Journal of Applied Science and Engineering A,
Volume 6, Issue 1, 2025, pp.38-52.
https://doi.org/10.5564/jase-a.v6i1.5392

Check for updates

Article

Using automated valuation models for real estate valuation

Enkhsuren Adilbish[1,*], Bayartugs Tamjav[2], Enkhtuya Bavuudorj[3], Zolzaya Dashdorj[4], Ankhtuya Dorjgotov[2], and Mungunsukh Shirnen[5]

[1] Department of Financial Management, University of Finance and Economics, Ulaanbaatar, Mongolia
[2] Department of Mathematics, School of Applied Sciences, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
[3] Department of Economics and Finance, School of Management, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
[4] Department of Information Technology, School of Information and Communication Technology, Mongolian University of Science and Technology, Ulaanbaatar, Mongolia
[5] Department of Financial and Economics, University of the Humanities, Ulaanbaatar, Mongolia

Corresponding author. Email: enkh112011@email.com

Abstract. Asset valuation is not only a fundamental tool in economic relations but also an integral part of asset utilization, management, and planning for all participants in these relationships. Until now, traditional approaches relying on 100% involvement of appraisers have remained dominant, but with technological advancement, there is an increasing trend toward automation. Automated valuation models offer professional appraisers advantages in reducing costs, improving predictive capabilities, and increasing computational precision, while mathematical calculations and statistical modeling enable the forecasting of real estate market prices based on historical transactions. This research examines the theoretical concepts, historical development, current state, and trends in real estate valuation and automated valuation models. The research has developed a four-stage process mapping that defines the boundaries between appraiser activities and automated operations in developing value predictions for real estate valuation (hereinafter REV). It also determines how to utilize values predicted by automated valuation models (hereinafter AVM). Using the AVM calculations, the appraiser makes decisions on selecting the actual value by applying coefficients that consider the property's specific characteristics and materiality. This research work is valuable for banks, insurance companies, financial institutions, valuation companies, and professional appraisers in preparing information for REV, developing pipeline processes for automated valuation models, and applying the proposed model for calculating real estate values, risk assessment, and calculating tax bases.

Keywords: data, price index, artificial neural network, information system, artificial intelligence

Received: 04 Aug. 2025; Revised: 10 Oct. 2025; Accepted: 26 Dec. 2025.

1 Introduction

As information technology is rapidly developing worldwide and digital transformation is penetrating all sectors, there is a pressing need to implement technological advancements in Mongolia's property valuation sector as well. In particular, it has become a crucial contemporary requirement to perform real estate valuation activities using data-driven, scientifically-based, and objective methods.

The International Association of Assessing Officers defines Automated Valuation Models (AVMs) as "computer software based on mathematical calculations that estimate market values based on market analysis of location, market conditions, and real estate characteristics from general and specially collected databases." This modeling provides an opportunity to introduce scientific foundations, stable methodologies, and specific standards to the valuation process.

In Mongolia, there is relatively abundant transaction data for residential real estate, and it is appropriate to develop and implement automated valuation modeling using this data and information as a foundation, employing econometric and artificial intelligence (AI) techniques. Such modeling will enable increased productivity in valuation activities, improved accuracy and objectivity of valuations, and savings in time and resources.

This research aims to study the characteristics of Mongolia's residential real estate market, data availability, and factors affecting prices, to identify theoretical and practical issues in the application of automated valuation modeling, and to explore the possibility of implementing digital transformation in valuation activities. This includes conducting comparative studies of techniques such as regression analysis, machine learning, and deep neural networks to determine the most suitable modeling for Mongolia's residential real estate market conditions.

As a result of this research, it will be possible to contribute to making digital transformation a reality in the property valuation sector by introducing data-driven, scientifically-based, and objective methods into the valuation activities of residential real estate.

2 Theoretical Approach to Automated Valuation Models

When valuing real estate, one should focus on relevant and appropriate approaches. The valuation approaches are:

- Market approach
- Income approach
- Cost approach

All these approaches are based on economic principles of price equilibrium, return expectations, and investment. These approaches are considered traditional approaches to property valuation. The application of property valuation methodologies is multifaceted.

The accuracy of value conclusions for real estate—whether calculated through income, cost, or adjustment of comparable property reference prices in an open market—depends on how thoroughly multiple valuation approaches have been considered.

Therefore, the simultaneous application of various valuation methodologies depends on the professional skills of the appraiser [7].

Table 1. Natural zones with a coefficient of variation in precipitation exceeding 33%. (Source: A.Enkhsuren, [4])

Market approach	Income approach	Cost approach
Definition		
The market approach is based on direct **comparison** with other comparable properties that have recently been sold or are being offered for sale. The unique characteristics of the subject property (such as location, size, quality, development potential, etc.) are the main limitations when applying this method. This approach provides accurate value assessments especially in active, open markets.	The income approach is a valuation methodology that makes value judgments based on the assessment of potential profits and cash flows. This approach recognizes market conditions, potential income, cash flow, and time value of money. The appraiser needs to consider the operating environment of the property to determine market income growth.	The cost approach is used when information about properties **with** similar highest and best use and similar improvements is scarce or unavailable in the open market. Using this approach, the appraiser calculates the depreciated replacement cost of the real estate and adds the value of the land on which the building is constructed to arrive at the total value of the property.
Properties suitable for valuation		
Land (urban planning and agricultural land) Residential real estate (private houses, apartment buildings) - Commercial real estate (shopping centers, office buildings, industrial/logistics centers, service buildings, hotels, resorts, restaurants, gas stations, etc.) - only in liquid markets with sufficient comparable evidence	• Commercial real estate (shopping centers, office buildings, industrial/logistics centers, service buildings, hotels, resorts, restaurants, gas stations, etc.) • Investment properties • Income-producing assets • Properties associated with business value	• Commercial real estate (shopping centers, office buildings, industrial/logistics centers, service buildings, hotels, resorts, restaurants, gas stations, etc.) - only when there are insufficient comparable properties in the open market • Special purpose properties (purpose-built hospitals, law enforcement buildings and facilities, special purpose industrial/logistics centers, specialized service buildings, etc.) • Newly constructed or fully renovated properties • Properties with unique characteristics for which there are no market comparables

Dr. S. Dorjsuren [7] noted in his work "Property Valuation Methodology" that valuation methodologies are developed based on research into the economic interests and behaviors of real estate owners and market participants, and are constantly being

enriched. This is also mentioned in the "Research on Modern Valuation Methods for Residential Real Estate" conducted by the European Automated Valuation Association in 2022 [6]. Therefore, we can consider the modern approach as one that develops traditional valuation approaches based on current economic conditions, behavioral patterns, advancements in information technology, artificial intelligence, and data analytics.

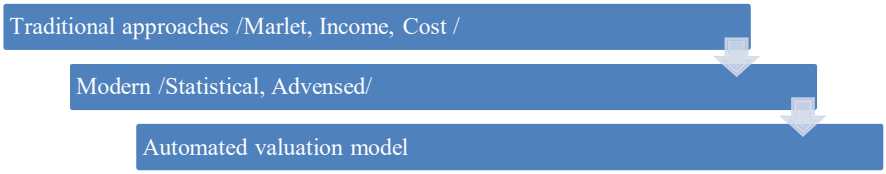

Figure 1. Real estate valuation approaches

The above valuation approaches have methods with different applications. In property valuation, the purpose of selecting the approaches and methods is to identify the most appropriate method for the given circumstances.

There is no single method that suits all possible conditions. The process of selecting a method should at least take the following into account, according to the International Valuation Standards [5]:

1. The appropriate basis (bases) for determining value and the preliminary conditions of value, as defined by the terms of the valuation contract and objectives.
2. The potential approaches and methods of valuation, along with their respective advantages and disadvantages.
3. The suitability of each method, determined by the characteristics of the asset, and the approaches or methods used by the relevant market participants.
4. The availability of reliable and accessible information required to apply the method(s).

The most widely used traditional valuation methodology is the market approach. The market-based valuation approach is most commonly used in the valuation of real estate and land. The market approach is considered the most suitable when it is possible to obtain comparable real estate sales prices in the region. In cases where the value of real estate can be determined based on income generated in future years, the income approach is considered appropriate. For real estate with rental income, such as business centers, offices, and apartments, the income approach is preferred. However, for real estate with uncertain income, special-purpose properties, or those not commonly sold in the open market, the cost approach is typically used for valuation.

We have conducted a detailed study of the traditional and modern approaches to property valuation methods and have provided a summary of which methodology is most suitable for use on different types of real estate.

Table 2. Comparison of Traditional and Modern Approaches to Real Estate Valuation

(Source: A.Enkhsuren, [4])

Traditional Valuation Approach	Modern approach
Definition	
Traditional real estate valuation methods are based on physical inspection of the property, comparative sales analysis, and valuation of land and buildings. Traditional valuation uses cost, market comparison, and income approaches to arrive at a value conclusion. Modern valuation approaches build upon the methodologies of traditional approaches but incorporate large datasets, information technology, and mathematical tools to derive value conclusions.	Modern valuation approaches are based on the methodologies of traditional approaches, utilizing large datasets, information technology, and mathematical tools to derive value conclusions.
Advantages	
The traditional approach allows for estimating value-enhancing improvements, physical depreciation, functional and economic obsolescence, and the impact of Environmental, Social, and Governance (ESG) factors through a physical inspection of the property. It also enables the incorporation of future market trends into property valuation, supporting investment planning and development strategies. Modern approaches improve the speed and accuracy of valuation activities. The modern approach provides access to large-scale, real-time data such as market trends, transaction history, and geospatial information. This enables more detailed and comprehensive analysis of the real estate market value. Modern real estate valuation approaches save time and resources compared to traditional methods, reduce human error, and provide more accurate valuations.	Modern approaches allow for increased speed and accuracy in valuation activities. Modern approaches provide access to large-scale, real-time data such as market trends, historical transaction records, and maps. This enables a more detailed and comprehensive analysis of the real estate market value. Modern valuation approaches save more time and resources compared to traditional methods, enable more accurate valuations, and reduce human error.
Disadvantages	
Although time-tested, these approaches are time-consuming and depend on the subjectivity of the appraisers.	Digitalizing the valuation process introduces the disadvantage that the accuracy and security of the data used during the valuation process may reduce the reliability of the valuation results. The reduction in the need for physical inspection of assets leads to the disadvantage of not accounting for improvements that increase the value of real estate, physical wear and tear, operational and economic obsolescence, and the impact of environmental and social factors.
Used Materials	
Based on the experience of the valuation expert, the valuation expert uses appropriate calculation programs.	Detailed algorithms are used to directly assess the value of real estate.

Table 3. Comparison of Modern Approaches in Property Valuation Methods

(Source: A.Enkhsuren, [4])

Nominal	Multiple Regression	Hedonic Valuation
	Advantages	
The values of the parameters that define the distribution of value are used. To calculate the values of the parameters, the factors influencing the value are formulated, and the coefficients are calculated for each real estate property. By using the indicator that contains the relative value of the real estate, the current value of the property is assessed.	This method uses the Gaussian "Least Squares Method." The multiple regression method applies a multiple regression model by including a large number of independent variables that influence the dependent variables in the analysis. These dependent variables are often the result of one or more causes. Multiple variables can combine to influence a single dependent variable. Therefore, it is recognized as a method that can be used in real estate valuation.	This method is based on the Hedonic Price Theory. According to the Hedonic Price Theory, if goods are considered heterogeneous, the price of each good is the sum of its independent characteristics or attributes. Each qualitative characteristic is considered as a good or service with its own value. Hedonic valuation is a model used to calculate the indirect price of factors such as structure, economy, environment, and social factors that affect real estate value. In the hedonic valuation equation, the degree and direction of influence of multiple factors are evaluated based on the characteristics and measures of the model's parameters.
	Common Use	
It is commonly used for bulk valuation of real estate for tax purposes.	It is frequently used for valuing real estate that is commonly traded on the open market.	It is often used for calculating real estate price indices.

Table 4. Comparison of Modern Approaches in Property Valuation Methods

(Source: A.Enkhsuren, [4])

Artificial Intelligence	Fuzzy logic	Spatial analysis
Advantages		
Artificial neural networks are computer systems that perform the core function of learning, which is a fundamental characteristic of the human brain. In this method, the learning process is calculated using examples (nodes). These networks consist of interconnected artificial neurons, with each connection having a weight. These weights contain information in the artificial neural network and are propagated through the network.	Fuzzy logic is a form of machine learning. It uses expressions such as partial "true" and partial "false" to represent the relationships between factors. In this method of valuation, the physical characteristics of the property are not required; instead, the economic and social interactions are considered important.	Spatial analysis is a form of machine learning. When using Geographic Information Systems (GIS), it improves the analytical capabilities of location measurements and variable relationships.
Common Use		
It is commonly used for valuing real estate that is frequently traded on the open market.	It is often used in real estate management decision-making, especially in markets with insufficient data or information sources.	It is commonly used in real estate brokerage activities, providing preliminary value estimates.

Table 5. Selection of Methods for Applying Traditional and Modern Approaches in Real Estate Valuation

(Source: A.Enkhsuren, [4])

Category	Modern approach					
	Nominal	**Multi regression**	**Hedonic valuation**	**Neural network**	**Fuzzy logic**	**Spatial analysis**
Agriculture /Pastureland, Agricultural land /	It is necessary to precisely identify the factors affecting value /For assets frequently traded on the market/	A regression model is used by incorporating multiple independent variables affecting the real estate-related variables. /When sufficient actual transaction and sales data is available/	It is used to estimate the implicit price of structural, economic, environmental, and social factors affecting the value of real estate.	A neural network expresses patterns by recognizing categories, filtering errors, and optimizing data. /When sufficient data is available/	Using approximate modeling with the help of uncertain logic, conclusions about the value are drawn through simple mathematical solutions. (In cases of limited data)	Location impact will be calculated based on Geographic Information Systems. (It will be based on maps.)

Residential /Public housing, Private housing/	There is a need to precisely identify the factors affecting the price /Property types commonly found in financial statements.	A regression model will be applied by including numerous independent variables that affect the variables related to financial statements.	It will be used to calculate indirect costs affecting the value of financial statements related to the structure, economy, environment, and social factors.	Artificial neural networks will be used to recognize categories, patterns, filter errors, and optimize data when data is sufficient.	Approximate modeling of any issue will be done using uncertain logic, and conclusions on value will be made through simple mathematical solutions when data is scarce.	Location impact will be calculated based on geographic information systems. /It will be based on maps.
Commercial Real Estate (Office) /Private office, Investment/unfinished, Rental sections/	It is necessary to precisely identify the factors affecting the price /Property types commonly found in financial statements.	A regression model will be applied by including numerous independent variables affecting the variables related to financial statements/ When real asset and sales data are sufficient/	*It will be used to calculate indirect costs affecting the value of financial statements related to structure, economy, environment, and social factors.*	Artificial neural networks will be used to recognize categories, patterns, filter errors, and optimize data when data is sufficient.	Approximate modeling of any issue will be done using uncertain logic, and conclusions on value will be made through simple mathematical solutions when data is scarce.	Location impact will be calculated based on geographic information systems. /It will be based on maps.
Commercial Real Estate (Services) /Sports facility, Cultural/arts, Retail, Education, Medical, Hotel, Restaurant, Bank/	It is necessary to precisely identify the factors affecting the price /Property types commonly found in financial statements.	A regression model will be applied by including numerous independent variables that affect the variables related to financial statements. / When real asset and sales data are sufficient/	It will be used to calculate indirect costs affecting the value of financial statements related to structure, economy, environment, and social factors.	Artificial neural networks will be used to recognize categories, patterns, filter errors, and optimize data when data is sufficient.	Approximate modeling of any issue will be done using uncertain logic, and conclusions on value will be made through simple mathematical solutions when data is scarce.	Location impact will be calculated based on geographic information systems. /It will be based on maps.

Commercial Real Estate (Industrial) /Auto garage, Light/heavy industrial, Mining, Mechanical repair, Outdoor roads/small buildings/	It is necessary to precisely identify the factors affecting the price /Property types commonly found in financial statements.	A regression model will be applied by including numerous independent variables that affect the variables related to financial statements.	It will be used to calculate indirect costs affecting the value of financial statements related to structure, economy, environment, and social factors.	Artificial neural networks will be used to recognize categories, patterns, filter errors, and optimize data when data is sufficient.	Approximate modeling of any issue will be done using uncertain logic, and conclusions on value will be made through simple mathematical solutions when data is scarce.	Location impact will be calculated based on geographic information systems. /It will be based on maps.
Outdoor roads, landscaping, small buildings /Fences/gates, Parking, Small warehouse, Landscaping, Embankment /	It is necessary to precisely identify the factors that affect the price /Properties commonly found in financial statements.	A regression model will be applied by including numerous independent variables that influence the variables related to financial statements. /When real asset and sales data are sufficient/	It will be used to calculate indirect costs affecting the value of financial statements related to structure, economy, environment, and social factors.	Artificial neural networks will be used to recognize categories, patterns, filter errors, and optimize data when data is sufficient.	Uncertain logic will be used to approximately model any issues, and value conclusions will be made through simple mathematical solutions when data is scarce.	Location impact will be calculated based on geographic information systems. /It will be based on maps.
Engineering systems and infrastructure /Heating/sewer/water/comm/electrical systems/	It is necessary to precisely identify the factors that affect the price /Properties commonly found in financial statements.	A regression model will be applied by including numerous independent variables that affect the variables related to financial statements.	It will be used to calculate indirect costs affecting the value of financial statements related to structure, economy, environment, and social factors.	Artificial neural networks will be used to recognize categories, patterns, filter errors, and optimize data when data is sufficient.	Uncertain logic will be used to approximately model any issues, and value conclusions will be made through simple mathematical solutions when data is scarce.	Location impact will be calculated based on geographic information systems. /It will be based on maps.

Explanation *: -The first recommended, - The second recommended, - Very few will be recommended

Most Automated Valuation Models (AVMs) calculate the value of a property for a specific period by analyzing the values of comparable characteristics. However, AVMs are unable to model in detail abnormal conditions such as significant property damage,

market crises, excessive profitability, or operational obsolescence. Therefore, when choosing to use this model, values can select the type of property it will be applied to, considering factors such as significant property damage, uniqueness, or complexities, and can make adjustments accordingly. AVMs are all based on the market comparison approach but are most effective when there is a large amount of sales data. The market comparison approach is based on the principle of substitution, which is central to identifying the difference between the best comparable market price and the utility of the property being valued. Recent research (867 studies) shows that the methodologies used in Automated Valuation Models build upon traditional valuation approaches, incorporating modern economic conditions, behaviors, information technology advancements, artificial intelligence, data, and modern valuation trends. These trends also include the use of machine learning methodologies in experimental studies.
The results of research studies using Automated Valuation Models (AVMs) show that machine learning algorithms and the outcomes of training demonstrate good performance in predicting real estate valuations and make it possible to explain the models clearly.

The results of research studies using Automated Valuation Models (AVMs) show that machine learning algorithms and the outcomes of training demonstrate good performance in predicting real estate valuations and make it possible to explain the models clearly.

Table 6. Overview of Automated Valuation Models and Real Estate Research Studies

(Source: A.Enkhsuren, [4])

№	Keywords	Research Object and Methodology
1	Artificial intelligence, residential real estate, data	**Object:** Urban and suburban real estate intended for residential purposes. **Research Methodology:** Multiple factor linear regression, decision trees, random forests, hedonic model, artificial neural networks.
2	Neural networks, commercial property, Real estate-Modern methods	**Object:** Commercial residential properties, smart real estate. **Research Methodology:** Multiple factor linear regression, artificial neural networks, fuzzy neural networks.
3	Big data, machine learning, real estate index, real estate valuation methods, information systems	**Object:** Urban and suburban real estate intended for residential purposes, commercial residential properties. **Research Methodology:** Linear regression, random forests, Support Vector Machine (SVM).
4	Mass valuation methods, valuation models, multiple regression	**Object:** Tax valuation, commercial residential properties, land valuation. **Research Methodology:** Linear regression, decision trees, random forests, hedonic model, Gradient Boosting.
5	Advanced valuation methods, statistics valuation methods, traditional valuation methods	**Object:** Residential real estate. **Research Methodology:** Multiple factor linear regression, decision trees, random forests, hedonic model, artificial neural networks.

3 Experiment and results of valuating residential real estate using an automated valuation model

In conducting the research experiment, the "Standard on Ratio Studies - 2013" approved by the International Association of Assessing Officers (IAAO) defined the acceptable margin of error for small residential real estate markets as 5-20% [1-3]. When the model's error results are within standardized limits, the assumption that the model can predict real estate values is made, with the goal of keeping the error probability below 10%.

To evaluate the model's prediction accuracy and overall performance, an independent test data set was used. The statistical results of the model's performance metrics are essential for assessing the accuracy of the model's predictions. We checked the errors and compiled statistical indicators for the aggregated data in order to evaluate residential real estate using an Automated Valuation Model. As of June 2024, the data for residential real estate consists of 7,499 units in Ulaanbaatar that are listed for sale.

Table 7. Statistical Indicators

Statistics

		Room	District	Balcony	Year of construction	Garage	Building floors	Area	Floor	Financial transaction	Window	ln total price
N	Valid	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Missing	7	7	7	7	7	7	7	7	7	7	7
Mean		2.62	3.12	1.06	2015.65	1.22	12.20	73.20	6.87	1.74	3.57	19.3145
Std. Error of Mean		.009	.016	.006	.109	.005	.053	.358	.051	.006	.017	.00684
Median		3.00	4.00	1.00	2018.00	1.00	12.00	68.00	6.00	2.00	3.00	19.2360
Mode		3	4	1	2024	1	16	80	3	2	3	18.83
Std. Deviation		.785	1.344	.485	9.414	.411	4.561	31.010	4.403	.480	1.459	.59196
Minimum		1	1	0	1980	1	1	17	1	1	1	17.37
Maximum		4	8	3	2024	2	25	312	25	3	8	21.29
Sum		19644	23367	7932	15115351	9112	91491	548946	51506	13062	26754	144839.78

***Explanation of the Used Indicators**

Total Price: The total price offered for the real estate on the market.

Area: The total area of the real estate.

Year of Construction: The year the real estate or building was put into operation.

Room Count: The total number of rooms in the real estate (including the kitchen).

District: The location is expressed by the "District," a unit of administrative division in Ulaanbaatar.

Balcony: The number of balconies in the real estate.

Garage: The number of heated or cold garages.
Floor Number: The floor number of the real estate or building.
Window Count: The number of windows in the real estate.
Financial Transaction: Bank long-term loans, loans between individuals or legal entities, and cash payments.
Building Floors: The total number of floors in the real estate or building.

Therefore, we conducted an experiment using Artificial Intelligence-enhanced AVMs for residential real estate in Ulaanbaatar and compared the performance of the models.

Table 8. Correlation Analysis

Correlations											
		ln_total_price	Room	Area	District	Window	Year of construc-tion	Garage	Financial transact-tion	Building floors	Balcony
Pearson Correlation	ln_total_price	1.000	.737	.866	.254	.002	.213	.447	.029	.315	.187
	Room	.737	1.000	.782	.159	.002	.087	.369	.045	.143	.271
	Area	.866	.782	1.000	.208	.006	.148	.502	.034	.121	.258
	District	.254	.159	.208	1.000	-.009	.145	.147	.022	.071	.065
	Window	.002	.002	.006	-.009	1.000	-.011	-.003	-.041	-.006	.014
	Year of construction	.213	.087	.148	.145	-.011	1.000	.095	-.008	.551	-.046
	Garage	.447	.369	.502	.147	-.003	.095	1.000	.006	.025	.163
	Financial transaction	.029	.045	.034	.022	-.041	-.008	.006	1.000	-.005	.027
	Building floors	.315	.143	.121	.071	-.006	.551	.025	-.005	1.000	-.022
	Balcony	.187	.271	.258	.065	.014	-.046	.163	.027	-.022	1.000
Sig. (1-tailed)	ln_total_price	.	.000	.000	.000	.443	.000	.000	.006	.000	.000
	Room	.000	.	.000	.000	.418	.000	.000	.000	.000	.000
	Area	.000	.000	.	.000	.309	.000	.000	.002	.000	.000
	District	.000	.000	.000	.	.227	.000	.000	.026	.000	.000
	Window	.443	.418	.309	.227	.	.163	.404	.000	.299	.119
	Year of construction	.000	.000	.000	.000	.163	.	.000	.246	.000	.000
	Garage	.000	.000	.000	.000	.404	.000	.	.311	.016	.000
	Financial transaction	.006	.000	.002	.026	.000	.246	.311	.	.320	.009
	Building floors	.000	.000	.000	.000	.299	.000	.016	.320	.	.026
	Balcony	.000	.000	.000	.000	.119	.000	.000	.009	.026	.

N	ln_total_price	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Room	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Area	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	District	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Window	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Year of construction	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Garage	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Financial transaction	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Building floors	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499
	Balcony	7499	7499	7499	7499	7499	7499	7499	7499	7499	7499

Strong Positive Correlations:

- There is a very strong positive correlation (0.866) between the total price and area.
- There is a strong positive correlation (0.737) between the total price and the number of rooms.
- There is also a strong positive correlation (0.782) between the area and the number of rooms.
- There is a moderate positive correlation (0.551) between the number of floors and the year of completion.
- There is a moderate positive correlation (0.447) between the garage and the total price.

Weak Correlations:

- The number of windows has very weak correlations with all other variables (values close to 0).
- Financial transactions have very weak correlations with most variables.
- The balcony has relatively weak correlations (between 0.187-0.271).

Statistical Significance Level (Sig.):

- Most correlations are statistically significant ($p < 0.001$).
- Some correlations related to the number of windows are not statistically significant ($p > 0.05$).

Sample Size:

- All variables have 7,499 observations.
  Since the sample size (N = 7499) is sufficiently large, the results can be considered reliable.

From the analysis, the factors that have the greatest impact on housing prices are the area and the number of rooms. However, indicators such as the number of windows and financial transactions do not have a significant impact on the price.

Experiment Results

We conducted the calculations of the experiment using Python programming and compared the performance between different models using the Root Mean Squared Error (RMSE) algorithm.

Table 9. Comparative Analysis of Validated Experimental Results

Tested Models	RMSE	Result
Stacking Regressor	0.0594	Combined advantages of multiple models to minimize error
Random Forest	0.0643	Best-performing individual model
Gradient Boosting	0.0682	Close to Random Forest in performance
Lasso Regression	0.2537	Potential overfitting issues
Ridge Regression	0.3299	Basic model with moderate performance
Linear Regression	0.3333	Baseline model with the weakest performance
SVR	0.4396	Poor suitability for this dataset
ANN	0.067	Shows promise with lower error in testing.

Summary of Results:

1. **Stacking Regressor (RMSE: 0.0594) –** Achieved the **lowest error** by combining the strengths of different models, resulting in superior predictive performance.
2. **Random Forest (RMSE: 0.0643) –** The **best-performing standalone model**, likely serving as a key base model in the Stacking Regressor ensemble.
3. **Gradient Boosting (RMSE: 0.0682) –** Delivered **comparable results** to Random Forest, demonstrating strong predictive capability.
4. **Lasso Regression (RMSE: 0.2537) –** While the **best among simple models**, it significantly underperformed compared to ensemble methods, possibly due to overfitting.
5. **Ridge Regression (RMSE: 0.3299) –** Produced **moderate results**, performing better than Linear Regression but still lagging behind advanced techniques.
6. **Linear Regression (RMSE: 0.3333) –** As a **basic linear model**, it failed to capture complex data relationships, resulting in the **poorest performance among regression models**.
7. **SVR (Support Vector Regression) (RMSE: 0.4396) –** The **worst-performing model**, indicating poor suitability for this dataset.
8. **ANN (Artificial Neural Network) (RMSE: 0.067) –** Showed **promising accuracy** in testing, suggesting potential for further optimization and application.

The **Stacking Regressor**—an **ensemble learning method** that integrates multiple regression models—effectively learned from the predictions of **Random Forest** and **Gradient Boosting**, adjusting weight coefficients to minimize errors and produce **optimal forecasts**.

Based on these findings, we recommend using:

1. **Stacking Regressor** for the highest accuracy,
2. **Random Forest** and **Gradient Boosting** as reliable standalone alternatives.

4 Conclusion

Internationally, the study and development of Automated Valuation Models (AVMs) have become increasingly active in recent years, as observed during the research process. For example, in highly developed countries, AVMs are widely used for valuing real estate for purposes such as collateral, taxation, transactions, and insurance, leading to significant time and cost savings. Meanwhile, for intellectual property, business assets, and extractive resources, specialized appraisers continue to play a key role, focusing on integrating these assets into economic circulation and attracting investment.

Globally, the next stage of AVM development is being defined as leveraging artificial intelligence to evaluate real estate based on photographic images, extracting data points from these images, and combining quantitative and qualitative indicators to estimate property values.

Developing and applying AVMs to certain types of valuations can offer economic and time-saving benefits. Examples include valuations for property taxes, collateral for bank loans, legal disputes, and insurance purposes.

In conclusion, establishing a property valuation database in Mongolia and utilizing AVMs and machine learning to assess property is feasible and recommended.

References

1. International Association of Assessing Officers, Standard on Automated Valuation Models. IAAO (2018)
2. International Association of Assessing Officers, Standard on Mass Appraisal of Real Property. IAAO (2024)
3. International Association of Assessing Officers, Standard on Ratio Studies. Chicago. IAAO (2013)
4. Enkhsuren, A., Using automated valuation models for real estate valuation. PhD dissertation, Business Administration (2025)
5. International Valuation Standards Council, International Valuation Standards. IVSC (2020)
6. Royal Institution of Chartered Surveyors: Automated valuation models (AVMs): Implications for the profession and their clients (2022)
7. Dorjsüren, S., Ül khödlökh khöröngiin ünelgee. [Real estate valuation] Ulaanbaatar (2005) (in Mongolian)