International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

A Comprehensive Review of Web Scraping and Machine Learning Techniques for City-Wise Rent and Living Cost Estimation

Ashwini Dhase[1], Vaishnavi Gawade[2], Pranali Gopane[3], Snehal Gavhane[4], Prof. Priyanka Bhore[5]

[1,2,3,4]Department of Artificial Intelligence and Data Science, Ajeenkya D.Y. Patil School of Engineering, Lohegaon, Pune.
[5]Assistant Professor, Ajeenkya D.Y. Patil School of Engineering, Lohegaon, Pune.

Abstract:
The rapid growth in urban populations, and the movement of individuals for education or job opportunities have put pressure on creating reliable methods for estimating rental costs. Current platforms that focus on real estate primarily use rental listings as a means of providing access to finding a place to live, however they do not generally include comprehensive living cost estimates, fairness analysis of prices, or personalised recommendations. This paper will address 36 research papers from 2001 to 2025, examining predictive rental price algorithms and web scraping techniques, big data analytics and ensemble machine learning models, deep learning methods, and spatial analysis methods, as well as using XAI in the housing domain. According to current research, ensemble-based learning methods—like Random Forest and boosting frameworks—perform better than conventional regression models in predicting rental and home prices. It was also shown that using geographic features, socio-economic indicators, multi-modal inputs, and explainability based on SHAP contribute to enhanced accuracy and transparency of the rental price prediction models. However, a unified model for web scraping in real time, estimating total living costs (rent, utilities, food, and commute), selecting and recommending AI based on total living costs, and creating a fair rent type for flatmates does not presently exist. The research presented provides a basis for the creation of an xAI-based living cost estimator using a unified, AI-driven rental cost estimator system.

Keywords: Rental Pricing Predictions, Cost-of-Living Estimations, Web Scraping Techniques, Machine Learning Algorithms, Ensemble Learning Methods, XGBoost Classifiers, Explainable AI (XAI) Solutions, SHAP (Shapley Additive Explanations) Values, Big Data Analytical Tools, Urban Real Estate Markets, Spatial Data Analysis Techniques, Recommendation Systems, Game Theory Principles, Shapley Values, Smart Homes / Housing Systems.

I. INTRODUCTION

The rapid growth of urban areas, along with student and professional relocation in and out of cities, has increased the demand for affordable and transparent systems of renting housing. The growing population is changing how people rent their living spaces by making it a complex, dynamic and data-driven market as city size continues to grow and wealth is created economically throughout the nation. Many of today's online housing listing sites lack comprehensive living cost estimates, evaluations of whether or not the price being asked for the home is fair, or any intelligent recommendation functions; they provide only basic information and don't provide much detail. Machine learning (ML) and big data analytics (BDA) have transformed the way people do business in the real estate industry. Numerous studies have demonstrated that ensemble ML approaches like Random Forest, Gradient Boosting, and XGBoost outperform traditional approaches to predicting value based on multiple linear regressions [1], [5], [15] .In addition, there is evidence that deep learning ML approaches, such as LSTM networks, help to understand the effects of time on housing price trends [22]. Incorporating spatial and socio-economic dimension features has been shown to improve the predictive ability of methods used to value real estate [18], [27].

Moreover, XAI techniques, such as SHAP (Shapley Additive Explanations), provide a way to enhance the transparency and interpretability of real estate predictive modelling [3], [4], [13]. In addition, web scraping and BDA have enabled researchers to collect vast quantities of rental data across different internet platforms, thereby increasing the robustness of the models and their real-time adaptability [8], [36]. Based on these examples of technological advances, most research has so far focused on how the data is used. An important gap exists between the prediction of rental fees and the expense of living fees that incorporate not just actual rent; however, but also expenses such as utilities, food, and commuting. Furthermore, there is little research regarding the use of individualised Artificial Intelligence (AI) to recommend an area for individuals; the equitable division of rent payments between roommates using game theory, and user-friendly visualisations of data. To this end, a comprehensive system must be developed utilising web-scraping, machine learning and/explanatory AI-based rental price prediction systems, as well as an analysis of total living costs (including all of the above-mentioned expenses).A systematic literature review was conducted on 36 research articles published during the years 2001 through 2025, to provide insight into the trends, methodologies, coding techniques and limitations associated with rental price prediction and housing analytics, so as to help identify and determine the current gaps in research, and serve as a foundation for constructing an integrated AI-based City Rent and Living Cost Estimation System Framework.

II. RELATED WORK

Machine Learning, Deep Learning, and Big Data Analytics have had a great effect on the real estate sector using Explanatory Artificial Intelligence Technologies. This section identifies existing research contributions under various themes related to Predicting Rental Prices and Performing Analysis on Housing Data.

A. Rental and Housing Price Prediction with Machine Learning Techniques

Many of the early research efforts into predicting housing prices were based on traditional regression and statistical methods. Comparative studies, however, show that sophisticated computational models—such as SVM, ensemble learning techniques, and neural networks—perform better in terms of predictive accuracy than conventional linear regression models.[30], [32], [33] In order to estimate land prices, recent research emphasizes the use of ensemble strategies, especially weighted average models, which have been shown to be more stable and dependable than conventional techniques. In addition to that, researchers have also shown that employing ensemble learning frameworks that employ boosting and bagging techniques has performed better than employing those models individually.[15],[23]Furthermore, employing Optimised XGBoost-based models has increased the accuracy of the models through Hyperparameter Tuning and Feature Selection [6], [7].

Deep Learning models have also been utilised to capture non-linear relationships and time-varying relationships in housing data. The use of LSTM models, which provide better forecasting capabilities, has improved time-series based property valuation. Hybrid machine learning frameworks which employ multiple machine learning-based methods have been proposed for balancing the Bias and Variance trade-off so that the models will achieve improved generalization performance.

B. Explainable AI in Real Estate

While machine learning models can increase accuracy, one of the biggest challenges with all machine-learning models continues to be the interpretability of the model. Recently developed works, however, have attempted to incorporate Explainable Artificial Intelligence (XAI) techniques into machine-learning models for greater transparency. In [3], SHAP (Shapley Additive Explanations) was applied to develop interpretability for XGBoost-based property valuation models. Further research in explored the value of Explainable Ensemble Models in maintaining stakeholder trust and confidence in AI-based decision-making. Research conducted in [12] and [13] emphasises the importance of the analysis of feature contribution in real estate pricing by identifying that a property's location, amenities, and socioeconomic characteristics are strong predictors of rental rates. In [11], the authors introduced Explainable Boosting Models, demonstrating that predictive performance can be achieved without sacrificing interpretability. The results of these studies show that when using machine learning techniques in real estate, it is crucial to strike an effective balance between predictive accuracy and model interpretability.

C. Big Data, Web Scraping, and Integration of Data

The rise of Web-based housing platforms has allowed the collection and processing of large amounts of rental data through web scraping techniques. The ROADRUNNER project Lane & MCKENNA published the first automatic web data extraction methods which were used to create modern scraping technologies. IN recent years, examples of web scraping being used for gathering rental data from multiple locations can be found and have been applied in order to improve the variety of datasets to be analysed [8]. Data fusion techniques using multiple sources were proposed in to integrate heterogeneous datasets for better predictions about housing prices [9]. The use of Big Data analysis methodologies was described to conduct comprehensive research on the large-scale urban housing market [10], [16], [17]. Furthermore, Machine Learning models based on Spatial and GIS Data were applied to model housing prices, and capture geographic dependencies [27], [29]. Performing Spatiotemporal Modelling on the data improved the quality of (i.e., accuracy) prediction by considering both time and place at the same time [18].These improvements indicate that combining Web-based real-time data with Spatial Analysis can assist in building accurate rental rate prediction systems.

D. Socioeconomic and urban factors in rental markets

There are external factors affecting housing prices, including accessibility, local amenities, and economic conditions. Some of these factors have been uncovered in research done on social economics in relation to the value of property as seen in [24]. Analysis of urban accessibility and housing prices was furthered in [25], which made use of machine learning to identify the correlation between the two research conducted in [31] on the rental markets of Indian cities showed the importance of conducting localised studies. The impact of pricing based on amenities was shown in [34], revealing that rental price differs greatly based on infrastructure and available facilities in the area. The results of this research illustrate the importance of incorporating contextual and environmental variables into the prediction of rental prices for greater accuracy.

Summary of Related Work

This body of literature shows vast improvements to rental and housing price prediction when employing ensemble learning, deep learning, big data analytics, and explainable AI. Most current systems only predict the price of rent, without taking into consideration the complete estimate for the cost of living, the use of personalised AI recommendations, real-time scraping pipelines, or the equitable distribution of rent. The existing gap motivates an integrated AI-based framework for estimating rental prices and housing cost of living.

III. METHODOLOGY

The integrated methodology presented in this section is designed to estimate rental costs and living expenses within cities. This framework incorporates concepts gained from previously reviewed sources and combines web scraping, machine learning to predict rent prices, explainable artificial intelligence (AI), and cost aggregation into one architecture.

A. Data Collection and Pre-processing

To begin with, automated methods of collecting rental data that have been developed over time can be achieved through web-scraping (also referred to as web-extraction). The first structured web-extraction techniques were developed and tested by research teams working on projects like the ROAD RUNNER project located in [36] and served to create an initial framework for the automatic collection of data from the web. Recent studies have used web scraping to collect thousands of rental listings from various market-based websites, which facilitated impartial and diverse datasets for determining the rental rates throughout the Midwest [8]. Merging different rental databases into a single database using multi-source data fusion techniques also adds to the reliability of the developed datasets [9].

To get more accurate and reliable results, it's critical to properly preprocess the data before applying analytical models. Collection of raw rental data is often incomplete, as there are missing, duplicate, out-of-format, or extreme outliers in the collection of rental data. Therefore, to ensure that the data analysed is accurate, all collected rental data sets are cleaned, normalised, encoded into categories and statistically examined to remove outliers before analytical model development. The accuracy of analytical models created in a big-data-based housing research is dependent on whether the collected data, through proper pre-processing and no bias was ever established [16], [17].

B. Feature Engineering and Spatial Integration

The Rental Prices of Housing are not only affected by the characteristics of the property but also by the geographic and socio-economic conditions of the area where the property is located. Spatial Machine Learning studies show that the inclusion of location-based characteristics can substantially improve the model's accuracy in predicting rental price [27]. Feature-based integration using GIS allows for the modelling of accessibility, proximity to public transportation, and nearby facilities such as hospitals and grocery stores [29].

Further, spatio-temporal modelling for the purpose of capturing both temporal and location-based fluctuations in rental prices improves the accuracy of rental price forecasting [18]. Additionally, socioeconomic factors such as developments in the neighbourhood, the quality of infrastructure and access to services are very important to the rental price estimated in the area [24], [25]. The inclusion of these contextual and environmental factors into the overall model for rental price provides a more comprehensive understanding of the dynamics of rental price determination.

C. Explainable AI and Machine Learning Framework

The recent body of literature indicates that the prediction of housing prices using ensemble models outperforms the use of traditional regression models. An example of using weighted ensembles to create robust and stable predictions has been confirmed in the literature [1], as well as using the XG boost boosting methodology to create accurate predictions using gradient substitution [6] and [15]. The findings from the previous research include the use of ensemble methods in the prediction of housing rents and the use of evaluation parameters such as MAE, RMSE, and R^2 to evaluate the performance of the models. [7]. Although predictive accuracy is an important criterion, it does not really help a model when deployed in the "real-world," therefore, explainable AI will also need to be utilized in conjunction with the machine learning models so that a user will understand how much influence each feature had on the final prediction, thus creating transparency by providing the user with information regarding which features impact the prediction of the price of a location, and why [3] and [4]. Therefore, when analysing the use of features such as location, amenity, and infrastructure to determine the amount of rent that should be charged, the user needs to have access to the explanation so they can trust the model and also assist in fairly determining the amount of rent to charge [12], [13].

D. Cost of Living Forecast and Intelligent Recommendations

Previous studies mainly focused on estimating rental values of property without including total property affordability factors. To overcome this limitation, this new solution estimated the monthly total cost of renting a property rather than estimating just the rental amount. The total monthly cost of renting would also include utilities, food, the cost of Internet connection, the cost of commuting, and would thus provide a more realistic way of looking at the cost of renting as an "affordability index." Some studies that studied urban housing markets have pointed out that it is important to study costs (to rent) in context rather than study prices in isolation [16], [21].

Equitable rent sharing between flatmates will also be modelled using the Shapley Value from Cooperative Game Theory, to ensure that each flatmate pays based on their room and amenities. While machine learning is used to provide an estimate of fair rental prices in price fairness [3]; incorporating a mechanism for fairly distributing rental costs, such as the Shapley Value from Cooperative Game Theory to flatmates makes the use of price fairness to flatmates much more practical. An AI-based recommendation engine will also analyse user preferences, e.g., budget, work location proximity, lifestyle needs, etc., to recommend personalised locations/areas. This integrated approach supports addressing four identified gaps in research by integrating predictive modelling, explainability, cost aggregation, and intelligent recommendations in one methodology.

IV. CHALLENGES AND RESEARCH GAPS

Machine learning and housing price prediction have made significant strides; however, there still exist many obstacles for the implementation of an all-inclusive rental and living cost estimation system. One of the biggest challenges of developing such a model is collecting rental data with both sufficient volume, and quality. Large-scale rental data collection has been made possible using web scraping techniques [8], [36] but most web-scraped datasets contain, non-conformity, missing data, duplicate listings, and unfair scrutiny. The use of big data frameworks is intended to address the scalability and integration issues associated with rental data [17] but real-time validation/standardisation of the data requires continuing to be addressed in rental analytics systems.

Incorporating spatial and contextual factors is another major challenge in developing an all-encompassing predictive rental/living cost model, although spatial modeling techniques such as Geographic Information Systems (GIS) can assist with improving prediction accuracy through geographic dependency modelling [27], [29] a large number of existing models utilize attributes at the property level to derive rental price predictions (e.g., optimum size and amenities). Spatio-Temporal modelling techniques improve forecasting results [18] but, as with all modelling techniques, require significant computational resources and updated data continuously to implement efficiently. Furthermore, socioeconomic indicators and accessibility indicators are often not utilised uniformly throughout the various studies as listed [24], [25] ultimately leaving the representation of the dynamics of rental markets incomplete.

The interpretability of the models has remained a major research focus. For instance, ensemble and boosting methods have been found effective in the prediction of rental prices[1], [15], [23] ; however, most complex systems are classified as black boxes, and therefore, there is no transparency in how those predictions are made. The employment of Explainable AI techniques (e.g., by developing SHAP) improves the application of transparency [3], [4] however, current rental platforms in the real world do not utilize/exhibit transparency through Explainable AI techniques.

A second research gap is the absence of holistic cost estimation frameworks. The majority of research has focused on estimating rent with no consideration for other living costs (such as utilities, food, internet, and transportation) [16], [21].Therefore, most affordability analyses are incomplete. There is a clear need for comprehensive, integrated systems (which combine rent predictions and a full evaluation of living costs) that can provide students and employed individuals with insights into the true cost of housing. In addition, fairness and personalisation remain largely unexplored factors within the current body of literature. Predictive models can be used to estimate property value [3]; however, there is limited research available regarding fair rent distribution methods among roommates, based on cooperative gaming theory principles. Furthermore, there is an absence of AI-generated personalised recommendations based on user preferences, workplace proximity, and lifestyle needs within the housing analytics market at this time. Most current systems only provide comparisons of multifamily listings, rather than providing intelligent tools for decision-making.

Finally, scalability and real-time responsiveness remain areas where current models really struggle, as they are all incredibly time-consuming to promote and require constant retraining and updating data. Although big data approaches can address scale to some extent [10], [16] integrated, automated, user-oriented property rental platforms, incorporating predictions, explanations, cost estimation and recommendations, remain largely absent. Thus, in conclusion, it may be said that although a significant amount of work has been done in machine learning-based housing price prediction, a gap still exists in the empirical validation of the performance of the model in different and practical environments.

V. FUTURE SCOPE

Future research involving rental cost estimation and living cost estimation systems should emphasise the integration of both live dynamic data pipelines and sophisticated predictive modelling techniques. While there is a large amount of literature that discusses using web scraping methods for collecting rental data from the internet [8], [36] many of these systems rely heavily on being provided data sets that are periodically updated, rather than on having access to continuous live data streams. If systems were designed to use real-time automated data synchronisation and incorporate other types of economic forecasts, such as seasonal variations in demand, the resulting level of reliability could be significantly increased. Similarly, there are many innovative spatio-temporal modelling approaches that allow for the integration of geographic dependencies and historical-based trends [18], [27] which will allow for more accurate and enhanced granular prediction of rental markets at the area level by improving the quality of data used in rental analytics.

An additional area of exploration would be expanding the usability of explainable and multimodal artificial intelligence frameworks. For example, ensemble learning models and boosting techniques have consistently produced high predictive performances [1], [15] and if these models were integrated with deep neural networks having multimodal inputs, such as urban infrastructure data sets and accessibility metrics, then the contextual understanding of all of the multipoint data could be improved as suggested by [5]. Meanwhile, explainable AI models such as SHAP [3], [4] can benefit from more robust implementations by extending to more dynamic forms of visualisation, such as interactive dashboards for allowing users' dynamic ability to explore both feature importance (with respect to pricing) as well as justification of pricing. As these improvements take hold, they will tend to create significantly stronger user trust, as well as develop greater practical adoption rates of AI-based rental estimation systems.

The overall focus for future systems should be on providing holistic affordability modelling and intelligent personalisation. Much of the current research has focused primarily on predicting rental prices without considering all the costs associated with living [16], [21]. Including utility forecasting, commute calculations, and inflation-based expense calculations to develop a more accurate assessment of the cost of living could provide a better measure of what is considered affordable. In addition, advanced recommendation algorithms and equitable rent distribution mechanisms could lead to greater personalisation and management of shared housing. Combining cloud technology, smart city data integration, and AI-driven decision-making will create advanced platforms for generating an estimate of rental and overall cost of living.

VI. CONCLUSION

The paper systematically analysed recent advances in rent and housing price prediction models published between 2001 and 2025. The review evaluated the development of machine learning, ensemble learning, deep learning, spatial analytics, and Explainable AI approaches to the real estate valuation process. The results indicate that ensemble methods such as Random Forest, Gradient Boosting and XGBoost consistently outperform traditional regression approaches in terms of predictive accuracy. The combination of spatial and spatio-temporal attributes also increases model performance because they capture geographic and temporal dependence within the housing market. In addition, Explainable AI approaches such as SHAP increase transparency and interpretability of pricing decisions.

Despite these advances in predictive modelling, there are still several limitations present in the current research. Firstly, most current studies examine rental price exclusively and do not consider all aspects of cost-of-living, which could include utilities, food, and commute costs. Secondly, while web scraping facilitates data acquisition at scale, challenges remain regarding data quality and the ability to provide real-time updates or obtain standardised data sets. Finally, there is limited use of fairness mechanisms regarding rent-sharing and Personalized Recommendations using AI in the current housing analytics literature.

With regard to identified research gaps, this article provides an argument for a single, AI-enabled rental and living costs evaluation system encompassing real-time web scraped information, ensemble-based rental price prediction, explainable artificial intelligence (AI) mechanisms, holistic cost of living models, and advanced recommendation engines. By creating an integrated framework together, these components will enable students, individuals seeking employment, and working-class individuals to have useful decision-making tools and enhance the transparency and data-driven nature of urban housing management. The outcomes of this analytical review establish a foundation for developing the next era of intelligent rental analytics systems.

REFERENCES:

1. S. Im, K. Kim, G. Lee, and H.-J. Lim, “Development of a weighted average ensemble model for predicting officially assessed land prices,” IEEE Access, vol. 13, pp. 96251–96260, 2025, doi: 10.1109/ACCESS.2025.3574698.
2. C. Stacy, T. R. Hodge, T. M. Komarek, C. Davis, A. Stern, O. Noble, J. Morales-Burnett, and A. Rogin, “Rent control and the supply of affordable housing,” Journal of Housing Economics, vol. 68, Art. no. 102063, 2025, doi: 10.1016/j.jhe.2025.102063. [3]
3. T. Kee and W. K. O. Ho, “eXplainable machine learning for real estate: XGBoost and Shapley values in price prediction,” Civil Engineering Journal, vol. 11, no. 5, pp. 2116–2133, 2025, doi: 10.28991/CEJ-2025-011-05-022.
4. E. Petrova and V. Ivanov, “Explainable ensemble models for urban housing price prediction,” Information Fusion, vol. 105, pp. 1–18, 2024, doi: 10.1016/j.inffus.2023.102214.
5. H. Tan, T. Su, X. Wu, P. Cheng, and T. Zheng, “A sustainable rental price prediction model based on multimodal input and deep learning—Evidence from Airbnb,” Sustainability, vol. 16, no. 15, Art. no. 6384, 2024, doi: 10.3390/su16156384.
6. H. Sharma, H. Harsora, and B. Ogunleye, “An optimal house price prediction algorithm: XGBoost,” Analytics, vol. 3, no. 1, pp. 30–45, 2024, doi: 10.3390/analytics3010003.
7. J. Zhang, “Predicting rental prices in India: A machine learning approach with hyperparameter optimization,” 2024.
8. C. Meyberg, U. Rendtel, and H. Leerhoff, “Flat rent price prediction in Berlin with web scraping,” AStA Wirtschafts- und Sozialstatistisches Archiv, vol. 18, no. 2, pp. 155–178, 2024, doi: 10.1007/s11943-024-00312-5.

9. Y. Zhao, J. Zhao, and E. Y. Lam, “House price prediction: A multi-source data fusion perspective,” Big Data Mining and Analytics, vol. 7, no. 1, pp. 120–134, 2024, doi: 10.26599/BDMA.2024.9020012.
10. W. Zhang and L. Chen, “Big data-driven housing market analysis using machine learning,” IEEE Access, vol. 11, pp. 98765–98780, 2023, doi: 10.1109/ACCESS.2023.3298745.
11. C. Fernández and A. Ruiz, “Explainable boosting models for real estate valuation,” Applied Artificial Intelligence, vol. 37, no. 1, pp. 1–18, 2023.
12. I. Rossi and M. Bianchi, “Explainable machine learning for property valuation,” Information Sciences, vol. 646, Art. no. 119242, 2023, doi: 10.1016/j.ins.2023.119242.
13. L. Zhao and M. Liu, “Explainable artificial intelligence in real estate pricing,” Applied Soft Computing, vol. 143, Art. no. 110382, 2023, doi: 10.1016/j.asoc.2023.110382.
14. N. Hassan and O. Khalil, “Deep neural networks for real estate market forecasting,” Applied Intelligence, vol. 52, no. 14, pp. 16543–16560, 2022, doi: 10.1007/s10489-022-03547-8.
15. X. Zhu, Y. Zhang, P. Nie, and X. Liu, “A Housing Price Prediction Method Based on Stacking Ensemble Learning Optimization Method,” IEEE Access, 2022, doi: 10.1109/CSCloud-EdgeCom58631.2023.00025.
16. R. Wang and X. Li, “Big data analytics for urban housing markets,” Sustainable Cities and Society, vol. 80, Art. no. 103789, 2022, doi: 10.1016/j.scs.2022.103789.
17. A. García and E. Martínez, “Big data approaches in real estate market analysis,” Journal of Big Data, vol. 9, no. 1, Art. no. 45, 2022
18. A. Soltani, M. Heydari, F. Aghaei, and C. J. Pettit, “Housing price prediction incorporating spatio-temporal dependency into machine learning algorithms,” Cities, vol. 130, Art. no. 103923, 2022, doi: 10.1016/j.cities.2022.103923.
19. D. Nguyen and E. Roberts, “Hybrid machine learning models for real estate price prediction,” Knowledge-Based Systems, vol. 245, Art. no. 108615, 2022, doi: 10.1016/j.knosys.2022.108615.
20. T. Suganya, G. Gowtham, and H. S. Raja, “Machine learning based prediction of house price,” International Journal of Health Sciences, vol. 6, no. S2, pp. 12890–12899, 2022, doi: 10.53730/ijhs.v6nS2.7985
21. T. Müller and A. Schneider, “Urban big data and predictive analytics in housing markets,” Cities, vol. 129, Art. no. 103876, 2022, doi: 10.1016/j.cities.2022.103876.
22. H. Kim, J. Lee, and H. Kim, “A deep learning model for real estate price prediction using LSTM networks,” IEEE Access, vol. 9, pp. 135789–135801, 2021, doi: 0.1109/ACCESS.2021.3112345.
23. A. Schmidt and L. Weber, “Ensemble learning approaches for residential property price prediction,” Expert Systems with Applications, vol. 178, Art. no. 115007, 2021, doi: 10.1016/j.eswa.2021.115007.
24. O. Martin and B. Clark, “The role of socioeconomic factors in housing price prediction using machine learning,” Habitat International, vol. 112, Art. no. 102373, 2021, doi: 10.1016/j.habitatint.2021.102373.
25. T. Yamamoto and H. Sato, “Urban accessibility and housing prices: A machine learning perspective,” Sustainable Cities and Society, vol. 72, Art. no. 103051, 2021, doi: 10.1016/j.scs.2021.103051.
26. A. L. Kokkinos and D. T. Koutmos, “Machine learning applications in real estate valuation,” Expert Systems with Applications, vol. 168, Art. no. 114301, 2021, doi: 10.1016/j.eswa.2020.114301.
27. J. Park and S. Lee, “Spatial machine learning for property valuation using GIS data,” Computers, Environment and Urban Systems, vol. 81, Art. no. 101485, 2020, doi: 10.1016/j.compenvurbsys.2020.101485.
28. D. McCluskey and P. McCord, “Spatial analysis of residential property markets,” Urban Studies, vol. 57, no. 8, pp. 1645–1663, 2020.

29. P. Silva and M. Costa, “GIS-based spatial modeling for housing price estimation,” Sustainable Cities and Society, vol. 54, Art. no. 101987, 2020, doi: 10.1016/j.scs.2020.101987.
30. L. Ferreira and S. Almeida, “Real estate price prediction using support vector machines,” Expert Systems with Applications, vol. 144, Art. no. 113113, 2020.
31. H. K. Singla and P. Bendigiri, “Factors affecting rentals of residential apartments in Pune, India,” International Journal of Housing Markets and Analysis, vol. 12, no. 4, pp. 699–715, 2019, doi: 10.1108/IJHMA-11-2018-0098.
32. M. Brown and S. Thompson, “Comparative study of machine learning algorithms for housing price prediction,” Expert Systems with Applications, vol. 129, pp. 456–468, 2019, doi: 10.1016/j.eswa.2019.03.045.
33. J. López and M. Sánchez, “Neural network-based real estate valuation models,” Applied Soft Computing, vol. 78, pp. 512–523, 2019, doi: 10.1016/j.asoc.2019.02.012.
34. W. Chen, Y. Deng, and Z. Zhang, “House prices and local amenities: Evidence from China,” Journal of Real Estate Finance and Economics, vol. 58, no. 3, pp. 367–389, 2019, doi: 10.1007/s11146-018-9675-4.
35. M. Pagourtzi and V. Assimakopoulos, “Real estate appraisal: A review of valuation methods,” Journal of Property Investment & Finance, vol. 36, no. 6, pp. 557–574, 2018, doi: 10.1108/JPIF-03-2018-0021.
36. V. Crescenzi, G. Mecca, and P. Merialdo, “The ROADRUNNER project: Towards automatic extraction of web data,” Università di Roma Tre, 2001.