International Journal of application or Innovation in engineering and management

Ijaiem.com/Mar 2025/ Volume 14/Issue 1/Article No-1/223-225
ISSN: 2319-4847

REAL ESTATE PRICE PREDICTION THROUGH MACHINE LEARNING MODELS

Jakkula Goutham [1], Vemula Bhoomika [2], Gudepu Timothi [3], A Sravani [4]

[1,2,3] UG Scholar, Department of IT,St. Martin's Engineering College, Secunderabad, Telangana,India – 500100

[4]Assistant Professor, Department of IT,St. Martin's Engineering College, Secunderabad, Telangana,India – 500100

vemulabhoomika3@gmail.com

Article Info

Received: 28-03-2025 Revised: 05 -04-2025 Accepted: 16-04-2025 Published:27/04/2025

Abstract:

In the dynamic realm of real estate, house prices became the crucial part to encounter the set of hurdles to make lives better. Through meticulous feature engineering, which includes advanced data cleansing and feature manipulation, coupled with robust machine learning techniques, our model offers a nuanced understanding of valuation dynamics. The optimization process involves hyper parameter tuning and cross-validation, employing cutting-edge methodologies to extract latent patterns and yield meaningful insights from the underlying data. Leveraging algorithms of supervised learning algorithms like linear regression and K-fold, chosen for their ability to discern intricate patterns within diverse datasets, our research pioneers a transformative approach to real estate valuation. Evaluation metrics such as Root Mean Squared Error (RMSE), Mean Squared Error (MSE) and R-squared were used to ensure a robust and accurate predictive framework which were promising.

Keywords: Real Estate, Data science, Machine learning, Predictive Modelling, Decentralization, System Valuation, Algorithm.

1. INTRODUCTION

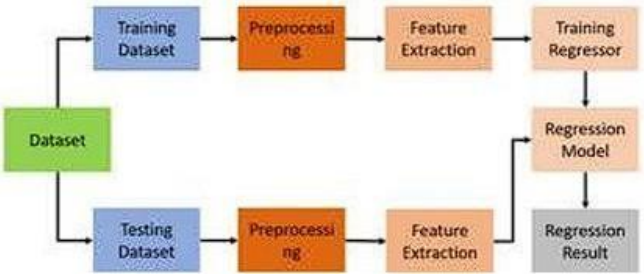
Determining home values in the ever-changing real estate market is a complex task with far-reaching consequences for people, neighborhoods, and larger economic systems. This task's intrinsic complexity highlights the necessity for a methodical, data-driven approach that goes beyond conventional valuation techniques to provide a comprehensive knowledge of the complex factors impacting real estate pricing [1]. Fundamental to this study is the understanding that accurate housing price prediction necessitates a deviation from traditional approaches [2]. The core issue statement centers on the necessity of accurately predicting home prices while taking into accounts variety of factors and changing market conditions. In order to meet this difficulty, our model uses a laborious process called feature engineering, which entails sophisticated data cleansing and manipulation techniques. Supported by strong machine learning techniques, the model is not only built to identify complex patterns in the data but is also optimized via cross- validation and hyper-parameter tweaking [3].

The fundamental conundrum is that property values are dynamic, impacted by a several factors depending upon market conditions. In order to tackle this, our model carefully designs features by utilizing cutting-edge methods for data cleaning and processing [4]. Along with revealing hidden patterns in the data, this strategy makes use of strong machine learning techniques to maximize prediction accuracy through cross- validation and hyper-parameter optimization [5]. The significance of this research lies in its commitment to address these challenges through a synergy of data science and web development. Real estate is a sector that thrives on information and informed decisions. Simultaneously, web development empowers users to interact with predictive models in real time, thus creating an accessible, transparent, and user-centric platform for academic exercise but a practical solution to a real-world problem. It seeks to empower property seekers with the tools and information necessary to make informed decisions [7]. By doing so, it aims to level the playing field, making property valuations more accessible, affordable, and transparent for abroader spectrum of individuals. This paper outlines the approach we have taken, which combines advanced data analysis, machine learning, and user-centric web development to offer a comprehensive, customizable, and informative solution to property valuation [8].

2. LITERATURE SURVEY

Z. F. Abut, H. Ş. Arlı, M. F. Akay and Y. Adıgüzel, "A New Hybrid Approach for Real Estate Price Prediction Using Outlier Detection, Feature Selection, and Clustering Techniques," 2023 8th International Conference on Computer Science and Engineering (UBMK), Burdur, Turkiye, 2023, pp. 1-6, doi: 10.1109/UBMK59864.2023.10286673. Economic risk is a probability that measures the possible alterations, as well as the uncertainty, generated by multiple internal or external factors. Global indicators are dynamic and sometimes the correlation is uncertain because they depend largely on a combination of economic, social, and environmental factors. Thus, our proposal consists

of a model for prediction and classification of multivariate risk factors such as birth rate and population growth, among others, using multiple neural networks and General Type-2 fuzzy system**S. Prongnuch, S. Sitjongsataporn, K. Intawichai and J. R. Kunkar, "Outcome-based Learning in Online STEM Activities for Robot and Real Estate Management Camp," 2022 7th International STEM Education Conference (iSTEM-Ed), Sukhothai, Thailand, 2022, pp. 1 4, doi: 10.1109/iSTEM-Ed55321.2022.9920828**.This paper presents the outcome-based learning STEM online activities for the robot & real estate and facility management camp (R^2 Camp) in order to develop the soft skills within knowledge management. The main objective is to provide knowledge on the real estate & facility management, the robotics engineering and how to apply robotics application to the real estate management, which can inspire young people to study in the robotics engineering and real estate & facility management. There are two sections of STEM online activities as: 1) Fundamental of real estate management and robotics engineering, 2) Robotics application in the real estate management. Pre-test and Post test about the both fundamental are used for outcome-based assessment. Results show that learners' outcomes can achieve the basic knowledge by 10.4%



A. Wandhe, L. Sehgal, H. Sumra, A. Choudhary and M. Dhone, "Real Estate Prediction System Using ML," 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India, 2023, pp. 1-4, doi: 10.1109/ICETET-SIP58143.2023.10151561.In recent years, machine learning has played a significant role in many aspects of our lives, including medical diagnosis, natural speech command, picture detection, product suggestion, spam recognition, and price prediction, etc. The desire to receive a profit on an investment property is a typical justification for home buying. They frequently want to know when and where to buy a property, thus they frequently ask the same questions. Current Real Estate Management System does not provide the prediction of the property price for users. In this project we offer the facility to the users to look for properties. This will provide the facility to view the system as an admin or user.

Y. Zheng, B. Yang, R. Zhang, Z. Bai and Y. Sun, "Mass Appraisal of Real Estate Prices Using Improved BP Neural Network with Policy Evaluation," 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Dalian, China, 2022, pp. 1036-1041, doi: 10.1109/TOCS56154.2022.10015915. This paper proposes a new policy index neural network (NN), which uses the back propagation (BP) NN model optimized by genetic algorithm (GA) and particle swarm algorithm and the policy modeling. consistency (PMC) index to quantify the relevant real estate policies. This study also trains and validates the big data of primary housing transactions of 33 properties in the main urban area of Weihai City.

3. PROPOSED METHODOLOGY

A proposed methodology for real estate price prediction using machine learning typically involves data collection, feature selection, model selection, and evaluation. Common models include support vector machines, random forests, and gradient boosting, which are trained on historical data to predict future prices are recorded and validated in a blockchain ledger to ensure transparency and security. The decentralized nature of blockchain ensures that every transaction, from user actions to vehicle simulations, is securely logged, minimizing the risk of tampering. Moreover, the use of advanced algorithms tailored for distributed systems ensures fast processing, real-time data retrieval, and highly efficient simulation outputs, which makes the system suitable for applications like connected and autonomous vehicles (CAVs) Significantly improved efficiency and speed in policy creation.

Figure 1: Proposed method of predicting house price.

Data Collection: The well-known dataset repository known as Kaggle is where the dataset used in this study was obtained. The dataset includes important columns that are essential for real estate value analysis, offering a full range of elements to investigate and precisely forecast property prices. The following columns are included in it: "area_type," "availability," "location," "size," "society," "total_sqft," "bath," "balcony," and "price." All of these columns combined offer a full range of featu3res required for real estate price analysis and forecasting: B. Data pre processing: Thorough pre-processing was performed on the dataset, which included imputation to fill up the gaps for the variables "availability," "balcony," and "society." One-hot encoding and label encoding were used to encode categorical variables such as "area_type" and "location." To ensure consistency for further research, numerical features most notably 'total_sqft' were also normalized using Standard and Min-Max scaling. C. Basic algorithm and background.

Linear regression:

International Journal of application or Innovation in engineering and management
Ijaiem.com/Mar 2025/ Volume 14/Issue 1/Article No-1/223-225
ISSN: 2319-4847

A fundamental supervised learning approach for predicting a continuous output variable based on one or more input data is called linear regression. The objective is to identify the best-fit line, usually shown by metrics like Mean Squared Error (MSE), that minimizes the deviation between the expected and actual values. Widely used in regression assignments, linear regression is an interpretable, computationally efficient approach that offers insights into the contribution of each feature to the anticipated outcome.

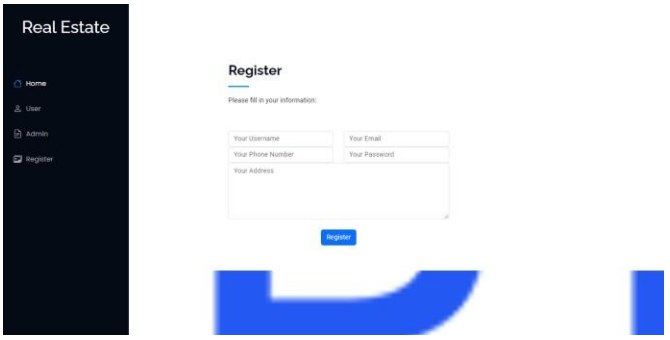
Random forest:

Using the power of several decision trees, Random Forest is an ensemble learning approach that improves prediction accuracy. A random subset of characteristics and data samples are used to build each tree in the forest, and the combined forecasts of all the individual trees yield the final prediction. By using an ensemble technique, overfitting is reduced and intricate correlations in the data are captured. for experimental results. It has been observed that the proposed system assumes the trustworthiness of TA, which is highly risky in a situation where almost every entity can be under attack. That means if the proposed system verifies the integrity of users, vehicles, and multimedia data, it must have to verify TA's integrity and authenticity. If the TA is under attack, all the data stored on it and the transactions it creates can become malicious, spoiling the whole Blockchain network.

Applications:

Property Valuation Tool: Automated systems that provide real-time property valuations based on various input features**.**

Investment Analysis Platforms: Tools that help investors assess potential returns on real estate investments by predicting future property values.



Real Estate Market Forecasting: Provides tamper-proof logs for accident analysis and system audits.

Comparative Market Analysis (CMA) Software: Verifies authenticity of updates and prevents malicious injections.

Property Recommendation Systems: Ensures authorized access to AV systems and services.

Risk Assessment Models: Enables safe and transparent transactions for tolls and services.

International Journal of application or Innovation in engineering and management
Ijaiem.com/Mar 2025/ Volume 14/Issue 1/Article No-1/223-225
ISSN: 2319-4847

Urban Development Planning: Prevents manipulation of sensor inputs and detects spoofing.

Real Estate Analytics Dashboards: Validates and aggregates map data from multiple AVs.

Mortgage Approval Systems: Protects EV charging stations from unauthorized access.

Real Estate Investment Trust (REIT) Analysis: Secures shared AI models and simulation data.

Advantages:

The ETC (Extra Tree Classifier) algorithm is commonly used in multi- armed bandit problems and online learning settings. Here are some advantages of the ETC algorithm:

Automated Learning: The system uses blockchain to decentralize user management, enhancing system efficiency and removing bottlenecks.

Adaptability: With blockchain encryption and hashing algorithms, data security and integrity are significantly enhanced.

Non-linear Relationships: The system supports real-time vehicle simulations with instant blockchain-based validation.

Feature Selection and Engineering:: The system can handle multiple users and transactions without compromise suitable for large-scale simulations.

Handling Large Datasets: Every transaction and user activity is logged in the blockchain, ensuring complete transparency and auditability.

Improved Accuracy: Provides a transparent and traceable record for legal and regulatory compliance. devices and users interact with AVs.

C**ost Efficiency:** Boosts user confidence by ensuring safety and privacy in autonomous transportation.

EXPERIMENTAL ANALYSIS

Experimental analysis of real estate price prediction using machine learning involves using various algorithms to predict prices, comparing their accuracy, and identifying key factors influencing prices, ultimately aiming to improve prediction accuracy.

Figure 1: Home Page

Figure 2: Registration Form

Figure 3: Admin Login Form

Figure 4: View Registered Users

International Journal of application or Innovation in engineering and management

Ijaiem.com/Mar 2025/ Volume 14/Issue 1/Article No-1/223-225
ISSN: 2319-4847

Figure 5: User Login Form

Figure 6: User Home Page

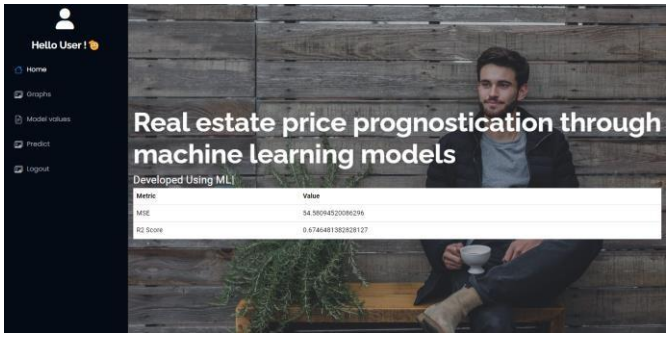


Figure 7: Model Evaluate Page

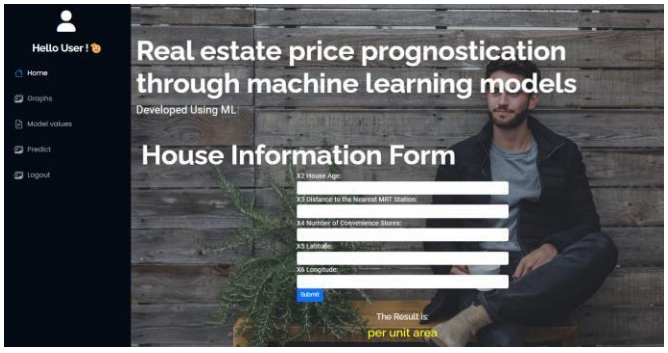


Figure 8: Prediction Form

International Journal of application or Innovation in engineering and management

Ijaiem.com/Mar 2025/ Volume 14/Issue 1/Article No-1/223-225
ISSN: 2319-4847

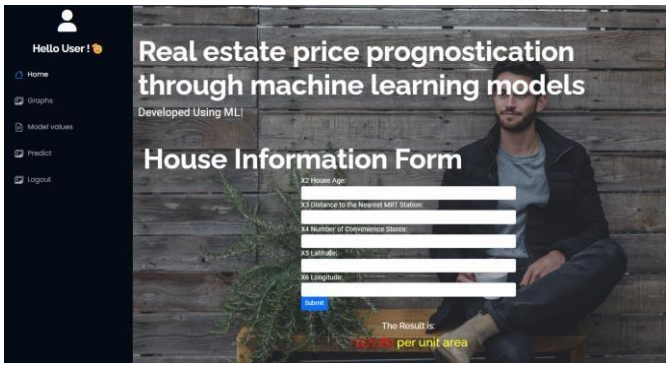


Figure 9: Output

4. CONCLUSION

The study presented here investigated on the use of Support Vector Machine, Random Forest, and Linear Regression models for real estate price prediction. The base was laid via meticulous preparation, which included data cleaning and outlier reduction. Evaluation metrics demonstrated the balanced performance of Linear Regression, while visual analysis demonstrated the robust predictive strength of the Random Forest model. The results highlight the fine balance that must be struck between interpretability and accuracy. Linear regression was found to be a reasonably balanced option, while Random Forest demonstrated exceptionally high accuracy. Although reasonable, Support Vector Machine displayed more mistakes. This study provides information to help with decision-making in the ever changing housing market. Subsequent improvements can concentrate on exploring features through feature engineering, improving the model, and using cutting-edge methods to improve prediction accuracy.

REFERENCES

[1] Z. F. Abut, H. Ş. Arlı, M. F. Akay and Y. Adıgüzel, "A New Hybrid Approach for Real Estate Price Prediction Using Outlier Detection, Feature Selection, and Clustering Techniques," 2023 8th International Conference on Computer Science and Engineering(UBMK),Burdur,Turkiye,2023,pp.16,doi:9/UBMK59864.2023.10286673.

[2] M. Arivukarasi, A. Manju, R. Kaladevi, S. Hariharan, M. Mahasree and S. Prongnuch, S. Sitjongsataporn, K. Intawichai and J. R. Kunkar, "Outcome-based Learning in Online STEM Activities for Robot and Real Estate Management Camp," 2022 7th International STEM EducationConference(iSTEMEd),Sukhothai,Thailand,2022,pp.14,doi:10.11

[3] A. Wandhe, L. Sehgal, H. Sumra, A. Choudhary and M. Dhone, "Real Estate Prediction System Using ML," 2023 11th International Conference on Emerging Trends in Engineering & Technology - Signal and Information Processing (ICETET - SIP), Nagpur, India, 2023, pp. 1-4, doi: 10.1109/ICETET-SIP58143.2023.10151561.

[4] A. Peter, A. A. Kumar, A. Rajeev, B. Baiju and V. S. Chooralil, "Real Estate Management System using Blockchain," 2023 International Conference on Innovations in Engineering and Technology (ICIET), Muvattupuzha,India,2023,pp.1-4,doi: 10.1109/ICIET57285.2023.10220623.

[5] M. S. Bennet Praba, U. Rajeev, A. Rathore and A. Kolangarath, "Real Time Automation on Real Estate using API," 2022 First International Conference on Electrical, Electronics, Information and CommunicationTechnologies (ICEEICT), Trichy,India,2022,pp.1-5,doi: 10.1109/ICEEICT53079.2022.9768428.

[6] Y. Zheng, B. Yang, R. Zhang, Z. Bai and Y. Sun, "Mass Appraisal of Real Estate Prices Using Improved BP Neural Network with Policy Evaluation," 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Dalian, China, 2022, pp. 1036-1041, doi: 10.1109/TOCS56154.2022.10015915.

[7] Y. Sun and G. Peng, "Developing Area Real Estate Valuation Based on Linear Regression and KNN Algorithm," 2022 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS), Changsha,China,2022,pp.3842,doi:10.1109/ICDSBA57203.2022.00014.

[8] S. Thokala, R.Jakkani, M.Alli and H. Shanmugasundaram, "Accident Prevention System using Machine Learning2023 IEEE Conference on Telecommunications, Optics and Computer Science (TOCS),Coimbatore,India,2023,pp.16,doi:10.1109/ICCCI56745.2023.10128202.

[9] H. A. V. P. U. Hapuarachchi, M. D. Manoratne, K. G. B. K. Gamlath, V. S. G. G, D. Sriyaratna and N. H. P. R. Supunya, "Realty Scout Smart System for Real Estate Analysis & Forecasting with Interactive User Interface," 2022 IEEE 7th International conference for Convergence in Technology(I2CT), Mumbai,India,2022,pp.16,doi:10.1109/I2CT54291.2022.9825335.

[10] Y. Zhao, X. Shen, X. Xu and Y. Xu, "Application of BP neural network in real estate batch assessment," 2022 41st Chinese Control Conference (CCC), Hefei, China, 2022, pp. 7018-7023, doi: 10.23919/CCC55666.2022.9901571.

[11] Y. Zhao, G. Chetty and D. Tran, "Deep Learning for Real Estate Trading," 2022 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE), Gold Coast,Australia,2022,pp.17,doi:10.1109/CSDE56538.2022.10089222.

[12] S. Wang, J. Zhu, Y. Yin, D. Wang, T. C. Edwin Cheng and Y. Wang, "Interpretable Multi-Modal Stacking-Based Ensemble Learning Method for Real Estate Appraisal," in IEEE Transactions on

:10.1109/TMM.2021. 3126153.

[13] T. Zhu, "The Impact of Non-immersive Virtual Reality Technologies on Consumers' Behaviors in real estate: A Website's Perspective," 2022 IEEE International Symposium on Mixed and Augmented Reality Adjunct (ISMAR-Adjunct), Singapore, Singapore, 2022, pp. 13-20, doi: 10.1109/ISMAR- Adjunct57072.2022.00013.

[14]L. Kong et al., "When permissioned blockchain meets IoT oracles: An on-chain quality assurance system for off-shore modular construction manufacture," 2022 IEEE 1st Global Emerging Technology Blockchain Forum: Blockchain & Beyond (iGETblockchain), Irvine, CA, USA, 2022, pp. 1-6, doi: 10.1109/iGETblockchain56591.2022.10087164.

[15] W. Coleman, B. Johann, N. Pasternak, J. Vellayan, N. Foutz and H. Shakeri, "Using Machine Learning to Evaluate Real Estate Prices Using Location Big Data," 2022 Systems and Information Engineering Design Symposium (SIEDS), Charlottesville, VA, USA, 2022,pp.168-172,doi: 10.1109/SIEDS55548.2022.9799393.

[16]R. Henker, D. Atzberger, W. Scheibel and J. Döllner, "Real Estate Tokenization in Germany: Market Analysis and Concept of a Regulatory and Technical Solution," 2023 IEEE International Conference on Blockchain and Cryptocurrency (ICBC), Dubai, United Arab Emirates, 2023, pp. 1-5, doi: 10.1109/ICBC56567.2023.10174954.

[17]T. Balasooriya et al., "Location Intelligence Based Smart E-Commerce Platform for Residential Real-Estate Industry," 2022 3rd International Conference on Smart Electronics and Communication (ICOSEC), Trichy, India, 2022,pp. 867-873, doi: 10.1109/ICOSEC54921.2022.9952023.

[18]Y. Zhao, X. Shen, X. Xu and Y. Xu, "Application of BP neural network in real estate batch assessment," 2022 41st Chinese Control Conference (CCC), Hefei, China, 2022, pp. 7018-7023, doi: 10.23919/CCC55666.2022.9901571.

[19]L. Kong et al., "When permissioned blockchain meets IoT oracles: An on-chain quality assurance system for off-shore modular construction manufacture," 2022 IEEE 1st Global Emerging Technology Blockchain Forum: Blockchain & Beyond (iGETblockchain), Irvine, CA, USA, 2022, pp. 1-6, doi: 10.1109/iGETblockchain56591.2022.10087164.

[20]M. Selim, M. R. Rabbani and A. Bashar, "Qard Hasan Based Cooperative Model for Home Financing and Its Effects in Home Ownership and Real Estate Development," 2022 International Conference on Sustainable Islamic Business and Finance (SIBF), Sakhir,Bahrain,2022,pp.48-52,doi:10.1109/SIBF56821.2022.994002.

[21] G. Wang, J. S. Suroso, D. Sanusi, J. A. Tanuwijaya and T. F. I. Theodora, "Applying Internet of Things Framework in Real Estate Business with Enterprise Architecture Approach," 2022 International Conference onInformation Management and Technology (ICIMTech),Semarang,Indonesia,2022,pp.219-224,doi:10.1109/ICIM Tech55957.2022.9915151.