A
Project Report
On
"House Price Prediction"

Prepared by
Maharshi Joshi (20DCS032)
Miral Joshi (20DCS033)
Anish Kabaria (20DCS034)

Under the guidance of

Prof. Parth Singh

A Report Submitted to

Charotar University of Science and Technology

for Partial Fulfillment of the Requirements for the

4th Semester Software Group Project-II (CE255)

Submitted at

CSE

DEPSTAR

At: Changa, Dist: Anand – 388421

April 2022

CERTIFICATE

This is to certify that the report entitled "**House Price Prediction**" is a bonafide work carried out by **Maharshi Joshi (20DCS032), Miral Joshi (20DCS033), Anish Kabaria (20DCS034)** under the guidance and supervision of **Prof. Parth Singh** for the subject CE255 - Software Group Project II (CSE) of 4th Semester of Bachelor of Technology in **DEPSTAR** at Faculty of Technology & Engineering – CHARUSAT, Gujarat.

To the best of my knowledge and belief, this work embodies the work of candidate himself, has duly been completed, and fulfills the requirement of the ordinance relating to the B.Tech. Degree of the University and is up to the standard in respect of content, presentation and language for being referred to the examiner.

Prof. Parth Singh

CSE

DEPSTAR, Changa, Gujarat.

Prof. Parth Goel

Head of Department Computer

Science and Engineering, DEPSTAR

CHARUSAT, Changa, Gujarat.

Dr. Amit P. Ganatra

Principal, DEPSTAR

Dean, FTE

CHARUSAT, Changa, Gujarat.

Devang Patel Institute of Advance Technology And Research At: Changa, Ta. Petlad, Dist. Anand, PIN: 388 421. Gujarat

ABSTRACT

The real estate market is a standout amongst the most focused regarding pricing and keeps fluctuating. It is one of the prime fields to apply the ideas of machine learning on how to enhance and foresee the costs with high accuracy. The objective of the project is the prediction of the market value of a real estate property. This system helps find a starting price for a property based on the geographical variables. By breaking down past market patterns and value ranges, and coming advancements future costs will be anticipated. This examination means to predict house prices in Metro cities with Decision tree regressor. It will help clients to put resources into a bequest without moving towards a broker. The result of this project proved that the Decision tree regressor gives an accuracy of 99%.

ACKNOWLEDGEMENT

We, the developer of "House Price Prediction System", with immense pleasure and commitment would like to present the project assignment. The development of this project has given us wide opportunity to think, implement and interact with various aspects of management skills as well as the new emerging technologies.

Every work that one completes successfully stands on the constant encouragement, good will and support of the people around. I hereby avail this opportunity to express our gratitude to number of people who extended their valuable time, full support and cooperation in developing the project.

I express deep sense of gratitude towards our Head of the Institute, Dr. Amit Ganatra , our Head of the Department, Prof. Parth Goel and project guide Prof. Parth Singh for the support during the whole session of study and development. It is because of them, that I was prompted to do hard work, adopting new technologies.

They altogether provided me favourable environment, and without them it would not have been possible to achieve my goal.

Thanks,
Maharshi Joshi
Miral Joshi
Anish Kabaria

DECLARATION BY THE CANDIDATE

I hereby declare that the project report entitled "House Price Prediction" submitted by me to Devang Patel Institute of Advanced Technology an Research, Changa in partial fulfilment of the requirement for the award of the degree of B.Tech in Computer Science and Engineering, from Department of Computer Science and Engineering, DEPSTAR/FTE, is a record of bonafide CE255 Software group Project (project work) carried out by me under the guidance of Prof. Parth Singh. I further declare that the work carried out and documented in this project report has not been submitted anywhere else either in part or in full and it is the original work, for the award of any other degree or diploma in this institute or any other institute or university.

Maharshi Joshi (20DCS032)

Miral Joshi (20DCS033)

Anish Kabaria(20DCS034)

Prof. Parth Singh

Teaching and Research Assistant

Department of Computer Science and Engineering,

DEPSTAR , CHARUSAT-CHANGA

TABLE OF CONTENTS

CHAPTER 1: INTRODUCTION

1.1 PROJECT OVERVIEW

The real estate sector is an important industry with many stakeholders ranging from regulatory bodies to private companies and investors. Among these stakeholders, there is a high demand for a better understanding of the industry operational mechanism and driving factors. Today there is a large amount of data available on relevant statistics as well as on additional contextual factors, and it is natural to try to make use of these in order to improve our understanding of the industry. Notably, this has been done on Kaggle's train.csv dataset on housing prices. This project can be considered as a further step towards more evidence-based decision making for the benefit of these stakeholders.

1.2 AIM AND OBJECTIVES

The aim is to predict the efficient house pricing for real estate customers with respect to their budgets and priorities. By analyzing previous market trends and price ranges, and also upcoming developments future prices will be predicted. The functioning involves a website which accepts customers specifications and then combines the application of Decision Tree algorithm of machine learning. This application will help customers to invest in an estate without approaching an agent. It also decreases the risk involved in the transaction. The current property buying or selling is hectic and expensive. As the customer has to roam places and has to pay commission to the Real estate agent. Also, the customer/buyer does not know whether the property is profitable in future or not. Hence, we design a website using data mining techniques to overcome the drawbacks of current system as everything is web based. We are implementing following :

1) User and device friendly web page

2) Location based search

3) Future estimate of property

1.3 FUTURE SCOPE

This research study looks into how machine learning models may be used to predict the price of a home using housing price data sets. Constructing our own regressors for the Decision tree regressor is outside the scope of the project. As a result, a library for machine learning models was utilised to concentrate on feature engineering and investigate how hyperparameters affect predictions. Another feature to include in this project is the use of Google Maps API, which eliminates the need for the user to manually type the address.

1.4 TOOLS & TECHNOLOGY USED

I. Jupyter Notebook
II. Visual Studio Code
III. Sublime Text3
IV. Python libraries
- Pandas
- Numpy
- Matplotlib
- Seaborn
- Sklearn
- Joblib
- Pickle
V. Flask Framework
VI. Html, Css and Bootstrap

CHAPTER 2: PROJECT MANAGEMENT

2.1 PROJECT PLANNING

2.1.1 Project Development Approach
This project consists of five different phases: initiation, planning, execution, monitoring and controlling, and closure.

Fig 2.1 Planning Model

2.2 PROJECT WORK SCHEDULING

Gantt Chart:

Fig 2.2 Gantt Chart

CHAPTER 3: SYSTEM REQUIREMENTS STUDY

3.1 USER CHARACTERISTICS

- End Users

  This product is designed mainly for END USERS. So, it is feasible for the end-usersto directly enter the features to web page and predict price.
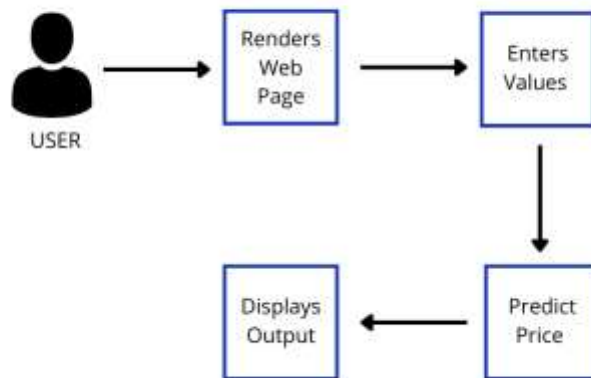
3.1.1 Use Case Diagram

Fig 3.1. Use Case Diagram

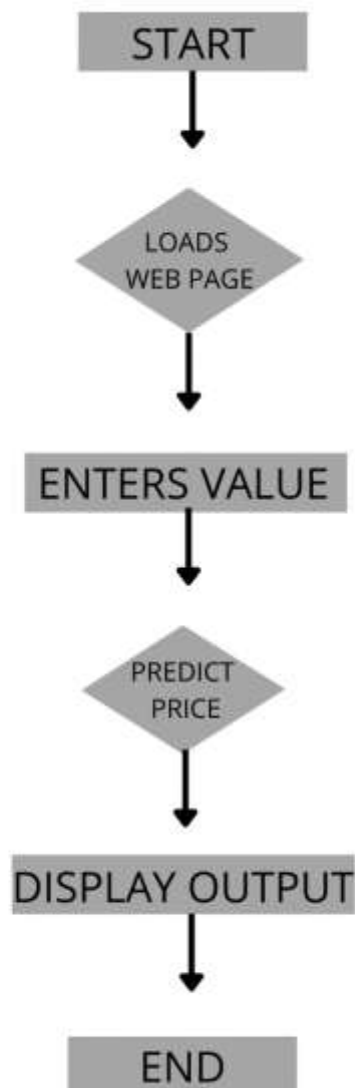## 3.2 HARDWARE AND SOFTWARE REQUIREMENTS

- Windows 7 or above:
- 2.5 GHz processor and 4 GB RAM
- Android Devices :
  - Memory 1-2GB Ram (dependent on model size)
  - Display between 4.7 and 6 inches

CHAPTER 4: SYSTEM FUNCTIONALITY

4.1 MAJOR FUNCTIONALITY

It provides a platform to predict price without creating an account. It asks for values for respective feature.

4.1 SYSTEM FLOW CHART

CHAPTER 5: SYSTEM ANALYSIS

5.1 STUDY OF PROPOSED SOLUTION

Nowadays, e-education and e-learning is highly influenced. Everything is shifting from manual to automated systems. The objective of this project is to predict the house prices so as to minimize the problems faced by the customer. The present method is that the customer approaches a real estate agent to manage his/her investments and suggest suitable estates for his investments. But this method is risky as the agent might predict wrong estates and thus leading to loss of the customers' investments. The manual method which is currently used in the market is out dated and has high risk. So as to overcome this fault, there is a need for an updated and automated system. Machine Learning algorithms can be used to help investors to invest in an appropriate estate according to their mentioned requirements. Also the new system will be cost and time efficient. This will have simple operations. The proposed system works on decision tree regressor.

5.2 REQUIREMENTS OF PROPOSED SYSTEM

5.2.1 Dataset

We have used dataset from Kaggle's Competitons named train.csv that stores data of metro cities. There are 11 features excluding the target and 12452 rows.

Features are:

1. POSTED_BY
2. UNDER_CONSTRUCTION
3. RERA
4. BHK_NO
5. BHK_OR_RK
6. SQUARE_FEET
7. READY_TO_MOVE
8. RESALE
9. CITY
10. LONGITUDE
11. LATITUDE
12. TARGET

5.2.2 Preprocessing Data

For preprocessing the data we have done following operations.

    a. Cleaning Data
    b. Normalizing Data
    c. Imputing Missing values
    d. Prediction and evaluation
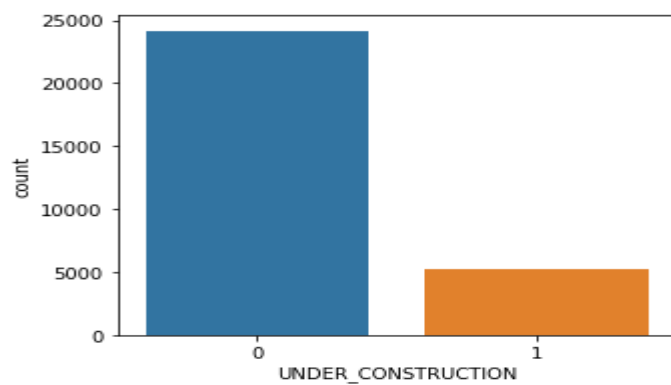   e. Spliting the data
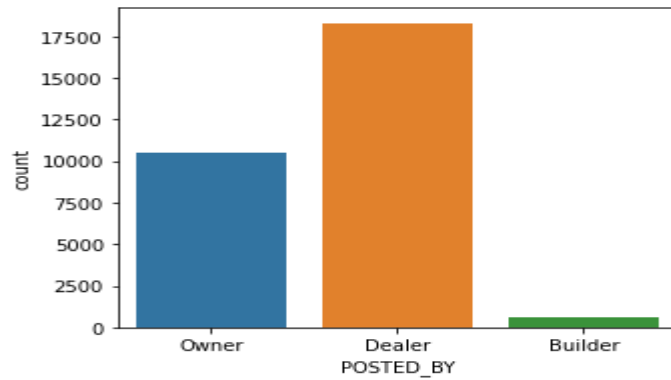
5.2.3 Decision Tree Regressor

Decision tree builds regression or classification models in the form of a tree structure. It breaks down a dataset into smaller and smaller subsets while at the same time an associated decision tree is incrementally developed. The final result is a tree with decision nodes and leaf nodes. A decision node (e.g., Outlook) has two or more branches (e.g., Sunny, Overcast and Rainy), each representing values for the attribute tested. Leaf node (e.g., Hours Played) represents a decision on the numerical target. The topmost decision node in a tree which corresponds to the best predictor called root node. Decision trees can handle both categorical and numerical data.
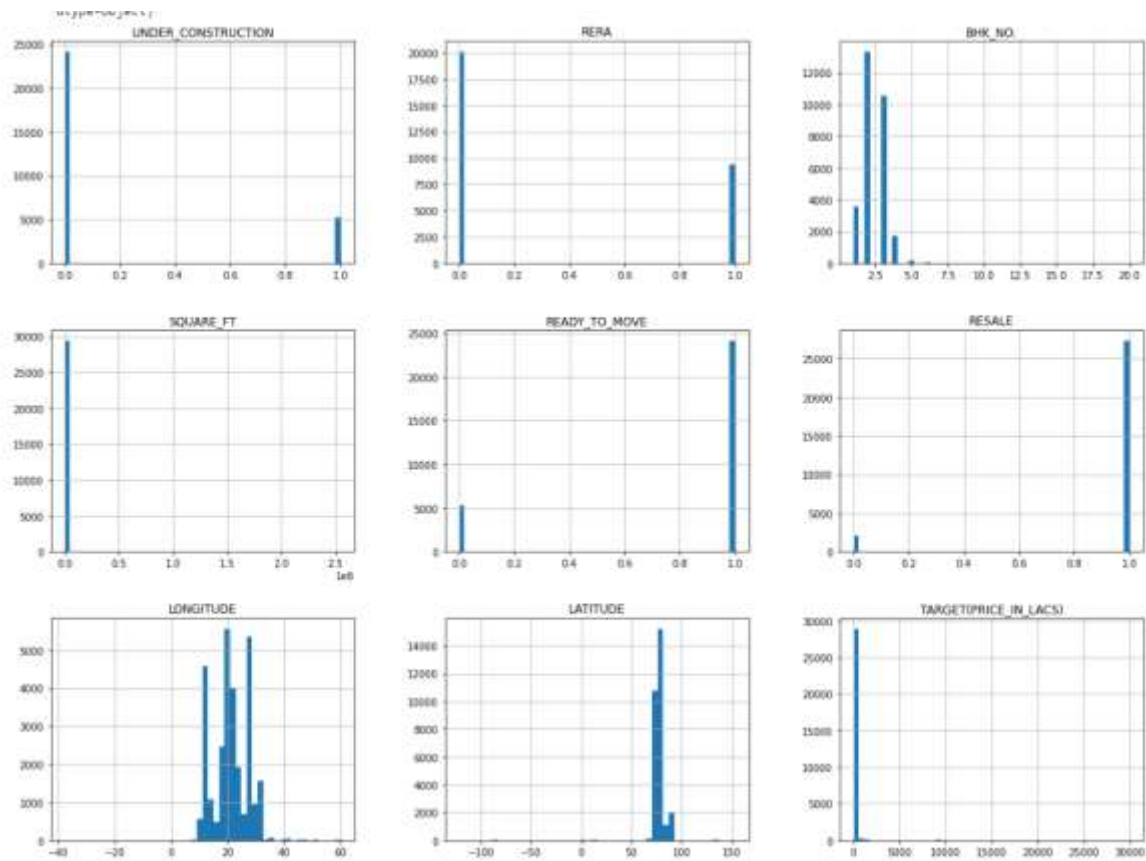
CHAPTER 6: SYSTEM IMPLEMENTATION

6.1 IMPLEMENTATION SCREENSHOTS:

```
housing.head()
```

	POSTED_BY	UNDER_CONSTRUCTION	RERA	BHK_NO.	BHK_OR_RK	SQUARE_FT	READY_TO_MOVE	RESALE	ADDRESS	LONGITUDE	LATITUDE	TARGET(PRICE_IN_LACS)
0	Owner	0	0	2	BHK	1300.236407	1	1	Ksfc Layout,Bangalore	12.969910	77.597960	55.0
1	Dealer	0	0	2	BHK	1275.000000	1	1	Vishweshwara Nagar,Mysore	12.274538	76.644605	51.0
2	Owner	0	0	2	BHK	933.159722	1	1	Jigani,Bangalore	12.778033	77.632191	43.0
3	Owner	0	1	2	BHK	929.921143	1	1	Sector-1 Vaishali,Ghaziabad	28.642300	77.344500	62.5
4	Dealer	1	0	2	BHK	999.009247	0	1	New Town,Kolkata	22.592200	88.484911	60.5

```
housing['LONGITUDE'] = np.log(housing['LONGITUDE'])
```

CHAPTER 7: CONCLUSION

7.1 SELF ANALYSIS OF PROJECT VIABILITIES

According to me, this project is completed with the primary functionalities asspecified earlier, but then again there is a lot more than this which can be done. So, then it is a challenge to further develop it in to full-fledged software as it was challenge to develop up to this very stage. Due to lack of skilled knowledge, the project cannot be fully completed so far. This technology will help a lot in the area of Machine Learning.

7.2 Summary

 The research question for this study is to study how well house prices can be predicted by using Decision tree regressor and Random forest regression. In this study we have found that the Decision tree regressor algorithm performs better at predicting house prices than the Random forest. However, there is still a difference between the actual prices in our testing data and the prices predicted by Decision tree regressor algorithm.

BIBLIOGRAPHY:

1. Wikipedia , Javapoint , Youtube. Random forest.

2. Wikipedia , scikit-learn . Decision Tree

3. HTML & CSS
   https://youtube.com/playlist?list=PL4cUxeGkcC9ivBf_eKCPIAYXWzLlPAm6G

4. https://www.geeksforgeeks.org

5. https://codewithharry.com/

6. https://docs.pyhton.org/3/

7. https://www.w3school.com/

8. https://flask.palletsprojects.com/en/2.1.x/