



House Price Prediction

Team Members:

Magesh S - CSE
Gunalakshmi N S - ECE
Pavithra P - ECE
Joselin Sophia Rani J - ECE

Guide:

Mr. Raja P

Disclaimer: The content is curated for educational purposes only.

OUTLINE

- Abstract
- Problem Statement
- Aims, Objective & Proposed System/Solution
- System Design/Architecture
- System Development Approach (Technology Used)
- Algorithm & Deployment
- Conclusion
- Future Scope
- References
- Video of the Project



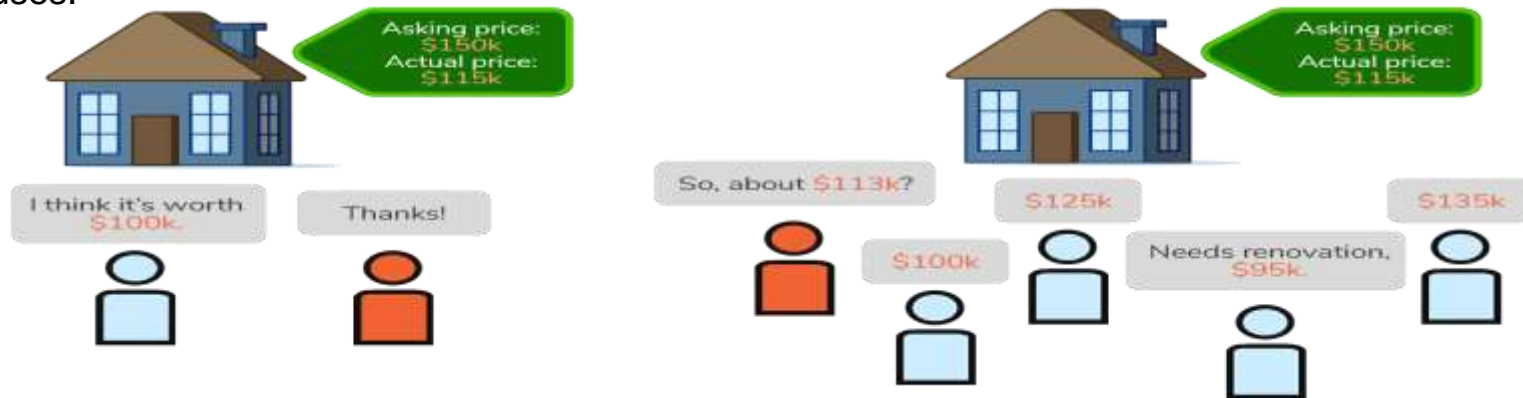
Abstract

- House Price Prediction (HPP) is commonly used to estimate the changes in housing price.
- Since housing price is strongly correlated to other factors such as location, area, population, it requires other information apart from HPP to predict individual housing price.
- House prices increase every year, so there is a need for a system to predict house prices in the future.
- We aim to make evaluations based on every basic parameter that is considered while determining the price.



Problem Statement

- Problem faced during buying a house
 Buying a house is a stressful thing.
 Buyers are generally not aware of factors that influence the house prices.
 Many problems are faced during buying a house.
- Hence real estate agents are trusted with the communication between buyers and sellers as well as laying down a legal contract for the transfer. This just create a middle man and increases the cost of houses.



Aim and Objective

- The main goal of the project is to predict the efficient house pricing for real estate customers with respect to their budgets and priorities.
- It is difficult to search house at places like Bangalore. Even if we find a house it is difficult to get a perfect price for the same.
- To overcome such problem Machine Learning technique can be used. Using Machine Learning Technique it will be easy to know the price of house based on the area available, number of bedrooms, facilities available.
- The focus is to create an “easy to use” website, which will allow to first time customer to complete their needs and ease.



Proposed Solution

- The proposed solution for the Bangalore House Price Predictor involves implementing a supervised machine learning algorithm, such as a regression model.
- The dataset will be preprocessed to handle missing values, outliers, and categorical variables.
- Feature engineering may be applied to extract relevant information from the available features.
- The dataset will then be split into training and testing sets to train and evaluate the model's performance.
- Various regression algorithms, such as linear regression, decision trees, or ensemble methods like Random Forest, will be tested to determine the most accurate model.
- Hyperparameter tuning and cross-validation techniques will be employed to enhance the model's robustness. The selected model will be deployed to predict house prices based on input features

System Architecture

- **Data Ingestion and Preprocessing:**

Raw housing data from various sources is ingested into the system.

Data preprocessing involves handling missing values, outliers, and encoding categorical variables.

Feature engineering may be performed to extract relevant information.

- **Training Pipeline:**

The preprocessed data is split into training and testing sets. The Ridge Regression model (or any chosen regression algorithm) is trained on the training dataset.

Hyperparameter tuning and cross-validation are performed to optimize model performance.

- **Model Persistence:**

The trained model is saved or serialized for future use. This enables quick deployment without the need for retraining each time.

- **Deployment:**

A deployment environment is set up, such as a cloud-based server or a web application framework.

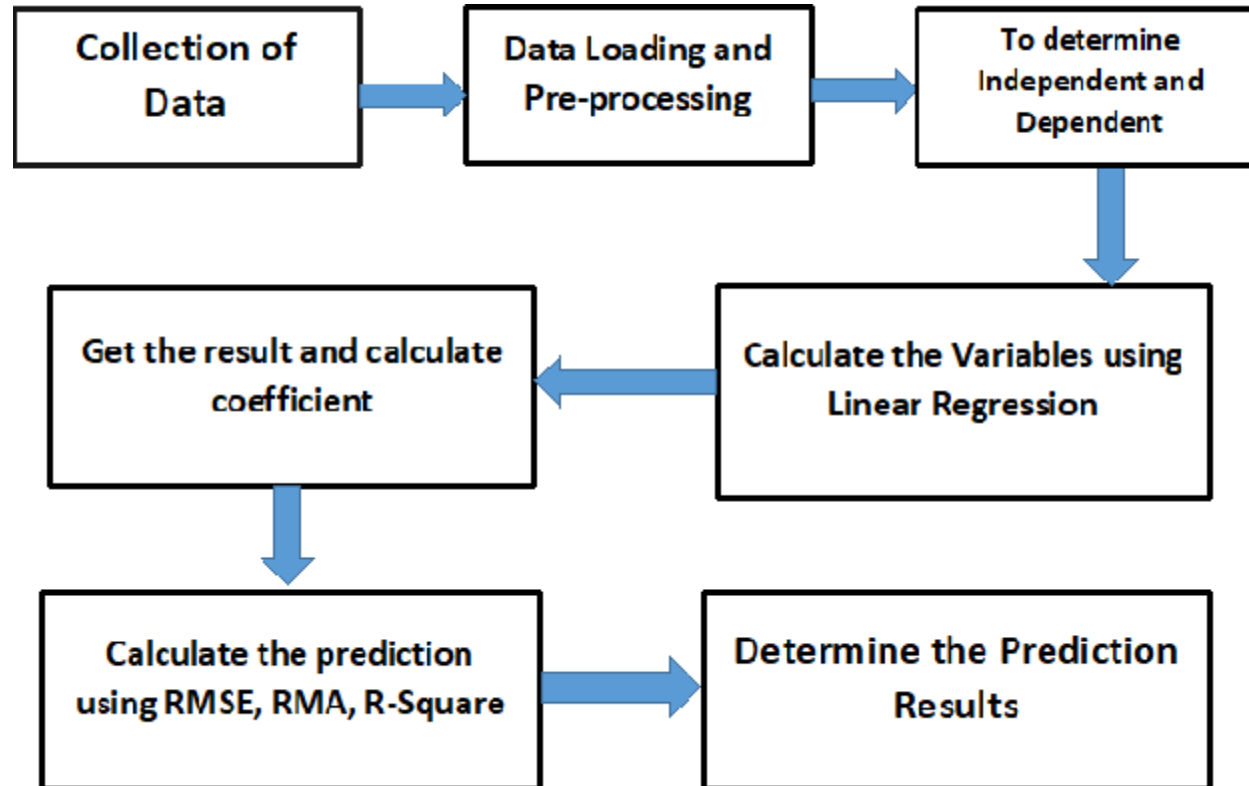
The Ridge Regression model is deployed, and APIs are created for model inference.

- **User Interface:**

A user-friendly interface (web or mobile application) allows users to input property details.



System Architecture



Algorithm & Deployment

- **Data Collection:**

Gather a comprehensive dataset with relevant features such as property size, location, amenities, and historical pricing.

- **Data Preprocessing:**

Handle missing values, outliers, and encode categorical variables. Explore feature engineering techniques to extract valuable information.

- **Split the Data:**

Divide the dataset into training and testing sets. Typically, 80-20 or 70-30 splits are common.

- **Train Ridge Regression Model:**

Implement Ridge Regression using a machine learning library like scikit-learn in Python. Tune hyperparameters, especially the regularization parameter (α), using cross-validation.

- **Model Evaluation:**

Assess the model's performance on the testing set using metrics like Mean Squared Error or R-squared.

- **Persist the Model:**

Save the trained Ridge Regression model to a file using serialization (e.g., joblib or pickle) for later deployment.

Conclusion

- This project has provided valuable insights into the complex world of real estate valuation.
- Through careful data collection, preprocessing, and model development, we have achieved significant progress in accurately predicting house prices.
- With several characteristics, the suggested method predicts the property price in Bangalore. We experimented with different Machine Learning algorithms to get the best model.
- When we give the property's square footage, the BHK, the number of bathrooms, and the location as input and then click 'Predict'. We can forecast the cost of what may be someone's ideal home.



Future Scope

- House price prediction can help the developer determine the selling price of a house and can help the customer to arrange the right time to purchase a house.
- In the future, the GUI can be made more attractive and interactive. It can also be turned into any real estate sale website where sellers can give the details and house for sale and buyers can contact according to the details given on the website
- To simplify it for the user, there can also be a recommending system to recommend real estate properties to the user based on the predicted price. The current dataset only includes a few locations of Bangalore city, expanding it to other cities and states of India is the future goal.
- To make the system even more informative and user-friendly, **Google maps** can also be included. This will show the neighborhood amenities such as hospitals, schools surrounding a region of 1 km from the given location. This can also be included in making predictions since the presence of such factors increases the price of real estate property.

Reference

- Hands-On Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems, Third Edition, October, 2022.
- Maharshi Modi, Ayush Sharma, Dr. P. Madhavan Applied Research on House Price Prediction Using Diverse Machine Learning Techniques, International Journal of Scientific & Technology Research Volume 9, Issue 04, April 2020.
- Dr. M. Thamarai, Dr. S P. Malarvizhi House Price Prediction Modeling Using Machine Learning, I.J. Information Engineering and Electronic Business, 2020, 2, 15-20.

GITHUB LINKS USED FOR REFERENCE:

- <https://github.com/Shreyas3108/house-price-prediction>
- <https://github.com/JangirSumit/BostonHousePricePrediction>

Thank you!