

HOUSE PRICE PREDICTION

Tejaswini Patil, Ruchira Ramteke, Harshada Kolhe, Sonali Mastud,
Pallavi Khedkar.



ruchiraramteke@gmail.com
sonalimastud99@gmail.com
pallavikhedkar07@gmail.com
harshada.kolhe1497@gmail.com

Department of Computer Engineering
Dr. D. Y. Patil College of Engineering Management and Research

ABSTRACT

House price forecasting is an important topic of real estate. The literature attempts to drive useful knowledge from historical data of property markets. Machine learning techniques are applied to analyze historical property transaction in India to discover useful models for house buyers and sellers. Revealed is the high discrepancy between house price in the most expensive and most affordable suburbs in the city. Moreover, experiment demonstrates that the multiple linear regression that is based on mean squared error measurement is a competitive approach. In this approach, with the help of random forest algorithm, we are going to predict House price. For this approach, we are going to consider different types of parameters which are area in square ft., Interior, Balcony and Construction year. With the help of machine learning techniques and Random Forest algorithm we are going to achieve our goal.

Keywords: ML, linear regression, Random Forest, House Price Prediction.

ARTICLE INFO

Article History

Received: 15th May 2022

Received in revised form :
15th May 2022

Accepted : 17th May 2022

Published online :

18th May 2022

I. INTRODUCTION

We all know that real estate property has always been a basic need for an individual. Due to this, every single organization in real estate business is trying to get an edge over other, which result in false price of the house, most of the time. As the population is also increasing day by day, so there arises a need to have an effective way to calculate or to predict the price of house. The buyers and the sellers, both are affected by the inconsistent and false prices, so the solution to this problem will give an idea about the price of a house on the basis of its features to both of them. Regression is best for prediction like these. Regression is a machine learning apparatus that encourages you to make expectations by taking in – from the current measurable information – the connections between your target parameter and a lot of different independent parameters. For the proposed model various regression algorithms like Multivariable Linear regression algorithm, Decision Tree regression and Random Forest regression are compared to get the most accurate results. Some machine learning methods Machine learning algorithms are often categorized as supervised or unsupervised.

Supervised machine learning algorithms can apply what has been learned in the past to new data using labeled examples to predict future events. Starting from the analysis of a

known training dataset, the learning algorithm produces an inferred function to make predictions about the output values. The system is able to provide targets for any new input after sufficient training[1].

In contrast, unsupervised machine learning algorithms are used when the information used to train is neither classified nor labeled. Unsupervised learning studies how systems can infer a function to describe a hidden structure from unlabeled data. The system doesn't figure out the right output, but it explores the data and can draw inferences from datasets to describe hidden structures from unlabeled data[1].

Semi-supervised machine learning algorithms fall somewhere in between supervised and unsupervised learning, since they use both labeled and unlabeled data for training – 11 typically a small amount of labeled data and a large amount of unlabeled data. The systems that use this method are able to considerably improve learning accuracy. Usually, semisupervised learning is chosen when the acquired labeled data requires skilled and relevant resources in order to train it / learn from it. Otherwise, acquiring unlabeled data generally doesn't require additional resources[1].

Reinforcement machine learning algorithms is a learning method that interacts with its environment by producing actions and discovers errors or rewards. Trial and error search and delayed reward are the most relevant characteristics of reinforcement learning. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal[1].

Motivation:

It helps to calculate price of house with the help of area, location and other features. To learn various machine learning algorithm.

Problem Definition:

Due to increase in urbanization, there is an increase in demand for renting houses and purchasing houses. Therefore, to determine a more effective way to calculate house price accurately is the need of the hour. So, an effort has been made to determine the most accurate way of predicting house price by using machine learning algorithms.

II. LITERATURE SURVEY

[1] Deep Learning Model for House Price Prediction Using Heterogeneous Data Analysis Along With Joint Self-Attention Mechanism 2021, To the best of our knowledge, Authors are the first research to incorporate attention mechanism and STN network to conduct house price prediction.

[2] Housing prices prediction with deep learning: an application for the real estate market in Taiwan 2020, CNN to predict housing prices. Experimental results show that CNN with housing features has the best prediction effect. This study can be used to develop targeted interventions aimed at the housing market.

[3] Prediction of House Pricing Using Machine Learning with Python, 2020 this paper contains what and how the house pricing model works with the help of machine learning and which dataset is used in our proposed model.

[4] Machine Learning based Predicting House Prices using Regression Techniques, 2020, The attempt is to construct a predictive model for evaluating the price based on factors that affects the price.

[5] A Gradient Boosting Method for Effective Prediction of Housing Prices in Complex Real Estate Systems, 2020, the experimental results show that the optimized GB (OGB) method can be used effectively for housing price prediction of real estate and achieves 0.01167 of the root mean square error; the lowest result compared to the other baseline machine learning models.

[6] House Price Prediction Using Regression Techniques: A Comparative Study, 2019, the motive of this paper is to help the seller to estimate the selling cost of a house perfectly and to help people to predict the exact time slap to accumulate a house. Some of the related factors that impact

the cost were also taken into considerations such as physical conditions, concept and location etc.

III. BLOCK DAIGRAMS

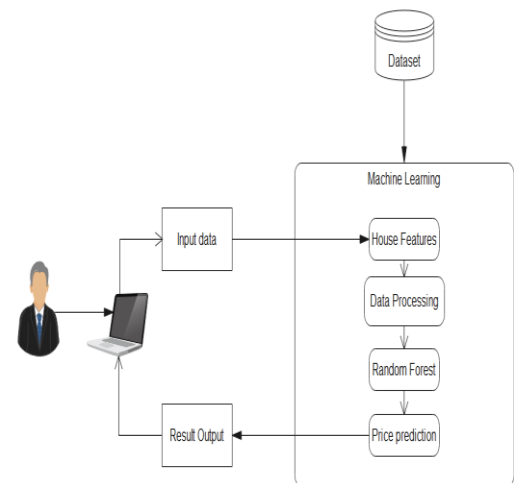


Fig 1. Block Diagram

Mathematical mode:

Let S be the Closed system defined as,

$$S = \{Ip, Op, A, Ss, Su, Fi\}$$

Where, Ip=Set of Input, Op=Set of Output, Su= Success State, Fi= Failure State and A= Set of actions, Ss= Set of user's states.

- Set of input=Ip={username, password, input text, dataset, random forest}
- Set of actions =A={F1,F2,F3,F4,F5,F6} Where,
F1= Register User
F2 = Authentcate User
F3 = Dataset Preprocessing
F4 = Random Forest Algorithm
F5= Classification
F6= Prediction
- Set of user's states=Ss={login state, input text, view records, view search results}
- Set of output=Op={House Price Prediction, Search results}
- Su=Success state={ Login Success, data process, house prediction, Search Results}
- Fi=Failure State={Invalid data, Login failed, Dataset read failure}
- Set of Exceptions= Ex ={NullPointerException, NullValues Exception, Exception }

IV. CONCLUSION

In this research paper, we have used machine learning algorithms to predict the house prices. We have mentioned the step by step procedure to analyze the dataset and finding the correlation between the parameters.

V. FUTURE SCOPE

For future work, we recommend that working on large dataset would yield a better and real picture about the model. We have undertaken only few Machine Learning algorithms that are actually classifiers but we need to train many other

classifiers and understand their predicting behavior for continuous values too. By improving the error values this research work can be useful for development of applications for various respective cities.

REFERENCE

- 1.] PEI-YING WANG¹, CHIAO-TING CHEN “Deep Learning Model for House Price Prediction Using Heterogeneous Data Analysis Along With Joint Self-Attention Mechanism” ACCESS.2021.3071306.
- 2.] Choujun Zhan, Zeqiong Wu, “Housing prices prediction with deep learning: an application for the real estate market in Taiwan” 2020 IEEE
- 3.] Mansi Jain, Himani Rajput, “Prediction of House Pricing Using Machine Learning with Python” 978-1-7281-4108-4/20/\$31.00 ©2020 IEEE.
- 4]. Manasa J “Machine Learning based Predicting House Prices using Regression Techniques” 978-1-7281-4167-1/20/\$31.00 ©2020 IEEE.
- 5]. Bandar Almaslukh, “A Gradient Boosting Method for Effective Prediction of Housing Prices in Complex Real Estate Systems” 978-1-6654-0380-1/20/\$31.00 ©2020 IEEE.