

# House Price Prediction Using Machine Learning

**Raghuram<sup>1</sup>, Kavya N<sup>2</sup>, S S Nagamuthu Krishnan<sup>3</sup>**

Student, Master of Computer Applications, RV College of Engineering, Bengaluru, India<sup>1</sup>

Student, Master of Computer Applications, RV College of Engineering, Bengaluru, India<sup>2</sup>

Assistant Professor, Master of Computer Applications, RV College of Engineering, Bengaluru, India<sup>3</sup>

**Abstract:** Real estate is not as transparent as other industries. It is difficult to predict the future price of houses due to the various factors that influence them. Houses regularly change in price and are hyped instead of being valued based on their valuation.

Predicting house prices can aid to find out the price of a house of specific region and will assist the people to get the right time to acquisition a home. This website will mainly help users to get the most precise house price of the relevant location with the assist of Machine Learning and AI Implementation, besides this website will be deployed in a cloud which makes accessibility of website to all kind of users all time. User can provide the needed features and get the most precise price of house in one click with no time.

**Keywords:** Flask, Heroku, Linear Regression algorithm, Lasso algorithm, Decision Tree algorithm, Bengaluru Dataset.

## I. INTRODUCTION

The real estate industry is a vital one, with many stakeholders ranging from government agencies to private businesses and investors. There is a growing inclination among these stakeholders for a better grasp of the industry's operational mechanisms and driving causes. Today, there is an abundance of data on pertinent statistics as well as ancillary contextual elements, and it is logical to seek to make use of them to advance industry expertise.

If there is one thing that is taken for granted after living in India for so long, is that housing and rental prices continue to grow. Property prices have rebounded amazingly well after the metropolitan crisis of 2008, particularly in major housing areas. However, it was surprising to learn in the fourth quarter of 2016 that Bombay home prices had dropped the most in the previous four years. In fact, median resale prices for condominiums and coops declined 6.3 percent, the first drop since Q1 of 2017. The fall has been linked in part to internal and international political uncertainty, as well as the 2014 election. As a result, it is possible to retain consumer openness while also making comparisons simple.

## II. EXISTING SYSTEM

There is a prominent amount of research have gone on house price prediction department, but every research has risen to any real-life solutions. For now, very few solutions are available those are:

- A. Buyers search for the houses with required features in websites and noting down the price of it manually.
- B. Contacting the agencies which help them to find the house which go with their requirements and budget.

## III. PROPOSED SYSTEM

Currently, e-learning and e-education are on the rise, with everything moving away from manual to automated processes. The purpose of this project is to evaluate home values in order to alleviate the inconveniences that individuals endure. The current way is for the user to consult a real estate agent to supervise his or her investments and recommend an appropriate house price. This strategy, however, is risky because the agent could forecast the wrong price, resulting in the user's assets being lost. The current manual process is obsolete and fraught with danger. As a result, an updated and automated approach for predicting housing prices is required to combat this.

Machine Learning algorithms can be used to estimate the price of a home in a certain location set of criteria stated. Furthermore, this approach is both economical and time effective, as the user receives the result in a single click and with no delay.

This study proposes the implementation of a house price prediction model for Bengaluru, India, which is a Machine Learning model combining Data Science and Web Development ideas. The online site is deployed as an app on the Heroku Cloud Platform, making it accessible from anywhere at any time. Housing prices fluctuate on a daily basis and are sometimes exaggerated rather than backed by value. This project's main focus is on projecting property prices using authentic factors like area size, location, number of bathrooms and number of bedrooms, that are designed to base estimation on every essential criterion taken into account when determining price.

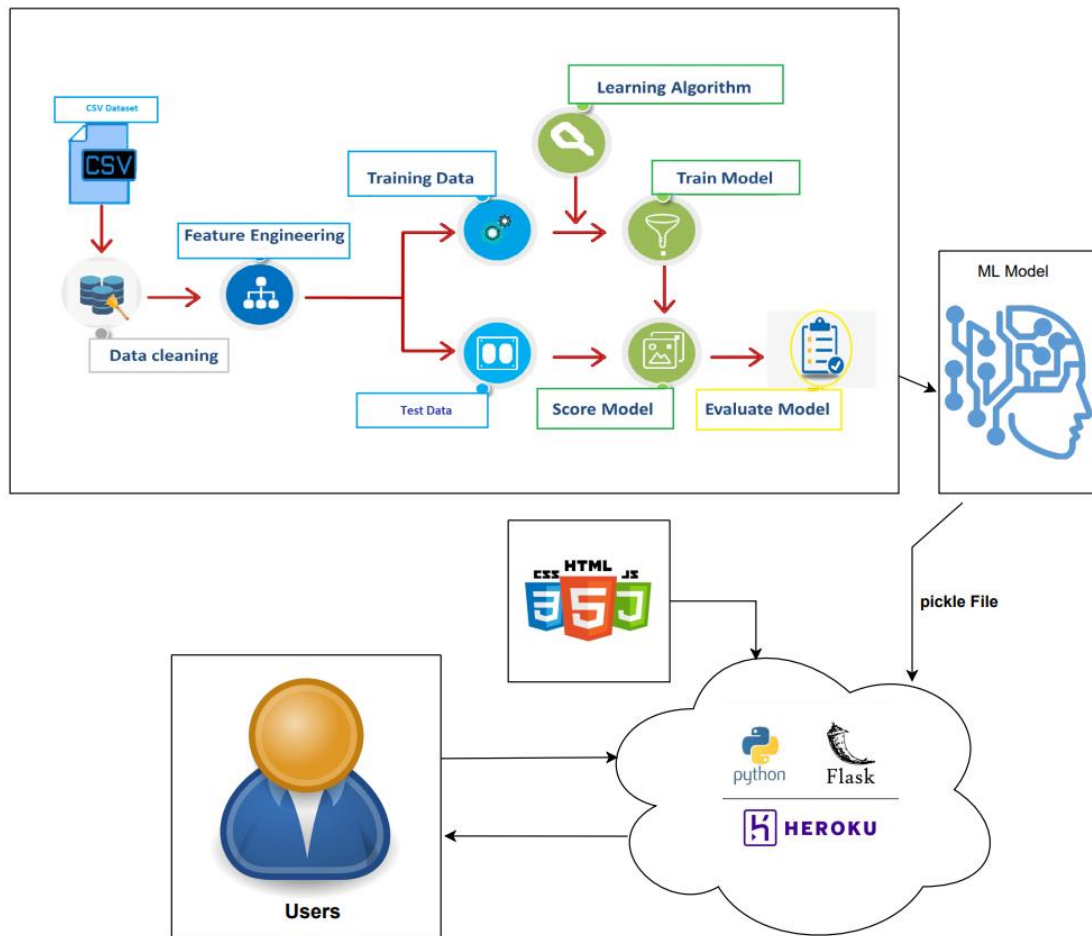


Fig. 1 Architecture Diagram

A. Machine Learning

Machine learning is a subfield of AI (Artificial Intelligence) that leverages algorithms and technologies to retrieve meaningful information from datasets. Machine learning approaches are well-suited to large datasets since manually interpreting massive amounts of data would be unfeasible without the expertise of machines. In computer science, machine learning intends to seek out solutions programmatically rather than mathematically, and so it is centered on designing algorithms that allow the machine to learn.

B. Flask

Flask is a Python web framework, essentially implies it provides tools, modules, and technologies to help users create web applications quickly and easily. This web application could be a set of interconnected pages, a blog, a Wikipedia, a browser datebook, or an official website.

C. Heroku

Heroku is a fully managed container-based cloud Platform as a Service (PaaS) that lets developers to focus on their core product without worried about maintenance servers, hardware, or infrastructure. The Heroku experience enables services, tools, workflows, and polyglot support available, all of which aid developers be much more productive.

IV. WORKFLOW OF THE APPLICATION

The DFD gives complete work flow of the application; at first the user has to enter the data, that data is converted in CSV format so that it can be loaded for data source; once the data is loaded it is preprocessed and cleaned, from the cleaned data the necessary features are extracted, on the extracted features the ml algorithm linear regression is applied in order to the price of house, once the price is predicted it will be sent as an output which can be visualized by the user.

Algorithms used to train model

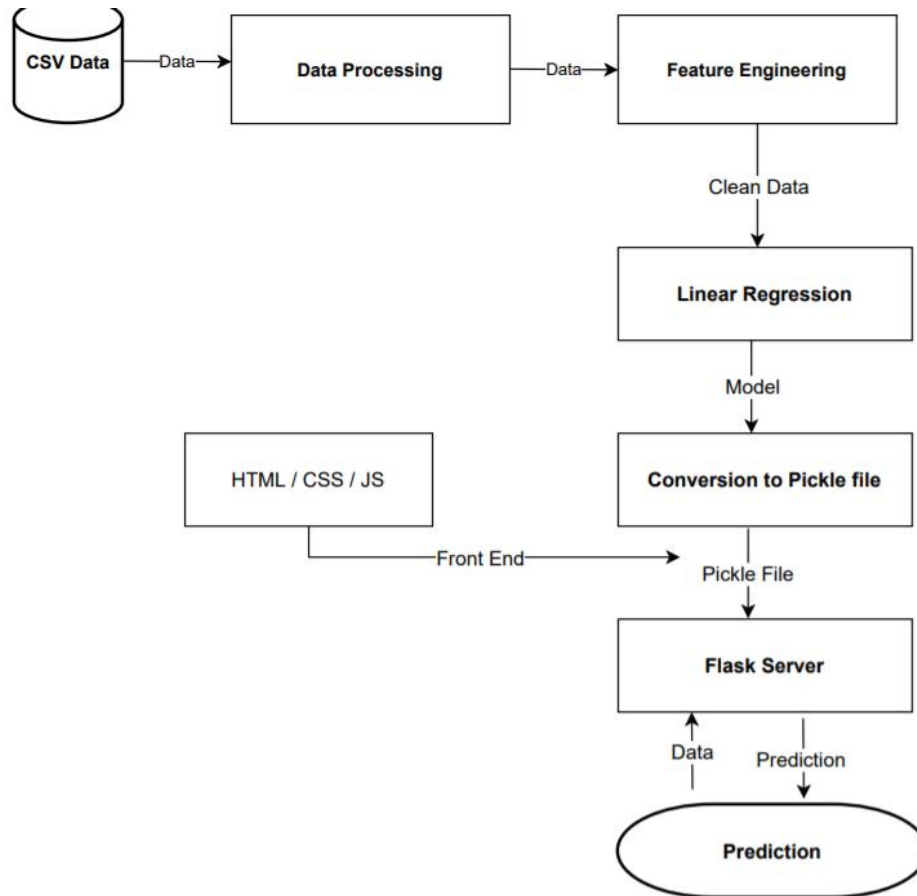


Fig. 2 Detailed Data Flow Diagram

**A. Lasso regression**

In machine learning as well as statistics, lasso which is a acronym for Least Absolute Shrinkage and Selection Operator; is a regression analysis method that performs variable selection and also regularization to improve the prediction accuracy and also interpretability of the resulting statistical model of the application.

**B. Decision Tree**

A decision tree is more precise type of probability tree which facilitates user to make a decision in the process.

**C. Linear Regression**

Linear regression is most popular ML algorithm. Basically, it is a statistical method which is used for predictive analysis. It predicts for continuous/real or numeric variables such as price, salary, age, sales, salary, age, etc.

This algorithm is evidence for linear relationship between a dependent variable-Y and one or more independent variables-X, therefore it is called as linear regression.

$$Y = a + bX$$

Here X -explanatory variable

Y- Dependent variable.

b- Slope of line

a- Intercept

In our project Area size, number of BHK, number of Bath, location are the independent variables(X), and price is the dependent variable.

**V. OUTCOME**

Among three algorithms the high accuracy of price prediction is obtained by applying linear regression algorithm there by it is used to train the model so that it can predict the prices automatically

```
***
  model  best_score  best_params
0  linear_regression  0.847796  {'normalize': False}
1      lasso  0.726738  {'alpha': 2, 'selection': 'cyclic'}
2  decision_tree  0.710426  {'criterion': 'mse', 'splitter': 'best'}
```

Fig. 3 Accuracy of the algorithms

## VI. CONCLUSION

This accurate ML model would allow investors and home buyers to estimate a property's realistic pricing, as well as home developers to calculate an economical house price. This model was created with several input attributes in mind, and it works well for house price prediction. And it has a easily operated interface. In the future, it could be used on a broad scale to evaluate house prices in many cities.

## VII. FUTURE ENHANCEMENT

In future this model can be applied to larger fields where now we have considered only Bengaluru (Toy)

- User can create can post data for selling a property.
- Direct interaction between buyer and seller.
- Model can be enhancing and can be used for rental purpose.

## ACKNOWLEDGMENT

The authors would like to thank the “RV College of Engineering” staff for their technical support towards this review paper.

## REFERENCES

Dataset link: [www.kaggle.com/bangalore-house-price-prediction](https://www.kaggle.com/bangalore-house-price-prediction)

- [1] Lu. Sifei et al, “A hybrid regression technique for house prices prediction”, In In proceedings of IEEE conference on Industrial Engineering and Engineering Management (IEEE), 2017.
- [2] 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), 2018.
- [3] B Jeevan, E Naresh, B P Vijaya Kumar, Prashanth Kambli,” Share Price Prediction using Machine Learning Technique”, In 3rd International Conference on Circuits, Control, Communication and Computing (I4C), 2018.
- [4] Nehal N. Ghosalkar, Sudhir Dhage, “Real Estate Value Prediction Using Linear Regression”, In International Conference on Computing Communication Control and Automation (ICCUBEA) – 2018
- [5] Furia, Palak, and Anand Khandare, "Real Estate Price Prediction Using Machine Learning Algorithm.", In conference on Data Science and Intelligent Computing, 2020.