Real Estate Price Prediction

Gudi sai Bharat

Syed Nakhi Ali

Sasi kumar, Kare

Abstract

In recent years, picture detection, normal speech command, email spam reorganisation, and medical diagnosis have all become more essential applications of machine learning. Machine learning algorithms are now used to help us with security alerts, public safety, and medicinal advancements. Machine learning allows us to give better customer service and create a safer vehicle system.

This study looks at how machine learning algorithms may be used to estimate real estate and housing values using datasets. This study looks at how machine learning algorithms are applied to datasets and how these predictions are applied to real-world user behaviour. This algorithm asserts that a faulty dataset can lead to negative predictions, as well as demonstrating which solution is optimal for this task.

Keywords:-Machine learning, Real estate , Price Prediction, House Price Algorithm.

INTRODUCTION

Machine learning is becoming increasingly important in today's industry and research.It uses algorithms and neural network models to continuously improve the performance of a computer system. It automatically creates a mathematical model based on sample data, commonly known as "training data," that will make decisions without being explicitly taught to do so

People and real estate companies both buy and sell properties these days; people buy houses to live in or as an investment, while agencies buy houses to run a business. On the other hand, we believe that everyone should receive exactly what they pay for. Due to a lack of adequate detection measures, under- and over-valuation in the property market is a major problem. House/real estate price to rent ratios are broad measurements. However, an in-depth analysis and judgement are required to understand this topic. That's why we're employing machine learning methods to find a good solution by training an ML model with thousands of data points. Which will be powerful enough to accurately estimate real estate values and will cater to everyone's needs?

The major goal of this research study is to use this machine learning technique to ML models that can be valuable to users. A buyer's main goal is to learn about their dream house that meets all of their needs. More purchasers are looking for a good deal, but there's no guarantee that they'll get a good deal on a product that isn't overpriced. In the same way, a seller seeks for a specific number that they can put on the real estate as a price tag, and this cannot be a guess; much researchis required to arrive at a valuation of a home.

LITERATURE REVIEW

Nowadays, real estate has evolved into much more than a necessity; it now represents much more, not only for those who purchase real estate, but also for those who sell it. Property, according toReal estate is not only a human's basic need, but it also represents a person's riches and status. Property values do not drop

rapidly, therefore investing in real estate appears to be advantageous in general. Real estate price variations, on the other hand, may have an impact on a wide range of investors, bankers, policymakers, and others. Investing in real estate may appear to be a lucrative financial alternative. Consequently, projecting real estate values is an essential economic indicator. It implies that today's real estate organisations are all working hard to gain a competitive advantage over their competitors. In order for humans to provide a desirable output, the process must be minimised. To design an algorithm that forecasts housing real estate prices based on certain input features utilising machine learning and artificial intelligence techniques. The business application of this algorithm is that classified web sites can use it to determine the prices of new properties that are about to be listed by taking some input variables and predicting the correct and justified price, ie; to take away by taking prize inputs from customers and thus not allowing any errors in the system.By using google colab \ jupyter IDE \ spider(anaconda).Jupyter IDE, on the other hand, is a free open-source tool that lets us share and create papers that include live ports, graphs, equations, and narrative text. It has data translation, data cleaning, numerical value simulation, and statistical modelling tools.To create a system, data visualisation and machine learning tools are used. This enables consumers to get a close estimate of the price of real estate. The user can specify the criteria by which they will be awarded awards for dream homes. The user can also acquire an example house plan to use as a reference for houses.Support vector machine and partial least square algorithms, as well as the corresponding features, are used to analyse and forecast the housing value of Bangalore. After the missing samples from the original data set have been removed, some samples are handled as training data and others as test data.

Methodologies

There has been some research in the real estate house price prediction field, but none of it has resulted in any practical solutions. There is a scarcity of data for anticipating house prices, which are determined by the firm. There are very few digital solutions available these days for such a broad market, and most consumers and businesses rely on the following ways.
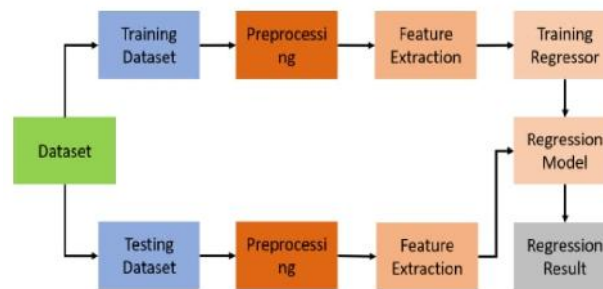

Buyers customer :-

1) When consumers first consider purchasing real estate, they are willing to go online and research trends and other information. They can hunt for a home that has everything they require, and by doing so, they can keep track of the prices associated with these homes. However, the average individual lacks thorough understanding and correct information about what the real price is, which could lead to inaccurate information if they assume the prices on the internet are accurate.

2) Another thought that comes to mind when looking for a home is to contact several real estate agents. The issue with agents is

that they are paid a fee just to look for a house and establish a price for us. People, in most circumstances, blindly believe these price tags because they have no other option. It's possible that there are cases involving hidden dealings with brokers and sellers. There could be an overprice without his or her knowledge.

Seller \Agencies:-

- When a person considers selling his or her home, they compare it to hundreds of other homes listed all over the world. Knowing the price by comparing it to the prices of several comparable estates. It takes a lot of time and there's a good chance you'll make a mistake with the pricing.
- Large real estate firms have a lot of things to offer, and they have to hire people to go through each of them. The bases anticipate the price tag on a human, so there's a risk for human error. These appointed individuals must also pay. Nonetheless, having a machine to execute these tasks for us by crunching enormous statistics can save us a lot of time, money, and effort that a human being cannot.
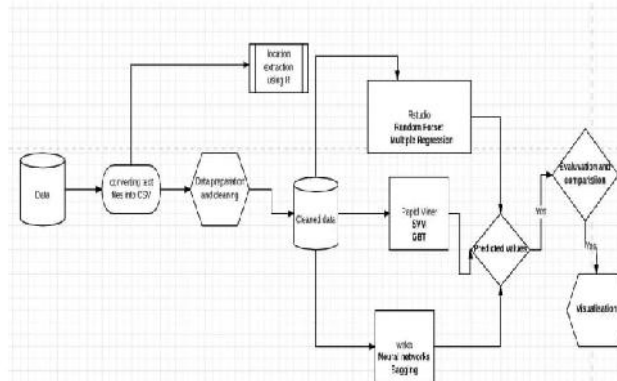
Description of datasets:-The UCI ML provides real estate housing data, which is used in these. To archive and the ageron, the data is split into rows and attributes. The following is a description of the dataset.

S.no	Variables	Integer type
	variable table	
1	Latitude	Real
2	Longitude	Real
3	Housing median age	Integer
4	Total Rooms	Integer
5	Total Bedrooms	Integer
6	Population	Integer
7	Households	Integer
8	Median Income	Real
9	Median House house(Price)	Integer
10	Ocean Proximity	Poly-nominal
11	Special Attribute Y	value to be pre-dicted

Algorithms:-

Lasso algorithm:-

• Lasso regression is a shrinkage-based kind of linear regression. Data values are shrunk towards a central point, such as the mean, in shrinking. The lasso technique encourages models to be simple and sparse (i.e. models with fewer parameters)

Decision tree:-

• A decision tree is a tool for making decisions that uses a tree-like model of actions and their possible outcomes, such as chance event outcomes, resource costs, and utility.

Multiple regression:-

It's a more advanced version of linear regression that takes into account numerous variables or features. It aids in the prediction of an attribute's unknown value based on the known value of two or more traits, also known as predictors.

OUT PUT/RESULT:-

When the code is run, the output plots appear first, followed by the forecast. These graphs aid in understanding the relationship between the target variable (price) and the various predictor variables.

This diagram shows a bar graph for the bedroom, carpet area, superpit area, bathroom, society, total sqft, and size. It can be observed in the datasets.
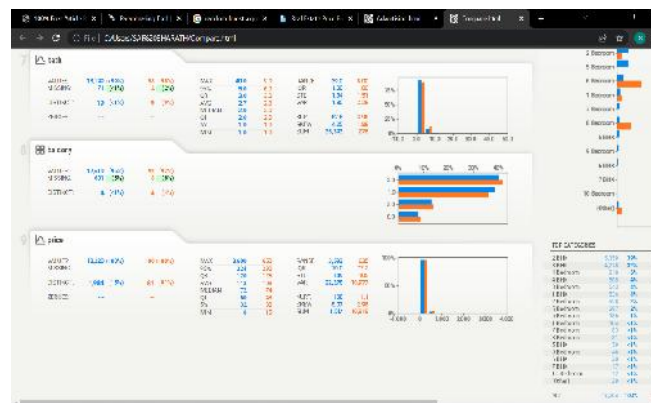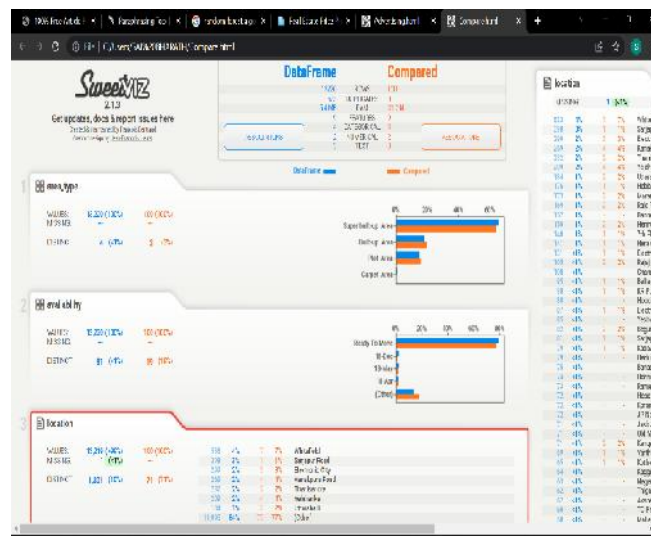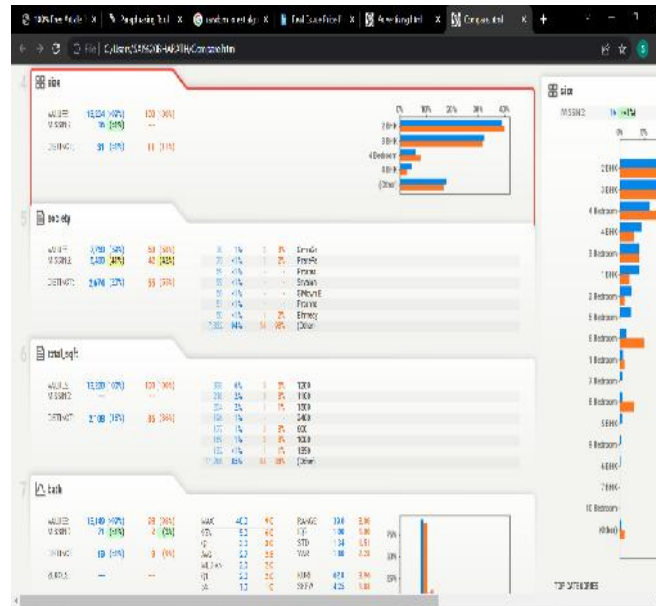
BANGALORE DATA SET			
ALGORITHM	RMSE mean	RMSE Std.Dev	MEAN CROSSval score
Decision tree	3.38	0.86	0.95
Linear regression	4.25	0.985	0.76
Lossoregression	4.30	0.953	0.84

Conclusion:-

The following table displays the RMSE Scores (mean and standard deviation) as well as Mean Cross-Validation Scores for four different machine learning algorithms in the Boston dataset and five different machine learning algorithms in the Melbourne dataset. Using these tables, we can deduce the accuracy of all of the methods for both datasets and find the best algorithm for price prediction.

Reference:-

- International journal of engineering research and technology(IJERT) ISSN:22780181
- International journal of Innovative technology and Exploring engineering(IJTIEEE) ISSN:22783075 Volume-8,issue 9