SEMINAR REPORT

ON

House Price Prediction

Name	Roll No.
Pranjali Khairnar	3301029
Apurva Mulik	3301038
Siddhi Phalle	3301054

Under the guidance

of

Ms. Seema Mane

DEPARTMENT OF INFORMATION TECHNOLOGY

Sinhgad Institute of Technology and Science, Narhe, Pune

Savitribai Phule Pune University

[Academic Year: 2022-2023]

Sinhgad Institutes

Department of Information Technology,Sinhgad Institute of Technology and Science, Narhe, Pune

CERTIFICATE

This is to certify that **Pranjali Khairnar** from Third Year having Roll Number**: 3301029 of Information Technology** has successfully completed his / her seminar on "House Price Prediction**"** at Sinhgad Institute of Technology and Science, Narhe in the partial fulfilment for the degree in Bachelor of Engineering.

Ms. Seema Mane	Prof. T.D. Khadtare	Dr. S.D. Markande
Guide	**Head of the Department**	**Principal**

ACKNOWLEDGEMENT

I take this opportunity to thank my internal guide **Ms. Seema Mane**, for giving me guidance and support throughout the Seminar. Her/his valuable guidelines and suggestions were very helpful.

I wish to express my thanks to **Prof. T.D. Khadtare,** Head of **Information Technology**, Sinhgad Institute of Technology and Science, Narhe for giving me all the help and important suggestions all over the Seminar Work. I thank all the teaching and non-teaching staff members, for their indispensable support and priceless suggestions.

I also thank my friends and family for their help in collecting data without which this Seminar report not have been Completed. At the end, my special thanks to **Dr. S.D. Markande,** Principal, Sinhgad Institute of Technology and Science, Narhe for providing ambience in the college, which motivate us to work.

ABSTRACT

This paper provides an overview about how to predict house costs utilizing different regression methods with the assistance of python libraries. The proposed technique considered the more refined aspects used for the calculation of house price and provide the more accurate prediction. It also provides a brief about various graphical and numerical techniques which will be required to predict the price of a house. 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.

CONTENTS

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

House/Home are a basic necessity for a person and their prices vary from location to location based on the facilities available like parking space, locality, etc. The house pricing is a point that worries a ton of residents whether rich or white collar class as one can never judge or gauge the valuing of a house based on area or offices accessible. Buying of a house is one of the greatest and significant choice of a family as it expends the entirety of their investment funds and now and again covers them under loans. It is the difficult task to predict the accurate valuesof house pricing. Our proposed model would make it possible to predict the exact prices of houses. 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. There are three factors that influence the price of a house which include physical conditions, concept and location. The value of a home is well known as a combination of a large variety of options. Therefore, the prediction of home price presents a novel set of challenges. Though an oversized variety ofspace units are dedicated to the present task, their performance and applications are restricted by a really long delay within the handling of data, the dearth of real-world settings and so the inadequacy of choices for housing. Our aim is to predict the value of the house marketing mistreatment classification techniques during this paper. Most of the present studies have concentrated on breakdown the distraction of the prediction of house costs. Several theories are born as an outcome of the analysis work committed by completely different researchers around the world. This paper picks up the most recent prediction analysis to assist economic predictors to use it. It provides a summary of the prediction markets and, together, the presentmarkets that build it easier to predict the market.

CHAPTER 2

LITERATURE REVIEW

1.1 Introduction

There are a couple of components that impact house costs. In this exploration, partition these components into three essential get-togethers, there are state of being, thought and territory. States of being are properties constrained by a house that can be seen by human recognizes, including the range of the house, the amount of rooms, the availability of kitchen and parking space, the openness of the yard nursery, the zone of land and structures, and the age of the house, while the thought is an idea offered by architects who can pull in potential buyers, for instance, the possibility of a moderate home, strong and green condition, and world class condition. Zone is a critical factor in shaping the expense of a house. This is in light of the fact that the zone chooses the normal land cost. Besides, the territory furthermore chooses the basic passage to open workplaces, for instance, schools, grounds, crisis facilities and prosperity centres, similarly as family preoccupation workplaces, for instance, strip malls, culinary visits, or much offer awesome landscape. Firstly, data collection is performed. Then data cleaning is performed to remove all the errors from the data and make it clean. Then data pre-processing is done. Then with the help of data visualization, different plots are created, which intends to depict the distribution of data in different forms. Towards the end, the business costs of the houses were determined with exactness and accuracy.

1.2 Summary of Literature Review

Sr.No.	Title of Paper	Conclusion
1.	Prediction of House Pricing Using Machine Learning with Python. Authors :- Mansi Jain, Himani Rajput, Neha Garg, Pronika Chawla Published :- July 2020	This paper provides an overview about how to predict house costs utilizing different regression methods. To improve the accuracies of the various regression algorithms that are applied on our house pricing dataset so that they would provide better results.

CHAPTER 3

PROBLEM STATEMENT AND QUESTIONS

1.3 Problem Statement

To predict house cost utilizing different regression methods with the assistance of python libraries.

1.4 Objective

1. To predict house prices and to get better and accurate results.
2. To the betterment of people this is helpful because the house pricing is a topic that concerns a lot of citizens whether rich or middle class.

1.5 Methodology

Linear Regression:

It is an algorithm of supervised machine learning in which the predicted output is continuous with having a constant slope. It is used to predict the values in a continuous range instead of classifying the values in the categories. Linear regression is used for performing different tasks like house price prediction. If you also want to know the method to find out house price prediction, then read our guide completely. This guide considers and explains every factor to perform the house price prediction in linear regression easily. To identify the variables affecting house prices, e.g. area, number of rooms, bathrooms, etc. To create a linear model that quantitatively relates house prices.

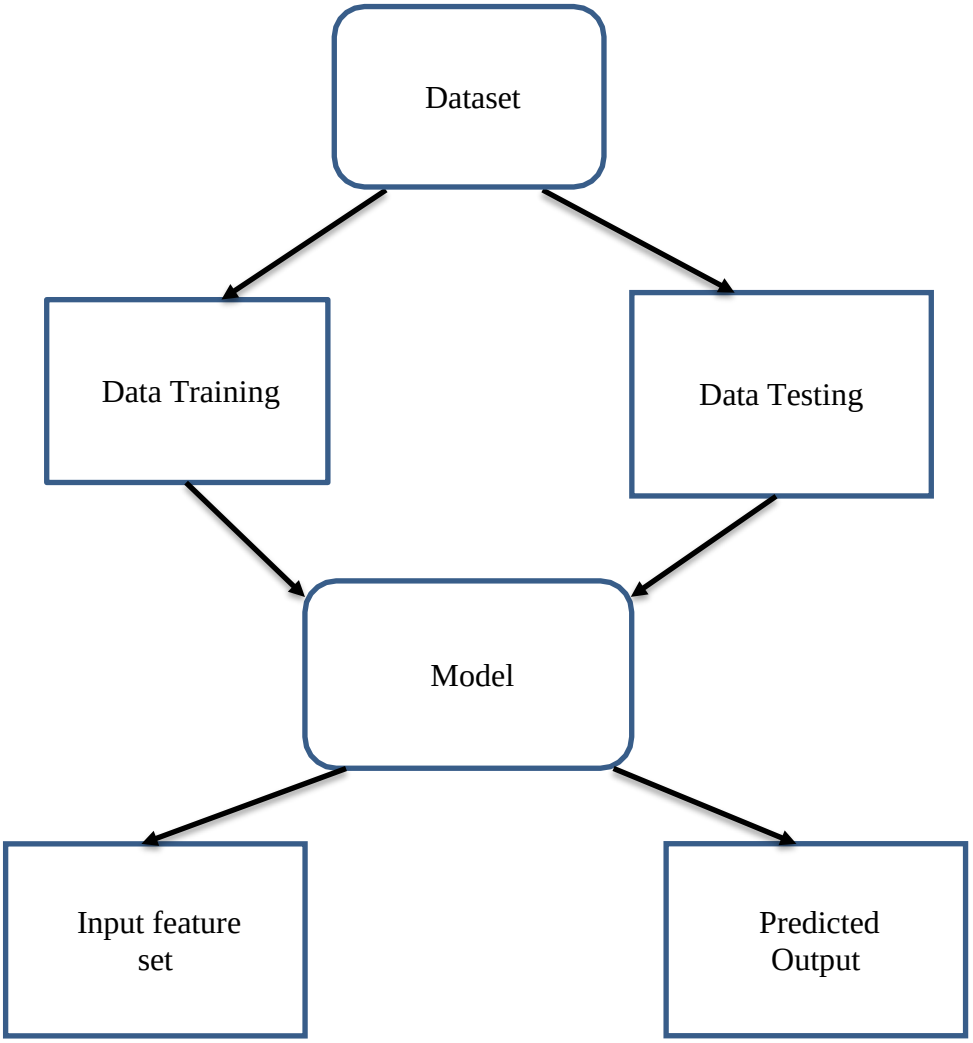


Fig. 3.1 Block diagram

Data Collection:

Data collection is the process of gathering information on variables in a systematic manner. This helps in finding answers too many questions, hypothesis and evaluate outcomes. Data Collection: There are 10 attributes and contains 20,640 instances. and it was downloaded from kaggle.com.

Data Visualization:

Data Visualization is the pictorial or graphical representation of information. It enables to grasp difficult concepts or identify new patterns. Data Visualization is seen by numerous orders as a cutting edge likeness visual correspondence. It includes the

creation and investigation of the visual portrayal of information. To impart data plainly and effectively, information representation utilizes measurable illustrations, plots, data designs and different apparatuses.

Data Pre-Processing:

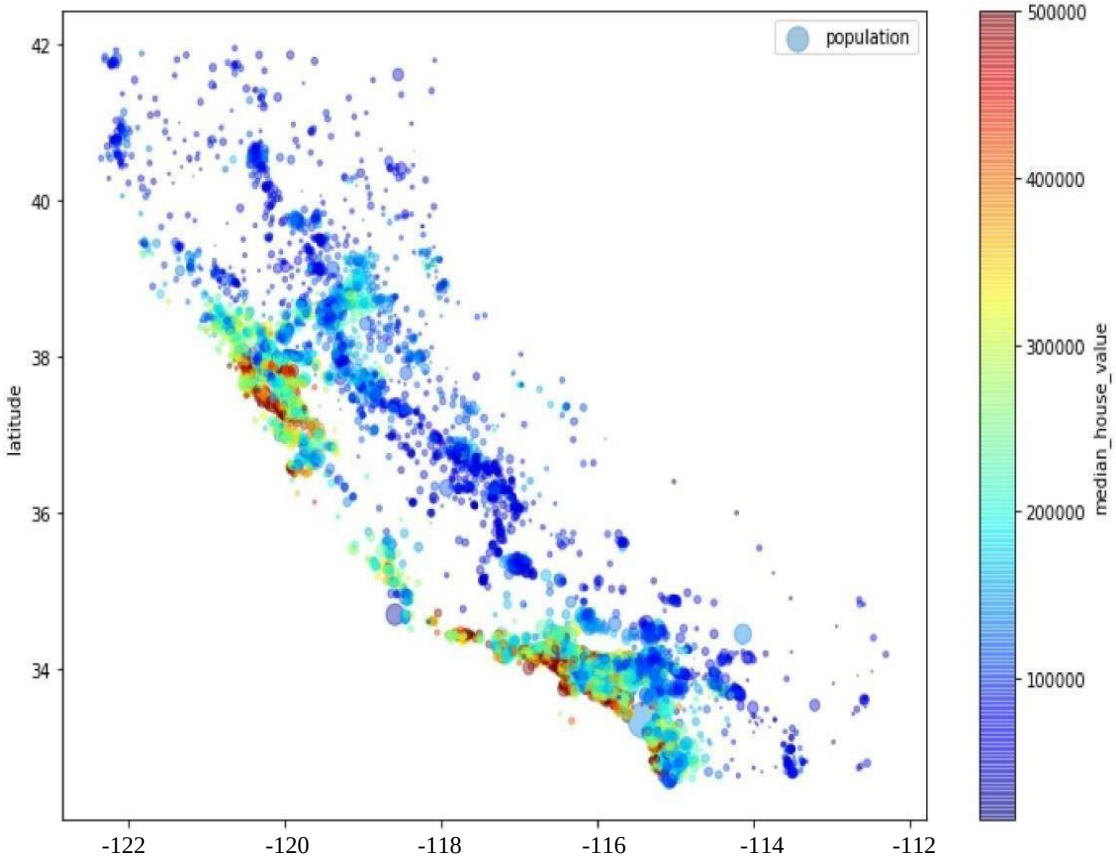
It is the process of transforming data before feeding it into the algorithm. It is utilized to change over crude information into a clean dataset. Data pre-processing is an information mining procedure which is utilized to change the crude information in a helpful and productive format. In any Machine Learning procedure, Data Pre-processing is that progression wherein the information gets changed, or Encoded, to carry it to such an express, that now the machine can without much of a stretch parse it. Data pre-processing is required errands for cleaning the information and making it appropriate for a Machine Learning model which likewise expands the precision and proficiency of a Machine Learning model.

Data Cleaning:

Data cleaning is the process of detecting and removing errors to increase the value of data. Data cleaning is carried out with the help of data wrangling tools. It is the way toward identifying and amending off base records from a record set, table or database. It finds the deficient information and replaces the messy information. The information is changed to ensure it is exact and right. Information cleaning is the way toward distinguishing and revising mistaken records from a record set, table or database.

CHAPTER 4

RESULT



The graph shows house prices in specific area where red is expensive, blue is cheap, larger circles indicate areas with a larger population. Various factors which affects the house pricing are considered and further worked upon them. Machine learning has been considered to complete out the desired task.

CHAPTER 5

FUTURE SCOPE

In future, many more algorithms can be applied on this dataset such as decision tree, Naïve Bayes, SVM etc. and find out their respective accuracies and use them to predict a better outcome and hence increase the accuracy. The KNN algorithm can also be applied to predict the accuracy. The k-means algorithm can also be applied. With the help of these algorithms, the house prices are accurately predicted. Hence, it would be of great help for the government and the people themselves. Artificial intelligence can also be applied to make our project more enhanced in the future. Sometimes people also prefer to stay near areas where basic facilities are easily available such as a general store, mother dairy, photocopy shop etc. This is also an extremely important factor that may affect the prices of houses and can be taken into consideration in the future. All the major factors that can affect the prices of houses in a particular area are almost covered and have worked upon them. In the future the model deployment of more algorithms can be performed to achieve accurate results.

CHAPTER 6

SUMMARY

The sales price for the houses are calculated using different algorithms. The sales prices have been calculated with better accuracy and precision. This would be of great help for the people. To achieve these results, various data mining techniques are utilized in python language. The various factors which affect the house pricing should be considered and work upon them. Machine learning has assisted to complete out task. Firstly, the data collection is performed. Then data cleaning is carried out to remove all the errors from the data and make it clean. Then the data pre-processing is done. Then with help of data visualization, different plots are created. This has depicted the distribution of data in different forms. Further, the preparation and testing of the model are performed. It is extremely important to improve the accuracy and precision of the algorithms in order to achieve better results. If the results are not accurate then they would be of no help to the people in predicting the sales prices of houses. It also made use of data visualization to achieve better accuracy and results. The sales price is calculated for the houses using different algorithms. The sales prices have been calculated with better accuracy and precision. This would be of great help for the people.

CHAPTER 7

REFERENCES

1. Jain, M., Rajput, H., Garg, N., & Chawla, P. (2020). Prediction of House Pricing usingMachine Learning with Python. 2020 International Conference on Electronics and Sustainable Communication Systems(ICESC). doi:10.1109/icesc48915.2020.9155839
2. https://raw.githubusercontent.com/ageron/handson-ml/master/datasets/housing/housing.csv

APPENDIX A

Seminar Log Record

Sr.No	Date	Discussion/Suggestion	Remark by guide
1	7/8/2022	Discussion about the overall subject. Selection a topic any 4 topics for the seminar. Selection to appropriate topic	
2	10/8/2022	Discussion about the first review ppt and contents included in it.	
3	17/8/2022	Submission of synopsis	
4	19/8/2022	Review I	
5	7/9/2022	Discussion about Review II and what should be included in presentation.	
6	19/9/2022	Review II	
7	15/10/2022	Discussion about the seminar report	
8	9/11/2022	Submission seminar report.	