INTERNATIONAL JOURNAL OF ADVANCE RESEARCH, IDEAS AND INNOVATIONS IN TECHNOLOGY

ISSN: 2454-132X
Impact Factor: 6.078
(Volume 7, Issue 3 - V7I3-2218)
Available online at: https://www.ijariit.com

House rate prediction system using Machine Learning

Pratyush Kumar Mishra
pratyush201998@gmail.com
Sai Vidya Institute of Technology, Rajanukunte, Karnataka

Richa Tiwary
richatiwary1911@gmail.com
Sai Vidya Institute of Technology, Rajanukunte, Karnataka

Harshitha T.
harshithat.17cs@saividya.ac.in
Sai Vidya Institute of Technology, Rajanukunte, Karnataka

Jagbeer Poonia
jagbeerpoonia.17cs@saividya.ac.in
Sai Vidya Institute of Technology, Rajanukunte, Karnataka

Dr. Shantakumar B. Patil
shantakumar.bpatil@saividya.ac.in
Sai Vidya Institute of Technology, Rajanukunte, Karnataka

ABSTRACT

Everyone's desire is to purchase a home. When it comes to getting the best price, some people spend more than the property is actually worth. To tackle this challenge, we devised a method for predicting the price of a property using Machine Learning's Linear Regression model. We took into account all of the factors that one considers when purchasing a home, such as the neighbourhood, the number of bedrooms, and the number of restrooms. It will undoubtedly assist the buyer in obtaining the best possible price.

Keywords—*Property, Buyer, Purchasing, Linear Regression*

1. INTRODUCTION

Buying a house is a difficult process. One must spend a lot of money and dedicate a lot of time, and then there's the matter of whether it's a good deal or not. The vast majority of purchasers are completely unaware of the factors that influence home pricing. Almost all of the houses are classified by the total square feet, the locality, and the number of bedrooms. Occasionally, houses are sold for X rupees per square foot. This provides the appearance that the above elements are virtually solely responsible for house prices. Real estate brokers are used to acquire the bulk of houses. People usually buy directly from sellers since there is a lot of legal terminology and paperwork that most people don't understand.

As a result, real estate agents are trusted to handle both buyer and seller contact as well as the creation of a legal contract for the transfer. This just introduces a middleman into the process, boosting the home's price. As a result, the residences are overpriced, and buyers should be aware of the real value of the properties. Various online tools, such as Zillow and Trulia, are accessible to help a person buy a house. These tools are completely free to use and provide a cost estimate for a variety of homes. In order to forecast property values, these technologies use weights to give weights to various criteria. A business, for example, uses millions of public and user-submitted data points to estimate the worth of a property three times a week. To achieve the best cost for the customer, we used all of the factors.

2. LITERATURE SURVEY

Bangalore is a popular investment destination in South India, attracting both locals and non-residents who come to work and subsequently opt to invest. Investors are interested in investing since the capital gain is considerable. Bangalore's property costs are relatively affordable, with yearly price rises of 442 percent. The top Bangalore localities are Koramangala, where average property prices are Rs 5,820 per sq. ft, Rajaji Nagar, where average prices are Rs 11,100 per sq ft, Ashok Nagar, where property rates are Rs 22,780 per sq. ft, and Malleshwaram, where property rates are Rs 12,200 per sq ft.

All of these locations have a livability quotient of 9.5 or above on a scale of one to 10. This implies that social infrastructure such as schools, universities, hospitals, shopping malls, hangouts, convenience stores, and other facilities are all within walking distance. Physical infrastructure, such as linking roads, travel and transit facilities, are also adequate in these locations with such a high livability score.

House Price Prediction using regression techniques is an article in conjunction with the most recent Forecast on Research Predictions, which considers trends in order to better manage their economics. The major goal of the project FORECASTING VARIATIONS ON Home PRICE was to discover the best algorithm for predicting house prices with a low error rate by employing appropriate algorithms. Because the majority of individuals will ultimately purchase or sell a property, this is an intriguing topic.

They attempt to present all feasible regression strategies that are applicable to our topic in this paper. The following is a synopsis

© 2021, www.IJARIIT.com All Rights Reserved

of all the sources used: Ridge regression regularises the [rc] regression coefficient by posing an interest on the size in [MLR]Multiple Linear Regression, which uses more than one attribute for prediction, and [RR]Ridge and [LR] LASSO Regressions, which use Ridge regression to regularise the [rc] regression coefficient by posing an interest on the size.

3. SYSTEM REQUIREMENT

We attempt to present all feasible regression strategies that are applicable to our topic in this paper. The following is a synopsis of all the sources used: Ridge regression regularises the [rc] regression coefficient by posing an interest on the size in [MLR]Multiple Linear Regression, which uses more than one attribute for prediction, and [RR]Ridge and [LR] LASSO Regressions, which use Ridge regression to regularise the [rc] regression coefficient by posing an interest on the size.

According to the 2019 survey, House planning and prediction using Machine Learning study finds suitable models for housing price prediction. It also gives information on the housing market in cities. The first stage is to prepare the original data, which is then turned into a cleaned dataset suitable for analysis. For data reduction and transformation, stepwise and PCA approaches are utilised. Data reduction is a technique for reducing data by eliminating redundant and noisy information. Data is converted into a single format in order to execute actions on it. After that, many strategies are applied and assessed in order to arrive at the best answer. The assessment phase is a competitive technique that combines a Step-wise and Support Vector Machine (SVM) model. As a result, subsequent deployments may make advantage of it. This study may also be applied to transactional datasets from the housing market in other parts of the country.

Housing price trends reflect the present economic environment and are a source of anxiety for both buyers and sellers. The cost of a house is also determined by its location. A house with excellent access to roads, schools, malls, and employment possibilities will command a higher price than one without such access. Manually predicting home prices is difficult and seldom correct, which is why several algorithms for home price prediction have been created. The goal of this system was to create a model that could accurately anticipate housing prices based on other inputs. For the Ames dataset, they utilised Linear Regression, which provided high accuracy. The Admin and User modules were used in the housing price forecast project. Locations can be added and viewed by the administrator. The majority of current software is designed to work with this version of Windows. Admin had the right to increase density on a per-unit-area basis. Users may look at the location and see what the expected house price is for that area.

The Hybrid Regression Technique for Housing Prices was proposed in this work. The focus of the prediction was on the use of creative feature engineering to identify the best features and their relationship to sales prices. Data normalcy and linearity were enhanced because to feature engineering.

Their approach demonstrated that dealing with the Housing dataset was straightforward, and that employing Hybrid techniques (65% Lasso and 35% Gradient Boost) produced better results in forecasting home values than using lasso, ridge, or gradient boost alone.

3.1 Windows 10

The Windows 10 operating system is a collection of Microsoft's Windows NT-based personal computer operating systems. The majority of current software is designed to work with this version of Windows.

3.2 Flask

Flask is a Python-based microweb framework. It's termed a micro framework since it doesn't necessitate the use of any other tools or libraries. There is no third-party database abstraction layer, form validation, or other components that rely on third-party libraries to perform routine operations.

3.3 Python

Python is an interpreter for object-oriented high-level programming languages that is dynamically semantic. It's really high. Python is an object-oriented high-level programming language with a dynamically semantic interpreter. Because of its high-level built-in data structures, dynamic type, and dynamic binding, it's well-suited for Rapid Application Development and as a scripting or glue language for linking existing components. The readability of Python's basic, easy-to-learn syntax is prioritised, lowering software maintenance costs. Python supports modules and packages, which helps with programme modularity and code reuse.

The Python language is administered by the Python Software Foundation, which provides an open source reference implementation of Python called CPython. You can also get the Python source code from this link. Apart from the Python implementation itself being open source, many open source projects utilise Python, and Python has a large number of open source libraries available to developers.

3.4 Anaconda

Anaconda is a scientific computing distribution for Python and other computer languages that aims to make package management and deployment easier. Package versions are tracked by the package management system. It's a scientific computing distribution for Python and other computer languages that aims to make package management and deployment easier. Package versions are tracked by Conda, the package management system.

3.5 SKLearn

Scikit-learn is a Python image processing package that is free and open source. Segmentation, geometric transformations, colour space manipulation, analysis, filtering, morphology, feature recognition, and other methods are included. It's built to work with Python's NumPy and SciPy numerical and scientific libraries.

3.6 VS Code

VS Code is a free source-code editor that includes features such as debugging, syntactic high lighting, intelligent code completion, snippets, code refactoring, and integrated Git.

4. SYSTEM ARCHITECTURE

This is just one of many case studies that have been conducted to better understand how Machine Learning may benefit businesses and address public issues. Figure 4.1 shows how to estimate property prices, measure sentiment from user evaluations, extract relevant documents, propose things, and search for pictures using house-level features. Machines will be able to grasp tasks of interest, match them to machine learning tools, and assess the output quality using these abstractions. The process of assessing models and algorithms, as well as the things described above, is known as the Machine Learning Pipeline.

We need to divide our initial dataset into two pieces now that we have it:

1. A testing set: Our classifier uses a training set to “learn” how each category appears by generating predictions on the

© 2021, www.IJARIIT.com All Rights Reserved

input data and then correcting itself when the predictions are incorrect. We can evaluate the classifier's performance on a testing set once it has been trained.

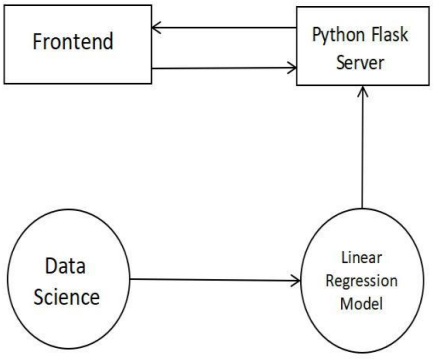

Fig 2 : Flow chart for proposed System

In successful applications, we may compute the risk-free and accurate amount for the residences using the model confidence level (probability). The goal of this project is to set up a machine learning model.

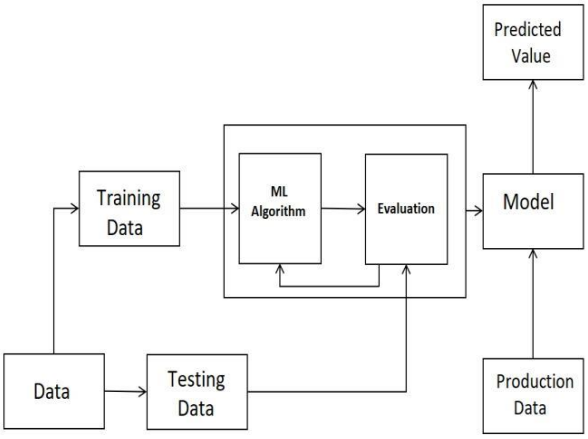

Fig 1 : Block Diagram for Proposed System

5. DEPLOYMENT DIAGRAM

The deployment of an ML-model is the integration of an ML-model into an existing production environment that can take in an input and provide an output that can be used to make real-world business decisions. Figure 4.2 depicts the Deployment Diagram. Consider developing a supervised machine learning (ML) model to help you decide whether or not to approve a home price prediction app.

6. IMPLEMENTATION AND RESULTS

We've loaded the data sets and then trained them using a Linear Regression model at the start of the implementation.

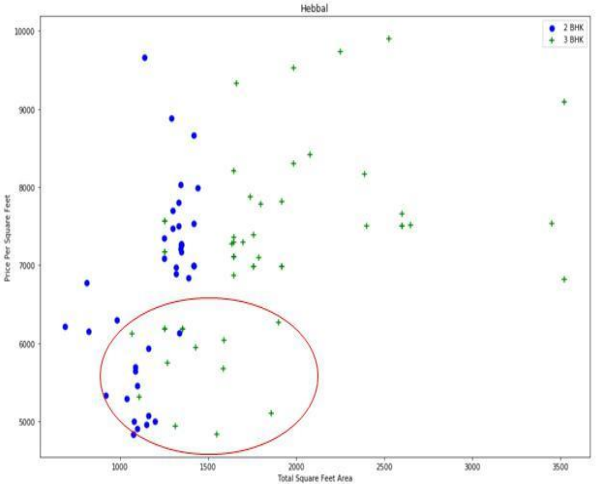

Fig 3 : Final Result

The outcome is depicted in the diagram below. The graph is displayed with Total Square Feet Area on the X-axis and Price on the Y-axis. We were able to correctly estimate the price based on the needed factors using the Linear Regression Algorithm.

7. CONCLUSION

Using Machine Learning, we created a method for predicting the price of homes. Using this technique, we can determine the appropriate price for any property, regardless of its size or location. The goal of creating this project is to spare the buyer from incurring additional costs as a result of the seller's advertised price.

The proposed approach, which is based on the Linear Regression algorithm, has been tested on a variety of locations and properties with a variety of parameters, demonstrating that the algorithm is extremely effective in predicting the price of a property. One of the key upcoming improvements is the addition of a home database of new metropolitan regions, allowing the client to examine more domains and make a more correct choice. More cost-influencing variables should be considered.

8. REFERENCES

[1] R Manjula, Shubham Jain, Sharad Srivastava and Pranav Rajiv Kher, "Real estate value prediction using multivariate regression models," IOP Conference Series: Materials Science and Engineering, 2017

[2] Eduard Hromada, "Mapping of real estate prices using data mining techniques," Czech Technical University, Czech Republic, 2015

[3] Adyan Nur Alfiyatin and Ruth Ema Febrita, "Modeling House Price Prediction using Regression Analysis and Particle Swarm Optimization," International Journal of Advanced Computer Science and Applications, 2017

[4] Li Li and Kai-Hsuan Chu, "Prediction of Real Estate Price Variation Based on Economic Parameters," Department of Financial Management, Business School, Nankai University, 2017.

[5] Nissan Pow, Emil Janulewicz and Liu Dave, "Applied Machine Learning Project 4 Prediction of real estate property prices in Montreal," 2016.

[6] Aminah Md Yusof and Syuhaida Ismail, Multiple Regressions in Analysing House Price Variations. IBIMA Publishing Communications of the IBIMA Vol. 2012 (2012), Article ID 383101, 9 pages DOI: 10.5171/2012.383101.

[7] Babyak, M. A. What you see may not be what you get: A brief, nontechnical introduction to over fitting regression-type models. Psychosomatic Medicine, 66(3), 411-421.

[8] Vasilios Plakandaras and Theophilos, Rangan Gupta*, Periklis Gogas "Forecasting the U.S. Real House Price Index"

© 2021, www.IJARIIT.com All Rights Reserved