International Journal of Scientific Research in Engineering and Management (IJSREM)
Volume: 07 Issue: 05 | May - 2023 SJIF 2023: 8.176 ISSN: 2582-3930

Real Estate Price Prediction Using Machine Learning Algorithm

Shreyash Rajendra Pawar

Pratik Sanjay Alai

Nikhil Madhav Taral

Yasim Mehboob Shaikh

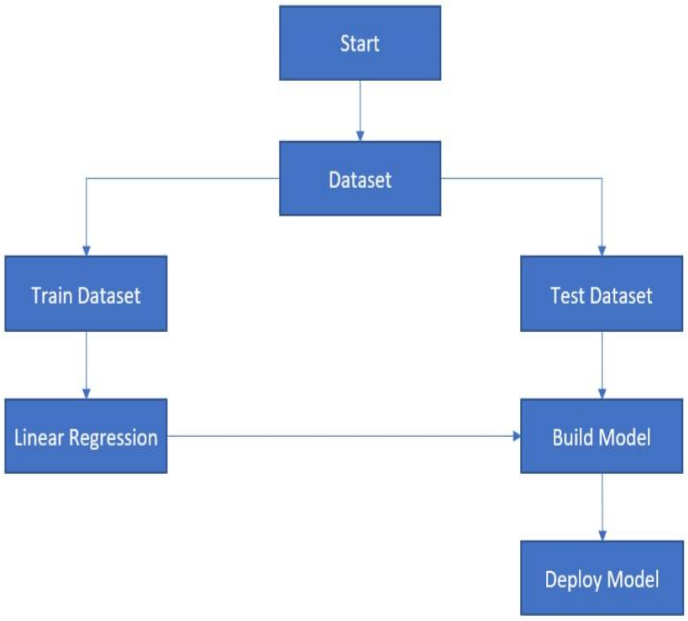
Abstract - Accurately forecasting prices carries immense significance across multiple industries, encompassing finance, real estate, and banking. Through the amalgamation of historical and present data, machine learning algorithms can be effectively employed to anticipate price trends. Notably, machine learning has demonstrated remarkable potential in the realm of predicting prices for diverse real estate assets, including houses and properties. By harnessing the capabilities of data and sophisticated algorithms, these models contribute to empowering individuals and enterprises to make well-informed choices while engaging in real estate transactions.

Given the significant fluctuations experienced in the real estate market, accurate price prediction has become essential for making informed decisions, whether it involves purchasing, selling, or investing in properties. Machine learning algorithms, such as linear regression, have emerged as powerful tools for real estate price prediction, capitalizing on technological advancements and the growing availability of data. Linear regression, a statistical approach that models the relationship between a dependent variable and one or more independent variables, has gained popularity as a suitable choice for regression tasks, including real estate price prediction.

1.INTRODUCTION

In the past, the task of finding suitable real estate properties within budgetary limitations posed considerable challenges for many individuals. This was mainly due to the time-consuming and arduous process of physically visiting each location and evaluating prices. However, the advent of machine learning models for real estate price prediction has alleviated this problem. These models leverage data for training purposes and empower buyers to access property prices conveniently through their mobile devices, from the comfort of their homes. As a result, individuals can now more easily identify and acquire properties that align with their budgetary constraints.

System Architecture



Methodology

Linear regression is a statistical method used to establish a relationship between a dependent variable (Y) and one or more independent variables (X) by fitting a linear equation to the data. The equation of a simple linear regression model can be represented as:

$Y = \beta_0 + \beta_1 X + \varepsilon$

where:

Y is the dependent variable (the variable you want to predict)

X is the independent variable (the variable you use to predict Y)

β_0 is the intercept (the value of Y when X = 0)

β_1 is the slope (the change in Y for every one-unit change in X)

ε is the error term (the difference between the predicted and actual Y value)

The goal of linear regression is to estimate the values of β_0 and β_1 such that the line fits the data as closely as possible. This is typically done using a technique called least squares regression, which minimizes the sum of the squared errors.

To estimate the values of β_0 and β_1, you can use the following formulas:

$\beta_1 = \sum(Xi - \bar{X})(Yi - \bar{Y}) / \sum(Xi - \bar{X})^2$

$\beta_0 = \bar{Y} - \beta_1\bar{X}$

where:

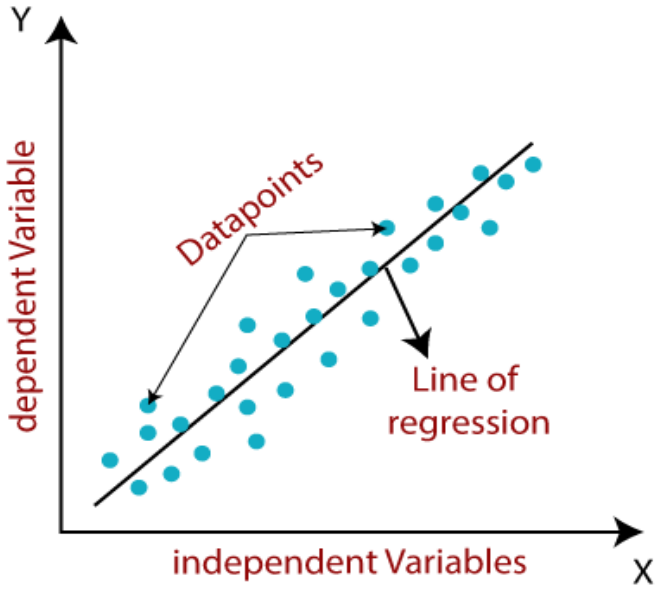
Xi is the ith value of X

$\bar{X}$ is the mean of X

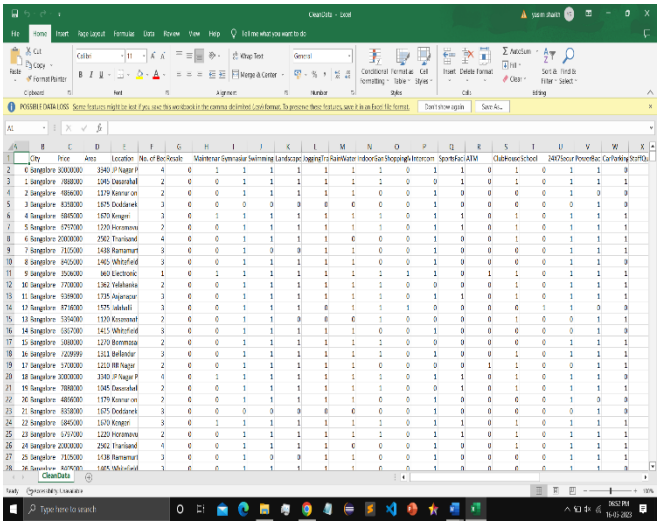
Yi is the ith value of Y

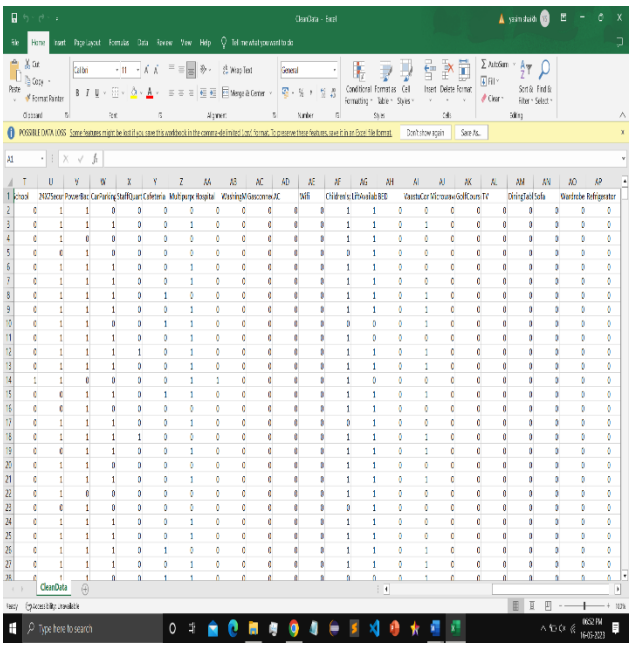
$\bar{Y}$ is the mean of Y

Linear Regression



Dataset Screenshots



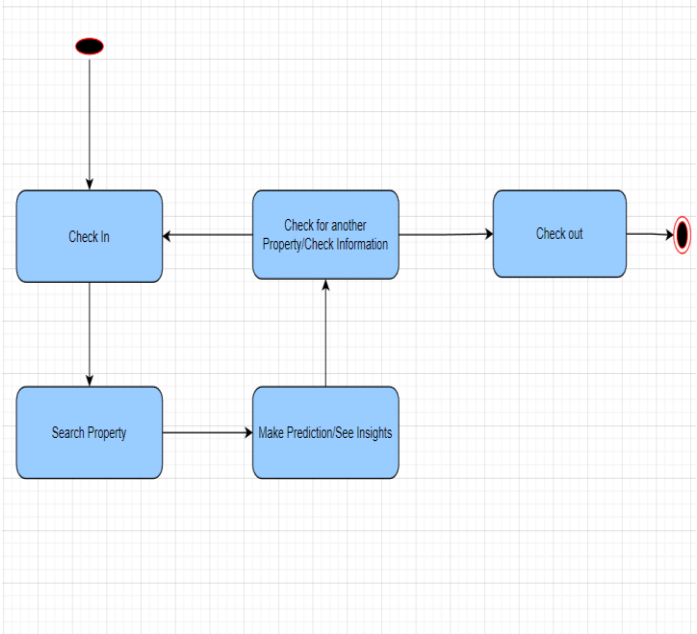


The main features of our dataset are city, price ,Area, Location and bhk.

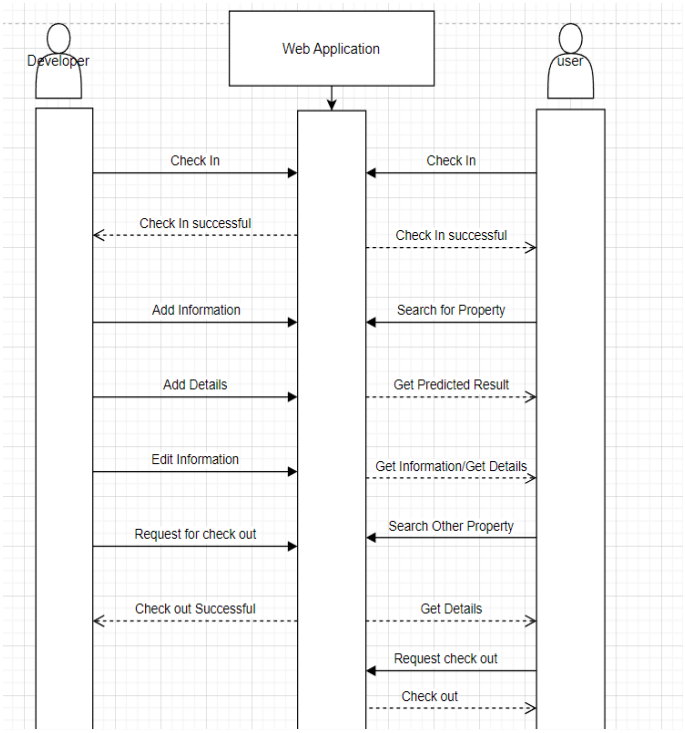
	City	Price	Area	Location	bhk
0	Bangalore	6700000.0	1180.0	Bangalore 5th Phase	2.0
1	Bangalore	5800000.0	1185.0	Bangalore 5th Phase	2.0
2	Bangalore	8500000.0	1430.0	Bangalore 5th Phase	3.0
3	Bangalore	8400000.0	1180.0	Bangalore 5th Phase	2.0
4	Bangalore	4876000.0	1185.0	Bangalore 5th Phase	2.0

UML Diagrams

State Diagram



Sequence Diagram



Steps

1. Data Collection: Gather data on the real estate market, including properties, their features, and their respective prices. This data can be collected from public sources such as real estate websites, government property listings, and real estate agents.

2. Data Preparation: Once data is collected, it needs to be preprocessed to remove any inconsistencies or outliers. This can be done by cleaning the data, handling missing values, and encoding categorical variables.

3. Feature Selection: Identify relevant features that can help predict the price of a property. These may include the location, size, number of bedrooms/bathrooms, amenities, and other factors that may affect the price.

4. Split Data: Split the dataset into training and testing sets. The training set is used to train the model, and the testing set is used to evaluate the model's performance.

5. Model Training: Apply the linear regression algorithm to the training set to train the model. The algorithm will attempt to learn the relationship between the features and the price of the property.

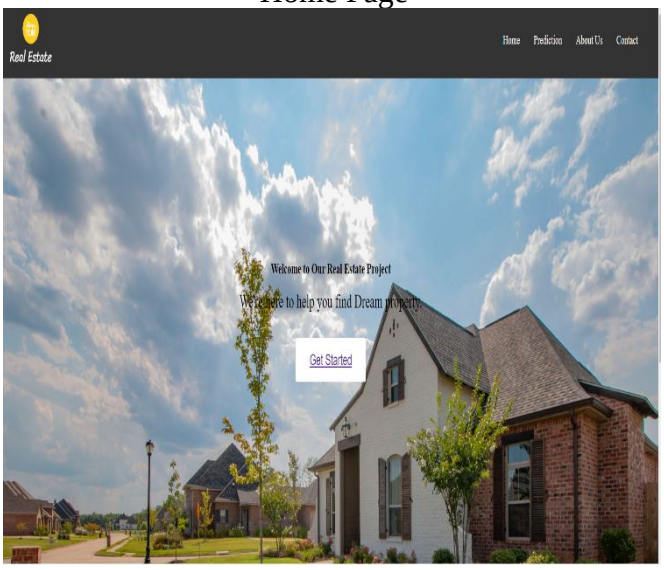
6. Model Evaluation: Evaluate the model's performance on the testing set. This can be done by computing performance metrics such as root mean squared error (RMSE), mean absolute error (MAE), and coefficient of determination (R-squared).

7. Model Improvement: If the model's performance is not satisfactory, try improving it by modifying the algorithm or adding more features. This process can be iterative, and multiple models can be trained and evaluated until satisfactory results are obtained.

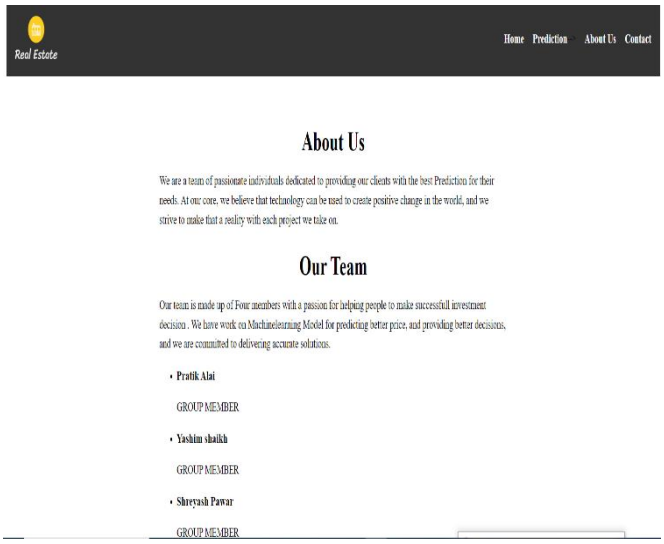
8. Model Deployment: Once the model is trained and validated, it can be deployed for use in predicting real estate prices. This can be done by providing the model with input data (property features), and it will output a predicted price.

Implementation

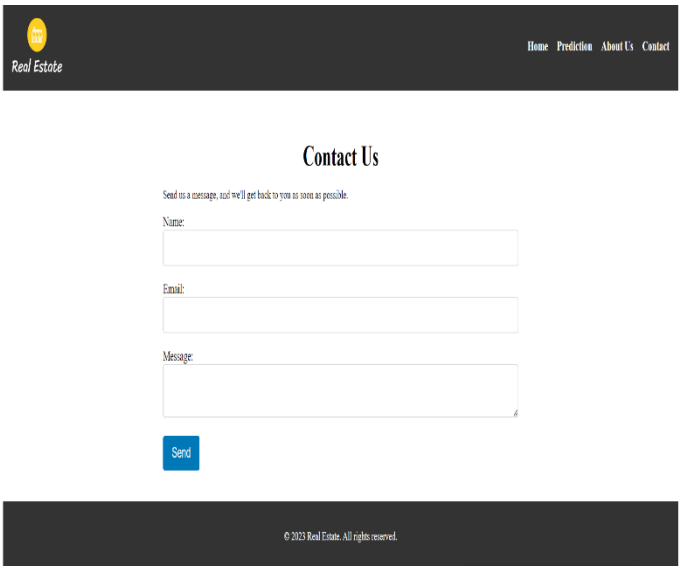
Home Page



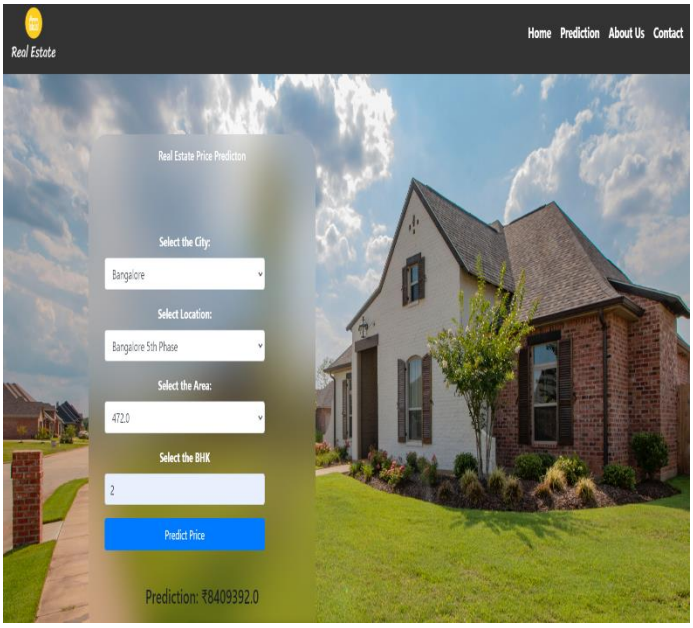
About Us Page



Contact Us Page



Prediction Page



3. CONCLUSION

In this study, our main focus has been on the development of a robust real estate price prediction system for multiple cities. The system we have designed enables users to conveniently search for property prices and houses in various cities, all from the comfort of their own homes.

ACKNOWLEDGEMENT

We extend our sincere appreciation to our supervisor for their invaluable feedback and guidance during the completion of this survey paper. Their expertise and support have been instrumental in shaping our work. Additionally, we would like to acknowledge the authors of the research papers we examined, as their contributions have significantly advanced the field of real estate price prediction through machine learning. Their insightful studies have provided us with a strong foundation for our research.

REFERENCES

[1] 1Ayush Varma, Abhijit Sarma, Sagar Doshi, Rohini Nair - "Housing Price Prediction Using Machine Learning and Neural Networks" 2018, IEEE.

[2] G.Naga Satish, Ch.V.Raghavendran, M.D.Sugnana Rao, Ch.Srinivasulu "House Price Prediction Using Machine Learning". IJITEE, 2019.

[3] CH. Raga Madhuri, G. Anuradha, M. Vani Pujitha -" House Price Prediction Using Regression Techniques: A Comparative Study" 2019 in (ICSSS),IEEE

[4] Sifei Lu, Zengxiang Li, Zheng Qin , Xulei Yang ,

[5] *Bharatiya, Dinesh, et al. "Stock market prediction using linear regression."Electronics, Communication, and Aerospace Technology (ICECA), 2017 International conference of. Vol. 2. IEEE, 2017*