GEOMATICS AND ENVIRONMENTAL ENGINEERING • Volume 20 • Number 4 • 2026

https://doi.org/10.7494/geom.2026.20.4.33

A Novel Housing Price Estimation Model Integrating GIS and DNN: A Case Study of Istanbul

Yakup Emre Çoruhlu[1], Mustafa Dihkan[1], Okan Yildiz[1], Mehmet Özgür Çelik[2]*, Hossein Azadi[3]

[1] Karadeniz Technical University, Geomatics Engineering, Trabzon, Türkiye
[2] Mersin University, Geomatics Engineering, Mersin, Türkiye
[3] University of Liège, Department of Geography, Liège, Belgium
* Corresponding author

Abstract: Housing valuation is a concrete reflection of socio-economic inequalities in urban space. Particularly in densely populated, spatially fragmented large cities like Istanbul, the current official mass appraisal system used for property taxation fails to reflect market reality. This situation results in revenue losses in property taxes and spatial injustices. This study develops a Deep Neural Network (DNN)-based model that integrates spatially derived variables from Geographic Information Systems (GIS) and incorporates 24 objective variables related to location, structure, and access in Istanbul. The model, trained on 3,757 samples created using open-source big data, estimated housing values with high accuracy (R^2 = 0.979). The findings show that spatial differences in housing values are strongly related to urban variables such as accessibility and proximity to infrastructure. This approach not only produces housing value estimates but also provides a theoretical and methodological framework for spatial analyses of how value is produced in urban space. The study has the potential to support the development of a fair, transparent, and updatable mass appraisal system, especially for developing cities.

Keywords: house price estimation model, mass appraisal, property tax, market value, real estate, DNN, GIS

Received: November 6, 2025; accepted: April 7, 2026

E-mails & ORCID iDs: yecoruhlu@ktu.edu.tr, https://orcid.org/0000-0002-8673-603X (Y.E.Ç.);
mdihkan@ktu.edu.tr, https://orcid.org/0000-0002-0027-236X (M.D.);
okan.yildiz@ktu.edu.tr, https://orcid.org/0000-0002-7664-5361 (O.Y.);
mozgurcelik@mersin.edu.tr, https://orcid.org/0000-0003-4569-888X (M.Ö.Ç.);
h.azadi@uliege.be, https://orcid.org/0000-0003-4363-9405 (H.A.).

1. Introduction

Urban housing markets play a central role in understanding spatial inequalities and socio-economic segregation within cities [1, 2]. Real estate valuation processes are decisive in shaping fiscal policies, land-use planning, and urban development strategies [3–7]. However, especially in rapidly urbanizing and developing economies, official valuation mechanisms often fail to reflect real market dynamics, leading to significant deviations between administrative and market values. This situation causes financial losses for municipalities due to undervalued properties and unequal tax burdens for citizens, resulting from inconsistent valuations [8, 9].

In response to this problem, new methods that integrate spatial data with machine learning (ML) are being developed to provide more accurate, geographically sensitive real estate valuations [10–15]. To contribute to this emerging body of literature, this study proposes a deep-learning-based valuation model integrated with Geographic (DNN)-based model integrated with GIS using 24 objectively measurable variables based on location, structure, and accessibility in Istanbul. The model aims to close the gap between existing official valuation systems and market realities while offering new possibilities for location-sensitive tax policies and planning practices.

In Türkiye, as in many developing countries, the basis for property tax is an official mass appraisal system administered by municipal and central government commissions. This system typically relies on static parameters such as average values per street or across broad neighborhood zones, along with standardized construction cost catalogs, and is updated infrequently (e.g., every four years). Consequently, it fails to capture the fine-grained spatial dynamics and real-time fluctuations in the housing market, leading to a substantial gap between taxable and market values. This gap results in significant fiscal losses for municipalities and raises concerns about spatial equity in taxation.

There are various methods and algorithms for estimating property values for taxation and transaction purposes. For example, advanced machine-learning algorithms have been widely used to improve the prediction of taxable asset values [16, 17]. The most prominent ones are support vector machines, random forests, decision trees, Bayesian methods, genetic algorithms, Markov prediction models, artificial neural networks, deep neural networks, and others [18–21]. In fact, some algorithms can achieve high levels of accuracy in predicting real estate prices [22], and higher levels of accuracy are achieved using artificial neural networks and deep neural networks [23]. Given spatial differences between regions, various criteria are used to create forecasting models at national or regional scales [24]. Neural-network algorithms can achieve high predictive accuracy when supported by a sufficient number of samples. In addition, to meet the large data requirements of these models, automated data collection algorithms that leverage web services used for real estate sales and rentals have been developed in various studies [25, 26]. With these online services, a large amount of sample data that can address the needs of network models is provided [27]. Moreover, more reliable housing-price forecasts can be obtained with larger sample sizes [18].

The inadequacy of the technical infrastructure in the real estate tax system indicates the need to establish public registries using information technology and to move towards mass property valuation [28]. Real estate valuation is an interdisciplinary process that requires considering multiple variables, depending on the property's location [29]. In this context, there are various strategies for estimating real estate prices. For example, large datasets are stored in web services that host real estate buying and selling advertisements. Data entry into these portals is performed by real estate agents, property owners, contractors, banks, or real estate brokers [30]. Although the prices entered into the system are not final transaction prices and reflect the seller's expectations, they are the most accurate data available for use in the valuation process.

Real estate valuation activities, carried out by many different institutions and organizations, may affect a wide range of transactions, ranging from foreclosures to easements, from leases to judicial proceedings. Although institutions seek to ensure internal consistency in their valuation standards, there is no legal standard in Türkiye governing real estate valuation criteria. Real estate valuation became a regulated profession in Türkiye with the Capital Markets Board's notification published in the Official Gazette on August 12, 2001. Later, with Law No. 5582 in 2007, the profession was further specialized under the Capital Markets Board. In 2014, the Turkish Appraisers Association (TAA) was established as a self-regulatory professional association with public legal personality in the field of real estate valuation. In the Capital Markets Board (CMB) circular, compliance with international standards is required under the *Statement of Valuation Standards in the Capital Markets* published by TAA and the Turkish Capital Markets Association. The content of the valuation report is also regulated [31]. When conducting and reviewing valuations, different results, even for similar real properties with the same characteristics, may also arise from the inherently subjective nature of the process. On the other hand, it may be possible to arrive at different real estate values in valuation processes conducted by different institutions for different purposes, based on different rules and criteria [32].

Based on studies [33, 34], housing-price estimation has been discussed in the literature using various criteria, methods, and approaches. The following notable studies and the criteria used in them are summarized in Table 1.

Table 1. Housing price estimation studies in the literature

Criteria	Representative variables	References
Age of the structure [years]	–	[18, 27, 34–37, 39, 40, 42–48]
Apartment gross area [m^2]	–	[18, 27, 34, 35, 36, 39, 41, 42, 45, 46, 48]
Apartment net area [m^2]	–	[17, 34, 41]
Cable TV	–	[36, 43]

Table 1. cont.

Criteria	Representative variables	References
Cooling type	split air conditioner	[37, 41]
Cooling type	central air conditioner	[37, 41]
Cooling type	heat pump	[37, 41]
Cooling type	ceiling fan	[37, 41]
Elevator (yes/no)	–	[17, 18, 36, 37, 39, 45, 46]
Floor type of the bathroom	screed	[36, 39, 43]
Floor type of the bathroom	floor ceramics	[36, 39, 43]
Floor type of the bathroom	vinyl flooring	[36, 39, 43]
Floor type of the bathroom	mosaic	[36, 39, 43]
Floor type of the living room	parquet floor	[36, 39]
Floor type of the living room	wood floor	[36, 39]
Floor type of the living room	floor ceramics	[36, 39]
Floor type of the living room	vinyl flooring	[36, 39]
Floor type of the living room	screed	[36, 39]
Floor type of the living room	carpet, mosaic, and marble	[36, 39]
Floor type of the living room	parquet floor	[36, 39]
Floor type of the living room	wood floor	[36, 39]
Floor type of the living room	floor ceramics	[36, 39]
Floor type of the living room	vinyl flooring	[36, 39]
Floor type of the living room	screed	[36, 39]
Floor type of the living room	carpet and marble	[36, 39]
Floor type of the living room	wood floor	[36, 39]
Floor type of the living room	floor ceramics	[36, 39]
Floor type of the living room	vinyl flooring	[36, 39]
Floor type of the living room	screed	[36, 39]
Floor type of the living room	carpet, mosaic, and marble	[36, 39]
Garage (yes/no)	–	[34, 36–38, 41, 45]
Garbage disposal	–	[36]
Heating system	bottled gas propane	[18, 41, 43]
Heating system	electricity	[18, 41, 43]
Heating system	natural gas	[18, 41, 43]

Table 1. cont.

Criteria	Representative variables	References
Hot water	–	[36]
Jacuzzi	–	
Locational feature	in urban area	[35, 36]
	in rural area	
Number of rooms [count]	–	[18, 27, 34, 36, 39, 40, 42, 43, 45, 46]
Sauna	–	[36]
Toilet	–	[36]
Type of house	independent (detached) house	[36–42]
	basement floor	
	apartment floor	
	slum flat	
	duplex	
Type of the structure	reinforced-concrete structure	[18, 34–36, 39, 41, 45]
	timber structure	
	briquette structure	
	stone structure	
	brick structure	
	adobe structure	
	electricity	
	natural gas	
Water system	–	[36]

Urban real estate valuation is at the intersection of spatial justice, fiscal governance, and the shaping of socio-economic inequalities in urban space [49, 50]. Especially in rapidly urbanizing developing countries, official valuation systems are often based on static, outdated, and spatially insensitive methods [51]. This leads to housing values failing to realistically reflect intra-city differences, such as differences associated with proximity to transportation infrastructure, public services, or environmental amenities [52–54]. In multi-centered, dense, and socially segregated metropolitan areas, such as Istanbul, the gap between official property values and market realities both deepens spatial inequalities and creates injustice and revenue loss in the tax system [55].

In Türkiye, the official property tax valuation system is based on mass appraisals conducted every four years, in which minimum unit values are assigned to streets or avenues without accounting for fine-grained spatial variations. This creates a significant gap between official tax values and actual market values, leading to inequitable tax burdens and revenue losses. This study focuses on estimating the market value of residential properties using a data-driven approach, aiming to provide an alternative to the traditional, spatially insensitive valuation methods.

The current official property tax system in Türkiye relies on street- or avenue-based general valuation definitions, assigning minimum unit values to entire streets or avenues while ignoring micro-locational variations. This approach assumes urban space to be homogeneous and fails to capture the relational and context-dependent nature of urban value, leading to both technical and social problems [49, 56, 57].

The aim of this study is to develop a method to estimate housing values in Istanbul with high spatial resolution and objectivity, using a deep learning model integrated with Geographic Information Systems (GIS) and leveraging big data as an alternative to traditional valuation approaches. It also seeks to provide a theoretical contribution to understanding how value is produced in urban space and to establish a methodological basis for a fair, transparent, and updatable valuation system for use in urban governance.

The originality of this study lies in integrating GIS-based spatial variables into a DNN model explicitly designed for mass housing valuation in a complex metropolitan environment such as Istanbul. Unlike previous research, which often relied on limited or subjective criteria, our model uses 24 objective spatial and structural variables, rigorously validated through GIS-based consistency checks. This approach achieves high prediction accuracy ($R^2 = 0.979$) and offers a scalable, updatable framework for practical applications in property taxation and urban governance, representing a significant advancement in the context of developing countries.

2. Background Information

Real estate and property taxes were used in Egypt, Babylon, Persia, China, and other parts of the ancient world [58]. The practice of collecting real estate tax in proportion to the size and yield of a property originated in earlier periods of history through the determination of property boundaries and owners and the calculation of their areas through land registry and cadastre studies [59]. Seljuk and Ottoman land systems also collected taxes directly or indirectly by identifying real estate in a similar manner [60]. Real estate taxation, a practice with historical roots, continues to be applied in nearly all countries today.

In Türkiye, the current property tax system operates on a street-by-street and plot-by-plot basis, with transactions overseen by property tax determination and assessment commissions [61]. However, the institutional and legal framework of Türkiye's real estate valuation system exhibits several shortcomings. The technical infrastructure of the property tax system is inadequate, indicating a clear need for modernized public registries using information technology and a shift toward mass property valuation [28]. These structural deficiencies result in significant discrepancies between official tax values and market values, especially for non-standard properties [62].

Under the current property tax system in Türkiye, transactions are conducted on a street-by-street and plot-by-plot basis, and the property-tax determination and assessment commissions face several operational problems [61]. In the valuation of non-standard properties for taxation purposes, substantial disparities arise between market value and assessed tax value [62]. Real estate appraisal, which involves measuring real estate in monetary terms, is a commercial activity. This activity is governed by the legal framework established under Law No. 6585 on the Regulation of Retail Trade. With this regulation, an important step has been taken to ensure that real estate trade is carried out safely and systematically by registered businesses [63]. Article 11 of the regulation aims to facilitate the supervision and control of real estate trade by establishing the Real Estate Transaction Information System (RETIS). There are deficiencies in the real estate appraisal sector in terms of law and management. Mass appraisal may be beneficial in ensuring the time, cost, and fairness of valuation.

As shown in Table 1, various criteria have been considered in many different studies, some of which are more objective and tangible, others more subjective and interpretable (based on the appraiser's opinion). Usually, property tax on land in Türkiye is assessed every four years based on the street or avenue where the parcel is located. In contrast, property tax value is calculated by determining the minimum unit value per square meter of streets and avenues, despite heterogeneity in value. All these transactions are carried out through commissions consisting mainly of civil servants from different institutions. The commissions define the minimum property tax value for streets and avenues. Therefore, the property tax base for each parcel can be calculated. If a building is on the parcel, the building's value, which is the basis for property tax, is calculated using annual catalogs based on information such as the type of building and the year of construction, and other relevant characteristics. Under this method of calculating current property tax, market values and publicly available property purchase and sale prices are not taken into account, resulting in significant losses in property tax collection.

In the globalization of valuation activities, the Royal Institution of Chartered Surveyors (RICS), the Appraisal Foundation (TAF) in the United States, the European Group of Valuers' Associations (TEGOVA), and the International Valuation Standards Council (IVSC) have become recognized standard-setting and professional organizations [64].

3. Methods

This chapter describes the study area, data collection process, and the analytical methods employed for model development and evaluation.

3.1. Study Area

The study area was determined as Istanbul, the city with the largest housing stock and the highest number of houses for sale in Türkiye. Information (not personal data) on the web pages of commercial enterprises that mediate the purchase, sale, and rental of real estate was used in the data collection phase. In these open-source databases, it was determined that there were approximately 700,000 houses for sale in Türkiye and 120,000 in Istanbul in 2024 [65]. From this large dataset, some districts of Istanbul and some neighborhoods within these districts were selected as the study area, and a non-homogeneous dataset was obtained through random sampling [66, 67] as shown in Figure 1.

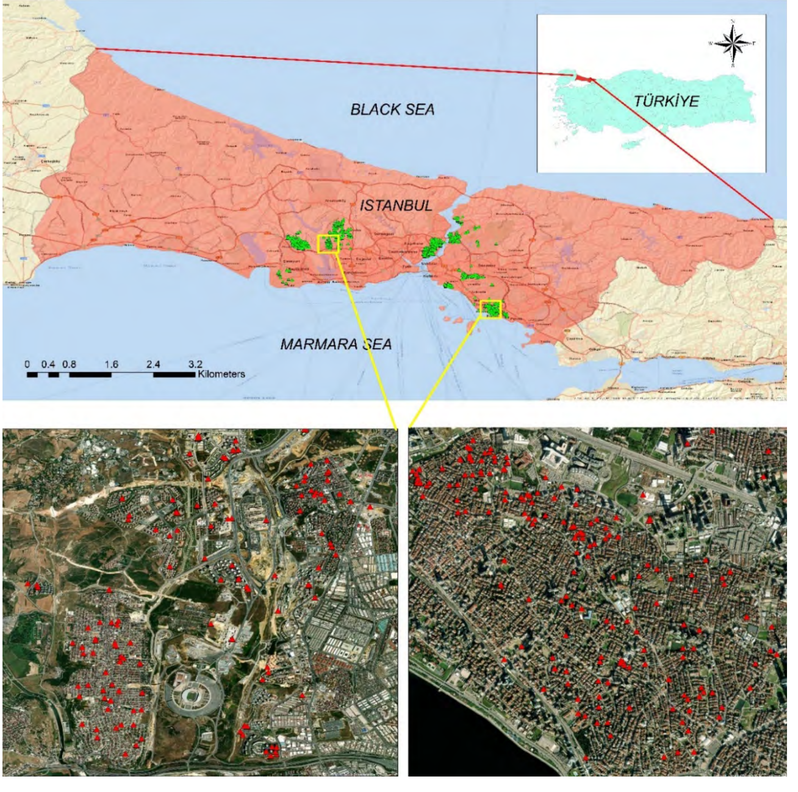

Fig. 1. Study area

This sampling approach ensured that our dataset was both spatially diverse and truly representative, thereby enhancing the quality of our analysis.

3.2. Data Collection

In the real estate valuation problem, property value is the dependent variable, and the criteria that constitute this value are the independent variables. In real estate valuation, the independent variables may vary depending on the characteristics of the property to be valued. What is important at this stage, in the context of this paper, is to eliminate subjective criteria and solve the valuation problem with purely metric and categorical variables. Several criteria are widely used in the preparation of real estate valuation reports, and there are no regulations that require the use of criteria. The most important consideration in real estate valuation is the correct and accurate determination of the property's value within a single valuation report. Therefore, in this study, the criteria presented in Table 2 were determined by considering the objective determination of real estate value through a method based on objective, metric criteria to the greatest extent possible. The selection was grounded on a comprehensive review of the real estate valuation literature and the authors' prior expertise, ensuring that only objective, measurable, and spatially explicit variables were included, thereby minimizing subjectivity.

Table 2. Criteria used for housing valuation

Metric criteria		Categorical criteria
housing locational / spatial features	Euclidean relations [m]	housing service features
– latitude [m] – longitude [m] – total number of floors – located floor – gross area [m^2] – net area [m^2]	– distance to religious facilities – distance to market / bazaar – distance to healthcare facilities – distance to educational facilities – distance to recreational areas – distance to urban transportation – distance to shopping centers / malls – distance to university – distance to intercity transportation	– within the site – security service – pool – fire escape – car park – sports facility – building age – elevator presence – number of rooms

The data utilized in this study were sourced from publicly accessible property listings on the Sahibinden.com website. The collection was conducted using web-scraping techniques for academic research purposes, adhering to the website's robots.txt guidelines. To uphold privacy and maintain ethical standards, all personal and sensitive data were excluded.

There are some potential inaccuracies (intentional or unintentional) in publicly accessible housing advertisements. In other words, in some advertisements, the property's location may be incorrectly reported, and in some advertisements, the property data may be manipulated. Moreover, in some advertisements, the visual appearance of the houses may be manipulated, and the property data may not be presented clearly. For this reason, in this research, advertisements created by property owners, contractors, or banks were considered, as they tend to be more transparent. On the other hand, the locations of the houses were checked using Google Earth tools to ensure the accuracy of the data contained in the advertisements. During data entry, the coordinate values provided by Google Earth were verified using the General Directorate of Land Registry and Cadastre (GDLRC) system, Parselsorgu [68]. By obtaining the house's actual location and the zoning or cadastral block information that overlaps with it, false or manipulated listings were removed from the dataset. The location of some houses overlapped with green spaces, roads, sports venues, etc., contrary to the actual situation, so these types of advertisements were examined and excluded from the study. To load points of attraction, which are metric evaluation criteria, into the GIS database, the determined criteria were first created as point data. Thus, the dependent variable, the price of houses for sale, and all independent variables were transferred to the GIS platform in the form of spatial, categorical, and textual data. Then, the Euclidean distances of the houses to the points of attraction were calculated. The Euclidean distances of all sample houses to the points of attraction were automatically calculated in the GIS platform, and the DNN method was used for analysis. In this study, we opted for Euclidean distance due to its simplicity and computational efficiency. However, future research could benefit from using network-based distance measures. This approach would provide a more nuanced understanding of real-world accessibility, particularly in topographically complex cities such as Istanbul, where the unique landscape and infrastructure may significantly influence how people navigate and access various locations.

3.3. Data Analysis

Artificial intelligence should be considered as a field that includes machine learning and deep learning [69]. These concepts are intertwined, and each has been popular in different periods. The goal of machine learning is to enable computers to learn from past events, in other words, from experience, and improve their performance for specific tasks [18]. Although there are various methods in property valuation, it is clear that the three principal methods are commonly recognized. These are the comparison method, the income method, and the cost method, and they are typically applied to land (in rural areas), plots (in urban areas), and buildings, respectively [19]. In various studies, advanced valuation methods based on statistics and machine learning are also used [27]. In recent years, advanced machine-learning algorithms have been widely used to improve performance in real estate price prediction [17]. For this purpose, many studies have analyzed the performance of current approaches such as support vector machines (SVM), random forest (RF), decision trees, Bayesian methods, genetic algorithms, Markov prediction models, artificial neural networks (ANN), and deep neural networks (DNN) [18, 19]. Therefore, various studies have shown that DNN models achieve higher levels of accuracy in real estate valuation [23]. For this reason, this study used the DNN method to estimate and predict real estate in Türkiye.

The integration of GIS and DNN was implemented through the following steps:

1. Spatial Data Collection & Validation: All spatial data (coordinates, parcel boundaries) were collected and rigorously validated using GIS tools to ensure locational accuracy.
2. Spatial Analysis: Euclidean distances from each house to various points of interest, such as schools, hospitals, and shopping malls, were calculated using the GIS platform. This methodology facilitated the generation of key locational variables, providing insights into the accessibility of these points of attraction.
3. Data Preprocessing for DNN: The processed spatial data, along with structural and categorical variables, were exported. Categorical variables were one-hot encoded, and numerical variables were normalized using z-score normalization.
4. DNN Model Training & Prediction: The preprocessed dataset was fed into the DNN model, implemented in TensorFlow/Keras.

3.4. DNN-Based Real Estate Price Prediction Method

Real estate price prediction at the city, region, or country scale is a complex problem associated with large datasets characterized by high complexity, uncertainty, and nonlinearity [24]. In recent years, many models based on ANN architecture have been designed to solve such problems, and the performance of different DNN designs has been examined [70, 71]. DNN, a supervised machine learning algorithm, can learn directly from data and successfully model non-linear relationships. In this process, it can model the input-output relationship for a very high number of samples without the need for explicit mathematical functions.

The basic learning rule of the DNN algorithm is the gradient steepest descent (GSD) method, which constantly updates the weights and bias terms of the network with the error-based backpropagation, thereby minimizing the overall error of the network [72]. Thus, when supported by sufficient samples, DNN can make predictions with high accuracy while minimizing user-induced bias compared to other advanced machine learning algorithms.

A dynamic data acquisition approach capable of handling the large number of samples required by the DNN during training is preferred. As a matter of fact, various studies have developed automated data collection algorithms that leverage web-based services widely used in real estate sales and rental processes [25, 26].

Using DNN provides several potential advantages [73]. DNNs learn from large datasets and therefore, provide higher accuracy rates than other methods. DNNs can automatically extract features from data, eliminating the need for manual feature extraction and helping reduce human error and improve performance. DNN can operate effectively even on large datasets, thereby increasing scalability. DNN can be adapted to various data types and problems. For example, a DNN trained for an image recognition problem can also be used for a language processing problem. DNN can automatically organize itself and detect relationships within data. It can make data-driven decisions and therefore produce more reliable decisions [74].

3.5. Housing Price Prediction with DNN-Based Approach

In this study, an approach for predicting housing prices using a DNN-based regression architecture is discussed. For this purpose, a total of 3,757 samples with 24 features, which are explained and found to be effective in housing price estimation, were used. In the proposed approach, the dataset was first subjected to various preprocessing stages.

In the first stage, the attributes of the data were analyzed, and numerical and categorical variables were identified. Observations containing missing values in the dataset were automatically deleted. Then, the categorical features were one-hot encoded, and the dataset (x) was updated to 3,757 × 61, with n = 3,757 observations and m = 61 variables.

Since the numerical variables in the dataset span wide numerical ranges, each variable was normalized using the z-score normalization method (1) and reduced to a narrow numerical range.

$$z_i = \frac{x_i - \mu}{\sigma} \tag{1}$$

where x_i denotes the original data, μ is the mean value, and σ is the standard deviation.

This normalization resulted in all numerical features having a mean of 0 and a standard deviation of 1, improving model convergence.

3.6. Preparation of DNN Architecture and Performance Evaluation

After the preprocessing phase was completed, the DNN architecture design, model training, and performance evaluation phases were carried out. All computational analyses were performed using Python. The DNN model was implemented using TensorFlow (backend) and the Keras library. Spatial analyses and data validation were conducted in QGIS 3.28, and the geospatial database was managed in PostgreSQL with the PostGIS extension. For this purpose, the Scikit-Learn and Keras libraries, which are widely employed in machine learning, were used alongside TensorFlow.

The Keras library includes many modules developed for different purposes, such as preprocessing, DNN layer creation, activation, and optimization [75]. In the design process of the DNN architecture, the selection of ideal values for various hyperparameters, such as the number of hidden layers, the number of units per hidden layer, the optimizer, the learning rate, etc., is very important for the model's prediction performance and stability. Therefore, based on prior experience and multiple experiments, the model was designed with the hyperparameters listed in Table 3.

The designed network contains three densely connected hidden layers with 128, 128, and 64 neurons, respectively. In the proposed model, the ReLU activation function was used in the input and hidden layers, and the linear activation function in the output layer. In addition, the ***Adam optimizer*** was selected for the model design.

To prevent the overfitting problem and to increase its generalization ability, the dropout technique was applied with a rate of 0.2 in the first and second hidden layers. Another hyperparameter affecting the learning capacity and speed of the designed DNN model is the ***learning rate.*** By selecting the parameter in question as ***large*** and keeping it constant throughout the training process, the model can initially be allowed to learn quickly, but unrealistically high accuracies can be obtained. On the other hand, large fluctuations may occur in the model's training error and accuracy during training. Therefore, instead of selecting a fixed learning rate during the training of the proposed DNN model, a variable learning rate strategy is used, reducing the initial learning rate using exponential decay. In this way, a very successful balanced trade-off was achieved between the learning rate and the model's stability.

Table 3. Hyperparameters of the DNN model

Model parameter	Feature
Number of layers	5
Number of hidden layers	3
Number of neurons in hidden layer 1	128
Dropout rate after hidden layer 1	0.2
Number of neurons in hidden layer 2	128
Dropout rate after hidden layer 2	0.2
Number of neurons in hidden layer 3	64
Activation function (except output layer)	ReLU
Activation function (output layer)	linear
Loss function	mean squared error
Optimizer	Adam
Initial learning rate	0.0001
Learning rate decay technique	exponential decay
Exponential decay steps	800
Exponential decay rate	0.7
Number of epochs	50
Batch size	16

To demonstrate that the designed DNN model can provide optimal solutions with the selected hyperparameters, the *k*-fold cross-validation technique was used. For this purpose, after the preprocessed dataset was shuffled, 20% was set aside as a hold-out set with random sampling, and the remaining 80% was subjected to *k*-fold cross-validation.

In *k*-fold cross-validation, *k* was set to 5, and training and validation were performed with different parts of the dataset. At this stage, performance metrics (MAE, RMSE, R^2, and adjusted R^2) were calculated for each fold and for overall cross-validation and presented in Table 4. The results in the table show that the designed model architecture achieves consistent accuracy across different subsets of the dataset under the selected hyperparameters, and that its generalization potential is at the desired level.

Table 4. *k*-fold cross-validation accuracy

Score	MAE	RMSE	R^2	Adjusted R^2
Fold 1 score	0.297	0.138	0.973	0.970
Fold 2 score	0.276	0.108	0.982	0.979
Fold 3 score	0.291	0.112	0.979	0.976
Fold 4 score	0.287	0.208	0.969	0.966
Fold 5 score	0.284	0.107	0.980	0.978
Cross-validated score	0.287	0.141	0.975	0.974
Holdout score	0.289	0.115	0.988	0.987

To evaluate the performance of the designed DNN model in the study, mean absolute error (MAE), root mean squared error (RMSE), R^2, and adjusted R^2 metrics were used. The formulas for these metrics are provided in Equations (2), (3), (4) and (5), respectively. MAE and RMSE are error measurements that reveal the magnitude of error between the estimated values of the model and the actual values.

In addition, R^2, another measure that reflects the predictive performance of the model, represents the ratio of the variance in the dependent variable (y) explained by the independent variables (x). R^2 expresses the extent to which the variance of one variable explains the variance of another variable. If there are more independent variables in the dataset, the R^2 value may improve unrealistically. The adjusted R^2 value can be used to avoid this situation. Adjusted R^2 provides a more realistic measure of model performance by considering the number of independent variables.

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|y_i - \hat{y}_i\right| \tag{2}$$

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

$$R^2 = 1 - \frac{\sum_{i=1}^{n}\left(y_i - \hat{y}_i\right)^2}{\sum_{i=1}^{n}\left(y_i - \overline{y}_i\right)^2} \tag{4}$$

$$Adj - R^2 = 1 - \frac{n-1}{n-p-1}(1-R^2) \tag{5}$$

In Equations (2)–(5), n is the number of samples, p is the number of independent variables, y_i is the actual price value for sample i, y_i is the predicted price value for sample i, $\overline{y}_i$ is the mean value of the dependent variable.

In the last stage, the dataset was shuffled again, and 80% was randomly set aside as training data and 20% as test data, and the final training of the designed model was performed for 50 epochs. In the final training stage of the designed DNN model, 25% of the training data was used for validation.

When Figure 2 is examined, it is observed that the validation and test values show a consistent distribution during training and that the generalization capability of the designed DNN model is at a good level.

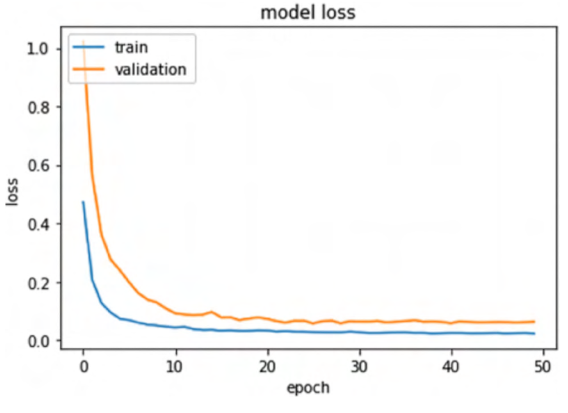

Fig. 2. Validation and test loss curves

In summary, the flowchart of the proposed approach for developing a new real estate valuation prediction model that can be the basis for property tax in the study area is shown in Figure 3.

Based on this process, the criteria derived from the literature and the authors' experience in previous studies were considered. Then, prior studies conducted by the authors and findings from the literature were used to select the criteria. Therefore, the categorical and metric variables used in the research were defined as independent variables. All the data were collected and stored in the Geographic Information Systems (GIS) database. From the obtained data, the samples used in the DNN training and testing process were determined by random sampling method. Then, the creation, training, and testing of the DNN were completed. In the step, the estimation results were imported into the GIS database. The DNN can now be used to estimate the prices of different houses.

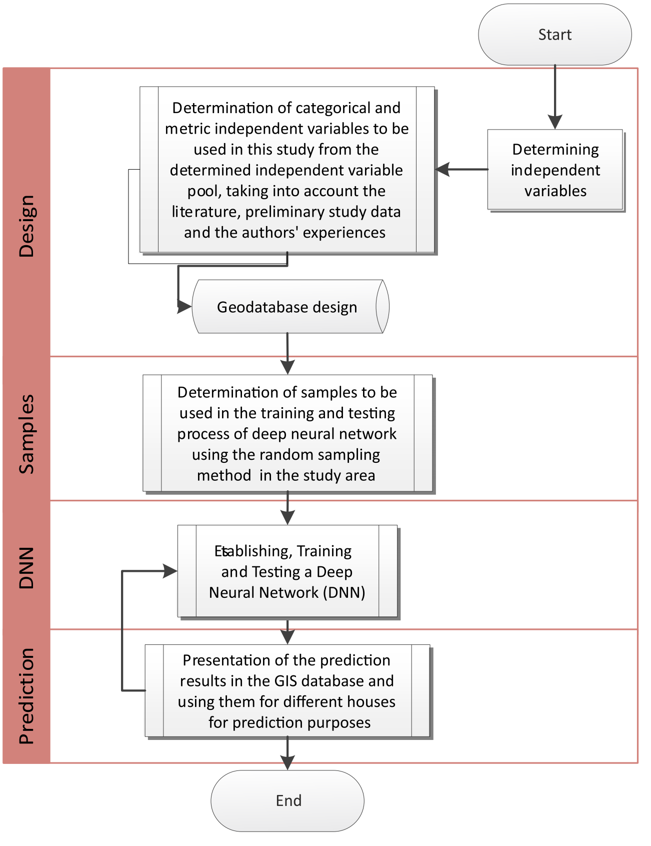

Fig. 3. Methodological process

4. Results

In this study, the data obtained from the portals were not directly included in the database during the automated data acquisition process; instead, they were first subjected to various consistency analyses based on location and data attributes. For example, in an advertisement where one or more of the data attributes, such as the neighborhood, street, block, parcel, site name, etc., were present, the relevant location area boundaries and the location data added to the advertisement were automatically checked using a GIS-based approach.

In addition, advertisements with similar features were evaluated together, and those with outlier location or price data were eliminated. As a result of all these processes, the descriptive statistics of the numerical variables are presented in Table 5.

Table 5. Descriptive statistics of numerical variables

Variable	mean	Std	min	25%	50%	75%	max
Price (TRY)	8,630,801	9,644,924	261,000	3,620,000	6,000,000	10,000,000	260,000,000
Latitude (LA) [°]	41.025501	0.056918	40.884831	40.990038	41.007811	41.073449	41.147863
Longitude (LO) [°]	28.985219	0.186826	28.618217	28.78736	29.062341	29.125365	29.218687
Net area [m^2] (A_N)	115.60	59.29	5.00	81.00	105.00	135.00	1,200.00
Gross area (A_G) [m^2]	141.10	66.62	20.00	100.00	130.00	165.00	1,250.00
Located floor (F_{LOC})	6.91	7.25	−2.00	2.00	4.00	9.00	35.00
Total number of floors (F_{TOT})	13.24	9.53	1.00	5.00	10.00	20.00	35.00
Distance to shopping centers / malls (D_{SCM}) [m]	2,892.31	2,152.19	102.00	1,366.00	2,209.00	4,287.00	9,483.00
Distance to religious facilities (D_{RF}) [m]	4,827.39	1,886.22	136.00	3,467.00	4,319.00	6,660.00	9,130.00
Distance to healthcare facilities (D_{HI}) [m]	1,666.78	1,115.21	19.00	997.00	1,515.00	2,040.00	8,101.00
Distance to intercity transportation (D_{ITO}) [m]	12,889.03	4,644.40	3,716.00	8,167.00	14,369.00	17,149.00	20,267.00
Distance to market / bazaar ($D_{M/B}$) [m]	457.92	418.94	6.00	198.00	325.00	579.00	4,242.00
Distance to educational facilities (D_{SM}) [m]	1,176.36	757.25	36.00	708.00	948.00	1,483.00	5,599.00
Distance to recreational areas (D_{RA}) [m]	2,580.65	2,183.72	36.00	1,302.00	1,665.00	2,627.00	10,073.00
Distance to university (D_U) [m]	9,135.80	5,558.90	24.00	3,055.00	9,771.00	12,790.00	24,444.00
Distance to urban transportation (D_{UTO}) [m]	841.96	527.37	8.00	461.00	740.00	1,091.00	4,158.00

These statistics illustrate the central tendency and spread of the key variables within the Istanbul housing dataset.

A boxplot was created to illustrate the distribution and dispersion of property prices within the study dataset (Fig. 4). The plot indicates a right-skewed distribution, showing that while the majority of housing prices in Istanbul cluster between 2 and 8 million TRY, there are several high-value outliers representing luxury properties, particularly in waterfront and central districts. An extreme value of approximately 260 million TRY marks the upper limit of the dataset. However, as evidenced by the high R^2 and low-error metrics, the DNN model's robustness was not significantly affected by this skewness, allowing it to effectively capture the underlying price patterns across the diverse range of values.

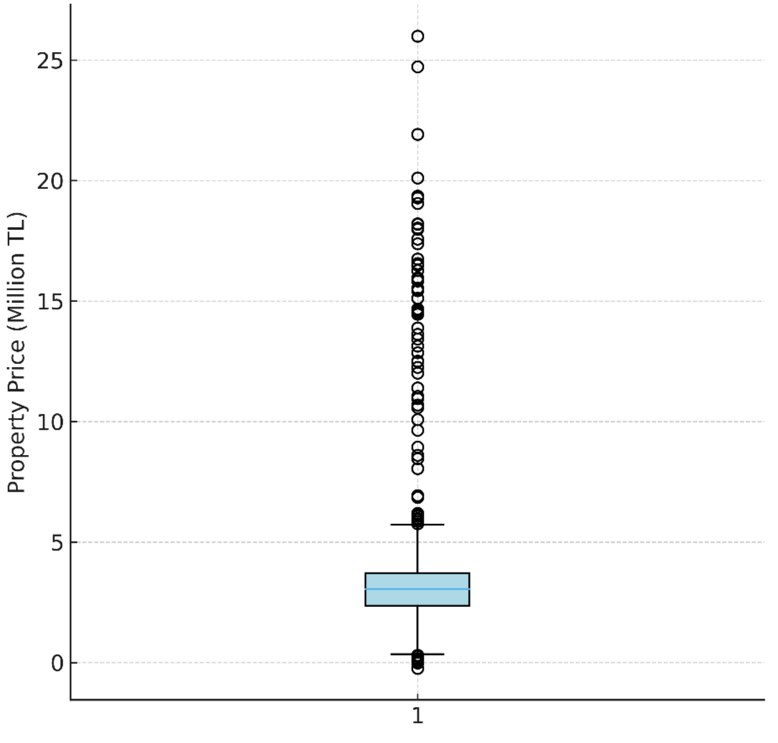

Fig. 4. Boxplot of housing prices in Istanbul

A heatmap of the correlations among the independent metric variables used in the study is presented in Figure 5. It shows the pairwise correlation coefficients between these variables. This visualization helps identify significant relationships, such as highly positive or negative correlations, and provides insights into potential multicollinearity or independence between variables.

Based on the results, the neighborhoods with concentrated housing sales activity in Istanbul, which has the highest population and the highest number of houses for sale in Türkiye, were selected as the study area. In this way, the largest possible number of samples was obtained in a short period using classical data acquisition methods and verified data sources. In addition, the study identified the necessary and sufficient criteria that can explain housing value at an optimal level, drawing on the housing valuation literature and accounting for the unique characteristics of web-based data. In the proposed approach, the reliability and quality of the sample data were assessed using various analyses, including GIS-based automatic spatial data analysis techniques during the preprocessing stages. Then, a DNN regression architecture capable of making predictions based on both geographical and non-geographical features was designed.

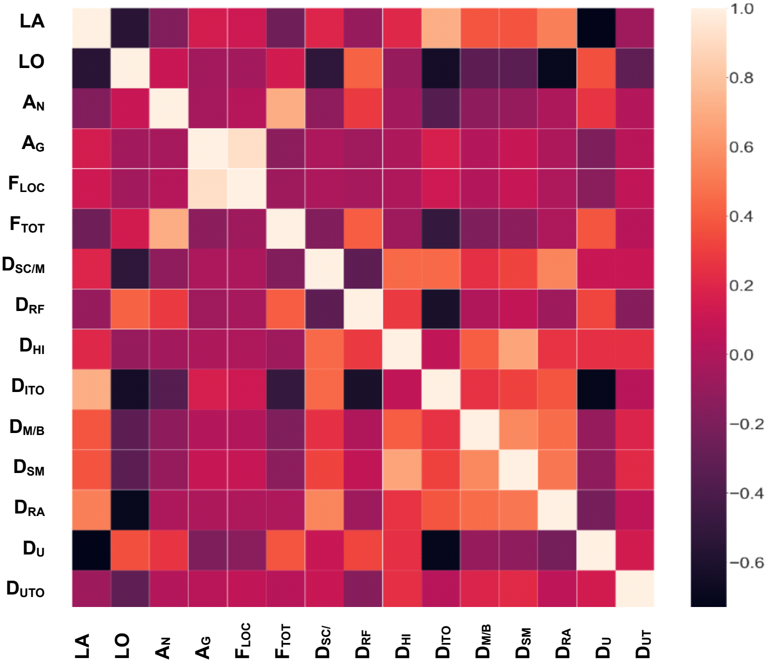

Fig. 5. Heatmap of correlation between numerical variables (variable definitions – see Table 5)

To evaluate the predictive performance of the designed DNN architecture on the test dataset, we calculated commonly used performance metrics in the literature, MAE and RMSE, which were 0.083 and 0.135, respectively. Low values indicate high predictive accuracy, whereas high values indicate low accuracy. Therefore, the obtained MAE and RMSE values indicate that the DNN prediction model can make predictions with very low errors in the proposed approach.

In addition, the R^2 and adjusted R^2 values, which are widely used in regression analysis, were calculated as 0.979 and 0.977, respectively. The R^2 value indicates the extent to which the dependent variable (estimated housing value) is explained by the independent variables. In other words, this value indicates that 97.9% of the variation in the estimated housing value in the proposed DNN model is explained by the 24 selected features.

In addition, in a prediction model with multiple independent variables, the adjusted R^2 metric uses the number of independent variables to counteract the unrealistically increasing R^2 that may result from adding more predictors. In this way, it provides a more realistic measurement of the model's goodness of fit.

In this study, the 97.7% adjusted R^2 indicates that the independent variables explain a significant portion of the variation in the dependent variable without overfitting. As a result, the relationship between the prices calculated for the 3,757 samples in the test dataset and the actual prices predicted by the designed DNN model is shown in Figure 6.

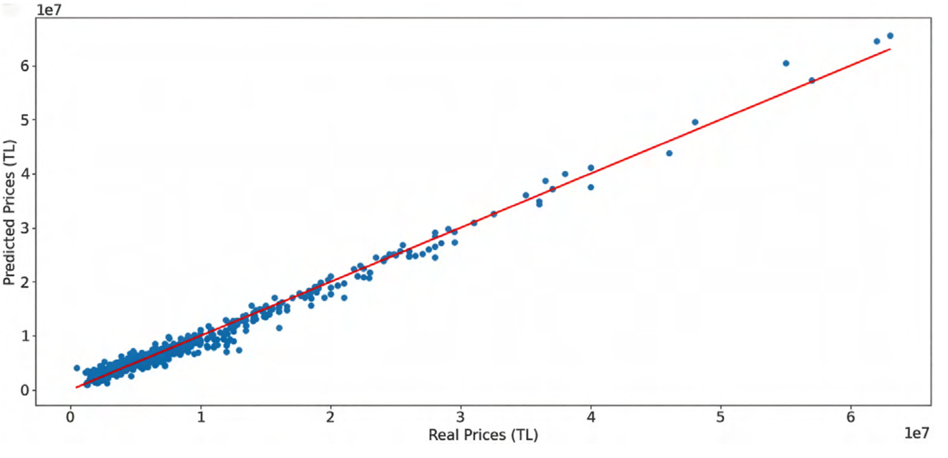

Fig. 6. Differences between actual prices and estimated values

In the study, the sales and purchase advertisements for houses on publicly accessible web services that mediate real estate were considered as a large data source. In the province of Istanbul, which was identified as the study area, advertisements for 3,757 houses for sale were taken into consideration, and all manually obtained data were entered into the GIS database.

The possibility that some criteria could negatively affect accuracy, be subjective, lack clear metric or categorical definitions, or be difficult to obtain was considered in detail during the selection of criteria. For example, criteria such as proximity to roads, proximity to major roads, maintenance cost information, view quality, and sunlight exposure, which posed challenges during data acquisition, were excluded from the evaluation.

Some criteria that were correlated with one another and thought to have very little effect on the house's value were not taken into consideration. Ultimately, the selected criteria achieved maximum accuracy with a minimal set of metrics. The R^2 values express the extent to which the predicted housing price (dependent variable) is explained by the independent variables. The DNN model developed in this study demonstrates strong predictive performance, with R^2 and adjusted R^2 values of 0.979 and 0.977, respectively.

In addition, the DNN algorithm using the selected criteria can predict housing prices with an MAE of 8.3% in the study area, lower than the legislative expectation of 20%. Thus, without the need for the classical property tax method applied in Türkiye, property tax can be calculated based on market value assessments from the legal and technical perspectives.

In calculating the real estate tax base values of houses with DNN, there is no need to determine minimum unit values per square meter for streets and avenues, conduct assessments by tax commissions every four years, use annual construction cost catalogs, or apply revaluation rate adjustments. Thus, in an approach that favors DNNs, dynamic estimation can be made without the need for traditional multivariate valuation models. The effect of the time value of money and inflation is eliminated thanks to the constantly updated dynamic database.

5. Discussion

Housing valuation is a concrete reflection of socio-economic inequalities in urban space. As noted by [76–78], in densely populated and rapidly changing cities such as Istanbul, current official valuation systems often fail to capture market realities. The study's findings show significant differences in housing values across accessibility, environmental factors, and spatial location. The development of digital data sources and the proliferation of web-based, publicly accessible advertisement portals have increased the usability of spatial big data in urban research. The study by [79] explored the opportunities and challenges presented by location-based big data (LocBigData) for smart cities. A study by Longley [80] highlighted the significance of data infrastructure for smart cities and pointed out the risks posed by uncertainties related to the origin, ownership, and control of data sources, particularly concerning scientific reproducibility and data reliability. However, these data sources are not suitable for direct use without verification and cleaning, as stated by [81, 82].

In this context, the data cleaning and location verification processes using GIS employed in the study make a significant contribution to ensuring reliability in spatial analysis. The proposed DNN-based model estimated housing values with lower error rates compared to classical valuation methods. In the study by [83], the DNN model demonstrated greater success in predicting housing values, achieving an accuracy rate of 90–95% compared to traditional estimation methods. Our study found an accuracy of 97.9%, indicating strong model performance. Similarly, [22] utilized machine learning, deep learning, and hedonic approaches for housing valuation, and the results indicated that both ML and DL methods outperformed traditional techniques. Another study [84] employed Geographically Neural Network

Weighted Regression (GNNWR), a deep learning-based method, and found that its estimation accuracy surpassed that of conventional methods.

These findings indicate that advanced machine learning and deep learning methods can be used effectively in urban data analytics. Moreover, as the study shows, spatial variation in housing values is directly related to variables specific to urban geography, such as proximity to transportation infrastructure and social amenities. Therefore, this study not only proposes a prediction model but also provides an innovative methodological framework for understanding and mapping intra-city value inequalities. This framework can also pave the way for the development of more equitable and spatially based tax policies for local governments.

Based on the results, in recent years, for large datasets, it is recommended to construct datasets using automated techniques whenever possible. However, such an approach could not be used in this study due to the security measures of the portals that provide open-source data: all data for this study was obtained manually. As [85] noted in studies using DNNs on large datasets with spatial and textual data, a large number of samples is essential for achieving higher prediction accuracy.

However, the accuracy of the sample data filtered from a data warehouse, the reliability of the data source, and the internal consistency of the data are also extremely important [85]. To ensure data reliability and consistency when estimating the value of the house, attention should be paid to the availability of samples in the geographical location of the house for sale or in areas very close to it. Thus, the sample can be internally validated. Moreover, by using the parcel where the house for sale is located, along with houses for sale adjacent to or near the relevant parcel as sample data, it is possible to perform internal consistency checks to determine whether the data is meaningful and reliable.

Therefore, when collecting open-source data, a sampling strategy that includes as many data sources as possible and that offers opportunities for control is useful to follow. Based on [83], developing effective housing value estimation approaches to large metropolitan cities or countries, and applying them quickly and with high accuracy in an ecosystem where instant purchase and sale processes occur, is very important. For this purpose, using data from web-based real estate purchase and sale portals in the development and testing of the proposed prediction models can provide significant benefits [86, 87]. These platforms can comprehensively collect spatial and descriptive data for houses as open-source inputs within voluntary GIS-based frameworks [88].

6. Conclusions

All immovable property as a source of wealth, including plots and land as well as constructed real estate such as shops and residences, is a basic tax source of state revenue in almost all countries. Taxation provides governments with the necessary resources to sustain public functions. One of the important parts of tax is real estate tax, which is generally calculated in proportion to the property's value. However, it may be calculated using prices below the real value, since the actual value of real estate may not be known.

Although the inadequacy of the real estate tax calculation method used in Türkiye is well recognized, the closeness of prices advertised on web-based real estate purchase, sale, and rental platforms to market values is also evident. Despite uncertainties in exact transaction prices, real estate portals in Türkiye provide large-scale datasets covering a wide range of property transactions. Despite this potential, the public authority responsible for calculating real estate tax has relied on the same method for about fifty years. Open-source, non-personal transaction data from these platforms is not currently utilized.

While one advantage of big data on these platforms is that public authorities can use it without high cost or effort, concerns regarding data reliability, quality, and the preprocessing required to make the data usable can be considered disadvantages. In other words, in addition to the inadequacy of the classical property tax system, the use of innovative housing price prediction algorithms may also create some problems. From this point of view, the performance metrics determined in deep learning and real estate valuation studies for integrating the classical property tax calculation method in Türkiye with modern data mining methods were considered jointly to develop an approach that addresses these limitations and aligns with the study's objectives.

In the study, the sales and purchase advertisements for houses in open-source web services that mediate real estate were considered as a large data source. In the province of Istanbul, which was designated as the study area, advertisements of 3,757 houses for sale were analyzed, and all the data obtained manually were transferred to the GIS database. The possibility that some criteria might negatively affect accuracy, be subjective, lack clear metric or categorical definitions, or be difficult to obtain was carefully evaluated during the selection process.

For example, criteria such as proximity to roads, proximity to major roads, maintenance costs, view quality, and sunlight exposure, which presented challenges during data acquisition, were excluded from the analysis. Some criteria that were correlated with one another and were thought to have little effect on the house's value were not taken into consideration. Ultimately, the selection process achieved maximum accuracy with a minimal set of objective metrics.

The R^2 values express the extent to which the predicted housing price (dependent variable) is explained by the independent variables. The DNN model developed in this study demonstrates strong predictive performance, with R^2 and adjusted R^2 values of 0.979 and 0.977, respectively. In addition, the DNN algorithm using the selected criteria can predict housing prices with an MAE of 8.3% in the study area, which is lower than the legislative expectation of 20%. Thus, property tax can be calculated based on market value estimation without relying on the classical methods applied in Türkiye, from both legal and technical perspectives.

In calculating the real estate tax base values of houses with DNN, there is no need to determine the minimum square meter unit values of streets and avenues, conduct assessments by real estate tax committees every four years, use annually published construction cost catalogs, or determine revaluation rates for updating real estate tax calculations in interim years. Therefore, a DNN-based approach enables dynamic estimation without relying on traditional multivariate valuation models. The effects of time value of money and inflation are eliminated by using a constantly updated dynamic database.

To develop the real estate tax system, it is recommended that public authorities create DNN-based models similar to the one proposed in the study and utilize data from web-based real estate platforms. This would enable continuous data provision and support a sustainable open-source and voluntary data infrastructure. It should also be considered that this database should be managed within a structure capable of operating in conjunction with other e-government platforms. Such a system would not only increase real estate tax collection but also enhance revenue from land registry transactions.

It is recommended that researchers working on real estate valuation using modern methods investigate various valuation criteria and real estate types, building on the findings of this paper. Collaboration between the public and private sectors to automate the acquisition of open-source data and accelerate data provision and verification processes should be considered a key priority. All these processes may be implemented through secure web-based databases integrated with e-government systems and operated under the supervision of responsible national institutions.

Funding

This research received no specific grant from any funding agency in the public, commercial, or not-for-profit sectors.

CRediT Author Contribution

Y.E.Ç.: conceptualization, methodology, investigation, resources, data curation, writing – original draft preparation, writing – review and editing, visualization.

M.D.: conceptualization, investigation, writing – review and editing.

O.Y.: conceptualization, writing – review and editing.

M.Ö.Ç.: conceptualization, writing – review and editing.

H.A.: conceptualization, writing – review and editing.

Declaration of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data Availability

Some or all data, models, or code that support the findings of this study are available from the corresponding author upon reasonable request.

Use of Generative AI and AI-Assisted Technologies

No generative AI or AI-assisted technologies were employed in the preparation of this manuscript.

References

[1] Ding J., Cen W., Wu S., Chen Y., Qi J., Huang B., Du Z.: *A neural network model to optimize the measure of spatial proximity in geographically weighted regression approach: A case study on house price in Wuhan*. International Journal of Geographical Information Science, vol. 38(7), 2024, pp. 1315–1335. https://doi.org/10.1080/13658816.2024.2343771.

[2] Liu A., Jia M., Chen C.: *From speculative dreams to ghost housing realities: Homeless homeowners and failed future-making in urban development*. Urban Geography, vol. 46(8), 2025, pp. 1821–1841. https://doi.org/10.1080/02723638.2025.2456545.

[3] Yesmaganbetova A., Dautkanova D., Duisenbekova O., Ismatullayev S., Rsymbetov B., Serikbaeva G.: *Quality of state real estate cadastre-maintenance process based on Ishikawa cause-and-effect diagram*. Geomatics and Environmental Engineering, vol. 19(5), 2025, pp. 103–123. https://doi.org/10.7494/geom.2025.19.5.103.

[4] Klimach A.: *Methodology for evaluating applicability of real estate information systems*. Geomatics and Environmental Engineering, vol. 19(4), 2025, pp. 99–118. https://doi.org/10.7494/geom.2025.19.4.99.

[5] Bogin A., Doerner W., Larson W.: *Local house price dynamics: New indices and stylized facts*. Real Estate Economics, vol. 47(2), 2019, pp. 365–398. https://doi.org/10.1111/1540-6229.12233.

[6] Genc N., Colak H.E., Ozbilgin F.: *Spatial performance approach to machine learning algorithms: A GIS-based comparison analysis for real estate valuation*. Transactions in GIS, vol. 29(1), 2025, e13303. https://doi.org/10.1111/tgis.13303.

[7] Robin E.: *Performing real estate value(s): Real estate developers, systems of expertise and the production of space*. Geoforum, vol. 134, 2022, pp. 205–215. https://doi.org/10.1016/j.geoforum.2018.05.006.

[8] Ansenberg U.: *Beyond displacement: The role of real-estate valuations in shaping urban displaceability*. Environment and Planning C: Politics and Space, vol. 43(8), 2024, pp. 1483–1505. https://doi.org/10.1177/23996544241259310.

[9] El Yamani S., Hajji R., Billen R., Arroyo Ohori K., van der Vaart J., Hakim A., Stoter J.: *Towards extending citygml for property valuation: Property valuation ADE*. ISPRS Annals of the Photogrammetry, Remote Sensing and Spatial Information Sciences, vol. X-4/W5-2024, 2024, pp. 127–135. https://doi.org/10.5194/isprs-annals-X-4-W5-2024-127-2024.

[10] Cellmer R., Kobylińska K.: *Housing price prediction – Machine learning and geostatistical methods.* Real Estate Management and Valuation, vol. 33(1), 2025, pp. 1–10. https://doi.org/10.2478/remav-2025-0001.

[11] Leow K., Lindenthal T.: *Enhancing real estate investment trust return forecasts using machine learning*. Real Estate Economics, vol. 53, 2025, pp. 574–606. https://doi.org/10.1111/1540-6229.12527.

[12] Mete M.O., Yomralioglu T.: *A hybrid approach for mass valuation of residential properties through geographic information systems and machine learning integration*. Geographical Analysis, vol. 55(4), 2023, pp. 535–559. https://doi.org/10.1111/gean.12350.

[13] Reyes-Carranza M., Mbugua Muthama D.: *Urban sprawl and the automation of building control in the peripheries of Nairobi*. Urban Geography, vol. 46(6), 2025, pp. 1344–1360. https://doi.org/10.1080/02723638.2025.2450307.

[14] Shatkin G., Braswell T.H., Martinus M.: *Mapping and the politics of informality in Jakarta*. Urban Geography, vol. 44(5), 2023, pp. 939–963. https://doi.org/10.1080/02723638.2022.2059321.

[15] Wan W.X., Lindenthal T.: *Testing machine learning systems in real estate*. Real Estate Economics, vol. 51(3), 2023, pp. 754–778. https://doi.org/10.1111/1540-6229.12416.

[16] Peter N.J., Okagbue H.I., Obasi E.C., Akinola A.O.: *Review on the application of artificial neural networks in real estate valuation.* International Journal of Advanced Trends in Computer Science and Engineering, vol. 9(3), 2020, pp. 2918–2925. https://doi.org/10.30534/ijatcse/2020/66932020.

[17] Varma A., Sarma A., Doshi S., Nair R.: *House price prediction using machine learning and neural networks*, [in:] *2018 Second International Conference on Inventive Communication and Computational Technologies (ICICCT 2018): 20–21 April 2018, Coimbatore, India*, IEEE, 2018, pp. 1936–1939. https://doi.org/10.1109/ICICCT.2018.8473231.

[18] Aydinoglu A.C., Bovkir R., Colkesen I.: *Implementing a mass valuation application on interoperable land valuation data model designed as an extension of the national GDI*. Survey Review, vol. 53(379), 2021, pp. 349–365. https://doi.org/10.1080/00396265.2020.1771967.

[19] Chen Y., Xue R., Zhang Y.: *House price prediction based on machine learning and deep learning methods*, [in:] *2021 International Conference on Electronic Information Engineering and Computer Science (EIECS 2021): 23–26 September 2021, Changchun, China*, IEEE, 2021, pp. 699–702. https://doi.org/10.1109/EIECS53707.2021.9587907.

[20] Chen K., Liu E., Deng M., Tan X., Wang J., Shi Y., Wang Z.: *DKNN: deep kriging neural network for interpretable geospatial interpolation*. International Journal of Geographical Information Science, vol. 38(8), 2024, pp. 1486–1530. https://doi.org/10.1080/13658816.2024.2347316.

[21] Wu S., Wang Z., Du Z., Huang B., Zhang F., Liu R.: *Geographically and temporally neural network weighted regression for modeling spatiotemporal non-stationary relationships*. International Journal of Geographical Information Science, vol. 35(3), 2021, pp. 582–608. https://doi.org/10.1080/13658816.2020.1775836.

[22] Yazdani M.: *Machine learning, deep learning, and hedonic methods for real estate price prediction*. arXiv, 2021, 2110.07151. https://doi.org/10.48550/arXiv.2110.07151.

[23] Khobragade A.N., Maheswari N., Sivagami M.: *Analyzing the housing rate in a real estate informative system: A prediction analysis*. International Journal of Civil Engineering and Technology, vol. 9(5), 2018, pp. 1156–1164.

[24] Contini C., Boncinelli F., Romano C., Scozzafava G., Casini L.: *Price vector issue in a choice experiment: A methodological proposal*. Food Quality and Preference, vol. 75, 2019. https://doi.org/10.1016/j.foodqual.2019.02.005.

[25] Jiang Z., Shen G.: *Prediction of house price based on the back propagation neural network in the keras deep learning framework*, [in:] *2019 6th International Conference on Systems and Informatics: ICSAI 2019: 2–4 November 2019, Vienna International Hotel, Shanghai, China*, IEEE, 2019, pp. 1408–1412. https://doi.org/10.1109/ICSAI48974.2019.9010071.

[26] Pang X., Zhou Y., Wang P., Lin W., Chang V.: *An innovative neural network approach for stock market prediction*. The Journal of Supercomputing, vol. 76(3), 2020, pp. 2098–2118. https://doi.org/10.1007/s11227-017-2228-y.

[27] Tabar M.E., Başara A.C., Şişman Y.: *Housing valuation study in Tokat province with multiple regression and artificial neural networks*. Turkish Journal of Land Management, vol. 3(1), 2021, pp. 1–7. https://doi.org/10.51765/tayod.832227.

[28] Çağdaş V.: *An application domain extension to CityGML for immovable property taxation: A Turkish case study*. International Journal of Applied Earth Observation and Geoinformation, vol. 21, 2013, pp. 545–555. https://doi.org/10.1016/j.jag.2012.07.013.

[29] Poon J., Hoxley M., Fuchs W.: *Real estate education: An investigation of multiple stakeholders*. Property Management, vol. 29(5), 2011, pp. 468–487. https://doi.org/10.1108/02637471111178146.

[30] Coruhlu Y.E., Toludan T.: *Data model development for 'buying and selling' transactions as a real estate acquisition method*. Survey Review, vol. 52(374), 2020, pp. 403–414. https://doi.org/10.1080/00396265.2019.1626585

[31] Kocabıyık İ., Öktem R.: *Multidimensional approaches in real estate valuation*. Research of Financial Economic and Social Studies, vol. 6(1), 2021, pp. 160–177. https://doi.org/10.29106/fesa.874843.

[32] Demirel B., Yelek A., Alağaş H.M., Eren T.: *Taşınmaz değerleme kriterlerinin belirlenmesi ve kriterlerin önem derecelerinin çok ölçütlü karar verme yöntemi ile hesaplanması* [Determination of criteria for immovable valuation and calculation weights of criteria with multicriteria decision making method]. Kırıkkale University Journal of Social Sciences (KUJSS), vol. 8(2), 2018, pp. 665–682. https://izlik.org/JA82FP95LE.

[33] Thamarai M., Malarvizhi S.P.: *House price prediction modeling using machine learning*. International Journal of Information Engineering & Electronic Business, vol. 12(2), 2020, pp. 15–20. https://doi.org/10.5815/ijieeb.2020.02.03.

[34] Awoamim Y.J.B.D.G.: *Vertical equity and uniformity in house price valuation*, [in:] *An Agenda for Repositioning African Property Market. Kaduna State University, Kafanchan Campus, Kaduna, Nigeria*, African Real Estate Society, 2021, pp. 38–50.

[35] Batista P., Marques J.L.: *Automated housing price valuation and spatial data*, [in:] Gervasi O. et al. (eds.), *Computational Science and Its Applications – ICCSA 2021: 21st International Conference, Cagliari, Italy, September 13–16, 2021, Proceedings, Part IV*, Lecture Notes in Computer Science, vol. 12952, Springer, Cham 2021, pp. 366–381. https://doi.org/10.1007/978-3-030-86973-1_26.

[36] Selim H.: *Determinants of house prices in Türkiye: Hedonic regression versus artificial neural network*. Expert Systems with Applications, vol. 36(2), 2009, pp. 2843–2852. https://doi.org/10.1016/j.eswa.2008.01.044.

[37] Montero J.M., Mínguez R., Fernández-Avilés G.: *Housing price prediction: parametric versus semi-parametric spatial hedonic models*. Journal of Geographical Systems, vol. 20(1), 2018, pp. 27–55. https://doi.org/10.1007/s10109-017-0257-y.

[38] Limsombunchao V.: *House price prediction: Hedonic price model vs. artificial neural network* [paper presentation]. New Zealand Agricultural and Resource Economics Society 2004 Conference, June 25–26, 2004, Blenheim, New Zealand. https://doi.org/10.22004/ag.econ.97781.

[39] Truong Q., Nguyen M., Dang H., Mei B.: *Housing price prediction via improved machine learning techniques*. Procedia Computer Science, vol. 174, 2020, pp. 433–442. https://doi.org/10.1016/j.procs.2020.06.111.

[40] Phan T.D.: *Housing price prediction using machine learning algorithms: the case of Melbourne city, Australia*, [in:] *2018 International Conference on Machine Learning and Data Engineering (iCMLDE 2018): 3–7 December 2018, Sydney, Australia*, IEEE, 2018, pp. 35–42. https://doi.org/10.1109/iCMLDE.2018.00017.

[41] Park B., Bae J.K.: *Using machine learning algorithms for housing price prediction: the case of Fairfax County, Virginia housing data*. Expert Systems with Applications, vol. 42(6), 2015, pp. 2928–2934. https://doi.org/10.1016/j.eswa.2014.11.040

[42] Liu X.: *Spatial and temporal dependence in house price prediction*. The Journal of Real Estate Finance and Economics, vol. 47(2), 2013, pp. 341–369. https://doi.org/10.1007/s11146-011-9359-3.

[43] Yalpır S., Bünyan Ünel F.: *Multiple Regression analysis for house valuation using GIS: A case study*, [in:] Wang Y. (ed.), *Electrical Engineering and Information Technology*, WIT Transactions on Information and Communication Technologies, vol. 63, WIT Press, 2014, pp. 353–359.

[44] Yasar A., Sarıtas İ.: *Use fuzzy logic approach for valuation of the house*. [in:] *Proceedings: The International Conference of Materials and Engineering Technology (TICMET'19), October 10–12, 2019, Gaziantep, Turkey*, Gaziantep Universit, 2019, pp. 833–836.

[45] Cheung W., Guo L., Kawaguchi Y.: *Automated valuation model for residential rental markets: Evidence from Japan*. Journal of Spatial Econometrics, vol. 2(1), 2021, pp. 1–34. https://doi.org/10.1007/s43071-021-00009-0.

[46] Şişman A., Şişman Y.: *Konutun değerine etki eden faktörlerin araştırılması* [An investigation of factors affecting housing value], [in:] *ISEAIA 2016: 4th International Symposium on Engineering, Artificial Intelligence and Applications: Proceedings: 2–4 November 2016, GAU – Girne, North Cyprus*, Girne American University, 2016, pp. 230–236.

[47] Sarip A.G., Hafez M.B.: *Fuzzy logic application for house price prediction*. International Journal of Property Sciences, vol. 5(1), 2015, pp. 24–30. https://doi.org/10.22452/ijps.vol5no1.3.

[48] Alfiyatin A.N., Febrita R.E., Taufiq H., Mahmudy W.F.: *Modeling house price prediction using regression analysis and particle swarm optimization*. International Journal of Advanced Computer Science and Applications, vol. 8(10), 2017, pp. 323–326. https://doi.org/10.14569/IJACSA.2017.081041.

[49] Barbosa V., Suárez M., Cerda J., Thoene U.: *Urban structure to determine equitable city growth for spatial justice: A case study of Chía-Bogotá, Colombia*. Journal of Urban Mobility, vol. 5, 2024, 100080. https://doi.org/10.1016/j.urbmob.2024.100080.

[50] Weber L.: *Maintaining Portland's "progressive dystopia": Housing, carcerality, and the real estate state*. Urban Geography, vol. 45(9), 2024, pp. 1575–1595. https://doi.org/10.1080/02723638.2024.2333210.

[51] Banabak S., Kadi J., Schneider A.E.: *Gentrification and the suburbanization of poverty: Evidence from a highly regulated housing system*. Urban Geography, vol. 45(9), 2024, pp. 1596–1618. https://doi.org/10.1080/02723638.2024.2325197.

[52] Guerrero O.A., Law S.: *The spatial structure of housing affordability and the impact of public infrastructure*. Journal of Simulation, vol. 18(6), 2024, pp. 973–987. https://doi.org/10.1080/17477778.2024.2325428.

[53] Lamas-Abraira L., Colombo A., Oliva-Codina J., Martín-Oriol J.: *Mapping unequal access to culture: A visual approach to culture-festival distribution and socio-economic asymmetries in Barcelona*. Urban Geography, vol. 46(7), 2025, pp. 1642–1649. https://doi.org/10.1080/02723638.2025.2453351.

[54] Tuan N.T., Lan N.T.: *The relationship between transport infrastructure and urban land prices: a study in Hanoi, Vietnam*. SAGE Open, vol. 15(1), 2025, pp. 1–13. https://doi.org/10.1177/21582440241309409.

[55] Dökmeci V., Erdoğan N.: *Spatial analysis of 2000–2018 residential prices in Istanbul*. KAPU Trakya Journal of Architecture and Design, vol. 1(1), 2021, pp. 61–77.

[56] Shakespeare R.M.: *Mapping the desirable and affordable city: Middle-income renter perceptions of housing opportunity*. Urban Geography, vol. 46(6), 2025, pp. 1408–1431. https://doi.org/10.1080/02723638.2024.2436831.

[57] Taşan-Kok T.: *Navigating the city: Role of property-market intelligence channels in urban governance networks*. European Urban and Regional Studies, vol. 32(2), 2025, pp. 197–220. https://doi.org/10.1177/09697764241266411.

[58] Alexander P., Balavac-Orlic M., Lymer A., Mukherjee S.: *Further improving our understanding of the tax awareness, tax literacy and tax morale of young adults*. JOTA: Journal of Tax Administration, vol. 8(1), 2023, pp. 193–220. https://jota.website/jota/article/view/17.

[59] Demir O., Uzun B., Çete M.: *Turkish cadastral system*. Survey Review, vol. 40(307), 2008, pp. 54–66. https://doi.org/10.1179/003962608X253484.

[60] Aydoğdu M.: *Ortaçağ Feodal Mülkiyet Anlayışı ve Osmanlı Hukukundaki Toprak Sisteminin Türk Hukukundaki Tarımsal İşletmelerin Mirasçılara Özgülenmesine Etkileri* [The concept of medieval feudal ownership and influences of the land system in Ottoman Law on the transfer and allocation of agricultural enterprises to heirs]. Dokuz Eylul University Law Review, cilt. 17(2), 2015, pp. 1–26. https://izlik.org/JA87UC25NB.

[61] Çoruhlu Y.E.: *Türk emlak vergi sisteminin irdelenmesi* [Investigation of Turkish real estate tax system], [in:] *İKSAD III – Uluslararasi Sosyal Bilimler Kongresi, Adana, Türkiye, 8–10 Mart 2019, cilt.1: Tam metin bildiri*, 2019, pp. 427–443.

[62] Aliefendioğlu Y.: *Emlak vergi değerlerinin takdirine ilişkin örgütlenme modelleri ve Türkiye için model önerisi üzerine tartışma* [paper presentation, summary text]. Tapu ve Kadastro Modernizasyon Projesi Ek Finansmanı Gayrimenkul Değerleme Çalıştayı, February 27 – March 03, 2017, İzmir, Türkiye.

[63] *Taşınmaz ticareti hakkında yönetmelik* [Regulation on real estate trade]. Resmî Gazete, no. 30442, June 5, 2018. https://www.mevzuat.gov.tr/mevzuat?MevzuatNo=24645&MevzuatTur=7&MevzuatTertip=5 [access: November 1, 2025].

[64] Uzun B., Yıldırım V., Çoruhlu Y.E., Yıldız O., Terzi F., Atasoy B.A.: *The process of transition to a value-based distribution model in the Turkish land readjustment system*. Land Use Policy, vol. 147, 2024, 107360. https://doi.org/10.1016/j.landusepol.2024.107360.

[65] Sahibinden. https://www.sahibinden.com/ [access: July 25, 2024].

[66] Cohen L., Manion M, Morrison K.: *Research Methods in Education*. Routledge, London 2000.

[67] Cohen J., Cohen P., West S.G., Aiken L.S.: *Applied Multiple Regression/Correlation Analysis for the Behavioral Sciences* (3rd ed.). Routledge, New York 2002.

[68] GDLRC [General Directorate of Land Registry and Cadastre]. https://parselsorgu.tkgm.gov.tr/#ara/cografi/40.91258663223653/39.017555326894886 [access: January 1, 2024].

[69] Gürsakal N.: *Makine Öğrenmesi*. Dora Yayıncılık, Bursa 2018.

[70] Lee C.: *Enhancing the performance of a neural network with entity embeddings: An application to real estate valuation*. Journal of Housing and the Built Environment, vol. 37(2), 2022, pp. 1057–1072. https://doi.org/10.1007/s10901-021-09885-2.

[71] Xu J.: *A Novel deep neural network based method for house price prediction*, [in:] *2021 International Conference of Social Computing and Digital Economy: ICSCDE 2021: Proceedings: 28–29 August 2021, Chongqing, China*, IEEE, 2021, pp. 12–16. https://doi.org/10.1109/ICSCDE54196.2021.00012.

[72] Heaton J.B., Polson N.G., Witte J.H.: *Deep learning for finance: Deep portfolios*. Applied Stochastic Models in Business and Industry, vol. 33(1), 2017, pp. 3–12. https://doi.org/10.1002/asmb.2209.

[73] Yu P., Yan X.: *Stock price prediction based on deep neural networks*. Neural Computing and Applications, vol. 32(6), 2020, pp. 1609–1628. https://doi.org/10.1007/s00521-019-04212-x.

[74] Birman R., Segal Y., Hadar O.: *Overview of research in the field of video compression using deep neural networks*. Multimedia Tools and Applications, vol. 79(17–18), 2020, pp. 11699–11722. https://doi.org/10.1007/s11042-019-08572-3.

[75] Gulli A., Pal S.: *Deep Learning with Keras: Implementing Deep Learning Models and Neural Networks with the Power of Python*. Packt Publishing, Birmingham 2017.

[76] Moralı O., Yılmaz N.: *Spatial heterogeneity in Istanbul housing market: A geographically weighted approach*. Journal of Economics Finance and Accounting, vol. 7(4), 2020, pp. 298–307. https://doi.org/10.17261/Pressacademia.2020.1304.

[77] Gunduz L., Çakır M., Cepni O.: *House price dynamics in Istanbul: District-level analysis of convergence clubs, spillovers, and key drivers*. Borsa Istanbul Review, vol. 25(1), 2025, pp. 195–205. https://doi.org/10.1016/j.bir.2024.12.017.

[78] Tekin M., Sari I.U.: *Real estate market price prediction model of Istanbul*. Real Estate Management and Valuation, vol. 30(4), 2022, pp. 1–16. https://doi.org/10.2478/remav-2022-0025.

[79] Huang H., Yao X.A., Krisp J.M., Jiang B.: *Analytics of location-based big data for smart cities: opportunities, challenges, and future directions*. Computers, Environment and Urban Systems, vol. 90, 2021, 101712. https://doi.org/10.1016/j.compenvurbsys.2021.101712.

[80] Longley P.A.: *Smart data infrastructure for urban analytics, city science, planning and design*. Environment and Planning B: Urban Analytics and City Science, vol. 51(5), 2024, pp. 1041–1044. https://doi.org/10.1177/23998083241246327.

[81] Chen X., Biljecki F.: *Mining real estate ads and property transactions for building and amenity data acquisition*. Urban Informatics, vol. 1(1), 2022, 12. https://doi.org/10.1007/s44212-022-00012-2.

[82] Yu D., Fang C.: *Urban remote sensing with spatial big data: a review and renewed perspective of urban studies in recent decades*. Remote Sensing, vol. 15(5), 2023, 1307. https://doi.org/10.3390/rs15051307.

[83] Mostofi F., Toğan V., Başağa H.B.: *Real-estate price prediction with deep neural network and principal component analysis*. Organization, Technology and Management in Construction, vol. 14(1), 2022, pp. 2741–2759. https://doi.org/10.2478/otmcj-2022-0016.

[84] Wang Z., Wang Y., Wu S., Du Z.: *House price valuation model based on geographically neural network weighted regression: The case study of Shenzhen, China*. ISPRS International Journal of Geo-Information, vol. 11(8), 2022, 450. https://doi.org/10.3390/ijgi11080450.

[85] Gudivada V., Apon A., Ding J.: *Data quality considerations for big data and machine learning: Going beyond data cleaning and transformations*. International Journal on Advances in Software, vol. 10(1–2), 2017, pp. 1–20.

[86] Gassar A.A.A., Cha S.H.: *Review of geographic information systems-based rooftop solar photovoltaic potential estimation approaches at urban scales*. Applied Energy, vol. 291, 2021, 116817. https://doi.org/10.1016/j.apenergy.2021.116817.

[87] Jain A., Verma C., Kumar N., Raboaca M.S., Baliya J.N., Suciu G.: *Image geosite estimation using convolutional auto-encoder and multi-label support vector machine*. Information, vol. 14(1), 2023, 29. https://doi.org/10.3390/info14010029.

[88] Chicco D., Warrens M.J., Jurman G.: *The coefficient of determination R-squared is more informative than SMAPE, MAE, MAPE, MSE and RMSE in regression analysis evaluation*. PeerJ Computer Science, vol. 7, 2021, e623. https://doi.org/10.7717/peerj-cs.623.