House Price Prediction

Ammar Anwar, Arbaz Ansari, Siddhesh Ghadigaonkar, Dr. Varsha Shah, Prof. Shiburaj Pappu

Student, Student, Student, Principal, Head of Department

Department of Computer Engineering,

Rizvi College of Engineering, Mumbai, India.

Abstract: Data mining is now commonly used in the real estate market. Real Estate is a clear industry in our ecosystem. The ability to extract data to extract relevant information from raw data makes it very useful to predict house prices, important housing features, and much more. Housing prices continue to change from day to day and are sometimes raised rather than based on calculations. Research has shown that fluctuations in housing prices often affect homeowners and the housing market. Literature research is done to analyze the relevant factors and the most effective models for predicting housing prices. The findings of this analysis confirmed the use of Artificial Neural Network, Support Vector Regression, and Linear Regression as the most efficient models compared to others. In addition, our findings also suggest that spatial and real estate agents are key factors in predicting house prices. This study will be of great benefit, especially to housing developers and researchers, to find the most important criteria for determining housing prices and identify the best machine learning model used to conduct research in this field.

Keywords: House price prediction, Machine Learning, Linear regression.

1. INTRODUCTION:

In this report, we propose our system "House price prediction". House is one of human life's most essential needs, along with other fundamental needs such as food, water, and much more. Demand for houses grew rapidly over the years as people's living standards improved. House price prediction can be done using multiple prediction models (Machine Learning Model) such as support vector regression, artificial neural network, etc. There are many benefits that home buyers, property investors, and housebuilders can reap from the house-price model. This model will provide a lot of information and knowledge to home buyers, property investors, and housebuilders, such as the valuation of house prices in the present market, which will help them determine house prices. The target feature in this proposed model is the price of the real estate property and the independent features are: no. of bedrooms, carpet area, the floor, car parking, and lift availability. The whole implementation is done using the python programming language.

2. PROBLEM STATEMENT:

The general and standardized real estate characteristics are often listed separately from the asking price and general description. Because these characteristics are separately listed in a structured way, they can be easily compared across the whole range of potential houses. Because every house also has its unique characteristics, such as a particular view or type of sink, house sellers can provide a summary of all the important features of the house in the description. All given real estate features can be considered by the potential buyers, but it is nearly impossible to provide an automated comparison of all variables due to the large diversity. This is also true in the other direction: house sellers have to estimate the value based on its features in comparison to the current market price of similar houses. The diversity of features makes it challenging to estimate an adequate market price. Apart from providing a summary of the important features of the house, the house description is also a means of raising curiosity in the reader, or in other words persuading the person. Housing prices are an important reflection of the economy, and housing price ranges are of great interest to both buyers and sellers. In this project, house prices will be predicted given explanatory variables that cover many aspects of residential houses. The goal of this project is to create a regression model that can accurately estimate the price of the house given the features.

3. LITERATURE REVIEW:

i. Survey Existing System:

Trends in housing prices indicate the current economic situation and also are a concern to the buyers and sellers. Many factors have an impact on house prices, such as the number of bedrooms and bathrooms. House price depends upon its location as well. A house with great accessibility to highways, schools, malls, and employment opportunities, would have a greater price as compared to a house with no such accessibility. Predicting house prices manually is a difficult task and generally not very accurate, hence there are many systems developed for house price prediction. Sifei Lu, Zengxiang Li, Zheng Qin, Xulei Yang, and Rick Siow Mong Goh had proposed an advanced house prediction system using linear regression. This system aimed to make a model that can give us a good house price prediction based on other variables. They used the Linear Regression for Ames dataset and hence it gave good accuracy.

ii. Limitation Existing system or research gap:

In the existing literature, a limited amount of work has been focused on the housing price prediction model, particularly, to solve the problem using machine learning approaches. A few identified papers were reported above. In addition, most of the past research considered the housing market problem as a classification problem to develop a classification model instead of a regression model. Therefore, the objective of the study is to predict the housing price valuation using machine learning techniques and considering competitive regression models. An improved ML-based algorithm is proposed, which includes the predicted target price binning variable as features in the model and improves the model accuracy significantly. More precisely, the model accuracy is increased by 10 percent compared to other contemporary machine learning techniques.

4. METHODOLOGY:

The proposed system face recognition-based attendance system can be divided into three main modules. The modules and their functions are defined as follows.

a. **Data Collection:**

For doing machine learning projects we need a vast amount of data which we have taken through the Kaggle website. The data consists of several factors in a house of a given locality such as the number of bedrooms, carpet area, location, etc. The dataset that we will be using will be batch i.e., static data and not dynamic

b. **Train the Dataset and Feature Extraction:**

The dataset will further be trained and processed using Linear Regression Algorithm which will predict the house price accurately. The algorithm used will be a regression since regression helps to get a value by analyzing a dataset, in this condition the value is the price of the house.

c. **Displaying the output:**

The project uses a web application to work with the users and give them accurate price predictions. The user will input features accordingly which they want on the frontend and the data will be processed on the backend using Linear Regression and the predicted output will be shown to the respective user

5. ALGORITHM:

1) Importing the required packages into our python environment
2) Importing the house price data
3) Data Visualization of the house price data
4) Modelling the data using the algorithms
5) Create a responsive website
6) Take inputs from the user and display the analyzed result

6. RESULT AND OUTPUT:

The result of our project has an accuracy of 86.45%, where the user has to enter the location, no. of bedrooms, lift availability, and car park to get the desired output.
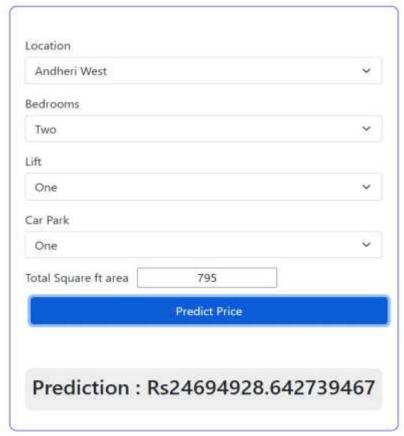
Fig 1. Result of a sample location

Tools required: -

1. Python
2. Jyupter Notebook
3. Scikit Learn.
4. NumPy

Conclusion:

This paper examined and analyzed the current research on the significant attributes of house prices and analyzed thedata mining techniques used to predict house prices. The accurate prediction model would allow investors or house buyers to determine the realistic price of a house as well as the house developers to decide the affordable house price. This paper discusses an overview of the concept of machine learning and its various applications. Taking the sample dataset for houses, and considering its various attributes, the prices for houses have been predicted by employing machine learning methods of regression for predicting the price of the estate using prior data.

References:

[1] Theobald, O. (2017). *Machine learning for absolute beginners*
[2] Mueller, J. P., & Massaron, L. (2016). *Machine learning for dummies*. John Wiley & sons
[3] G. Gao *et al.*, "Location-Centered House Price Prediction: A Multi-Task Learning Approach,"
    pp. 1–14, 2019, [Online]. Available: http://arxiv.org/abs/1901.01774.

[4] T. D. Phan, "Housing price prediction using machine learning algorithms: The case of Melbourne city, Australia," *Proc. - Int. Conf. Mach. Learn. Data Eng. made 2018*, pp. 8–13, 2019, DOI: 10.1109/iCMLDE.2018.00017.

[5] Park, B., & Bae, J. K. (2015). Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data. Expert Systems with Applications.

[6] Y. Zhou, "Housing Sale Price Prediction Using Machine Learning Algorithms," 2020.

[7] T. Mohd, S. Masrom, and N. Johari, "Machine learning housing price prediction in petaling Jaya, Selangor, Malaysia," *Int. J. Recent Technol. Eng.*, vol. 8, no. 2 Special Issue 11, pp. 542–546, 2019, Doi: 10.35940/ijrte. B1084.0982S1119.

[8] A. Varma, A. Sarma, S. Doshi, and R. Nair, "House Price Prediction Using Machine Learning and Neural Networks," *Proc. Int. Conf. Inven. Commun. Comput. Technol. ICICCT 2018*, pp. 1936–1939, 2018, Doi: 10.1109/ICICCT.2018.8473231.

[9] H. Wu *et al.*, "Influence factors and regression model of urban housing prices based on internet open access data," *Sustain.*, vol. 10, no. 5, pp. 1–17, 2018, Doi: 10.3390/su10051676.