

REAL ESTATE PRICE PREDICTION

Yash Sheth¹, Sahil Morudkar², Palak Nayak³, Abhay Patil⁴

^{1,2,3}UG student, Dept of Information Technology, Mumbai University, Maharashtra, India

⁴Assistant Professor, Dept of Information Technology, Mumbai University, Maharashtra, India

Abstract -The traditional approach of the sales and marketing goals no longer help the companies to cope up with the pace of the competitive market, as they are carried out with no insights to customers purchasing patterns. Major transformations can be easily seen in the domain of sales and marketing as a result of Machine Learning advancements. Owing to such advancements, various critical aspects such as consumers' purchase patterns, target audience, and predicting sales for the recent years to come can be easily determined, thus helping the sales team in formulating plans for a boost in their business. The aim of this paper is to predict the price of Real Estates (Houses) in India using some different Machine Learning Algorithms and to see which one has the most accuracy. The buyers are just not concerned about the size(square feet) of the house and there are various other factors that play a key role to decide the price of a house/property. A comprehensive study of sales prediction is done using Machine Learning models such as Linear Regression, Decision Trees Regression, Gradient Boost & Random Forest Regressor.

Key Words: Random Forest Regressor, Regression, Algorithms, Real Estate, Price Prediction, Data mining, Machine Learning.

1. INTRODUCTION

Real Estate Property is not only the basic need of a man but today it also represents the richness and prestige of a person. Investment in the real estate generally seems to be profitable because their property values do not decline rapidly. Changes in real estate price can affect various household investors, bankers, policy makers and many. Sales forecasting has always been a very significant area to concentrate upon. Manual infestation of being able to predict House Prices could lead to drastic errors leading to poor management of the organization and most importantly would be time consuming, which is something not desirable in today's expedited world. A major part of the global economy relies upon the business sectors, which are expected to produce appropriate quantities of products to meet the overall needs. The forecasting process can be used for many purposes, including: predicting the future demand of the products or service and predicting how much of the product will be sold in a given period.

In our paper we have proposed the machine learning algorithms towards the data collected across various property aggregators across India. The objective here is to predict the price of Houses in India using three different algorithms and then comparing them to see which one gives a more accurate result based on some key features gathered from the raw data we have. Accurately predicting house prices can be a daunting task. Analysis and exploration of the collected data has also been done to gain a complete insight of the data. Analysis of the data would help the business organizations to make a probabilistic decision at each important stage of marketing strategy.

1.1 Problem Statement

Housing prices in any city are an important reflection of the economy, and housing price ranges are of great interest for both buyers and sellers. Ask any home buyer to describe their dream house, and they probably won't begin with the height of the basement ceiling or the proximity to an east-west railroad. But in reality there is much more that influences price negotiations than the number of bedrooms or a white-picket fence. Prices of real estate properties are sophisticatedly linked with our economy.

Despite all of this, we do not have accurate measures of housing prices based on the vast amount of data available. Simulation results show that the FLSR provides a superior prediction function as compared to ANN and FIS in capturing the functional relationship between dependent and independent variables and has the lowest computational complexity. Therefore, the goal of this project is to use machine learning algorithms to predict the selling prices of houses based on many economic factors.

2. LITERATURE SURVEY

- [1] Nihar Bhagat, Ankit Mohokar, Shreyash published in the International Journal of Computer Applications. This work aim towards the forecasting of house prices using Data Mining.
- [2] Sunitha Cheriyan, Shaniba Ibrahim, Saju Mohanan, Susan Treesa published in the year 2018. This study briefly analyzes the concept of sales data and sales forecast to predict the sales of any store using the previous data.

[3] Akshay Krishna, Akhilesh V, Animikh Aich, Chetana Hegde published in the year 2018. This paper has tried to predict the sales of a retail store using different machine learning techniques and tried to determine the best algorithm suited to a particular problem statement.

[4] Purvika Bajaj, Renesa Ray, Shivani Shedge, Shravani Vidhate published in the year 2020 in the International Research Journal of Engineering and Technology. This paper has aimed to propose a dimension for predicting the future sales of Big Mart Companies keeping in view the sales of previous years. A comprehensive study of sales prediction is done using Machine Learning models such as Linear Regression, K-Neighbors Regressor, XGBoost Regressor and Random Forest Regressor.

[5] Sekban, Judi published in the year 2019 on 'Applying machine learning algorithms in sales prediction.' This is the thesis in which various distinct procedures of machine learning algorithms are utilized to get better and optimal results, which are further examined for the prediction task. It has used of four algorithms, an ensemble technique etc. Feature selection has also been implemented using several different tactics..

[6] Panjwani, Mansi, Rahul Ramrakhiani, Hitesh Jumnani, Krishna Zanwar and Rupali Hande published in the year 2020. In this paper, the objective is to get proper results for predicting the future sales or demands of a firm by applying techniques like Clustering Models and measures for sales predictions. The potential of the algorithmic methods are estimated and accordingly used in further research.

- Anticipate the future costs by breaking down past market patterns and value ranges, and coming advancements.
- To make it user friendly and free of cost for the users.

4. METHODOLOGY

The methodology of predicting the price of real estate proceeds by obtaining & preprocessing the dataset, then applying various algorithms/regression techniques to find out the best suitable algorithm for the project.

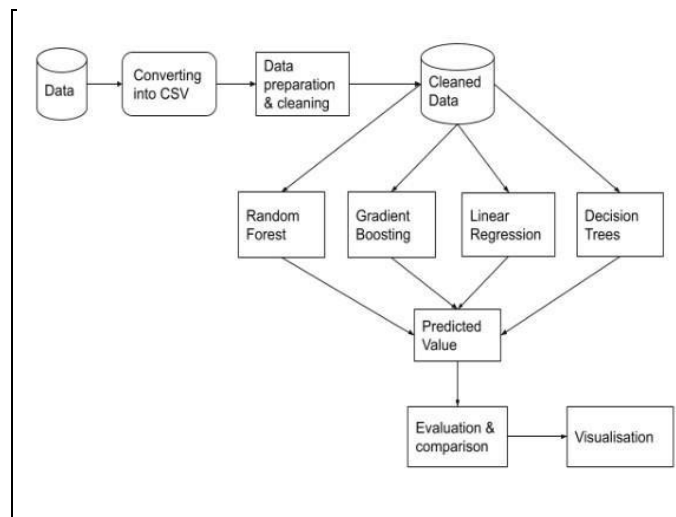


Fig-1: Block Diagram for Price Prediction System

3. PROPOSED SYSTEM

3.1 Aim and Objectives

The main aim of the project is to help people who plan to buy a house so they can know the price range in the future, then they can plan their finances well.

3.2 Objective

- To help all the users that are trying to find out the cost of any real estate property.
- To predict the market value of a real estate property.
- Help to find a starting price for a property based on the geographical variables.

4.1 Working

In this project, first the imported dataset is preprocessed. Data preprocessing includes converting the data into an understandable form, cleaning the data & detecting and eliminating any kind of outliers. Outliers are noisy data that they do have abnormal behavior comparing with the rest of the data in the same dataset. Now as the data is cleaned, the regression techniques that we are using i.e. Random Forest, Linear Regression, Gradient Boosting & Decision Trees are applied. We get an output from all the above tried algorithms. These outputs are then compared and the most accurate algorithm is finally selected and used for the project.

4.2 : Data Visualization

The correlation between target variable and other attributes is found by using this color encoded matrix from a library named Seaborn called Heat Map.

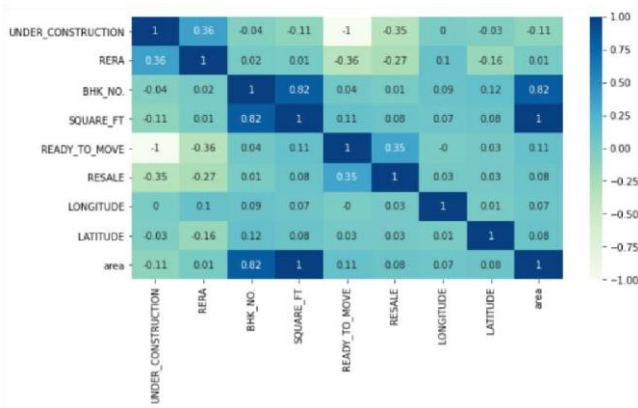


Fig-2: Heatmap for correlation between attributes

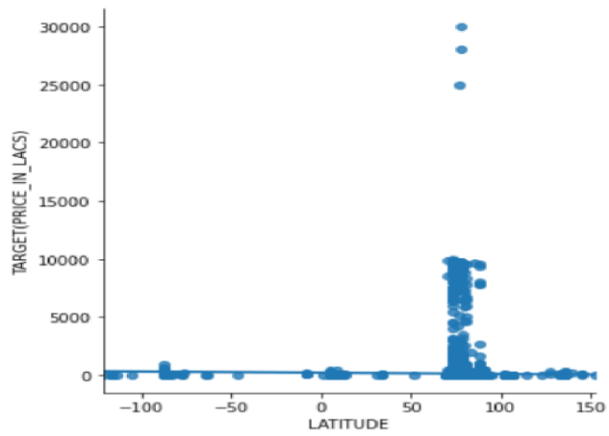


Fig-5: Line-plot (lmlot) of Latitude Parameter

4.3 : Graph Plots

Using the box-plot to visualize the distribution of data into quartiles. It shows us the minimum, maximum, median and the third quartile in data using the Seaborn Library.

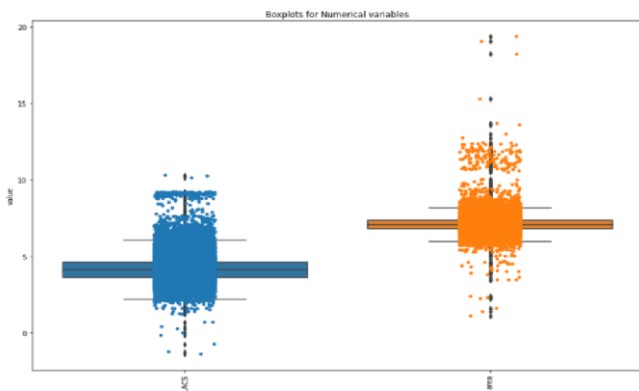


Fig-3 : Box-Plot

Line-plot also called as lmlots shows a line on 2 dimensional plane. Here we have used it to see the distribution of Longitude and Latitude Parameters used in our dataset.

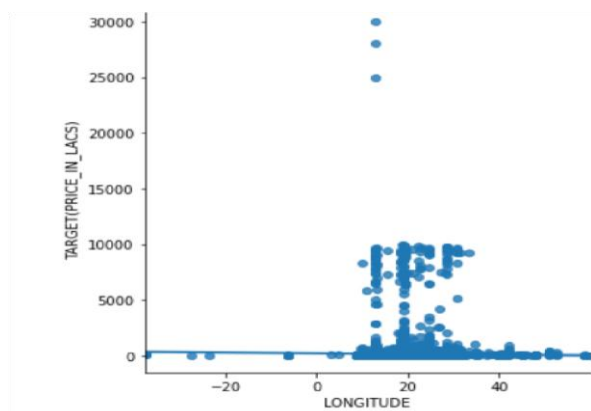


Fig-4: Line-plot (lmlot) of Longitude Parameter

5. EXPERIMENTAL RESULTS

The predicted output after entering all the inputs.

Fig-6 : Predicted Output

CONCLUSION

In this report, we have discussed a flexible solution for estimating the price of any real estate property on the go rather than the other techniques with less accurate solutions. We have also made a comparative study of various available technologies to solve and their feasibility

and advantages/disadvantages. Our project deals with developing a better and accurate prediction to overcome the shortcomings of the existing system that predicts the price of a real estate property.

ACKNOWLEDGEMENT

We would wish to thank our parents and friends for their support and encouragement throughout the progress of the project work and we would also like to thank our guide, Mr. Abhay Patil for guiding throughout the project work.

REFERENCE

- [1] Nihar Bhagat, Ankit Mohokar, Shreyash House Price Forecasting using Data Mining. International Journal of Computer Applications.
- [2] Sunitha Cheriyan,Shaniba Ibrahim,Saju Mohanan,Susan Treesa Intelligent Sales Prediction Using Machine Learning Techniques-cheriyan2018. [3] Akshay Krishna, Akhilesh V, Animikh Aich, Chetana Hegde Sales-forecasting of Retail Stores using Machine Learning Techniques 3rd IEEE International Conference on Computational Systems and Information Technology for Sustainable Solutions 2018.
- [3] Purvika Bajaj, Renesa Ray, Shivani Shedge, Shravani Vidhate, Prof. Dr. Nikhil Kumar Shardoor IRJET- SALES PREDICTION USING MACHINE LEARNING ALGORITHMS Volume: 07 Issue: 06 | June 2020.
- [4] Sekban, Judi. "Applying machine learning algorithms in sales prediction." (2019).
- [5] Panjwani, Mansi, Rahul Ramrakhiani, Hitesh Jumrani, Krishna Zanwar, and Rupali Hande. Sales Prediction System Using Machine Learning. No. 3243. EasyChair, 2020.
- [6] Steven C. Bourassa, Eva Cantoni, Martin Edward Ralph Hoesli,Spatial Dependence, Housing Submarkets and House Price Prediction The Journal of Real Estate Finance and Economics.
- [7] Ahmed Khalafallah, Neural Network Based Model for Predicting Housing Market Performance, Tsinghua Science & Technology 13(S1):325-328, October 2008.
- [8] Ahmad Abdulal,Nawar Aghi House Price Prediction Independent project (degree project), 15 credits, for the degree of Bachelor of Science (180 credits) with a major in Computer Science Spring Semester 2020.
- [9] ALISHA KUVALEKAR SHIVANI MANCHEWAR SIDHIKA MAHADIK HOUSE PRICE FORECASTING USING MACHINE LEARNING April 2020. [11] Ayush Varma,Abhijit Sarma,Rohini Nair,Sagar Doshi House Price Prediction Using Machine Learning and Neural Networks 2018 Second International
- [10]Conference on Inventive Communication and Computational Technologies (ICICCT). [12] Smith Dabreo, Shaleel Rodrigues, Valiant Rodrigues, Parshvi Shah Real Estate Price Prediction
- [11] IJERT Volume 10, Issue 04 (April 2021).
- [12] Aswin Sivam Ravikumar Real Estate Price Prediction Using Machine Learning December 2017.