

# House Price Prediction using Machine Learning

Mahesh Yadav<sup>1</sup>, Nageshwari Gaudgaonkar<sup>2</sup>, Simran Shaikh<sup>3</sup>, Tukaram Kore<sup>4</sup>,  
Suhass Salgar<sup>5</sup>, Prof. A. S. Shimpi<sup>6</sup>

Brahmdevdada Mane Institute of Technology, Belati, Solapur, India<sup>1-6</sup>

mahi224224@gmail.com, nageshwarigaudgaonkar@gmail.com, simranshaikh2026@gmail.com

tukaramkore0984@gmail.com, suhassalgar46@gmail.com

**Abstract:** Machine Learning plays a virtual role from past years in normal speech command, product recommendation. This all of things shows that ML is trending technology in almost all fields so we are trying to coined up ML in our project. Nowadays the real estate market is a standout amongst the most focused regarding pricing and keep fluctuating. People are looking to buy a new home with their budgets and by analyzing market strategies. But main disadvantage of current system is to calculate a price of house without necessary prediction about future market trends and result is price increase. So, the main aim of our project is to predict accurate price of house without any loss. There are many factors that have to be taken into consideration for predicting house price and try to predict efficient house pricing for customers with respect to their budget as well as also according to their priorities. So, we are creating a housing cost prediction model. By using Machine learning algorithms like Linear Regression. This model will help people to put resources into a bequest without moving towards a broker. The result of this research gives maximum accuracy.

**Keywords:** Linear Regression (LR), Data Pre-processing, Machine Learning, Classification etc

## I. INTRODUCTION

One of the basic requirements of livelihood in the recent world is to buy a house of your own. The price of the house may depend on various factors. Real estate agents and many who are involved in selling the house want a price tag on the house which would be the real worth of buying the house. The prediction of the price of the house is often very hard for the inexperienced.

Here in this project, we are going to use linear regression algorithm (a supervised learning approach) in machine learning to build a predictive model for estimation of the house price for real estate customers.

In this project we are going to build the Machine learning model by using python programming and other python tools like NumPy, pandas, matplotlib etc. We are also using scikit-learn library in our approach of this project. For going further into this project, we prepare dataset which consist of features like number of bedrooms, area of house, locality, etc. After preparing the dataset we will use 80 percentage of data for training the ML model and 20 percentage of data for the testing the ML model.

This project's objective is to explore different machine learning algorithms and identify the most effective approach for predicting house prices accurately. Machine learning algorithms are purely based on data. Machine Learning algorithms are an advanced version of the regular algorithm. It makes programs smarter by allowing them to automatically learn from the data provided by us.

The results of this study can help real estate agents, property appraisers, and investors make informed decisions about buying, selling, and investing in properties.

The aim of this system is to create a website through which the user can give his house requirements as input which is then passed on to the linear regression model for predicting the house price. The website also allows user to forecast the predicted house price to a particular date which is also specified by the user.

## **II. LITERRATURE REVIEW**

### **2.1 House Price Forecasting**

People looking to buy a new home tend to be more conservative with their budgets and market strategies. The existing system involves calculation of house prices without the necessary prediction about future market trends and price increase. The goal of the paper is to predict the efficient house pricing for real estate customers with respect to their budgets and priorities. By analyzing previous market trends and price ranges, and also upcoming developments future prices will be predicted. The functioning of this paper involves a website which accepts customer's specifications and then combines the application of multiple linear regression algorithm of data mining. This application will help customers to invest in an estate without approaching an agent

### **2.2 Applied Research On House Price Prediction Using Diverse Machine Learning Techniques**

With the booming civilization and ever-changing market requirements, it is essential to know the market drifts. Today prediction of house prices according to the trends is the principal essence of the study. It is imperative for an individual to understand the business trends so that he can prepare his budgetary needs according to his requirements. Real Estate is an ever-growing enterprise with an expanding society. For an investor, it is essential to comprehend the business drifts, which can assist him to underwrite in the right way and augment his business throughput. Sometimes clients get dupe by the hoax market rate set up the agent due to which the real estate industry is less translucent these days. With an uptick in convince of the dataset, it's viable for a researcher to develop a model with high accuracy. The previous model with decreased accuracy and overfitting of data reduces the efficiency, whereas the proposed system resolves such issues and provides a better and enhanced model with a rich user interface.

### **2.3 House Price Prediction using a Machine Learning Model: A Survey of Literature**

Research has stated that the fluctuations in house prices are often a concern for house owners and the real estate market. A survey of literature is carried out to analyze the relevant attributes and the most efficient models to forecast the house prices. The findings of this analysis verified the use of the Artificial Neural Network, Support Vector Regression and XGBoost as the most efficient models compared to others. Moreover, our findings also suggest that locational attributes and structural attributes are prominent factors in predicting house prices. This study will be of tremendous benefit, especially to housing developers and researchers, to ascertain the most significant attributes to determine house prices and to acknowledge the best machine learning model to be used to conduct a study in this field.

### **2.4 House Price Prediction Modeling Using Machine Learning**

Machine Learning is seeing its growth more rapidly in this decade. Many applications and algorithms evolve in Machine Learning day to day. One such application found in journals is house price prediction. House prices are increasing every year which has necessitated the modeling of house price prediction. These models constructed, help the customers to purchase a house suitable for their need. Proposed work makes use of the attributes or features of the houses such as number of bedrooms available in the house, age of the house, travelling facility from the location, school facility available nearby the houses and shopping malls available nearby the house location. House availability based on desired features of the house and house price prediction are modeled in the proposed work and the model is constructed for a small town in West Godavari district of Andhrapradesh. The work involves decision tree classification, decision tree regression and multiple linear regression and is implemented using Scikit-Learn Machine Learning Tool.

## **III. WORK CARRIED OUT**

The main purpose of this project is to develop a reliable model that can accurately estimate the prices of house based on their characteristics. Examples a homebuyer is interested in purchasing a house in a particular area. They input various features of the house, such as its size, number of bedrooms, bathrooms, and location, into the predictive model. The model then processes this information and provides an estimated price for the house based on historical data and patterns learned during training.

### 3.1 Algorithm

The Linear Regression (LR) technique fits a straight line to the data points to represent the relationship between the independent variables (properties such as square footage, number of bedrooms, and so on) and the dependent variable (house pricing). The algorithm finds the best-fitting line by reducing the sum of squared discrepancies between predicted and actual values. This line shows the linear relationship between attributes and housing prices, allowing for precise forecasts depending on new input data. LR essentially evaluates how changes in the independent factors affect the variable that is dependent for the purpose to make accurate house price predictions.

## IV. IMPLEMENTATION DETAILS

### 4.1 Data Collection:

Gather a data on various features that includes relevant features of houses such as location, number of rooms, square feet, and sale prices. Ensure the dataset that has the features that you are about to consider.

### 4.2 Data Pre-processing

Clean and prepare the collected data for model training. Handle missing values, perform feature scaling to bring features to a similar range, encode categorical variables, and address outliers. Additionally, one can explore feature engineering techniques to create new meaningful features.

### 4.3 Model Training:

In this phase, data is broken down into two parts: Training and Testing. There are 80% of data is used for training purpose and remaining 20% used for testing purpose. The training set include target variable. The model is trained by using various machine learning algorithms and getting the result.

### 4.4 Model Evaluation:

Assess the model's performance using metrics like Mean Squared Error or R-squared to measure how well the model fits the data. Additionally, techniques like cross-validation help validate model robustness and prevent overfitting by splitting the dataset into training and testing subsets multiple times. These evaluation methods ensure the reliability and effectiveness of the predictive models in real-world applications.

### 4.5 Prediction:

The prediction process revolves around leveraging historical data and various features of properties to forecast the prices of unseen properties accurately. Machine learning algorithms, such as linear regression, are trained on datasets containing information like square footage, location, number of bedrooms/bathrooms, and other relevant factors. During prediction, the trained model takes input features of a new property and produces an estimated price based on patterns learned from the training data. Continuous refinement and optimization of the model ensure it captures intricate relationships between features and prices, enabling more precise predictions.

## V. RESULTS





## VI. CONCLUSION

In this paper we have successfully developed a machine learning web solution to predict house prices based on various features. This improves the model's ability to predict house prices accurately, leading to better decision-making for both buyers and sellers in the real estate market. By implementing the model in a web-based solution, users can input data on a house, and the solution will provide an estimated price based on the model's predictions. This makes it easier for buyers and sellers to obtain a rough estimate of a property's value without the need for extensive research. Overall, this machine learning web solution for house price prediction provides a valuable tool for the real estate industry and can aid in making more informed decisions regarding property values.

## REFERENCES

- [1] Ms. Ankita Gupta, Yashwant Jangid, Tushar Tiwari, Saurabh Jain, Rushab Sawant A Multi Feature-Based Housing Price Prediction for Indian Market Using Machine Learning||, ISSN 2347 8527, IJCMS December 2017
- [2] William M. Doerner, Alexander N. Bogin, Property Renovations and Their Impact on House Price Index Construction||, <https://www.fhfa.gov/PolicyProgramsResearch/Research/PaperDocuments/wp1702.pdf>
- [3] 3 KANG Ling-Wei and D. X. Zhu, The Land Prices and Housing Prices Empirical Research Based on Panel Data of 11 Provinces and Municipalities in Eastern China||, 2013 International Conference on Management Science Engineering (20th) July 17- 19, 2013.
- [4] Susmita Ray, A Quick Review of Machine Learning Algorithms, ||. International Conference on Machine Learning, Big Data, Cloud and Parallel Computing, India, 2019
- [5] Suchibrotta Dutta, Debanjan Banerjee, Predicting the Housing Price Direction using Machine Learning Techniques||, 2017 IEEE ICPCSI