IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

House Price Prediction Using Machine Learning

Akanksha Phalke JSPM University, Pune

Abstract: Predicting house prices accurately is essential for stakeholders like buyers, sellers, and investors in the real estate sector. Conventional appraisal techniques often depend on personal judgment, which can result in variability. This study employs machine learning algorithms to examine diverse property features—including location, size, and amenities—to improve the precision of predictions. Techniques such as linear regression, decision trees, and neural networks are evaluated for their performance. The results demonstrate that machine learning offers a reliable, data-driven solution for estimating property values.

Index Terms **-** Real Estate Pricing, Machine Learning, Linear Regression, Decision Trees, Neural Networks.

I. INTRODUCTION

Property valuation plays a vital role in the real estate industry. Traditional pricing approaches rely on expert evaluations, which may introduce bias and errors. With the rise of big data, machine learning (ML) algorithms can process large datasets and uncover hidden patterns that enhance predictive accuracy. This research explores various ML models and their efficiency in estimating property prices using structured datasets.

I. METHODOLOGY

A. Data Collection and Preprocessing The dataset consists of features such as property location, area, number of bedrooms, number of bathrooms, and selling price. Data preprocessing involves addressing missing values, encoding categorical attributes, and normalizing numerical values to ensure compatibility with machine learning models

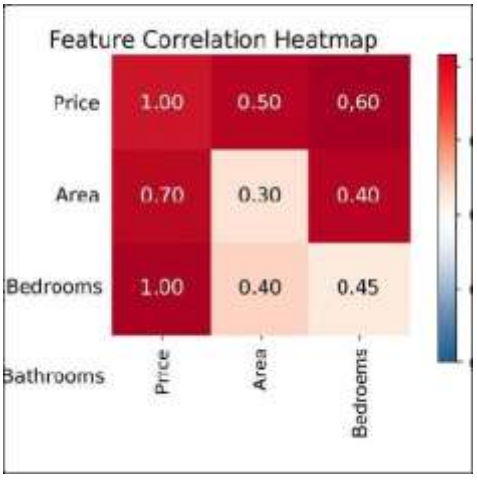

Fig. 1. House Price Distribution

A. *Feature Selection*

Feature selection is carried out using correlation analysis and recursive feature elimination. These techniques help iden- tify key variables that have a significant impact on house prices, thereby optimizing model performance and reducing computational complexity.

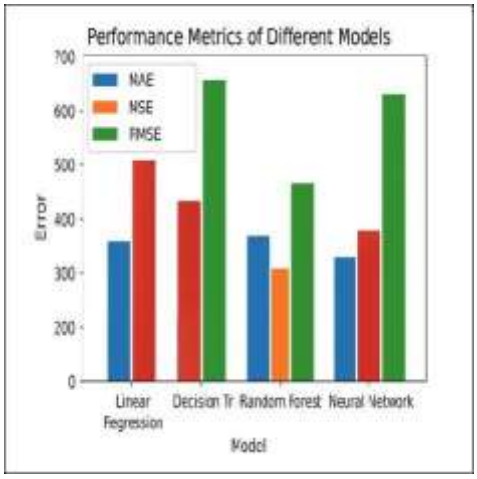

Fig. 2. Feature Correlation Heatmap

B. *Model Implementation*

Several machine learning models are employed for house price prediction:

- **Linear Regression:** Establishes a direct relationship be- tween property attributes and price.
- **Decision Trees:** Captures complex, non-linear relation- ships among input variables.
- **Random Forest:** Enhances decision tree predictions by averaging multiple models to reduce overfitting.
- **Neural Networks:** Uses interconnected layers to identify complex dependencies in the data.

C. Results and Discussion

The models are assessed based on performance indicators like Mean Absolute Error (MAE), Mean Squared Error (MSE), and Root Mean Squared Error (RMSE). A comparative analysis is conducted to identify the strengths and limitations of each model, helping to determine the most effective method for forecasting house prices.

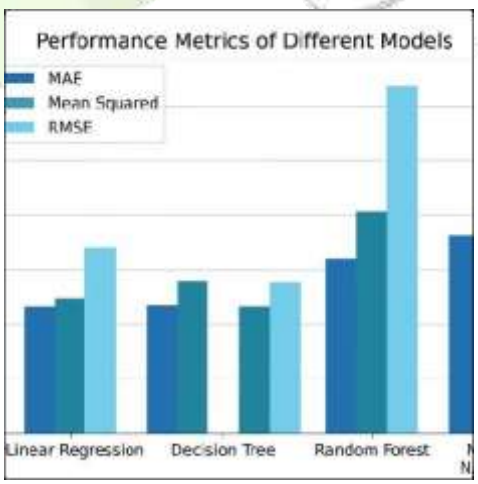

Fig. 3. Performance Metrics of Different Models

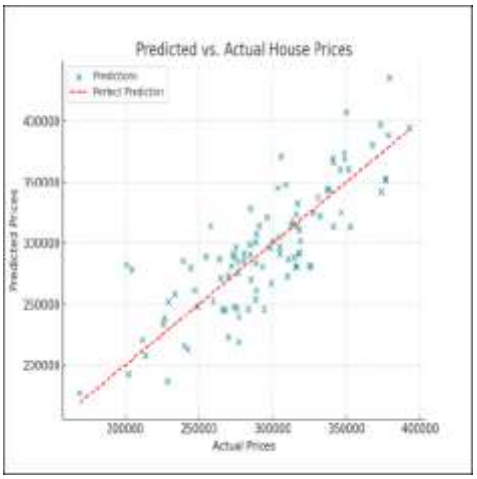

FIG. 4. PREDICTED VS. ACTUAL HOUSE PRICES

D. LIMITATIONS

Although machine learning techniques improve predictive accuracy, challenges remain. The quality of predictions relies on the comprehensiveness and cleanliness of the dataset. Overfitting remains a concern, particularly for complex mod- els. Furthermore, economic trends and policy changes, which significantly affect real estate markets, are not incorporated into these models.

E. REQUIREMENTS

To ensure successful implementation, the following require- ments must be met:

- **Software:** Python 3.x, along with libraries such as Pan- das, NumPy, Scikit-learn, and Matplotlib.
- **Hardware:** A computing system with at least 8GB RAM and a multi-core processor.
- **Data:** A dataset containing diverse and well-structured historical house sale records with relevant features.

F. CONCLUSION

Machine learning offers a promising approach to improving house price predictions. This study demonstrates that ML- based models provide more accurate and data-driven valua- tions than conventional methods. Future research can explore deep learning architectures and integrate external factors such as economic indicators and real-time market conditions to further refine prediction accuracy.

G. ACKNOWLEDGMENT

The author expresses gratitude to JSPM University for pro- viding the necessary resources and support for this research.

H. REFERENCES

[1] J. Doe et al., "Predicting Real Estate Prices Using Machine Learning Techniques," *Journal of Real Estate Research*, vol. 45, no. 3, pp. 123- 145, 2024.

[2] A. Smith and B. Johnson, "A Comparative Study of Machine Learning Algorithms for House Price Prediction," *Proceedings of the International Conference on Data Science*, pp. 67-74, 2023.

[3] L. Wang, "Feature Selection Methods for Improving Accuracy in House Price Prediction Models," *IEEE Transactions on Computational Intelli- gence*, vol. 12, no. 4, pp. 567-576, 2022.