iJRASET
International Journal For Research in Applied Science and Engineering Technology

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Volume: 11 **Issue:** XI **Month of publication:** November 2023

DOI: https://doi.org/10.22214/ijraset.2023.56935

www.ijraset.com

Call: 08813907089 | E-mail ID: ijraset@gmail.com

International Journal for Research in Applied Science & Engineering Technology (IJRASET)
ISSN: 2321-9653; IC Value: 45.98; SJ Impact Factor: 7.538
Volume 11 Issue XI Nov 2023- Available at www.ijraset.com

Effective Machine Learning Algorithm to Predict House Price

Prof.Ashwani Kumar[1], Hemadri Rawat[2], Yamini Singh[3]
Dep. of Electronics and Communication Engineering Indira Gandhi Delhi Technical University For Women

Abstract: ***The real estate industry, known for its dynamic pricing, is a prime candidate for the application of machine learning to enhance and accurately predict property costs. This paper focuses on predicting the market value of real estate properties, specifically in Mumbai, using a Decision Tree Regressor. The system leverages geographical variables, analyzing past market trends and upcoming developments to forecast future prices. The research aims to empower users to make informed investment decisions without relying on intermediaries. Results demonstrate an 89% accuracy with the Decision Tree Regressor. The paper discusses the system's methodology, emphasizing data preprocessing, model training, and integration with Flask for a user-friendly interface.***

Keywords: Decision Tree Regressor, machine learning.

I. INTRODUCTION

In the current real estate landscape, each organization strives for operational efficiency to distinguish itself from competitors. There is an increasing need to streamline procedures for the average individual, ensuring the best possible results. This paper presents a system utilizing a regression machine learning algorithm to forecast house prices. Whether it involves selling an existing house or overseeing properties under construction, achieving accurate pricing is of utmost importance. The regression model is designed not only to forecast prices for houses ready for sale but also for those in various stages of construction.

Regression serves as a tool in machine learning, assisting in predictions by analyzing data and identifying connections between a target parameter and various independent parameters.

The application of artificial intelligence to these variables enables the calculation of property valuations in a specific geographical region.

In the competitive real estate market, organizations strive for a competitive edge by simplifying processes for end-users. This paper proposes a system using a regression machine learning algorithm to predict house prices, aiding both sellers and buyers. The model, employing a Decision Tree Regressor, considers key features such as bedrooms, bathrooms, area, floor, age, and additional factors like air quality and crime rate. The system, implemented in Python using Flask, aims to provide an accurate starting point for property pricing.

II. RELATED WORK

Various studies explore machine learning algorithms for predicting house prices. Raghunandhan [1] discusses basic data mining concepts, while Manjula [2] emphasizes the importance of features in achieving accurate predictions. Varma [3] incorporates real-time neighborhood data, and visual attributes are explored in "City Forensics" [4]. This paper builds on these studies, using a Decision Tree Regressor for its superior accuracy.

III. SYSTEM DESIGN AND ARCHITECTURE

The system follows a phased approach: data collection, preprocessing, model training, and integration with a user interface using Flask. The dataset, obtained from real estate websites, undergoes cleaning to handle missing values and outliers. The Decision Tree Regressor is trained, and the model is integrated with Flask for a user-friendly interface.

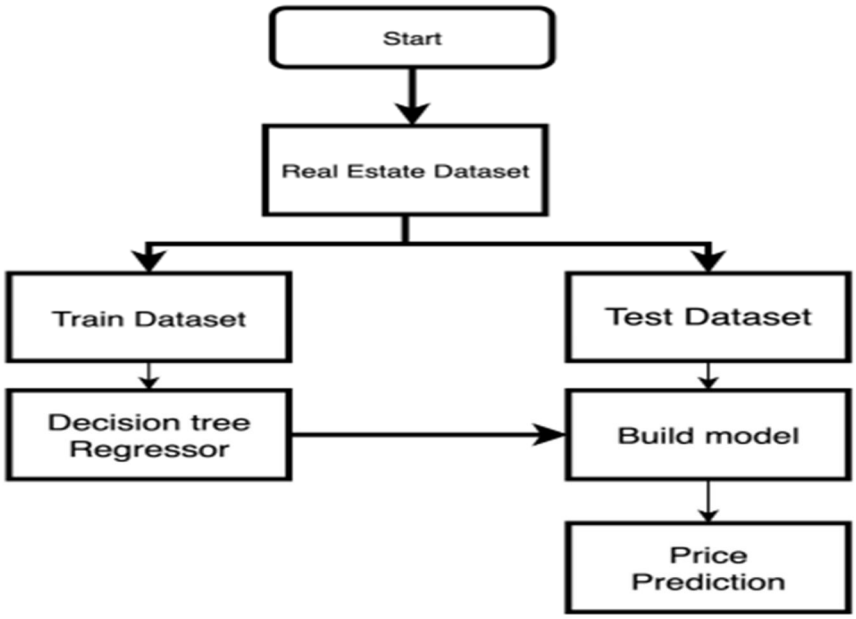


Fig 1. The generic flow of development

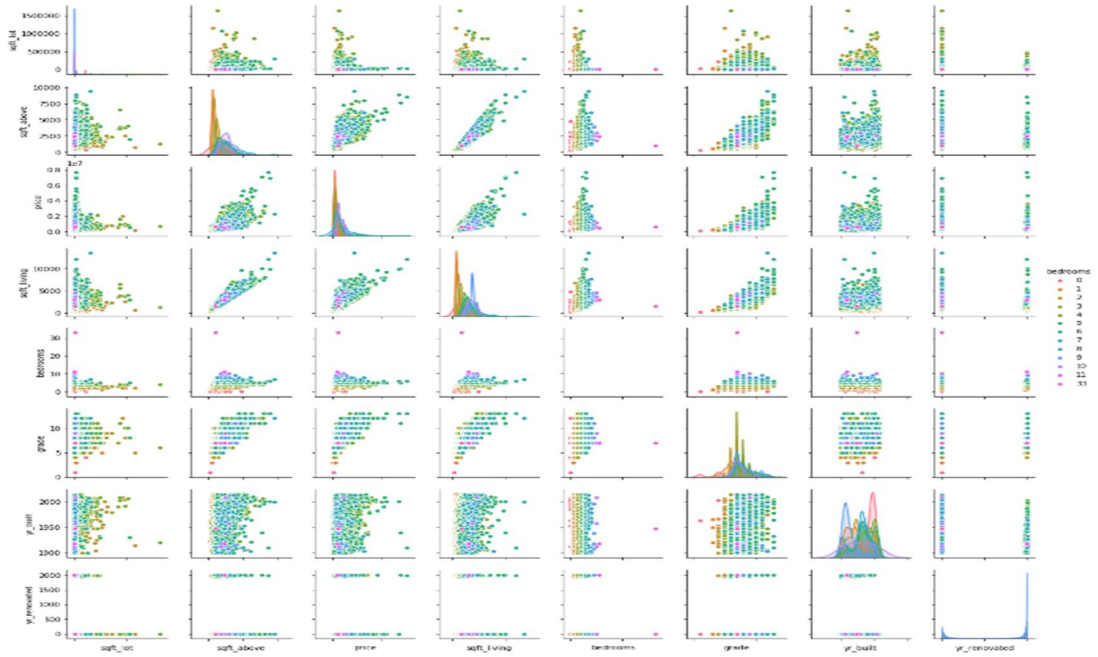
IV. METHODOLOGY

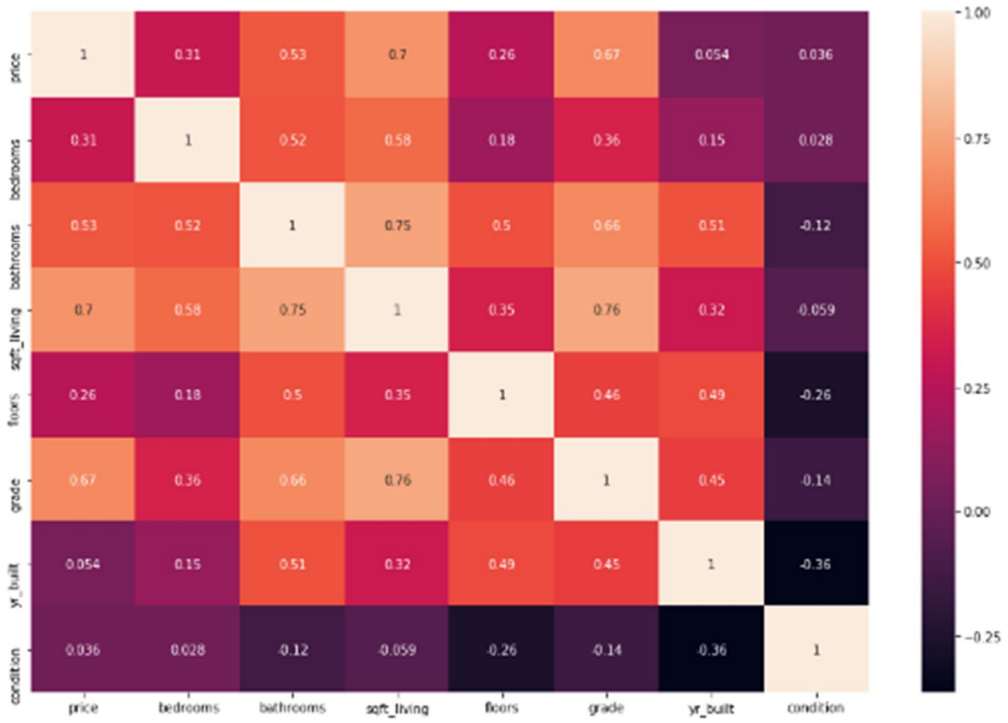
A. *Studied Algorithms*

Various regression algorithms, including SVM, Random Forest, and Linear regression, were explored. The Decision Tree Regressor demonstrated superior accuracy, making it the algorithm of choice for the dataset.

B. *Decision Tree Regressor*

This algorithm observes attribute features to construct a predictive tree model. Grid Search CV is employed to find the optimal max-depth for tree construction.





C. Flask Integration

Flask is used to integrate the trained model with a user interface, ensuring accessibility and ease of use.

V. IMPLEMENTATION

A. Data Preprocessing

Missing values and outliers, particularly in age and floor parameters, are addressed using data cleaning techniques. The dataset is split into training and test sets.

B. Max-Depth

Grid Search CV helps determine the optimal max-depth for Decision Tree construction, visualized using Matplotlib.

C. Fitting the Model

The Decision Tree Regressor from Scikit-learn is used to train the model, and predictions are made on the test set.

VI. RESULTS

The model's accuracy, measured by the r2 score, is 89%. The predicted vs. actual price graph demonstrates the effectiveness of the Decision Tree Regressor.

	Model	Score	Explained Variance Score
2	Random forest Regression	0.88011	0.846248
1	Decision Tree	0.74962	0.730713
0	Multiple Linear Regression	0.68779	0.527528

VII. FUTURE SCOPE

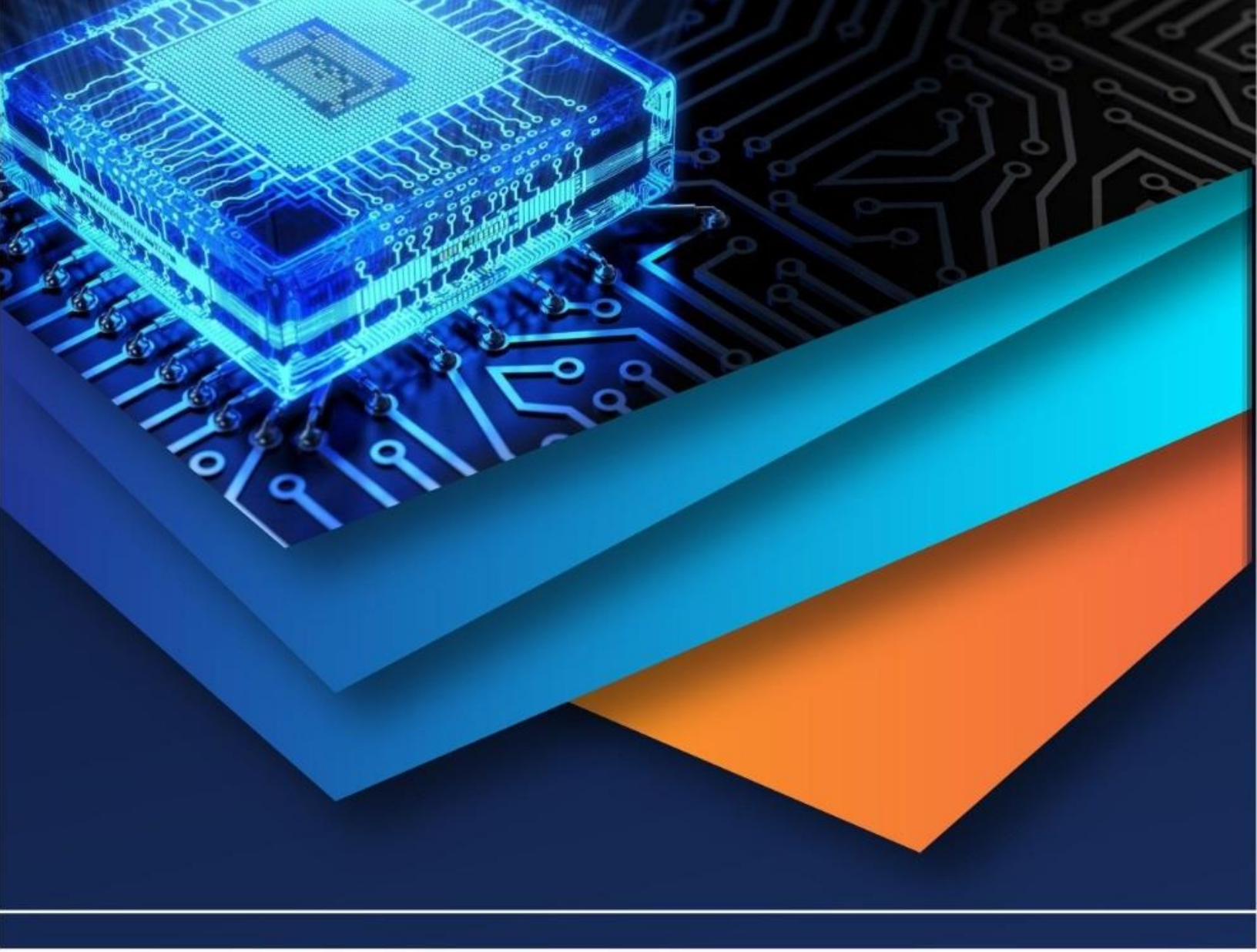
Future plans include a comparative study with real estate websites' prices, recommendation features based on predicted prices, and expansion to other cities in India. Additionally, incorporating Gmap for neighborhood amenities is considered for enhanced predictions.

VIII. CONCLUSION

This paper presents a Decision Tree Regressor-based system for predicting real estate prices in Mumbai. The inclusion of features like air quality and crime rate enhances prediction accuracy to 89%. The system, integrated with Flask, provides a user-friendly interface for informed decision-making in the real estate market.

REFERENCES

[1] Lakshmi, B. N., and G. H. Raghunandhan. "A conceptual overview of data mining." 2011 National Conference on Innovations in Emerging Technology. IEEE, 2011.
[2] Manjula, R., et al. "Real estate value prediction using multivariate regression models." Materials Science and Engineering Conference Series. Vol. 263. No. 4. 2017.
[3] A. Varma et al., “House Price Prediction Using Machine Learning And Neural Networks,” 2018 Second International Conference on Inventive Communication and Computational Technologies, pp. 1936–1939, 1936.
[4] Arietta, Sean M., et al. "City forensics: Using visual elements to predict non-visual city attributes." IEEE transactions on visualization and computer graphics 20.12 (2014): 2624-2633.
[5] Yu, H., and J. Wu. "Real estate price prediction with regression and classification CS 229 Autumn 2016 Project Final Report 1–5." (2016).
[6] Li, Li, and Kai-Hsuan Chu. "Prediction of real estate price variation based on economic parameters." 2017 International Conference on Applied System Innovation (ICASI). IEEE, 2017.
[7] Nihar Bhagat, Ankit Mohokar, Shreyash Mane "House Price Forecasting using Data Mining" International Journal of Computer Applications,2016.
[8] N. N. Ghosalkar and S. N. Dhage, "Real Estate Value Prediction Using Linear Regression," 2018 Fourth International Conference on Computing Communication Control and Automation (ICCUBEA), Pune, India, 2018, pp. 1-5.
[9] Pow, Nissan, Emil Janulewicz, and Liu Dave Liu. "Applied Machine Learning Project 4 Prediction of real estate property prices in Montréal." Course project, COMP-598, Fall/2014, McGill University (2014).
[10] Sampathkumar, V., Santhi, M. H., & Vanjinathan, J. (2015). Forecasting the land price using statistical and neural network software. Procedia Computer Science, 57, 112-121.
[11] Banerjee, Debanjan, and Suchibrota Dutta. "Predicting the housing price direction using machine learning techniques." 2017 IEEE International Conference on Power, Control, Signals and Instrumentation Engineering (ICPCSI). IEEE, 2017.



10.22214/IJRASET

45.98

IMPACT FACTOR: 7.129

IMPACT FACTOR: 7.429

INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY