ISSN: 2582-7219

International Journal of Multidisciplinary Research in Science, Engineering and Technology

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

Impact Factor: 8.206

Volume 8, Issue 8, August 2025

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808079

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

HOUSEMINT-AI

Gunasekaran K, Karan Bharadwaj B

Assistant Professor, Department of MCA, AMC Engineering College, Bengaluru, India

Student, Department of MCA, AMC Engineering College, Bengaluru, India

ABSTRACT: In In the modern real estate market, accurate property valuation remains a significant challenge due to the influence of multiple dynamic factors such as location, property size, amenities, and market trends. Traditional valuation methods often rely on real estate agents, public listings, or personal estimates, which can be outdated, biased, or manipulated, leading to uninformed decisions. In order to provide real-time, data-driven house price predictions, the suggested system, HOUSEMINT-AI, incorporates machine learning into a web application built with Django. Reliability and accuracy are ensured by training the model on datasets that have been cleaned and free of outliers. Users can input details such as location, number of bedrooms, bathrooms, and property area to obtain instant price predictions. Unlike conventional platforms that focus primarily on property listings, HOUSEMINT-AI emphasizes accuracy, transparency, and usability. By providing predictive insights, the system assists buyers in avoiding overpaying, sellers in setting competitive prices, & investors in making informed decisions. The platform's flexible model & lightweight architecture guarantee accessibility in a range of real estate markets while preserving scalability for future improvements. Keywords – Real Estate Price Prediction, Machine Learning, Django Framework, Linear Regression, Data Preprocessing, Predictive Analytics, Housing Market

I. INTRODUCTION

A crucial component of the real estate market has always been the appraisal of residential properties. Accurate pricing not only benefits buyers and sellers but also ensures transparency & trust within the housing sector. In recent years, rapid urbanization, fluctuating market conditions, & evolving buyer preferences have made property valuation increasingly complex. Prices are no longer influenced solely by property size or location; instead, a combination of factors such as connectivity, local amenities, infrastructure development, and neighborhood growth patterns plays a crucial role in determining the market value.

Traditionally, property prices have been estimated using manual methods, consultations with real estate agents, or by referencing public listings. While these approaches are widely used, they often lack precision & may be subject to bias or outdated market information. This can lead to significant disparities between the asking price & the actual market value, resulting in overpayment, underpricing, or missed investment opportunities.

Although the emergence of digital platforms in the real estate industry has made it easier to find properties, the problem of precise valuation has not always been resolved. Many popular platforms focus primarily on property listings and advertisements rather than providing precise, data-driven price estimations. This gap highlights the need for an intelligent, transparent, & reliable prediction system that can process real-time data inputs & deliver accurate valuations.

This need is met by the suggested system, HOUSEMINT-AI, which uses machine learning methods built into a web application built with Django. The solution employs a Linear Regression model trained on cleaned and outlier-free housing datasets to generate price predictions based on user-provided parameters such as location, built-up area, number of bedrooms, & bathrooms. The aim is to equip users—buyers, sellers, & investors—with reliable insights that facilitate better decision-making, minimize reliance on biased intermediaries, and promote fair pricing in the housing market.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808079

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

II. LITERATURE SYRVEY

In early studies in this area, linear regression was the main model utilized because of its ease of use and interpretability. But because of its shortcomings in identifying non-linear patterns, more sophisticated models like Random Forests, Decision Trees, and Gradient Boosting algorithms like XGBoost were adopted. Particularly when used on properly preprocessed datasets, these models have shown increased accuracy. Property valuation has historically depended on expert opinion and comparative market analysis, both of which are subject to subjectivity and have little flexibility in response to shifting market trends. A promising method to automate and increase the precision of real estate price forecasts is the emergence of machine learning. Because it was simple to use and interpret, linear regression was primarily used in early research in this area. However, because of its shortcomings in identifying non-linear patterns, more sophisticated models such as Random Forests, Decision Trees, and Gradient Boosting algorithms like XGBoost were adopted. Particularly when used on well-preprocessed datasets, these models have demonstrated increased accuracy. These models are still not commonly applied in the Indian housing market, despite promising scholarly findings. The distinctive features of Indian cities are not sufficiently captured by the Western datasets that are the subject of many studies. Furthermore, because these models are not commonly used in the real world, there is a disconnect between theoretical advancement and real-world implementation.

Existing System

Current digital platforms catering to the Indian real estate market—such as MagicBricks, NoBroker, and 99acres—primarily function as property listing portals. These systems facilitate basic search and filter functionalities and offer details such as amenities, location, and seller contact options. However, they typically rely on seller-quoted prices or aggregated historical data, which lack objectivity and real-time analytical capability. Despite extensive research in machine learning-based price prediction models showing superior accuracy, many of these models remain confined to academic experimentation and are rarely translated into publicly usable applications. Consequently, there exists a substantial disconnect between research developments in predictive modeling and their practical deployment, limiting the impact of such innovations in real-world scenarios.

Proposed System

HOUSEMINT offers reasonable and accurate price estimates based on factors like location, size, and the number of bedrooms and bathrooms, as opposed to depending solely on conjecture or seller-quoted prices. The platform provides a more accurate alternative to conventional property listing websites because it was created especially with the Indian housing market in mind. HOUSEMINT helps users learn more about the buying and selling process and lessens the chance of overpricing or undervaluing a home by providing them with current, objective price recommendations.

III. SYSTEM ARCHITECTURE

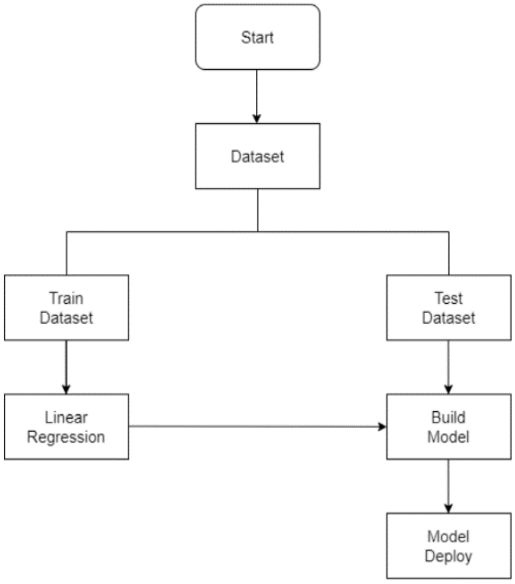
The presentation layer, application layer, and data layer make up the three tiers of the HOUSEMINT-AI system architecture. Efficiency, maintainability, and scalability are improved by this modular framework. The user interface is made up of the presentation layer, which is constructed with HTML, CSS, JavaScript, and Bootstrap. User input, including property details, is captured and sent to the backend. User authentication, model inference, and input validation are managed by the application layer, which was created with the Django framework. Prior to prediction, this layer also carries out crucial preprocessing operations like one-hot encoding and outlier removal. The data layer manages contact details, user-submitted feedback, and prediction history using SQLite3 as the backend database. Each component can function independently thanks to this layered architecture, which also ensures seamless integration for a positive user experience.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808079

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)



IV. METHODOLOGY

A thorough methodological pipeline that includes data collection, preprocessing, model training, and web integration is used in the development of HOUSEMINT-AI. A carefully selected dataset of Bengaluru housing records, including details like location, size, square footage, number of bathrooms, and cost, forms the basis of the system. To guarantee dependability, this CSV-formatted dataset underwent extensive preprocessing. Numerical data was extracted from textual fields using feature engineering, and missing values were handled by imputation or elimination. Extreme entries that defied logical property constraints were eliminated as outliers, and total square footage values represented as ranges were standardized. To guarantee compatibility with regression models, one-hot encoding was used to encode categorical variables, especially location.

The linear regression model at the heart of the machine learning component was selected due to its interpretability and low computational cost. Relationships between different housing features and the target price variable were captured during training using the cleaned dataset. Experiments with more complex models, like XGBoost, produced even higher predictive accuracy even though Linear Regression was the main model. Real-time prediction features were then made possible by embedding the trained models into the Django web framework. Through a responsive front end, the web interface receives user inputs and forwards them to the backend, where dynamic preprocessing and model inference take place. For administrative control, the predicted prices are kept in the backend database and displayed in the Indian currency format.

V. DESIGN AND IMPLEMENTATION

In order to create a unified system for property valuation, HOUSEMINT-AI's design is centered on integrating machine learning capabilities into a web-based framework. Standard web technologies like HTML, CSS, Bootstrap, and JavaScript are used in the development of the front end to guarantee responsiveness and accessibility on a variety of devices. Users can specify property attributes through the interface's user-friendly form inputs, which the backend processes.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808079

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

The Django framework, which is used to implement the backend, offers strong support for data validation, routing, and model interaction. Before being sent to the Linear Regression model, input data obtained via the web forms is preprocessed to standardize formats and encode categorical values.

The primary data repository is the SQLite3 database. Prediction logs, contact form submissions, feedback entries, and administrator credentials are just a few of the different types of user data it keeps. Administrators can view prediction history, manage records, and examine feedback to further improve the system through the admin interface, which is accessible through secure authentication. To protect sensitive data, security features like session management and login protection have been incorporated.

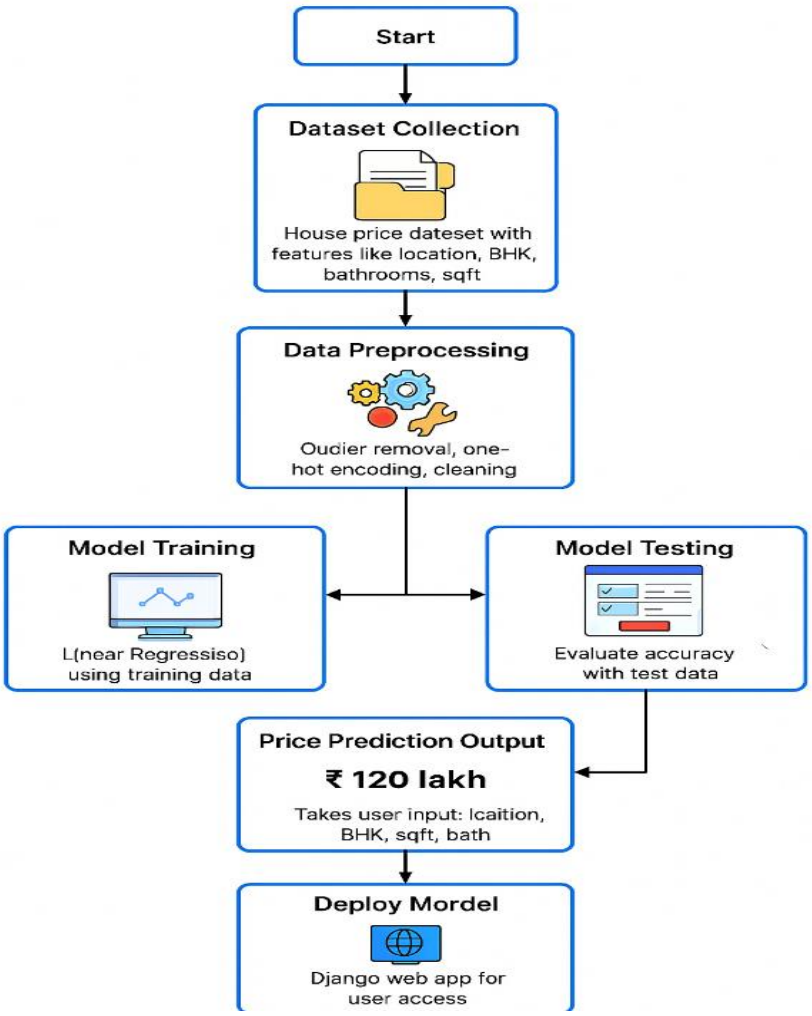


Fig 5.1 Flowchart of Working System

VI. OUTCOME OF RESEARCH

Evaluation revealed that the XGBoost model continuously outperformed 90%, highlighting its superior ability to capture intricate feature relationships, while the Linear Regression model performed well, with an R2 score above 85%. The web interface of the system allowed for quick, accurate, and user-friendly property valuation. The method was validated and the system's practicality in real estate valuation was established by the accuracy attained through careful preprocessing and model tuning.

© 2025 IJMRSET | Volume 8, Issue 8, August 2025| DOI:10.15680/IJMRSET.2025.0808079

ISSN: 2582-7219 | www.ijmrset.com | Impact Factor: 8.206| ESTD Year: 2018|

International Journal of Multidisciplinary Research in Science, Engineering and Technology (IJMRSET)

(A Monthly, Peer Reviewed, Refereed, Scholarly Indexed, Open Access Journal)

VII. RESULT AND DISCUSSION

The model produced more accurate price estimates by eliminating inconsistent entries and making sure all property details were consistent. In order to eliminate noise that might otherwise distort the results, this step proved particularly crucial.

Performance was compared using two distinct models: a simpler linear regression model and a more intricate gradient boosting technique. The linear model performed well in scenarios requiring prompt responses, like live platform inquiries, due to its speed and simple output. Users and developers also found it easier to understand how the predictions were generated because of its transparency.

However, the advanced model (XGBoost) produced better results when the properties were more complicated or uncommon. It addressed anomalous trends in the data, like properties with non-standard configurations or those in quickly developing regions. It was therefore particularly useful for determining fair prices in situations where more conventional estimation techniques might not be sufficient.

The system continuously generated precise estimates that closely matched market trends during testing in several Bengaluru regions.

VIII. CONCLUSION

The viability and efficiency of incorporating machine learning-based prediction models into a useful, web-accessible real estate valuation platform are demonstrated by HOUSEMINT-AI. It can adjust to new developments and offers precise, up-to-date price projections specific to the Indian housing market.

Important features like administrator tools, user history tracking, and feedback collection are supported by the current implementation. Future developments will concentrate on adding real-time market feeds, integrating location intelligence (such as the accessibility of schools, transit, and other facilities), and growing the platform to support more cities. With these enhancements, HOUSEMINT-AI will be positioned as a complete decision-support tool for investors, buyers, and sellers.

REFERENCES

[1] 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, 2016, pp. 785–794.
[2] An Introduction to Statistical Learning, G. James, D. Witten, T. Hastie, and R. Tibshirani, Springer, 2013.
[3] Packt Publishing, 2019; S. Raschka and V. Mirjalili, Python Machine Learning, 3rd ed.

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

INTERNATIONAL JOURNAL OF MULTIDISCIPLINARY RESEARCH

IN SCIENCE, ENGINEERING AND TECHNOLOGY

| Mobile No: +91-6381907438 | Whatsapp: +91-6381907438 | ijmrset@gmail.com |

www.ijmrset.com