Asian Journal of Computer Science and Technology
ISSN: 2249-0701 (P); 2583-7907 (E)
Vol.12 No.1, 2023, pp.31-34

DOI: https://doi.org/10.51983/ajcst-2023.12.1.3592

# An Algorithm to Predict Real Estate Price using Machine Learning

**Ruchi Vyas and Jitendra Sharma**
Assistant Professor, Department of Computer Science and Engineering,
Geetanjali Institute of Technical Studies, Rajasthan, India
E-mail: ruchi.vyas@gits.ac.in, jitendra.sharma@gits.ac.in
(Received 30 April 2023; Accepted 8 May 2023; Available online 12 May 2023)

*Abstract* - **Real Estate Pricing has always been a varying from time to time keeping a buyer confuse on how to get exact price of the property at certain point of time. A machine learning algorithm is used to predict the exact price of the property. Predicting the accurate amount of the real estate is very much a matter of concern as people are investing too much now a days in property where a property dealer could easily charge more amount if the buyer is not knowing the market price of the property. The major focus of this research paper is to predict the accurate price of the property without a hassle. Apart from that it also focuses on increasing the accuracy of the already existing system. There is various machine learning algorithm available for prediction such as Naive Bayes, Logistic Regression, Classification, Random Forest KNN, Support Vector Machine, Lasso, Linear Regression etc. The aim of this research is to predict the market value of real estate properties based on geological location. By analyzing previous market patterns, value ranges, and upcoming developments, we can determine a starting price for a property based on geological variables. Our study is applicable to any location, and we have utilized three machine learning algorithms to make predictions. Our findings indicate that linear regression provides the most accurate predictions, with an accuracy rate of 85%. This system eliminates the need for clients to rely on brokers and provides them with the confidence to invest in real estate. The accuracy of our system surpasses that of previous methods used in the industry.**
*Keywords:* **Price Prediction, Machine Learning Algorithm, Real Estate**

## I. INTRODUCTION

In the competitive world of real estate, organizations strive to gain an edge over their rivals by streamlining processes and optimizing outputs. This study proposes a linear regression machine learning algorithm to predict the prices of houses, both those already on the market and those still under construction. The goal is to determine the appropriate price tag for a house before putting it up for sale. Using regression analysis, the model examines the relationships between the target parameter (house price) and various independent parameters, such as the number of bedrooms, bathrooms, area, and location. These features are essential for predicting property values, as higher values can lead to a decrease in house prices. The implementation is done in Python, with the Scikit-learn library providing the linear regression model and Grid Search CV identifying the best max-depth value. The user interface is created using Flask, along with HTML, CSS, and Javascript for webpage presentation.

The objective of this project is to provide real estate customers with an accurate forecast of economic house pricing based on their available resources. To achieve this, we will analyze past market trends and price ranges, as well as upcoming developments, in order to determine future prices. By using this application, customers can invest in an estate without the need for a real estate agent, thereby reducing the associated transactional risks. Currently, the process of buying and selling property is both stressful and expensive, as customers must physically visit multiple locations and pay commissions to agents. Furthermore, customers/buyers are often unsure whether the property will be profitable in the future. In response, we have developed a website that utilizes data mining techniques to overcome the limitations of the current system. Our website provides users with a web-based platform that offers future estimates of property values and approximate costs based on a variety of factors.

## II. LITERATURE SURVEY

The global financial crisis that occurred in recent years has sparked a renewed interest in understanding the relationship between asset prices, particularly housing prices, and economic activity. Lamer (2007) has noted that the housing market predicted eight out of the ten post-World War II recessions, which emphasizes its importance as a leading indicator of the real economy. Lamer has gone so far as to suggest that “housing is the business cycle.” Vargas and Silva (2008) support this notion by stating that changes in housing prices are crucial in determining the phase of the business cycle. During an economic boom, the housing sector responds to excess demand by expanding rapidly, leading to an increase in both construction and employment, and driving up nominal house prices. Conversely, during the contraction phase, the decline in private finance reduces demand, resulting in a decrease in nominal house prices. However, nominal house prices typically fall slowly since homeowners are hesitant to lower their prices. Most of the adjustment is achieved through declines in sales volume, leading to a decline in the construction sector and housing-related occupations. Additionally, during a contraction or recession, real house prices fall rapidly due to general inflationary trends that decrease real house prices, which are perceived as having sticky values.

The present system that contains the involvement of human interference in making the document handwritten and saving the document may lead to many drawbacks and contain many loopholes.

Some of them are,

1. The current system requires a significant amount of manual work, from filling out forms to preparing documents. This can create stress and anxiety among workers and often results in suboptimal outcomes.
2. The current system is prone to errors and requires extensive manual work to make even minor changes, which can be time-consuming and frustrating for employees.
3. Given that the current system is managed by human workers, it is inevitable that errors may occur from time to time.

## III. STATEMENT OF PROBLEM

A Web Based and Mobile App Based Solution to Predict Real Estate Price using Machine Learning.

## IV. ALGORITHM

To complete a comprehensive Machine Learning project, we need to follow a series of steps:

1. Clearly comprehend the business requirements.
2. Obtain the dataset, either by creating one or updating an existing one.
3. Visualize the data to gain deeper insights and develop a better understanding of it.
4. Preprocess the data so that it can be fed into our machine learning model, ensuring we use the model with the highest accuracy.
5. Execute different models and start training them. Now select that model whose performance is best.
6. Next step is to fine tune all those models who have good performance by just by updating the hyperparameter.
7. The result needs to be published between team members.
8. After results system need to be launched and must be maintained over time.

In the process of developing the model, various regression algorithms were studied. Linear regression, Random Forest, lasso, etc. all algorithms tested on the training dataset. However, the Linear regression provided the highest accuracy of 84% in terms of predicting the house prices. After finalized the dataset we can easily apply the machine learning algorithms for the best output in this project.

In this project, we train our dataset with linear regression algorithms. After successfully dealing with the model, it produces a better output, the next step is to fusion with UI, for this fusion flask is used. Flask is a web framework. This means Flask offers different tools, technologies, and libraries that allow it to build a web application.

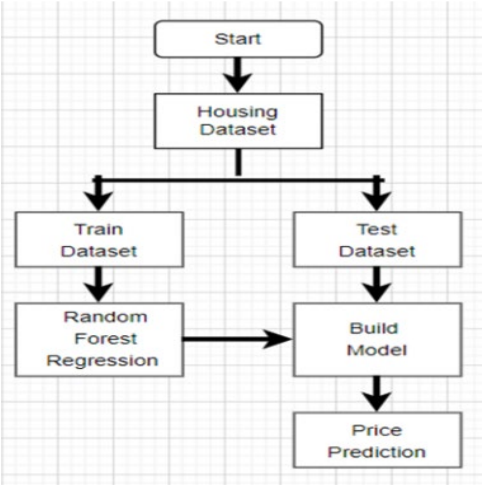


Fig. 1 Process diagram

## V. REQUIREMENTS

*A. Software Requirements*

1. Python Envt.
2. Anaconda
3. Python Flask
4. HTML CSS
5. Javascript
6. Panda Library
7. Matplotlib for analyzing and visualization
8. SCIKIT library

*B. Hardware Requirements*

1. 8 GB RAM.
2. 512 GB HDD.
3. Intel Processor I5, I7.

## VI. IMPLEMENTATION

The missing values for the rooms and bathrooms in the dataset were dealt with using the Pandas library, and the target attribute was dropped. To gain insight from the given data set we can find min max and std deviation & mean value of the target attribute after the calculation. During a conversation with a real estate expert business manager, it was suggested that an apartment with 2 bedrooms should have a minimum of 600 sqft (i.e., 300 sqft per bedroom). Any apartment with less than 600 sqft but with 2 bedrooms was deemed suspicious and removed as an outlier. For instance, a 6 BHK apartment with only 1020 sqft or an 8 BHK apartment with a total of 600 sqft were considered data errors and were safely discarded. Usually, a dataset is being divided into two where first consist of 80% of data set which is basically used for training purpose. Other part is 20% consist of test dataset.

Initially we have already mentioned that a grid search CV will help us to find the maximum depth of a tree through the algorithm then we have used matplotlib library for the graph plotting which shows data visualization and graphs so obtained could be used for data analysis.

1. Since 2BHK Apartments could be straight away removed who have price per sqft less than 1 BHK Apartments.
2. Depending upon the graph obtained we could see the data highlighted as they are being removed by using the function outliers.

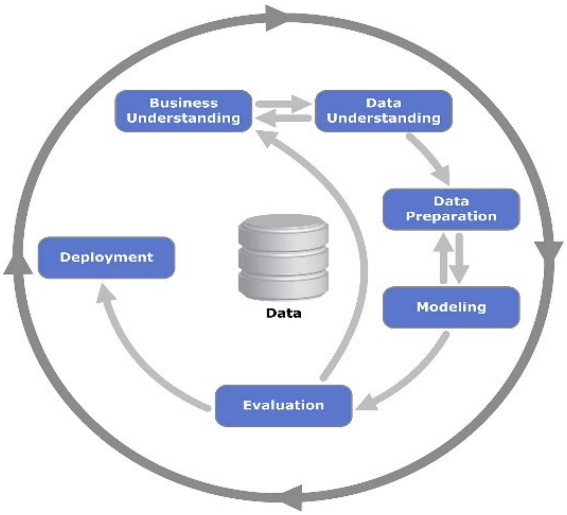


Fig. 2 Methodology Used

## VII. RESULTS OF THE STUDY

The Output so obtained after using pycharm and jupyter notebook.

TABLE I RESULTS

| Sl. No. | Value |
|---|---|
| 1 | 53.86 |
| 2 | 48.42 |
| 3 | 163.93 |
| 4 | 167.2 |
| 5 | 23.87 |

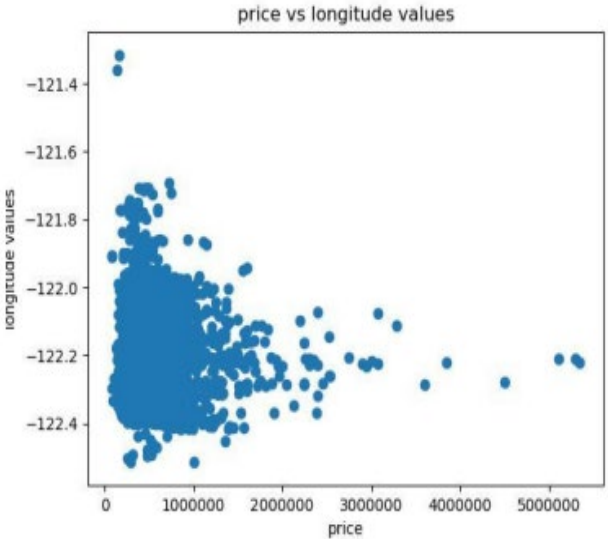


Fig. 3 Sites plotting before and after changing

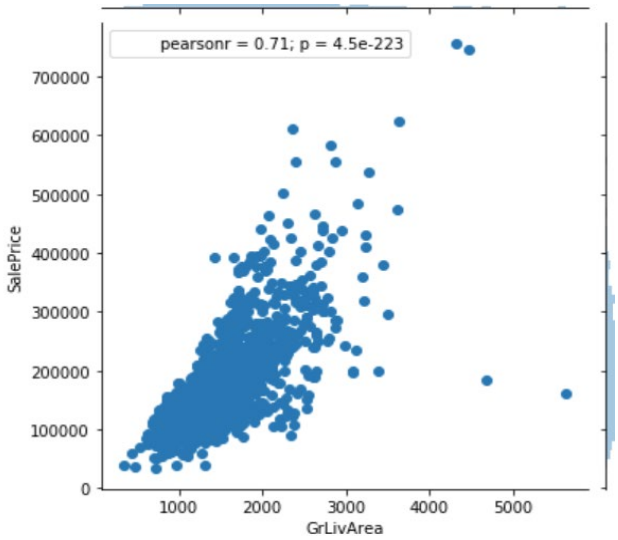


Fig. 4 Plot of site area for a location

## VIII. CONCLUSION AND FUTURE SCOPE

In the current real estate industry, managing and extracting large amounts of data has become a challenging task. Moreover, it is essential to ensure that the extracted data is useful for one's requirements. The proposed system effectively utilizes a data mining algorithm to optimize data usage. However, the system can be enhanced by including additional attributes to increase its applicability. One potential future scope is to expand the estate database to include more cities, which would provide users with more options to explore and make informed decisions. Although the system identified most residential areas, it does not evolve complexes or multi story houses & apartment located near industrial area. Such properties can be added in the future to provide more accurate results. The demand for housing in metropolitan cities has led to an increase in the number of private builders offering real estate with additional features to attract more customers. Our proposal is a solution which is Web Based as well as mobile application which could eventually help the buyer as well as seller and builders. As a builder can advertise their real estate on our Web Based solution or Mobile App Based solution. From a buyer's point of view, it can view multiple options making it a more convenient solution for all having their needs fulfilled.

## REFERENCES

[1] A. P. Singh, K. Rastogi and S. Rajpoot, "House Price Prediction Using Machine Learning," in *3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, Greater Noida, India, pp. 203-206, 2021, DOI: 10.1109/ICAC3N53548.2021.9725552.

[2] A. Varma, A. Sarma, S. Doshi and R. Nair, "House Price Prediction Using Machine Learning and Neural Networks," in *Second International Conference on Inventive Communication and Computational Technologies (ICICCT)*, Coimbatore, India, pp. 1936-1939, 2018, DOI: 10.1109/ICICCT.2018.8473231.

[3] S. Anand, P. Yadav, A. Gaur and I. Kashyap, "Real Estate Price Prediction Model," in *3rd International Conference on Advances in Computing, Communication Control and Networking (ICAC3N)*, Greater Noida, India, pp. 541-543, 2021, DOI: 10.1109/ICAC3N53548.2021.9725772.

[4] M. Jain, H. Rajput, N. Garg and P. Chawla, "Prediction of House Pricing using Machine Learning with Python," in *International Conference on Electronics and Sustainable Communication Systems (ICESC),* Coimbatore, India, pp. 570-574, 2020, DOI: 10.1109/ICESC 48915.2020.9155839.

[5] V. Gampala, N. Y. Sai and T. N. Sai Bhavya, "Real-Estate Price Prediction System using Machine Learning," in *International Conference on Applied Artificial Intelligence and Computing (ICAAIC),* Salem, India, pp. 533-538, 2022, DOI: 10.1109/ICAAIC 53929.2022.9793177.

[6] Y. Panchal, M. Mer and A. Ghosh, "Residential Property Price Prediction Using Machine Learning: Makan SETU," in *International Seminar on Application for Technology of Information and Communication (iSemantic),* Semarang, Indonesia, pp. 257-262, 2022, DOI: 10.1109/iSemantic55962.2022.9920395

[7] T. D. Phan, "Housing Price Prediction Using Machine Learning Algorithms: The Case of Melbourne City, Australia," in *International Conference on Machine Learning and Data Engineering (iCMLDE),* Sydney, NSW, Australia, pp. 35-42, 2018, DOI: 10.1109/iCMLDE .2018.00017.

[8] C. Chee Kin, Z. Arabee Bin Abdul Salam and K. Batcha Nowshath, "Machine Learning based House Price Prediction Model," in *International Conference on Edge Computing and Applications (ICECAA),* Tamilnadu, India, pp. 1423-1426, 2022, DOI: 10.1109/ICE CAA55415.2022.9936336.

[9] G. K. Kumar, D. M. Rani, N. Koppula and S. Ashraf, "Prediction of House Price Using Machine Learning Algorithms," in *5th International Conference on Trends in Electronics and Informatics (ICOEI),* Tirunelveli, India, pp. 1268-1271, 2021, DOI: 10.1109/ICO EI51242.2021.9452820.

[10] D. Vathana, R. Patel and M. Bargoti, "Real Estate Price Prediction Using Machine Learning Algorithm," In *Mohanty, M.N., Das, S. (eds) Advances in Intelligent Computing and Communication,* Lecture Notes in Networks and Systems, Springer, Singapore, Vol. 430, 2022, DOI: https://doi.org/10.1007/978-981-19-0825-5_47.