MAPE). The choice of multiple metrics follows recommendations that reliance on a single error measure may obscure weaknesses in model robustness (Shcherbakov et al., 2013). Three robust procedures were implemented. First, temporal sub-sampling was applied by dividing the dataset into early (2022-2023) and late (2024) subsamples to assess model stability over time. Second, models were re-estimated using a reduced feature set that excluded locational variables to test sensitivity to omitted-variable bias. Third, variance inflation factors (VIF) were calculated for linear models to detect multicollinearity, while permutation-based feature importance tests were conducted for ML models to ensure interpretability (Renigier-Biłozor et al., 2022; Iban, 2022). The best-performing model was containerized into a Docker-based web application prototype. This proof-of-concept tool enables real-time estimation by allowing users to input apartment characteristics and receive predicted valuations. Containerization was chosen for its scalability and replicability across environments, an approach recommended in recent ML-valuation frameworks that emphasize bridging methodological contributions with practical deployment (Calainho et al., 2024; Deppner et al., 2025).

4. Results

This study evaluated second-hand apartment prices in Ulaanbaatar, Mongolia, using a hedonic regression model and nine machine learning algorithms on a dataset of 9,326 listings from www.unegui.mn, collected as of October 6, 2024. The dataset was refined from an initial 9,641 listings to correct seller-reported inaccuracies, resulting in 9,326 eligible units, comprising 9,274 from Ulaanbaatar's six central districts (Khan-Uul, Bayanzurkh, Bayangol, Sukhbaatar, Songinokhairkhan, and Chingeltei) and 52 from outlying areas (Nalaikh, Baganuur, and villages). The dataset comprised 13 attributes, including location, building age, floor number, total area, window count, construction quality, material age, sale conditions, and garage availability (see Methodology).

Housing Market Characteristics: The distribution of listings showed Khan-Uul district (KhUD) as the most active market, accounting for 46.2% of the sample (4,306 units), followed by Bayanzurkh (BZD, 25%) and Bayangol (BGD, 11.4%). The average selling price in KhUD and Sukhbaatar (SBD) was 4.4 million tugriks per square meter, significantly higher than Songinokhairkhan (SKhD, 2.97 million tugriks) and BGD (3.56 million tugriks). By construction year, 65.9% of units (6,148) were built within the last decade, while 7.6% (707 units) were pre-2000, with an average price of 3.8 million tugriks per square meter, consistent with the sample mean.

Housing prices in Ulaanbaatar exhibit substantial variation. As illustrated in Fig. 1, the average price per square meter is approximately 4.0 million tugriks, with a median of 3.7 million. The price range extends from 0.27 million to 9.9 million tugriks, encompassing both affordable and high-end apartments, particularly in central districts. The existence of extreme values suggests that simple linear models may be insufficient to capture the market's complexity, thereby justifying the application of advanced machine learning methods for a more comprehensive analysis.

Hedonic Regression Analysis: A hedonic regression model was employed to examine the influence of independent factors on housing prices, as presented in Table 1.

The hedonic regression results confirm that location, particularly KhUD (25.1% dependence) and SBD (7.4%), significantly drives price premiums, while SKhD (-18.3%) has a negligible impact. Structural features, such as total area (39.7%), number of floors (34.7%), and garage availability (22.2%), are major price determinants. Floor materials (parquet, cement, tile, laminate) and certain door types (quality, vacuum) can negatively affect prices, reflecting a preference for modern designs. Financial factors, including bank loans and barter options, positively influence value.

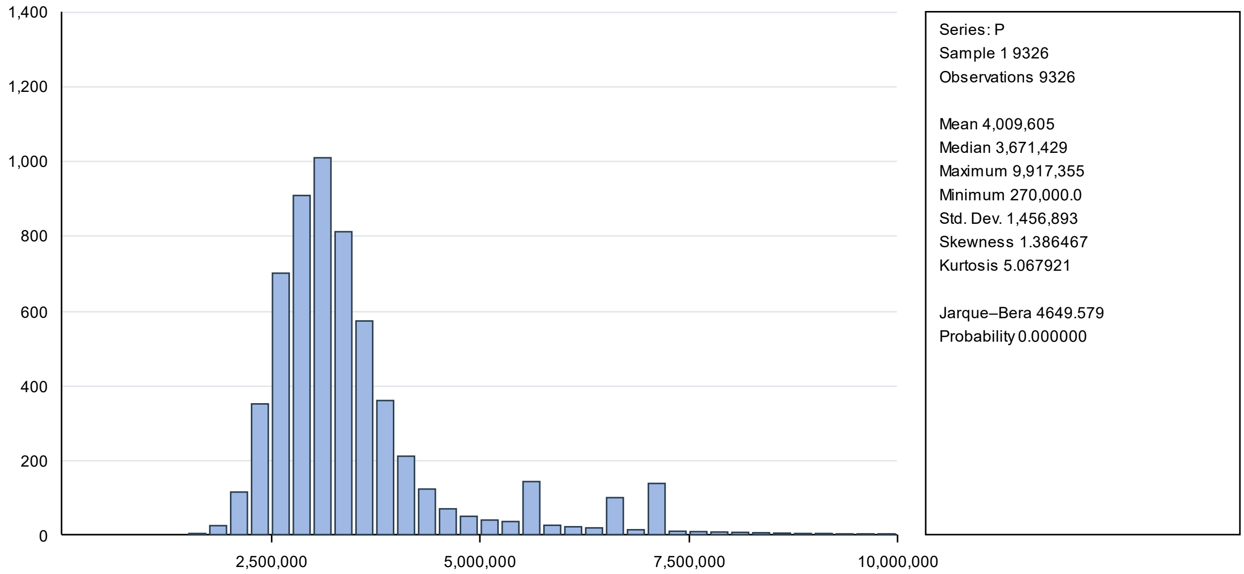

Fig. 1: Price histogram

Table 1: The impact of independent factors

Independent variables	Parameter	Standard error	t-statistic	Probability	Dependence
BGD (=1 if yes, 0 otherwise)	354360.2***	170120.4	2.082997	0.0373	-11.0%
BZD (=1 if yes, 0 otherwise)	323638.0**	168195.8	1.924174	0.0544	-13.9%
SKHD (=1 if yes, 0 otherwise)	-15108.90	173261.0	-0.087203	0.9305	-18.3%
CHD (=1 if yes, 0 otherwise)	832075.5***	178669.5	4.657066	0.0000	-0.2%
KHUD (=1 if yes, 0 otherwise)	727901.3***	167881.2	4.335813	0.0000	25.1%
SBD (=1 if yes, 0 otherwise)	890534.3***	172740.0	5.155344	0.0000	7.4%
Wooden floor (=1)	-661512.4***	349221.4	-1.894249	0.0582	3.1%
Parquet flooring (=1)	-1031876***	331770.1	-3.110214	0.0019	-5.3%
Cement floor (=1)	-788391.2***	368747.8	-2.138023	0.0325	1.7%
Tile flooring (=1)	-884742.7***	400699.4	-2.207996	0.0273	0.7%
Laminate flooring (=1)	-933362.5***	346832.4	-2.691105	0.0071	3.2%
No balcony (=1)	552197.1***	109329.9	5.050741	0.0000	3.3%
One balcony (=1)	407495.0***	102785.8	3.964508	0.0001	-7.6%
Two balconies (=1)	193731.6**	103504.6	1.871721	0.0613	4.7%
Currently in use (=1)	-19712.64***	1798.341	-10.96157	0.0000	13.9%
Garage available (=1)	303249.9***	32735.16	9.263736	0.0000	22.2%
Vacuum window (=1)	191517.6	138894.8	1.378868	0.1680	-10.2%
Iron-frame vacuum window (=1)	584227.7***	156726.5	3.727690	0.0002	11.7%
Wooden window (=1)	550881.1***	182313.5	3.021614	0.0025	4.4%
Number of floors in the building	111813.9***	3350.400	33.37328	0.0000	34.7%
High-quality door (=1)	-426024.6***	126258.6	-3.374223	0.0007	-12.9%
Iron door (=1)	-300242.4***	128013.3	-2.345400	0.0190	9.0%
Vacuum door (=1)	-409510.1***	152561.6	-2.684227	0.0073	0.9%
Iron vacuum door (=1)	-286905.3***	137977.4	-2.079364	0.0376	7.8%
Total area (m^2)	8973.384***	426.4453	21.04228	0.0000	39.7%
Floor level of the apartment	5901.312**	3158.324	1.868495	0.0617	19.6%
Leasehold (=1)	333351.2***	79704.03	4.182363	0.0000	-2.6%
Bank loan available (=1)	345436.5***	83024.94	4.160635	0.0000	4.9%
Barter option available (=1)	361567.9***	102471.3	3.528479	0.0004	1.8%
Total number of windows	23674.19**	12956.12	1.827259	0.0677	30.6%
Occupied (=1)	492941.3***	48631.47	10.13626	0.0000	11.1%
Constant	40914576***	3624397	11.28866	0.0000	

**: $p < 0.05$; ***: $p < 0.01$

Machine Learning Model Performance: The machine learning algorithms were implemented using Python 3.13: Linear Regression, Lasso Regression, Ridge Regression, K-Nearest Neighbors (KNN), Gaussian Process Regression, Support Vector Regressor (SVR), Random Forest, XGBoost, and Deep Neural Network (DNN). After hyperparameter tuning and 5-fold cross-validation (see Methodology), models were evaluated on a held-out test set (20% of data, random state=42) using Mean Squared Error (MSE), R-squared (R^2), and Mean Absolute Error (MAE). Table 2 summarizes the test-set performance, with Fig. 2 and Fig. 3 visualizing R^2 and MSE, respectively. The deep neural network (DNN), tuned with two hidden layers of 64 neurons each, a learning rate of 0.001, and 100 epochs with early stopping, achieved the highest predictive accuracy among all models.

On the test set, it recorded an R^2 of 0.918 and an MSE of 0.051. Compared to XGBoost (R^2 = 0.880, MSE = 0.058), the DNN exhibited a 4.3% higher R^2 and a 12.1% lower MSE, while Random Forest followed with R^2 = 0.850 and MSE = 0.065. Support Vector Regression (SVR) performed the weakest (R^2 = 0.610, MSE = 0.092), reflecting its limited ability to capture nonlinear dynamics in the heterogeneous housing dataset. MAE results reinforced these findings, with the DNN achieving the lowest value (0.140), outperforming XGBoost (0.160) and Random Forest (0.180).

Table 2: Test set performance comparison of machine learning algorithms

Model	MSE	R^2	MAE
Deep neural network	0.051	0.918	0.140
XGBoost	0.058	0.880	0.160
Random forest	0.065	0.850	0.180
Gaussian process	0.070	0.750	0.190
K-nearest neighbors	0.075	0.720	0.200
Ridge regression	0.079	0.680	0.210
Lasso regression	0.080	0.670	0.220
Linear regression	0.082	0.660	0.210
Support vector regressor	0.092	0.610	0.250

Fig. 2 compares the R^2 scores of the models, clearly illustrating the superior accuracy of the DNN. Fig. 3 presents the corresponding MSE values, where the DNN again outperforms alternative approaches. Rather than duplicating the numerical results, Figs. 2 and 3 highlight the performance gap across models, supporting the conclusion that advanced nonlinear methods substantially outperform traditional linear regression models. Bar plot comparing R^2 scores across all models, with DNN achieving the highest value (0.918).

Importantly, these machine learning outcomes align with the hedonic regression findings, confirming that location (particularly KhUD and SBD), total area, and structural features, such as garage availability and the number of floors, are the most significant determinants of housing prices. The DNN's superior performance demonstrates its ability to capture complex, nonlinear interactions across these factors, consistent with prior research (Choy and Ho, 2023).

A Docker-based web application was developed using the DNN model to deliver real-time price estimates. This tool, which is being adapted for iOS and Android, is designed to enhance transparency in Mongolia's secondary housing market and meet the need for data-driven valuation solutions.

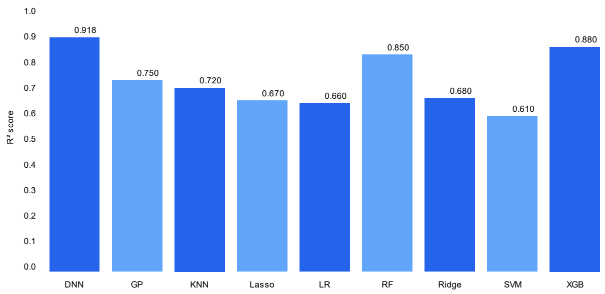

Fig. 2: R-scores of machine learning algorithms on the test set

5. Discussion

This study compared nine machine learning algorithms for predicting apartment prices in Ulaanbaatar's secondary housing market. The DNN achieved the strongest performance with an R^2 of 0.918 and an MSE of 0.051, confirming its ability to capture non-linear interactions and heterogeneous features more effectively than linear regression and support vector regression. These findings are consistent with prior research showing the superiority of ensemble and deep learning models in data-scarce or transitional markets (Choy and Ho, 2023; Gao et al., 2022). The use of systematic hyperparameter tuning, cross-validation, and robustness checks further supports the reliability of the results and provides practical guidance for model selection (Shcherbakov et al., 2013; Hernes et al., 2024).

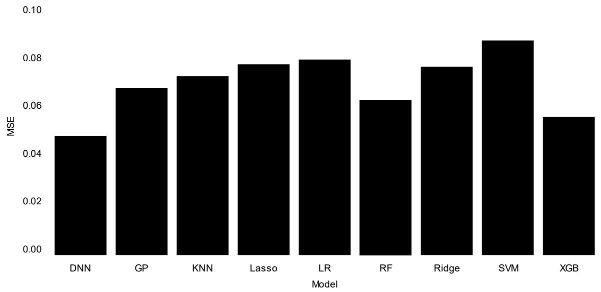

Fig. 3: MSE of machine learning algorithms on the test set

Determinants of price variation were also identified. Location proved to be the most influential factor, with apartments in Khan-Uul and Sukhbaatar districts commanding premiums due to better infrastructure, environmental quality, and amenities. Structural characteristics, including floor area, building level, number of rooms, and parking availability, were positively associated with price, consistent with hedonic pricing theory (Potrawa and Tetereva, 2022; Abidoye and Chan, 2018). Interestingly, certain interior features, such as parquet or tiled flooring, were negatively correlated with prices, suggesting that buyers prioritize modern design over traditional materials. This highlights the behavioral dimension of real estate markets, where subjective perceptions significantly shape value (Renigier-Biłozor et al., 2022; Baur et al., 2023).

The results also revealed spatial inequalities in Ulaanbaatar's housing market. Higher prices in central and southern districts indicate concentrated investment and demand, while peripheral areas lag. These patterns highlight the importance of incorporating geolocation variables into valuation models and underscore the need for policies that address urban disparities through balanced development (Zhang et al., 2021; Soltani et al., 2022). Beyond academic contributions, the study demonstrated practical relevance by developing a Docker-based prototype for real-time price estimation. While preliminary, this tool illustrates how machine learning valuation frameworks can enhance transparency and reduce informational asymmetries in transitional economies, bridging methodological advances with actionable decision support (Calainho et al., 2024; Deppner et al., 2025).

6. Conclusion

This study sets out to address the challenges of real estate valuation in transitional economies by applying machine learning methods to Ulaanbaatar's secondary housing market. Using a dataset of 9,326 apartment listings from 2022 to 2024, nine algorithms were systematically compared, ranging from linear regression to deep neural networks. The analysis demonstrated that advanced machine learning techniques, particularly ensemble models and deep learning, substantially outperform traditional econometric methods in capturing complex, non-linear relationships and handling heterogeneous data. The Deep Neural Network achieved the highest predictive accuracy (R^2 = 0.918), underscoring the potential of data-driven approaches in improving property valuation in volatile contexts.

Beyond predictive performance, the study identified key determinants of price variation, with location, floor area, number of rooms, and building level exerting the strongest positive effects, while some interior features, such as parquet or tiled flooring, were negatively correlated with value. These results not only confirm the relevance of hedonic pricing theory but also highlight the behavioral and perceptual dimensions of housing markets. Furthermore, the analysis revealed significant spatial inequalities across Ulaanbaatar, highlighting the importance of geolocation variables and underscoring the need for balanced urban development policies.

A significant contribution to this research is its emphasis on practical application. The Docker-based web prototype demonstrates that advanced machine learning methods can be implemented as real-time tools for end users.

These applications enhance market transparency, reduce information asymmetries, and provide effective support for households, investors, and policymakers in contexts where formal appraisals are limited. This study also offers a comparative evaluation of multiple algorithms with rigorous validation, providing guidance for model selection under uncertainty. Furthermore, it bridges academic research with practical deployment, illustrating a pathway toward sustainable and transparent valuation systems. Future research could incorporate additional spatiotemporal data, integrate unstructured information such as property descriptions or images, and test the adaptability of the proposed framework in other emerging urban environments.

List of abbreviations

AI	Artificial intelligence
BGD	Bayangol district
BIM	Building information modeling
BZD	Bayanzurkh district
CHD	Chingeltei district
DNN	Deep neural network
KhUD	Khan-Uul district
KNN	K-nearest neighbors
MAE	Mean absolute error
MAPE	Mean absolute percentage error
ML	Machine learning
MSE	Mean squared error
RMSE	Root mean squared error
SBD	Sukhbaatar district
SKhD	Songinokhairkhan district
SVR	Support vector regression
VIF	Variance inflation factor

Acknowledgment

The authors would like to express their sincere gratitude to Mandakh University for providing financial and institutional support for this research.

Compliance with ethical standards

Conflict of interest

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

References

Abidoye RB and Chan APC (2018). Improving property valuation accuracy: A comparison of hedonic pricing model and artificial neural network. Pacific Rim Property Research Journal, 24(1): 71-83. https://doi.org/10.1080/14445921.2018.1436306

Atoum I (2025). The critical role of evaluation metrics in handling missing data in machine learning. International Journal of Advanced and Applied Sciences, 12(1): 112-124. https://doi.org/10.21833/ijaas.2025.01.011

Baldominos A, Blanco I, Moreno AJ, Iturrarte R, Bernárdez Ó, and Afonso C (2018). Identifying real estate opportunities using machine learning. Applied Sciences, 8(11): 2321. https://doi.org/10.3390/app8112321

Baur K, Rosenfelder M, and Lutz B (2023). Automated real estate valuation with machine learning models using property descriptions. Expert Systems with Applications, 213: 119147. https://doi.org/10.1016/j.eswa.2022.119147

Calainho FD, van de Minne AM, and Francke MK (2024). A machine learning approach to price indices: Applications in commercial real estate. The Journal of Real Estate Finance and Economics, 68: 624-653. https://doi.org/10.1007/s11146-022-09893-1

Choy LHT and Ho WKO (2023). The use of machine learning in real estate research. Land, 12(4): 740. https://doi.org/10.3390/land12040740

Deppner J, von Ahlefeldt-Dehn B, Beracha E, and Schaefers W (2025). Boosting the accuracy of commercial real estate appraisals: An interpretable machine learning approach. The Journal of Real Estate Finance and Economics, 71: 314-351. https://doi.org/10.1007/s11146-023-09944-1 **PMid:38625136 PMCid:PMC10031694**

Gao Q, Shi V, Pettit C, and Han H (2022). Property valuation using machine learning algorithms on statistical areas in Greater Sydney, Australia. Land Use Policy, 123: 106409. https://doi.org/10.1016/j.landusepol.2022.106409

Ghysels E, Plazzi A, Valkanov R, and Torous W (2013). Forecasting real estate prices. Handbook of Economic Forecasting, 2: 509-580. https://doi.org/10.1016/B978-0-444-53683-9.00009-8

Hernes M, Tutak P, Nadolny M, and Mazurek A (2024). Real estate valuation using machine learning. Procedia Computer Science, 246: 4592-4599. https://doi.org/10.1016/j.procs.2024.09.323

Ho WKO, Tang BS, and Wong SW (2021). Predicting property prices with machine learning algorithms. Journal of Property Research, 38(1): 48-70. https://doi.org/10.1080/09599916.2020.1832558

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

Jin B and Xu X (2025). Predictions of residential property price indices for China via machine learning models. Quality & Quantity, 59: 1481-1513. https://doi.org/10.1007/s11135-025-02080-3

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

Potrawa T and Tetereva A (2022). How much is the view from the window worth? Machine learning-driven hedonic pricing model of the real estate market. Journal of Business Research, 144: 50-65. https://doi.org/10.1016/j.jbusres.2022.01.027

Renigier-Biłozor M, Źróbek S, Walacik M, Borst R, Grover R, and d'Amato M (2022). International acceptance of automated modern tools use must-have for sustainable real estate market development. Land Use Policy, 113: 105876. https://doi.org/10.1016/j.landusepol.2021.105876

Shah SZA, Abdulkader OA, Jan S, Shah MA, and Anwar M (2025). A holistic evaluation of machine learning algorithms for text-based emotion detection. International Journal of Advanced and Applied Sciences, 12(7): 55–75. https://doi.org/10.21833/ijaas.2025.07.006

Shcherbakov MV, Brebels A, Shcherbakova NL, Tyukov AP, Janovsky TA, and Kamaev VAE (2013). A survey of forecast error measures. World Applied Sciences Journal, 24: 171-176.

Soltani A, Heydari M, Aghaei F, and Pettit CJ (2022). Housing price prediction incorporating spatio-temporal dependency into machine learning algorithms. Cities, 131: 103941. https://doi.org/10.1016/j.cities.2022.103941

Tegshjargal S, Tsolmon S, and Namuun A (2025). Possibility of predicting inflation: Using machine learning model. International Journal of Social Science and Humanities Research-MIYR, 5(2): 49–64. https://doi.org/10.53468/mifyr.2025.05.02.49

Zhang P, Hu S, Li W, Zhang C, Yang S, and Qu S (2021). Modeling fine-scale residential land price distribution: An experimental study using open data and machine learning. Applied Geography, 129: 102442. https://doi.org/10.1016/j.apgeog.2021.102442