Global Journal of Advanced Engineering Technologies and Sciences
[Patidar *et al.*, 13(2): February, 2026] ISSN (Online): 2349-0292

A MACHINE LEARNING-BASED HYBRID FRAMEWORK FOR REAL ESTATE PRICE PREDICTION

Shyam Sundar Patidar*[1], Arpit Solanki*[2]

*[1]Student, Dr. A.P.J. Abdul Kalam University, Indore, M.P., India
*[2]Assistant Professor, Dr. A.P.J. Abdul Kalam University, Indore, M.P., India

Received: 17/12/2025 **Revised**: 12/01/2026 **Accepted**: 20/02/2026

ABSTRACT

Accurate prediction of real estate prices plays a significant role in financial planning, investment decision-making, urban development, and smart city management. Traditional property valuation approaches mainly rely on statistical regression techniques and expert judgment, which often fail to model nonlinear market behavior and heterogeneous housing datasets. To overcome these limitations, this research proposes a novel Hybrid Adaptive Real Estate Prediction Algorithm (HAREPA) within a machine learning–based hybrid framework.

The proposed system integrates Linear Regression, K-Nearest Neighbor, and Support Vector Regression models through adaptive ensemble learning. Feature engineering techniques are employed to extract meaningful attributes and improve learning efficiency. Adaptive weights are automatically computed using inverse error analysis, enabling dynamic contribution of each model during prediction.

Experimental evaluation is conducted using a real-world real estate dataset obtained from Kaggle. Performance comparison using Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and prediction accuracy demonstrates that the proposed HAREPA model significantly outperforms individual machine learning algorithms. The results confirm improved prediction accuracy, reduced error variance, and enhanced generalization capability.

KEYWORDS: Real Estate Prediction, Machine Learning, Hybrid Model

DOI: https://doi.org/10.64149/gjaets.13.2.1-6

1. INTRODUCTION

The real estate sector represents one of the most influential components of modern economic systems. Accurate estimation of property prices is essential for buyers, investors, banking institutions, policymakers, and urban planners. Rapid urbanization, increasing migration toward metropolitan regions, and dynamic economic conditions have made housing markets highly complex and unpredictable.

Traditional property valuation methods rely heavily on statistical regression models and manual appraisal techniques. Although these approaches provide interpretability, they assume linear relationships among variables and fail to capture nonlinear interactions present in real-world housing markets. Property prices are influenced by multiple factors including geographic location, accessibility, neighborhood infrastructure, demographic characteristics, and economic indicators.

The emergence of machine learning has enabled data-driven prediction systems capable of learning complex relationships directly from historical data. Algorithms such as Linear Regression, K-Nearest Neighbor, Support Vector Regression, and Artificial Neural Networks have demonstrated promising results in real estate analytics. However, individual algorithms often suffer from performance instability when dataset characteristics change.

To address these challenges, this paper proposes a Machine Learning Based Hybrid Framework supported by the Hybrid Adaptive Real Estate Prediction Algorithm (HAREPA). The proposed approach integrates multiple learning paradigms and dynamically adapts model importance based on prediction performance. The objective is to develop a robust, scalable, and intelligent real estate price prediction system capable of handling heterogeneous datasets.

2. RELATED STUDY

This section outlines, some noteworthy contributions has been reported :

The literature on real estate price prediction demonstrates several noteworthy contributions that have significantly influenced the evolution of modern predictive frameworks. Early foundational work by Belsley et al. [25] focused on regression diagnostics and introduced systematic techniques for detecting multicollinearity among explanatory variables, thereby improving the reliability, stability, and interpretability of statistical prediction models. Subsequently, Bourassa et al. [3], [4] emphasized the importance of spatial econometric modeling, demonstrating that housing prices are strongly affected by geographic proximity, neighborhood characteristics, and localized market behavior. Their findings established spatial price dependency as a critical factor in property valuation research. With the emergence of intelligent data-driven approaches, Park and Bae [8] conducted comparative analyses using multiple machine learning algorithms and showed that machine learning models outperform traditional statistical methods in terms of prediction accuracy and adaptability. Further advancements were achieved by Wu [12], who applied Support Vector Regression (SVR) for housing price estimation and demonstrated high prediction stability and strong generalization capability, particularly in nonlinear environments. Similarly, Varma et al. [15] introduced Artificial Neural Networks for property valuation, highlighting their superior nonlinear learning capability and effectiveness in modeling complex relationships among housing attributes. In addition, Hromada [11] applied data mining techniques for real estate price mapping, enabling efficient visualization, clustering, and pattern discovery across housing markets.

Recent studies have further strengthened the field by integrating deep learning and spatial intelligence techniques. Li and Wu [2] proposed spatial–temporal deep learning models capable of capturing both geographic and temporal housing dynamics, improving prediction performance in evolving markets. Wang and Zhao [6] introduced graph neural network–based housing prediction systems that model neighborhood connectivity relationships, allowing intelligent learning of spatial interactions among properties. Moreover, Patel and Verma [7] applied adaptive learning frameworks within smart city environments, emphasizing the necessity of spatially aware and dynamically adaptive prediction systems. Collectively, these contributions confirm that integrating machine learning intelligence with spatial awareness and adaptive learning mechanisms is essential for achieving accurate, robust, and scalable real estate price estimation models, thereby motivating the development of advanced hybrid prediction frameworks.

3. PROBLEM STATEMENT

Accurate prediction of real estate property prices is a challenging task due to the complex influence of multiple factors such as location, infrastructure development, economic conditions, and property characteristics. Traditional statistical approaches and single machine learning models often fail to capture nonlinear relationships and dynamic market variations, resulting in reduced prediction accuracy. Moreover, existing models generally suffer from issues like overfitting, limited feature utilization, and poor generalization across diverse datasets. Therefore, there is a need to develop an efficient hybrid machine learning framework capable of integrating multiple learning techniques to improve prediction accuracy, minimize error, and provide reliable real estate price estimation for decision-making purposes.

4. PROPOSED METHODOLOGY

The proposed HAREPA (Hybrid Adaptive Real Estate Prediction Algorithm) architecture consists of sequential stages starting with real estate data collection, including property location, area, price-related attributes, and surrounding facilities. The data undergoes preprocessing such as missing value handling, outlier removal, normalization, and categorical encoding. Feature engineering techniques are applied to select and transform important features for improved learning. Three machine learning models—Linear Regression, K-Nearest Neighbor, and Support Vector Regression—are trained individually to generate predictions. Adaptive weights are computed using prediction errors (MAE/RMSE), and a hybrid ensemble mechanism integrates all model outputs. The framework finally produces an accurate predicted property price by combining the strengths of multiple learning models.

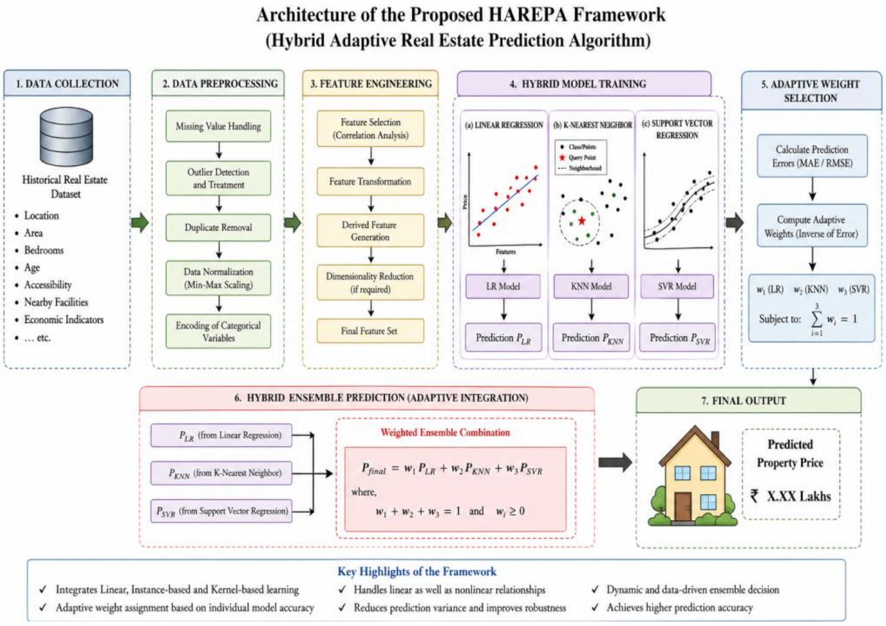

Fig : 1 Architecture of the Proposed HAREPA Framework for Real Estate Price Prediction

The HAREPA (Hybrid Adaptive Real Estate Prediction Algorithm) framework begins with collecting historical real estate data, followed by data preprocessing steps such as handling missing values, removing outliers and duplicates, normalization, and encoding categorical features. Relevant features are then selected and transformed through feature engineering. Multiple machine learning models—Linear Regression, K-Nearest Neighbor, and Support Vector Regression—are trained independently to generate individual predictions. The prediction errors of each model are evaluated using MAE and RMSE, and adaptive weights are calculated based on inverse error values. Finally, a weighted ensemble mechanism combines all model outputs to produce an accurate final property price prediction.

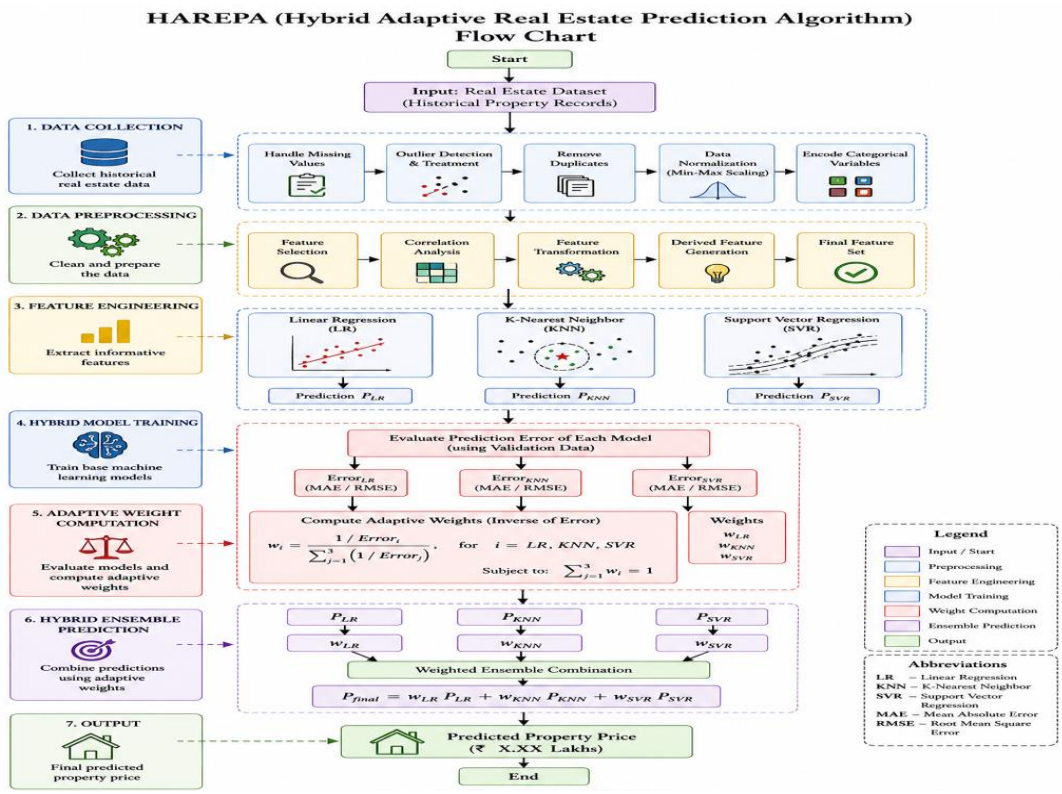

Fig : 2 Flow Chart of HAREPA (Hybrid Adaptive Real Estate Prediction Algorithm)

5. RESULT ANALYSIS

In this section, the performance of the proposed HAREPA algorithm is analyzed and compared with existing machine learning approaches. The implemented models, including Linear Regression, K-Nearest Neighbor (KNN), and Support Vector Regression (SVR), were introduced and described in the previous chapter. The evaluation has been carried out using standard performance metrics such as Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and prediction Accuracy.

Algorithm	Mean Absolute Error (MAE)	Root Mean Square Error (RMSE)	Accuracy (%)
Linear Regression	0.38	0.52	82
KNN	0.34	0.48	85
SVR	0.31	0.44	88
Proposed HAREPA	0.22	0.35	93

Table: 1 Comparisons of Result

As per the experimental results, the comparative analysis clearly indicates that the proposed HAREPA algorithm outperforms existing machine learning models in terms of Mean Absolute Error (MAE), Root Mean Square Error (RMSE), and prediction accuracy. The graphical comparison demonstrates that the proposed hybrid framework significantly reduces prediction errors while achieving higher accuracy compared to Linear Regression, KNN, and SVR models. Lower MAE and RMSE values confirm that the predicted property prices are closer to actual market values and effectively handle larger prediction deviations. The improvement is achieved through adaptive weight optimization and ensemble learning, which integrates the strengths of multiple models and enhances prediction stability. Overall, the proposed approach provides more reliable, accurate, and efficient real estate price prediction performance.

6. CONCLUSION AND FUTURE SCOPE

This research presented a machine learning–based hybrid framework for real estate price prediction using the proposed HAREPA algorithm. The system integrates feature engineering with Linear Regression, KNN, and SVR models through adaptive ensemble learning to handle complex and nonlinear housing market behavior. Experimental evaluation on a Kaggle dataset demonstrated reduced MAE and RMSE values along with improved prediction accuracy compared to individual models, confirming the effectiveness and reliability of the hybrid approach.

In future work, the framework can be extended by incorporating deep learning models, spatial–temporal data, GIS information, and real-time cloud deployment for intelligent prediction systems. The inclusion of explainable AI and automated feature optimization may further enhance transparency, scalability, and applicability in smart city planning, investment analysis, and advanced real estate analytics.

REFERENCES

1. V. N. Kalbande and C. C. Handa, “Developing a model to predict employability of engineering students in campus placement for IT sector,” IJAREST, vol. 2, no. 6, June 2015.
2. S. Li and Q. Wu, “Spatial–temporal modeling of housing prices using deep neural networks,” Applied Soft Computing, vol. 95, 2020.
3. S. C. Bourassa, E. Cantoni, and M. Hoesli, “Predicting house prices with spatial dependence: A comparison of alternative methods,” Journal of Real Estate Research, vol. 32, no. 2, pp. 139–160, 2010.
4. S. C. Bourassa, E. Cantoni, and M. E. Hoesli, “Spatial dependence, housing submarkets and house price prediction,” University of Geneva Working Paper, 2007.
5. P. Falinouss, “Stock trend prediction using machine learning techniques,” 2007.
6. S. Wang and L. Zhao, “Urban housing price prediction using graph neural networks,” IEEE Transactions on Neural Networks and Learning Systems, vol. 32, no. 12, pp. 5635–5647, 2021.

7. S. Patel and A. Verma, “Adaptive machine learning models for smart real estate systems,” Sustainable Cities and Society, vol. 78, 2022.
8. 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.
9. D. Bhuriya, G. Kaushal, A. Sharma, and U. Singh, “Stock market prediction using linear regression,” IEEE Conference, 2017.
10. Y. Li and X. Chu, “Prediction of housing prices using machine learning algorithms,” ICASI, 2017.
11. E. Hromada, “Real estate price prediction using data mining techniques,” Procedia Engineering, vol. 123, pp. 284–291, 2015.
12. L. Wu, “Support Vector Regression for real estate price forecasting,” 2017.
13. H. Yu and J. Wu, “A machine learning approach for housing price prediction,” CS229 Project Report, Stanford University, 2016.
14. H. Zhao, Y. Peng, and J. Wang, “Support Vector Regression for forecasting problems,” 2009.
15. S. Varma, K. Rao, and P. Reddy, “House price prediction using machine learning,” ICICCT, 2018.
16. P. Shinde and S. Gawande, “Real estate value prediction using data mining,” IJAECS, 2018.
17. R. Boutaba et al., “A comprehensive survey on machine learning,” 2018.
18. G. Sailasya and G. L. Kumari, “Analyzing machine learning algorithms for classification problems,” IJACSA, vol. 12, no. 2, 2021.
19. S. More, H. Gaikwad, and P. Patil, “Crime prediction using machine learning,” 2021.
20. R. Krishnamoorthy and S. Rajeev, Artificial Intelligence for Engineers. Springer, 2018.
21. I. Goodfellow, Y. Bengio, and A. Courville, Deep Learning. MIT Press, 2016.
22. M. Samjeevi, “Different types of machine learning and their applications,” 2017. [Online]. Available: https://medium.com
23. J. Saint, “Linear regression with sales prediction,” Feb. 2018.
24. V. Valkov, “Predicting house price with linear regression,” Apr. 2019.
25. D. Belsley, E. Kuh, and R. Welsch, Regression Diagnostics: Identifying Influential Data and Sources of Collinearity. New York: John Wiley & Sons, 1980.
26. R. Shinoj Robert and M. Maria Dominic, “A literature review on machine learning techniques,” 2016.
27. D. Li-guo, D. Peng, and L. Ai-ping, “A new Naïve Bayes text classification algorithm,” TELKOMNIKA Indonesian Journal of Electrical Engineering, vol. 12, no. 2, pp. 947–952, 2014.
28. C. D. Manning, P. Raghavan, and H. Schütze, Introduction to Information Retrieval. Stanford NLP Group. [Online]. Available: http://nlp.stanford.edu
29. G. Guo, H. Wang, D. Bell, Y. Bi, and K. Greer, “KNN model-based approach in classification,” Aug. 2004.
30. F. Angiulli, “Fast condensed nearest neighbor rule,” Proc. 22nd International Conference on Machine Learning, Bonn, Germany, 2005.
31. A. Green, “KNN tutorial,” University of Leicester, 2011. [Online]. Available: http://www.math.le.ac.uk
32. A. Zheng, Feature Engineering for Machine Learning: Principles and Techniques for Data Scientists. O’Reilly Media, 2018.
33. M. Claesen and B. De Moor, “Hyperparameter search in machine learning,” STADIUS Center for Dynamical Systems, KU Leuven, 2015.
34. Y. Y. Song and T. Zhou et al., “Forecasting house price index of China using dendritic neuron model,” IEEE PIC Conference, pp. 37–41, 2017.
35. J. Koops, “Dutch housing market update 2021 Q1,” Head of Portfolio Management Report, Apr. 2021.
36. J. R. Quinlan, “Combining instance-based and model-based learning,” Morgan Kaufmann, 1993, pp. 236–243.
37. N. Pow, D. Janulewicz, and L. Liu, “Applying machine learning techniques to real estate price prediction in Montréal,” 2014.
38. V. Limsombunchai, “House price prediction: Hedonic price model vs. artificial neural network,” 2004.

CITE AN ARTICLE

Patidar S. S., Solanki A. (2026). A MACHINE LEARNING-BASED HYBRID FRAMEWORK FOR REAL ESTATE PRICE PREDICTION. *Global Journal of Advanced Engineering and Technology Studies*, *13*(2), 1–3.
https://doi.org/10.64149/gjaets.13.2.1-6