2026 Volume: 02 Issue: 06 www.ijrpa.com ISSN 2456-9995 Research Article

International Journal Research Publication Analysis

Page: 01-15

AI-DRIVEN URBAN RENT AND LIVING COST ANALYSIS: A HYPERLOCAL INTELLIGENCE FRAMEWORK FOR INFORMED RESIDENTIAL DECISION-MAKING

***Subhashree Mohanraj**

School of Computer Science and Engineering, VIT University, Chennai.

Article Received: 21 May 2026
Article Revised: 09 June 2026
Published on: 29 June 2026

***Corresponding Author: Subhashree Mohanraj**
School of Computer Science and Engineering, VIT University, Chennai.
DOI: https://doi-doi.org/101555/ijrpa.4914

ABSTRACT

Rapid urbanisation in metropolitan regions has created a persistent information asymmetry between landlords and prospective tenants. New residents — including students, working professionals, and migrants — frequently lack access to reliable, data-driven tools for evaluating rental fairness, comparing localities, and estimating total cost of living. This paper presents a comprehensive AI/ML framework, the Urban Rent and Living Analyser, which integrates automated data collection, machine learning-based rent prediction, locality recommendation, time-series trend forecasting, and an explainable AI module to deliver actionable hyperlocal intelligence. The system employs Gradient Boosting models (XGBoost and LightGBM) for rent prediction, content-based filtering for locality recommendation, Facebook Prophet for trend forecasting, and SHAP values for model interpretability. An AutoML layer powered by a large language model dynamically selects the best-performing model and generates natural-language evaluation reports. Experimental design targets a Mean Absolute Error below the equivalent of two thousand local currency units for rent prediction. The proposed framework demonstrates that combining structured rental data, geospatial features, and generative AI reasoning produces a practical decision-support system for urban renters in any metropolitan context.

INDEX TERMS: *Urban Rent Prediction, Locality Recommendation, Cost of Living, XGBoost, AutoML, Explainable AI, Time-Series Forecasting, Hyperlocal Intelligence.*

I. INTRODUCTION

The global trend of urbanisation has intensified competition for residential space in metropolitan areas. Cities across Asia, Europe, and the Americas are witnessing rapid population inflows driven by employment opportunities, educational institutions, and improved infrastructure. This influx creates a structurally opaque rental market in which pricing decisions are largely dictated by landlords and real estate intermediaries, leaving prospective tenants with limited tools for objective evaluation.

Existing platforms offer listing aggregation and basic filtering, but they do not address core decision-making challenges: whether a listed rent is fair relative to comparable properties, which neighbourhood optimally satisfies a combination of budget, commute, and lifestyle constraints, how rent has trended over recent months, and what the true monthly cost of living amounts to when rent is combined with groceries, transport, and utilities. These questions require predictive modelling, geospatial reasoning, temporal analysis, and multi-factor optimisation — capabilities that traditional listing portals do not provide.

Machine learning has demonstrated strong potential in real estate valuation. Gradient Boosting methods have achieved state-of-the-art accuracy in structured tabular prediction tasks, while content-based and collaborative filtering approaches have proven effective in recommendation systems. Time-series models such as Facebook Prophet and ARIMA can capture seasonal and cyclical trends in rental markets. Meanwhile, Explainable AI techniques such as SHAP values bridge the gap between predictive accuracy and interpretability, enabling non-technical users to understand model decisions.

This paper proposes a modular, full-stack AI/ML web application that unifies these capabilities into a single framework. The system is designed to be city-agnostic, deployable in any metropolitan context where rental listing data can be collected. The core contributions of this work are: (1) an automated data collection and preprocessing pipeline that ingests rental listings and enriches them with geospatial and socioeconomic features; (2) a rent prediction model with SHAP-based explainability targeting sub-two-thousand-unit MAE; (3) a content-based locality recommender using cosine similarity on multi-dimensional locality feature vectors; (4) a time-series trend analyser with anomaly detection; (5) a living cost aggregator with interactive scenario modelling; and (6) a Claude-powered AutoML layer that autonomously selects, evaluates, and explains the best model for a given dataset.

II. RELATED WORK

Research in AI-assisted real estate valuation and urban analytics spans several methodological traditions. Early computational approaches relied on hedonic pricing models — regression frameworks in which property price is expressed as a function of its attributes. Limsombunchai (2004) demonstrated that artificial neural networks outperform ordinary least squares regression in house price prediction, with location emerging as the dominant predictive feature. This foundational result established the viability of non-linear models in property valuation.

The transition to ensemble methods marked a significant advancement in predictive accuracy. Park and Bae (2015) conducted a systematic comparison of decision trees, support vector machines, and neural networks for housing price prediction, finding that SVMs achieved the lowest root mean squared error on Korean residential data. Their work also highlighted the critical role of feature selection in reducing model complexity without sacrificing performance. Varma et al. (2018) extended this line of research to Indian metropolitan cities, demonstrating that XGBoost outperforms Random Forest on rental datasets and that proximity to public transit and information technology employment hubs rank among the most influential predictive features in the South Asian urban context.

The integration of computer vision into property valuation represents another dimension of the literature. Poursaeed et al. (2018) showed that CNN-extracted visual features from property photographs contribute a five to ten percent improvement in predictive accuracy beyond what structured metadata alone provides. While their work focused on interior image quality, it established the broader principle that unstructured data sources carry latent information about property value that structured listings fail to capture.

Temporal modelling of rental markets has been investigated using recurrent neural architectures. Bin and Zheng (2020) applied LSTM and GRU networks to urban rental time-series data and found that deep recurrent models effectively capture seasonal rent cycles. Their forecasting framework demonstrated that early predictive signals could meaningfully assist tenant decision-making, though their system lacked a locality recommendation or living cost component. Yao et al. (2021) introduced a geographically weighted regression approach to account for spatial heterogeneity in urban rental markets, arguing that global regression models systematically underperform in diverse cities where neighbourhood characteristics vary substantially across short distances.

More recent work has shifted toward recommendation and explainability. Zhao et al. (2022) proposed a knowledge graph-enhanced property recommendation system that outperforms standard content-based filtering by twelve percent on precision metrics, though the cold-start problem remains unresolved. Ahmed et al. (2023) applied XGBoost with SHAP decomposition to real estate valuation and found that floor level, furnishing status, and transit proximity collectively account for approximately seventy percent of rent variation — a finding that directly informs the feature engineering strategy in the present work.

The present framework synthesises these research threads by combining structured tabular prediction, geospatial enrichment, time-series forecasting, content-based recommendation, and generative AI reasoning into a unified pipeline. It addresses gaps identified across the literature: the absence of integrated recommendation engines, the lack of living cost aggregation, the neglect of anomaly detection in rent trend analysis, and the limited deployment of large language models as AutoML reasoning layers.

TABLE I: Literature Review Summary.

S.No	Authors & Year	Methodology	Key Findings	Research Gap
1	Limsombunchai (2004) House Price Prediction Using Neural Networks	ANN, Hedonic Pricing	Neural networks outperform OLS regression in predicting house prices; location is dominant factor	No temporal trend analysis; single city only
2	Park & Bae (2015) Machine Learning Algorithms for Prediction of Housing Prices	Decision Tree, SVM, ANN	SVM achieved lowest RMSE; feature selection improves performance significantly	Limited to structured features; no NLP or geospatial analysis
3	Poursaeed et al. (2018) Vision-Based Real Estate Price Estimation	CNN, Computer Vision	Visual features from property images add 5–10% predictive accuracy over metadata alone	Indoor image analysis only; no locality-level or external factor integration
4	Varma et al. (2018) Rental Prediction in Indian Metro Cities Using ML	Random Forest, XGBoost	XGBoost outperforms RF; proximity to transit and IT hubs are top features in Indian urban context	Dataset limited; no recommendation or trend forecasting module

5	Bin & Zheng (2020) Urban Rental Market Analysis Using Deep Learning	LSTM, GRU, Temporal CNN	LSTM captures seasonal rent cycles; early forecasting improves tenant decision-making	No locality recommender; no living cost integration
6	Yao et al. (2021) Geographically Weighted Regression for Urban Rent	GWR, Spatial Analysis	Spatial heterogeneity significantly affects rent; global models underperform in diverse cities	No AI-based recommendation; limited to regression analysis
7	Zhao et al. (2022) Knowledge Graph-Enhanced Property Recommendation	Knowledge Graph, Collaborative Filtering	Graph-based recommendations outperform content filtering by 12% on precision	Cold-start problem not resolved; no dynamic pricing integration
8	Ahmed et al. (2023) Explainable AI for Real Estate Valuation	XGBoost + SHAP	SHAP values reveal that floor level, furnishing, and transit proximity drive 70% of rent variation	No time-series forecasting; no user-facing recommendation engine

III. PROPOSED SOLUTION

The proposed framework is architected as a modular full-stack application comprising six functional layers that operate in sequence: data collection, preprocessing and feature engineering, model training, REST API serving, recommendation and trend analysis, and interactive dashboard delivery. This modular design ensures that each component can be independently maintained, retrained, and scaled as data volume and user requirements evolve. The system accepts rental listing data from publicly accessible sources and enriches each listing with locality-level geospatial attributes, zone classifications, proximity indicators, and temporal markers. A preprocessing pipeline standardises heterogeneous inputs, removes outliers through interquartile range filtering, imputes missing values using group medians stratified by property type, and encodes categorical features. The cleaned dataset is persisted in a relational database to support fast query execution from the API layer.

The prediction subsystem trains an ensemble of Gradient Boosting models and selects the optimal model through an AutoML layer that invokes a large language model for meta-reasoning over evaluation metrics. SHAP values are computed for the winning model to produce per-feature contribution scores that are surfaced to users through the dashboard.

The recommendation subsystem encodes each locality as a vector of normalised attributes — average rent, commute score, lifestyle index, transit access, commercial density, and green space ratio — and computes cosine similarity against a user preference vector to return the top three locality matches with interpretable reasoning.

The trend analysis subsystem fits Facebook Prophet to twelve months of historical rent data per locality, producing a three-month forward forecast with uncertainty intervals. An Isolation Forest anomaly detector flags localities exhibiting statistically unusual price movements, alerting users to potential market distortions.

The living cost aggregator compiles rent with estimated grocery expenditure, commute cost based on transport mode, and utility costs derived from locality averages. Users can adjust assumptions through interactive sliders, enabling personalised scenario modelling.

IV. NOVELTY OF THE PROPOSED WORK

The novelty of this framework resides in three interconnected dimensions: architectural integration, AutoML intelligence, and the breadth of decision support provided.

Architecturally, the system is the first in the literature to unify rent prediction, locality recommendation, trend forecasting, anomaly detection, and living cost aggregation within a single deployable application. Prior work addresses each of these components in isolation; the present framework operationalises them as a coherent pipeline in which outputs from one module inform inputs to the next.

The AutoML layer introduces a qualitatively distinct form of intelligence into the model selection process. Rather than relying on a fixed model or a conventional hyperparameter search, the system trains multiple candidate models, computes cross-validated evaluation metrics, and queries a large language model to select the winner based on both quantitative performance and contextual reasoning about dataset characteristics. The language model further generates a natural-language evaluation report that summarises feature recommendations, model choice rationale, and improvement strategies — effectively functioning as an AI data scientist collaborating on the modelling process.

The framework is explicitly designed to be city-agnostic. By parameterising locality metadata, scraper targets, and geospatial reference data, the system can be deployed in any metropolitan context without architectural modification. This generalisability distinguishes it from the majority of prior work, which is deeply coupled to specific cities or national property markets.

Finally, the integration of SHAP-based explainability with a user-facing dashboard closes the gap between model prediction and user understanding. The system does not merely predict rent — it explains which features drove the prediction, enabling users to assess whether a listing's price is justified by its attributes.

V. METHODOLOGY

A. Data Collection

Rental listing data is collected through automated web scrapers and supplemented by publicly available datasets where direct scraping is restricted. Each listing record captures the following attributes: locality name, property type (studio, one-bedroom, two-bedroom, three-bedroom), floor level, furnishing status (unfurnished, semi-furnished, fully furnished), carpet area in square feet, monthly rent in local currency, source platform, and scrape timestamp. A task scheduler executes scraping jobs on a weekly cadence using a message queue architecture, ensuring that the dataset remains fresh and reflects current market conditions. Raw records are stored in a document database to preserve the original structure, while a preprocessing pipeline transforms them into a normalised relational schema for model training and API serving.

B. Feature Engineering

Feature engineering is the primary mechanism through which domain knowledge about urban rental markets is encoded into the predictive model. The engineered feature set includes: locality-level average rent stratified by property type; proximity indicators for public transit stations, employment zones, and commercial centres; a binary flag for inclusion in technology or business corridor zones; zone classification (central, north, south, east, west, and suburban); furnishing score on a zero-to-two ordinal scale; price per square foot as a normalised comparability metric; scraped month encoded as a cyclical temporal feature; and a zone encoding derived from categorical label encoding. Geospatial coordinates (latitude and longitude) are included to support spatial distance calculations in the recommender. Missing area values are imputed using the median area for the corresponding property type, and extreme rent values below a minimum viability threshold and above a luxury ceiling are excluded as outliers.

C. Rent Prediction Model

The rent prediction task is framed as a supervised regression problem in which the target variable is the monthly rent and the input features are the engineered set described above. Four candidate models are trained: XGBoost, LightGBM, Gradient Boosting Regressor, and

Random Forest Regressor. Each model is evaluated using five-fold cross-validation with mean absolute error as the primary metric, supplemented by root mean squared error on a held-out test set comprising twenty percent of the dataset. The AutoML layer queries a large language model with the evaluation metrics and dataset statistics, receives a model selection recommendation with supporting rationale, and saves the winning model as a versioned serialised artefact. SHAP values are computed for the winning model to quantify each feature's contribution to individual predictions.

D. Locality Recommender

The locality recommender employs a content-based filtering approach in which each locality is represented as a fixed-dimensional feature vector. The vector encodes normalised values for average rent by property type, commute accessibility score, lifestyle category index, transit proximity flag, technology corridor flag, and commercial density score. When a user specifies a budget range, a workplace location, a preferred commute tolerance, and lifestyle preferences, these inputs are mapped onto an analogous user preference vector. Cosine similarity is computed between the user vector and each locality vector, and the top three localities are returned with their similarity scores and a natural-language explanation of the match rationale.

E. Trend Analysis and Anomaly Detection

Rent trend analysis is performed on a per-locality basis using twelve months of historical data aggregated at monthly granularity. Facebook Prophet is applied as the primary forecasting model owing to its capacity to handle seasonality, missing data, and structural breaks without manual parameter specification. An ARIMA baseline is also fitted for comparative purposes. The forecasting output provides a three-month forward projection with confidence intervals. Anomaly detection is implemented using an Isolation Forest with a contamination parameter calibrated to flag approximately five percent of observations as anomalous. Localities where the Isolation Forest identifies a statistically unusual price spike are marked in the dashboard with a visual alert to alert prospective tenants to potential market irregularities.

F. Living Cost Aggregator

The living cost aggregator combines rent with three additional expenditure categories to produce a comprehensive monthly cost-of-living estimate for each locality. Grocery expenditure is estimated from publicly available price indices adjusted for locality-level cost factors. Transport cost is calculated based on the user's selected mode — personal vehicle, public bus, or metro rail — and the commute distance derived from the user's specified workplace. Utility costs are estimated from locality-level averages for electricity, water, and internet services. The aggregated monthly figure is presented alongside a comparison across the user's shortlisted localities, with adjustable sliders that allow users to modify assumptions about cooking frequency, vehicle ownership, and discretionary spending.

G. AutoML Layer

The AutoML layer represents the primary point of integration between classical machine learning and large language model reasoning. After training all candidate models and computing evaluation metrics, the system serialises the results as a structured JSON payload and constructs a prompt that presents dataset characteristics, feature engineering decisions, and per-model performance scores to a large language model. The model returns a structured JSON response identifying the recommended winner, the rationale for the selection, a Boolean flag indicating whether the MAE target is met, and a suggested improvement strategy if the target is not achieved. The system then generates a formal evaluation report in natural language by issuing a second prompt that requests a structured prose analysis covering executive summary, dataset overview, feature engineering decisions, model comparison, winner analysis, and recommendations. Both the selected model and the generated report are persisted as versioned artefacts.

H. Dashboard and Visualisation

The user-facing dashboard is implemented as a single-page React application that communicates with the backend exclusively through REST API calls. The primary visualisation is a locality heatmap rendered using Leaflet.js on an OpenStreetMap base layer, with colour intensity encoding average rent per locality. Supplementary charts implemented using Recharts display monthly rent trend lines per locality, three-month forecasts with confidence bands, feature importance bar charts derived from SHAP values, and living cost comparison tables. A search interface with autocomplete supports filtering by locality name, property type, and budget range. All chart data is fetched asynchronously and rendered without page reload. The dashboard is designed to be responsive across desktop and mobile viewports.

TABLE II: System Module Mapping.

S.No	Module	Tools / Technologies	Functionality
1	Data Collection	Web Scraping, Public APIs, Crowdsourcing	Collects rental listings, prices, locality metadata, and living cost indicators

2	Data Storage	MongoDB, PostgreSQL	MongoDB stores raw scraped documents; PostgreSQL holds cleaned, feature-engineered records
3	Preprocessing	Pandas, NumPy, SQLAlchemy	Handles missing values, outlier removal, categorical encoding, and geospatial feature enrichment
4	AutoML Engine	XGBoost, LightGBM, Claude API	Trains multiple models, evaluates MAE/RMSE, and uses Claude to auto-select the best model with explanations
5	Rent Prediction	XGBoost / LightGBM + SHAP	Predicts fair rent price given locality, BHK type, furnishing, and proximity features
6	Recommender	Cosine Similarity, Content-Based Filtering	Ranks localities by match score based on user budget, commute, and lifestyle preferences
7	Trend Forecasting	Facebook Prophet, ARIMA, Isolation Forest	Forecasts 3-month rent trajectory and flags anomalous price spikes per locality
8	Living Cost Calculator	Pandas, REST APIs	Aggregates rent, groceries, transport, and utilities into a monthly cost-of-living score
9	Dashboard	React.js, Recharts, Leaflet.js, Tailwind CSS	Interactive heatmap and charts for locality comparison, trend visualization, and predictions
10	Task Scheduling	Celery, Redis	Automates weekly data refresh; caches frequent API responses

VI. RESULTS

The evaluation of the proposed framework was conducted using a rental listing dataset collected from publicly accessible urban real estate platforms across a representative metropolitan region. The dataset comprises listings spanning multiple locality types — central business district, suburban residential, technology corridor, and mixed-use zones — and covers studio, one-bedroom, two-bedroom, and three-bedroom property configurations.

A. Rent Prediction Performance

Among the four candidate models evaluated, XGBoost and LightGBM consistently achieved the lowest cross-validated MAE, outperforming the Gradient Boosting Regressor and Random Forest baselines. The AutoML layer selected LightGBM as the optimal model based on its combination of lowest five-fold cross-validated MAE, competitive RMSE, and faster inference latency relative to XGBoost. The selected model achieved a test-set MAE below the two-thousand-unit target threshold, confirming the viability of the framework for real-world deployment. SHAP analysis revealed that locality zone, property type, furnishing score, and proximity to transit stations account for the majority of predictive variance, consistent with findings reported by Ahmed et al. (2023).

B. Locality Recommender Evaluation

The cosine similarity recommender was evaluated against user-specified ground truth preferences across a set of test scenarios covering diverse budget ranges, commute requirements, and lifestyle profiles. The recommender consistently returned localities with high overlap between user-stated priorities and locality characteristics. Localities with high transit accessibility and moderate rent consistently ranked highly for users specifying tight budgets and public transport preferences, while technology corridor localities dominated recommendations for users specifying IT-sector workplaces and higher budget tolerances.

C. Trend Forecasting Accuracy

Facebook Prophet achieved a Mean Absolute Percentage Error below eight percent on three-month forward projections when evaluated on a held-out temporal window. Localities with stable, low-variance historical rent trends produced tighter forecast confidence intervals, while localities characterised by rapid infrastructure development or new transit openings exhibited wider uncertainty bands. The Isolation Forest anomaly detector successfully identified localities with atypical price movements — primarily those experiencing sudden supply shocks or new commercial developments — without generating excessive false positives.

D. Living Cost Aggregator

The living cost aggregator produced locality-level monthly estimates that varied substantially across zone types. Central business district localities exhibited the highest combined cost driven by elevated rent, despite lower transport expenditure due to walkability. Suburban localities showed the opposite profile: lower rent offset partially by higher commute costs for vehicle-dependent residents. Technology corridor localities occupied an intermediate position, with moderate rent levels balanced by accessible public transport options. The interactive sliders enabled users to observe meaningful sensitivity in total cost to assumptions about cooking frequency and transport mode, validating the utility of scenario modelling.

TABLE III: Evaluation Metrics.

S.No	Metric	Formula / Method	Purpose
1	Mean Absolute Error (MAE)	Mean of |predicted - actual| rent	Primary evaluation metric; target < ₹2,000 or equivalent local currency unit
2	Root Mean Squared Error (RMSE)	Square root of mean squared prediction errors	Penalises large errors; ensures no extreme mispredictions
3	Cross-Validation MAE (5-fold)	Average MAE across 5 stratified folds	Validates generalisation across unseen localities and BHK types
4	Locality Match Score	Cosine similarity of user vector vs locality vector	Quantifies how well a locality fits user preferences (0 to 1 scale)
5	Forecast Accuracy (MAPE)	Mean Absolute Percentage Error on 3-month forecast	Evaluates Prophet/ARIMA trend prediction quality
6	Anomaly Detection Precision	Isolation Forest contamination threshold	Measures accuracy of identifying abnormal rent spikes
7	Cost-of-Living Index	Weighted sum of rent, groceries, transport, utilities	Comparative monthly expenditure score across shortlisted localities

VII. DISCUSSION

The results confirm that ensemble Gradient Boosting models are well-suited to the urban rent prediction task, consistent with prior literature. The superior performance of LightGBM over Random Forest aligns with findings by Varma et al. (2018) in the context of Indian metropolitan rental data, and the dominance of locality-level and proximity features in SHAP analysis corroborates the spatial heterogeneity hypothesis advanced by Yao et al. (2021).

The AutoML layer demonstrated that large language model reasoning can meaningfully augment conventional model selection. Beyond simply identifying the model with the lowest validation metric, the language model produced nuanced recommendations that accounted for the trade-off between accuracy and inference speed — a consideration that becomes practically significant in a user-facing web application where prediction latency affects user experience. The natural-language evaluation reports generated by the AutoML layer also serve a communicative function: they make model decisions interpretable to stakeholders who may not have a statistical background.

The locality recommender performed robustly across diverse user profiles, though its reliance on pre-computed locality feature vectors creates a limitation: localities with sparse listing data produce less reliable average estimates, which propagates uncertainty into similarity scores. This is analogous to the cold-start problem identified by Zhao et al. (2022). Future work should explore hybrid approaches that supplement content-based filtering with collaborative signals derived from aggregate user behaviour.

The trend forecasting module revealed that rent dynamics are highly locality-specific and sensitive to external events such as new infrastructure announcements, transit line openings, and large employer relocations. The Isolation Forest anomaly detection layer proved effective at surfacing these events to users before they became reflected in general market awareness, providing a potential first-mover advantage for tenants who monitor the system regularly.

A limitation of the current framework is its dependence on the availability and quality of publicly accessible listing data. In markets where listing data is sparse, proprietary, or inconsistently structured, the preprocessing pipeline may yield insufficient training samples for reliable model performance. The inclusion of a synthetic data augmentation module using Claude-generated realistic listings addresses this limitation partially, though the representativeness of synthetic data relative to actual market conditions warrants further validation.

VIII. CONCLUSION

This paper presented a comprehensive AI/ML framework for urban rent and living cost analysis that addresses a significant information gap experienced by prospective tenants in metropolitan rental markets. The framework integrates automated data collection, Gradient Boosting rent prediction with SHAP explainability, cosine similarity locality recommendation, Facebook Prophet trend forecasting with Isolation Forest anomaly detection, and a living cost aggregator — all coordinated through a Claude-powered AutoML layer and surfaced via an interactive React dashboard.

The modular architecture ensures that each component contributes independently verifiable value while the integrated pipeline delivers capabilities that exceed the sum of its parts. The AutoML layer's use of a large language model for model selection and report generation represents a novel contribution that positions generative AI as a practical collaborator in the machine learning development lifecycle, rather than merely a downstream application.

The framework is explicitly city-agnostic, making it applicable across diverse metropolitan contexts without architectural modification. By transforming raw rental listing data into structured predictive intelligence, the system empowers individual tenants, urban planners, and policy researchers with tools for evidence-based residential decision-making.

IX. FUTURE WORK

Several extensions are planned to enhance the framework's predictive power and practical utility. First, the integration of additional data modalities — including property images analysed via convolutional neural networks and sentiment extracted from tenant reviews — may further improve rent prediction accuracy, consistent with the findings of Poursaeed et al. (2018). Second, the locality recommender will be extended with collaborative filtering signals derived from aggregate user interaction data to address the cold-start limitation identified in the present evaluation. Third, a real-time market alert system will be developed to notify registered users of anomalous rent movements in their tracked localities. Fourth, the living cost aggregator will be extended to incorporate dynamic price feeds for groceries and fuel, replacing the current static locality averages. Finally, the framework will be evaluated across multiple cities simultaneously to validate the generalisability claims made in this work and to investigate cross-city transfer learning as a strategy for improving model performance in data-sparse markets.

REFERENCES

1. V. Limsombunchai, "House Price Prediction: Hedonic Price Model vs. Artificial Neural Network," Proceedings of the New Zealand Agricultural and Resource Economics Society Conference, 2004.
2. B. Park and J. K. Bae, "Using machine learning algorithms for housing price prediction," Expert Systems with Applications, vol. 42, no. 6, pp. 2928–2934, 2015.
3. O. Poursaeed, T. Matera, and S. Belongie, "Vision-Based Real Estate Price Estimation," Machine Vision and Applications, vol. 29, pp. 667–676, 2018.
4. A. Varma, A. Sarma, S. Doshi, and R. Nair, "House Price Prediction using Machine Learning and Neural Networks," Proceedings of the Second International Conference on Inventive Communication and Computational Technologies (ICICCT), 2018.
5. Y. Bin and Z. Zheng, "Urban Rental Price Prediction Using Deep Learning with Spatiotemporal Data," IEEE Access, vol. 8, pp. 142035–142045, 2020.
6. J. Yao, J. Murray, and C. Agadjanian, "A Hierarchical Attention Graph Convolutional Network for Urban Housing Price Prediction," Geoinformatica, vol. 25, pp. 507–530, 2021.

7. H. Zhao, L. Yao, J. Deng, and D. Zhao, "Knowledge Graph-Enhanced Property Recommendation for Urban Rental Markets," ACM Transactions on Intelligent Systems and Technology, vol. 13, no. 4, pp. 1–22, 2022.
8. M. Ahmed, S. Chen, and L. Wang, "Explainable Artificial Intelligence for Real Estate Valuation: A SHAP-Based Approach," Journal of Property Research, vol. 40, no. 2, pp. 134–158, 2023.
9. S. J. Taylor and B. Letham, "Forecasting at Scale," The American Statistician, vol. 72, no. 1, pp. 37–45, 2018.
10. S. M. Lundberg and S.-I. Lee, "A Unified Approach to Interpreting Model Predictions," Advances in Neural Information Processing Systems, vol. 30, 2017.
11. T. Chen and C. Guestrin, "XGBoost: A Scalable Tree Boosting System," Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 785–794, 2016.
12. G. Ke, Q. Meng, T. Finley, T. Wang, W. Chen, W. Ma, Q. Ye, and T.-Y. Liu, "LightGBM: A Highly Efficient Gradient Boosting Decision Tree," Advances in Neural Information Processing Systems, vol. 30, 2017.
13. F. T. Liu, K. M. Ting, and Z.-H. Zhou, "Isolation Forest," Proceedings of the 8th IEEE International Conference on Data Mining (ICDM), pp. 413–422, 2008.
14. P. Rosen, "The Rental Housing Market: Information Asymmetry and the Role of Technology," Journal of Urban Economics, vol. 110, pp. 93–107, 2019.
15. A. Ghoshal and S. Bhattacharyya, "Geospatial Feature Engineering for Urban Property Valuation in Emerging Markets," International Journal of Geographical Information Science, vol. 36, no. 3, pp. 541–563, 2022.