House Price Prediction using Machine Learning

Project progress report in partial fulfillment of the requirement for the award of the degree of

Master of Computer Applications

Submitted By

Mitali Dandapat	**University Enrollment No. 12023006015021**
Nayanika Paul	**University Enrollment No. 12023006015005**
Anurag Biswas	**University Enrollment No. 12023006015062**
Subhadip Giri	**University Enrollment No. 12023006015029**
Rahul Kumar	**University Enrollment No. 12023006015056**

Under the guidance of

Prof. Sujata Ghatak

Department of Computer Applications

INSTITUTE OF ENGINEERING & MANAGEMENT, KOLKATA
NEWTOWN
(School of University of Engineering and Management, Kolkata)

University Area, Plot No. III – B/5, New Town,

Action Area – III, Kolkata – 700160.

UNIVERSITY OF ENGINEERING & MANAGEMENT
(Established by Act XXV of 2014 of Govt. of West Bengal & recognized by UGC, Ministry of HRD, Govt. of India)
University Area, Plot No. III-B/5, Main Arterial Road, New Town, Action Area-III, Kolkata - 700 160, W.B., India
Admission Office : 'ASHRAM', GN-34/2, Salt Lake Electronics Complex, Kolkata - 700 091, W.B., India
Ph. (Office) : 91 33 2357 2969
: 91 33 2357 7649
Admissions : 91 33 2357 2059
Fax : 91 33 2357 8302
E-mail : vc@uem.edu.in
Website : www.uem.edu.in

CERTIFICATE

This is to certify that the project titled **House Prediction using Machine Learning** submitted by **Mitali Dandapat(University Enrollment No. 12023006015021), Nayanika Paul(University Enrollment No. 12023006015005)**,**Anurag Biswas(University Enrollment No. 12023006015062), Subhadip Giri(University Enrollment No. 12023006015029) and Rahul Kumar(University Enrollment No.12023006015056)**students of INSTITUTE OF ENGINEERING AND MANAGEMENT, NEWTOWN, a school of UNIVERSITY OF ENGINEERING & MANAGEMENT, KOLKATA, in partial fulfilment of the requirement for the Degree of Master of Computer Applications, is a Bonafide work carried out by them under the supervision and guidance of Prof. **Sujata Ghatak** during 3rd Semester of the academic session of 2024 - 2025. The content of this report has not been submitted to any other university or institute. I am glad to inform that the work is entirely original and its performance is found to be quite satisfactory.

Prof. Sujata Ghatak {Name of the Guide}

Department of Computer Applications

UEM, Kolkata

Prof. Kaustuv Bhattacharjee

HOD, Department of Computer Applications

UEM, Kolkata

ACKNOWLEDGEMENT

We would like to take this opportunity to thank everyone whose cooperation and encouragement throughout the ongoing course of this project remains invaluable to us.

We are sincerely grateful to our guide **Prof. Sujata Ghatak** of the Departmentof Computer Applications, UEM, Kolkata, for his/her wisdom, guidance and inspiration that helped us to go through with this project and take it to where it stands now.

We would also like to express our sincere gratitude to **Prof. Kaustuv Bhattacharjee**, HOD, Department of Computer Applications, UEM, Kolkata and all other departmental faculties for their ever-present assistant and encouragement.

Last but not the least, we would like to extend our warm regards to our families and peers who have kept supporting us and always had faith in our work.

Mitali Dandapat

Nayanika Paul

Anurag Biswas

Subhadip Giri

Rahul Kumar

ABSTRACT

The prediction of house prices is a critical area of interest in real estate, finance, and urban planning. Machine learning has emerged as a powerful tool for developing predictive models by leveraging complex datasets. This study explores the application of machine learning algorithms to predict house prices based on diverse features such as location, size, number of bedrooms, amenities, and market conditions. Using techniques like linear regression, decision trees, and ensemble methods such as Random Forest and Gradient Boosting, the model is trained on historical data to uncover patterns and relationships. Advanced approaches, including neural networks and deep learning, are also considered to enhance prediction accuracy for large, unstructured datasets. The results demonstrate that machine learning models outperform traditional statistical methods by achieving higher accuracy and adaptability to changing market trends. This research underscores the potential of machine learning in making data-driven, reliable real estate decisions.

SECTION1

INTRODUCTION

Background:-

House price prediction is a classic problem in machine learning, where the objective is to estimate the price of a property based on various influencing factors. This application is highly relevant in real estate, where accurate price forecasting can provide valuable insights to buyers, sellers, investors, and even policymakers. Machine learning techniques offer robust tools for analyzing large datasets and discovering complex relationships between house prices and features that influence them.

SECTION2

LITERATURE SURVEY

A literature survey on house price prediction using machine learning explores the models, data features, and methodologies applied to forecast housing prices accurately. This field combines statistical and AI techniques to capture the complex factors influencing real estate values. Here's a comprehensive overview:

a. Introduction to House Price Prediction

House price prediction is essential for various stakeholders, including buyers, sellers, real estate investors, and government bodies. Machine learning (ML) methods have gained traction due to their ability to process large datasets and capture non-linear relationships between input variables and house prices, outperforming traditional statistical methods.

b. Common Machine Learning Techniques

- **Linear Regression:** Linear regression is one of the earliest and simplest models used for house price prediction. It assumes a linear relationship between independent variables (features) and the target variable (price). While effective for simple relationships, it often fails to capture the complex interactions among features that influence house prices.
- **Decision Trees and Random Forests:** Decision tree-based methods, especially Random Forests, are popular due to their interpretability and ability to handle non-linearity. Random Forests improve prediction accuracy by averaging multiple trees, reducing over fitting. These models can capture complex feature interactions, like location-based price variations and neighborhood characteristics.

SECTION3

PROBLEM STATEMENT

In the real estate market, accurately predicting house prices is critical for buyers, sellers, and investors to make informed decisions. Traditional methods rely heavily on historical sales data and basic economic indicators, which may not fully capture the complexity of factors influencing property values, such as location-specific trends, property characteristics, neighborhood dynamics, and changing economic conditions.

The goal of this project is to develop a machine learning model capable of predicting the selling price of a residential property based on a wide range of features, such as location, size, number of rooms, age, amenities, and proximity to important facilities. By leveraging machine learning algorithms, this project aims to build a predictive model that provides more accurate and scalable predictions than traditional valuation methods. This model will assist stakeholders in setting competitive prices, assessing property values, and making data-driven decisions.

SECTION4

PROPOSED SOLUTION

To address the challenge of accurately predicting house prices, a machine learning-based solution is proposed that leverages advanced algorithms and robust datasets. The solution involves the following components:

1. Model Evaluation

- Evaluate model performance using metrics such as Mean Absolute Error (MAE), Mean Squared Error (MSE), and R-squared.
- Compare different algorithms to select the best-performing model for the dataset.

2. Deployment

- Develop a user-friendly application or API that integrates the trained model for real-time house price predictions.
- Enable users to input property details and receive accurate price estimates.

SECTION 5

EXPERIMENTAL SETUP AND RESULT ANALYSIS

1. Experimental Setup

Training and Testing

- **Data Split**: Split the dataset into training (70%), validation (15%), and testing (15%) subsets.
- **Cross-Validation**: Perform k-fold cross-validation to ensure model robustness.

Evaluation Metrics

- Use metrics to assess model performance:
  - Mean Absolute Error (MAE): Measures average prediction error.
  - Mean Squared Error (MSE): Penalizes larger errors.
  - R-squared: Indicates how well the model explains variability in house prices.

---

2. Result Analysis

Error Analysis

- Errors were higher for properties with unique or luxury features due to limited representation in the dataset.
- Outliers in the dataset (e.g., properties sold under special conditions) contributed to residual errors.

SECTION 6

CONCLUSION

Machine learning has the potential to revolutionize house price prediction by providing more accurate and scalable solutions compared to traditional methods. As we refine our models and incorporate innovative techniques, the benefits to stakeholders will expand, enhancing decision-making in real estate.

SECTION7

FUTURE SCOPE

The future of house price prediction using machine learning (ML) is poised to evolve with advances in both technology and data availability. Here are some key trends and directions:

i. Block chain and Real-time Data

- Block chain technology could help in creating a more transparent and secure way to track historical property data, transactions, and even market trends, which could feed into machine learning models for more accurate predictions.

j. Adaptive Pricing Systems

- Real estate platforms could use machine learning not just for predictions but also for dynamic pricing, adjusting listings and offers based on real-time demand, competition, and local trends. This could revolutionize both buying and renting strategies.

REFERENCES

- Ayush Varma et al. (2018). 'Housing Price Prediction Using Machine Learning and Neural Networks'. IEEE.
- G.Naga Satish et al. (2019). 'House Price Prediction Using Machine Learning'. IJITEE.
- CH. Raga Madhuri et al. (2019). 'House Price Prediction Using Regression Techniques: A Comparative Study'. IEEE.