4th National Conference on Green Technology and Science for Sustainable Development
© 2023 IJSRST | Volume 10 | Issue 10 | Print ISSN: 2395-6011 | Online ISSN: 2395-602X
International Journal of Scientific Research in Science and Technology

Real Estate Price Prediction Using ML

MS. Megha A. Patil, Payal Vasant Patel, Trupti Sanjay Pawar
Department of Computer Science and Engineering, Anuradha Engineering College, Chikhli, Maharashtra, India

ABSTRACT

The real estate market is one of the price orientedsector which is constantly changing due to high rate of civilization. Using Machine Learning techniques to predict costs with high efficiency is one of the key area.Our article estimates the market value of real estate. The system will help people to find the starting price of the property based on different areas.By analysing past business models, price ranges and future progress, future prices can be predicted. We used many variations on this approach, and our results are not the exclusive judgment of a single technique, but aweighted average of several techniques to give the most accurate results. The most successful Machine Learning algorithm proved is the random forest that has better compatibility to the data situation. The article also has several validation techniques on regression models that provide useful results for predicting home prices. The planning process will take into account the base factors used to calculate the home price and will provide a more accurate estimate.

Keyword: Machine Learning, real estate, random forest, regression.

I. INTRODUCTION

Housing is one of the essential factors to measure the success rate of a country's economy. As the economy grows, people tend to move form urban to rural areas,and as a result, the population of urban dwellers is increases, the demand for accommodation facilities will increase.Therefore house price also increase.In addition to these infrastructures development in an area can cause sudden increase in housing prices in a particular area.House Price Index (HPI)often used to calculatehouse price increases. .How're, studies show that the use of HPI is insufficient in the 21st century.House prices are affected by many variables.Like size of real estate,numbers of rooms,usable gardens,size of land and property,numbers of kitchen and bath rooms etc.{from 1-s2}Therefore Machine Learning is technologies have grown and raised its capabilities across a suited of application. Mlis a computer program andbranch of artificial intelligence that learn from experiences to predict current performance and improve future data.There are different types of learning in machine learning,supervise and unsupervised learning. The machine learning works by learning past patterns using selection algorithms and predicting future outcomes.This article focuses on the literature review of real estate price prediction based on machine learning models and based real estate price prediction based on factors affecting real estate prices, which have been used frequently in past.

II. LITERATURE REVIEW

In the 21st century, this century housing is redundant, it represents a lot today. is not only for those who want to buy real estate, there are also companies selling these properties. According to [4], houses are not only needed by people, but today also represents wealth and fame. Investing in products often looks profitable because values do not drop quickly. Changes in real estate prices affect many home investors, investors, policy makers and more. Investing in real estate seems like a good investment option. It is seen that the estimated real estate value of is an important economic indicator. The [3] shows that every organization in the real industry today is working hard to gain a competitive advantage over their other competitors. The need to simplify the process for ordinary people while providing the best possible results. [6] proposed using machine learning and artificial intelligence to create an algorithm that can predict real estate prices based on some input. The commercial application of this algorithm is that distribution websites directly use this method to estimate the value of new properties to be listed soon, and it takes about different ideas to estimate the correct and reasonable price. Home value estimation is a big topic, done with a lot of computer science methods. Use Google Colab/Jupiter IDE. Jupiter IDE is an open source web application that helps us share and create files with LiveCode, views, equations and annotations. It has tools for data cleaning, data transformation, numerical simulation, statistical models, data visualization and machine learning tools. [10] developed a system called to help people know the true value of the attribute. Users can get the price of the house they want by giving the code according to their needs. Users can also get the home plan to use the homes.A holds information about values. According to the findings of [1], the best accuracy is provided by the random forest regressor, followed by the decision tree regressor. Similar results are the produced by Ridge and Linear Regression with a slight on Lasso. In all selected groups, there is no difference between all , regardless of the strong and weak groups. It gives a good indication that the match value can be used alone to estimate sales without considering other features to reveal an overfit of model . Also, in the very weak group, the accuracy drops. For each particular option, the same outcome pattern can be seen in Root Mean Square Error (RMSE) . Like Machine Learning, Outbound, Decision Trees, Network Learning, Fuzzy Logic, ANFIS (Adaptive Neuro-Fuzzy Inference System) and Linear Performance Pricing. In machine learning models, data is divided into two parts, training and testing. 80% of the data is used for educational purposes and 20% is used for testing. Training methods involve different goals. The model was trained using various machine learning algorithms with random forest regression predicting better results. They use the NumPy and Pandas Python libraries to implement the algorithm. [1] Other forms of machine learning use various regression models for prediction. In addition, is compared with other learning models and decision tree regressors such as Lasso, LassoCV, Ridge, RidgeCV. Multivariate Linear Regression and Lasso CV performed best out of 84.

III. METHODOLOGY

A. Data Collection and Exploration

The first step is to identify machine learning-related articles in real estate price prediction,thusresearching for good literature reviewed on the topic in two licensed database, scopes and web of science.An advanced search using the query string revealed records,over 2,254 available items.The second stage is the review of statistical data;The is a selection of articles on the research topic. The third level is attribution and exclusion by the

author with various attributions and executions.First and foremost,authors will select articles from journals with empirical and theoretical data, that is, they will not include research articles,book chapters and book series by text type. Then social science indexed article is selected.Finally,the author alsofocuses on machine learning research in real estate forecasting, making the research paper applicable to all fields.Fourth is data abstraction and analysis. This level contains important information and analysed to focus on a particular study in response to question of the information in the article and the objectives of the method. Find and analyse the ML distribution for real estate price estimation from the collected data. Select an appropriate dataset on which our machine learning model will build.

B. Data Preprocessing

The first is to investigate missing data fromselecteddataset.Any observation which had missing value was removed from dataset. Drop features that are not required to build our model.

C. Feature Engineering

Remove features that are not required to build our model. Any location having less than 10 data points should be tagged as "other"location.Removeoutlier using business logics.We should remove outliers per location using mean and one standard deviation.

D. Model Selection

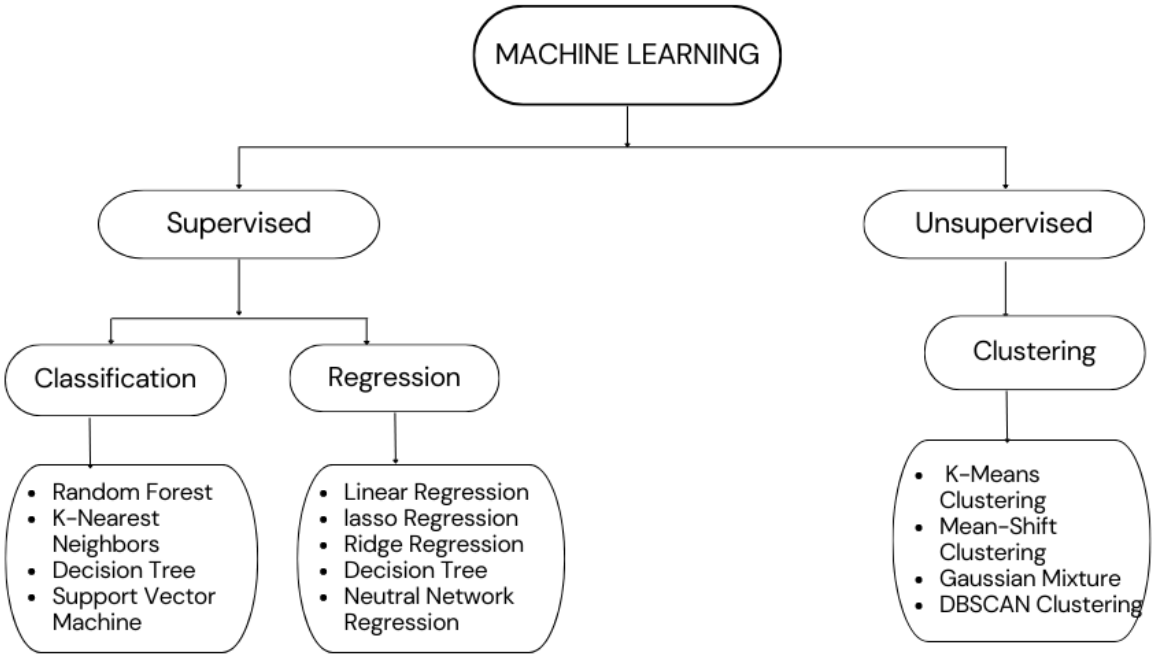

Fig no :1 Types of machine learning

Model selection is the way of selecting machine learning model from different types of machine learning models for a training dataset. There are different types of machine learning models Support vector regression (SVR),multiple linear regression, decision tree , etc.

- **support vector regression.**

There are two types of machine learning supervise andunsupervised. Support Vector Regression is a supervised learning algorithm used to predict random variables.SVR is inspired by Support vector machines(SVMs). SVMs is able to process non linear results and overcome small sample learning issues. Its ability to generate economic forecasts in a variety of markets, including real estate, shows that the model can overcome non-horizontal problems and few learning problems. In addition, this model is often used for building standards, as it does not rely on the ability of the model to disperse thoughts and express ideas (linear or non-linear) [22]. Support Vector Regression has huge advantages over models, because it avoids a lot of trouble while increasing agreement by reducing both risk and empirical.

- **Multiple linear regression**

Multiple linear regression is a regression model that uses a straight line to predict the relationship between a variable and two or more variables. We can find relationship between two or more independent variables and one dependent variable by using multiple linear regression The strength of the multiple regression models can be seen when evaluating the value of the relationship between the dependent variable in andthe independent variable in . [28] used a multiple regression model to explain the evolution of independent variables by variable. This model can be used independently and for the price of the house, the size of the house, the number of apartments, bedrooms, etc. It can be used to predict house price based on variables. Therefore, house price was determined as the target or living variable, other features were determined as a different variable, and the main variable was determined by , which determines the correlation coefficient of each feature.

- **decision tree**

A decision tree regressor identifies features and trains a tree-like model to predict future data to provide large outcomes. The highest and lowest value of the graph is obtained by the decision tree regressor and then splits the data by as seen by the system. The Web Search CV is a strategy for tracking changes that will generate values and learn patterns for each combination of the parameters shown in the subsection. Search System In this calculation, CV is used to determine the best value of the most important used to construct the decision tree.

E. Training the model

In this step we divide our model into two modules: a training set and test set In training, you import the prepared data into a machine learning model to see patterns and make predictions. It allows the model to learn from the data so that it can complete the task. Over time, through training, the model gets better at predicting.

F. Testing and Integrating with UI.

Apply the training model to the test data and estimate the cost of generating the . Then the training model is included in the front end using Flask in python .

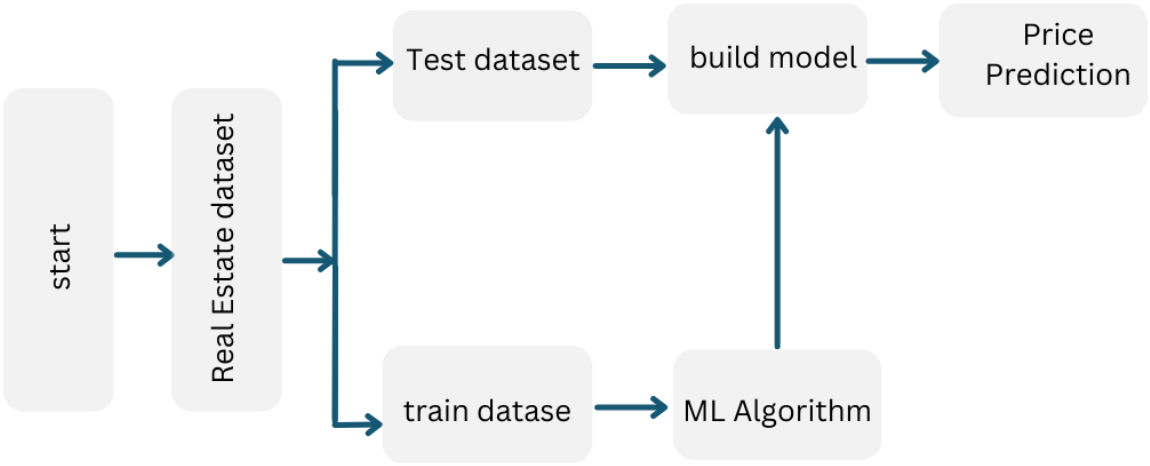

Fig no:2 Flow diagram

IV. PURPOSE WORK

The purpose of the system is to determine the value of the building by looking at the various features entered by the user. These features are fed into the ML model and give the label How does this feature affect the bid according to . This will be done by first searching for a suitable file that fits the developer's and user's requirements. Also, after processing the file, it goes through a process called file cleanup ; where all junk files are removed and the file raw text is converted to a .csvfile.In addition, the data will be passed through the file first, where the missing data will be corrected and if necessary, the tag will be coded. Also, this will go through a data transformation to where it will be converted to a NumPy array so that it can be sent to train the model. During the training, various machine learning algorithms will be used to train the models and their errors. ratios will be subtracted so algorithms and models will complement to produce accurate predictions. Users and companies will be able to log in and then fill out a form specifying the various characteristics of the property they want to estimate the price of . In addition, the form will be sent after the selection of all features is completed.User enters their purchase information into the template and within seconds, user can see the estimated price for the product they have entered.

V. FUTURE WORK

The main purpose of this study is to analyze the value estimation we obtained using different learning systems such as random forest, multiple regression, support vector machine, gradient boosted trees, neural networks and bagging, i.e. it is obviously Random. forest has higher accuracy in prediction than forest, and my research also provides a way to find actions involving in prediction. Therefore, this research will be useful for people, government and future studies are as follows: Every new system and software can help estimate the cost of . Price Forecast This can be improved by adding many elements to the building such as: environment, market and many other variables. Estimated information can be stored in a file and an application can be created for people so that they have a short-term idea and deposit their money safely. If there is real-time data capability,

the data can be connected to the H2O, the machine learning algorithm can be directly connected to the interface, and the application environment can be created.

VI. CONCLUSION

Buying your own home is everyone's dream. Using this model of , we want people to buy houses and real estate at reasonable prices and not be scammed by sloppy agents just for their money. Additionally, the model will also help large companies set their prices against the and provide them with accurate forecasts, saving the a lot of intervention and saving a lot of time and money. Accurate real estate prices are inherent in the market and we want to ensure that using this model. Buying your own home is everyone's dream.Using this model of , we want people to buy houses and real estate at reasonable prices and not be scammed by sloppy agents just for their money. Additionally, the model will also help large companies set their prices against the and provide them with accurate forecasts, saving the a lot of intervention and saving a lot of time and money. Accurate property prices are important in the market, and we hope to achieve that using this model. One more thing can add is that the uses the entire user interface, so can have many functions for the user to use in many places with the ML model. Finally, create an integrated Webapplication that can estimate cost when users need to complete a project.

VII. REFERENCES

[1]. D. Phan, "Housing Price Prediction Using Machine Learning Algorithms: The Case of Melbourne City, Australia," 2018 International Conference on Machine Learning and Data Engineering (iCMLDE), Sydney, NSW, Australia, 2018, pp. 35-42, doi: 10.1109/iCMLDE.2018.00017.

[2]. J. Modern Education and Computer Science, 2020, 6, 46-54 Published Online December 2020 in MECS (http://www.mecs-press.org/) DOI: 10.5815/ijmecs.2020.06.04

[3]. Byeonghwa Park, Jae Kwon Bae,Using machine learning algorithms for housing price prediction: The case of Fairfax County, Virginia housing data,Expert Systems with Applications,Volume 42, Issue 6,2015,Pages 2928-2934,ISSN 09574174,https://doi.org/10.1016/j.eswa.2014.11.040.

[4]. International Journal of Innovative Technology and Exploring Engineering (IJITEE) ISSN: 2278-3075, Volume-8 Issue-9, July 2019

[5]. Rawool, A. G., Rogye, D. V., Rane, S. G., &Vinayk, A. (2021). House price prediction using machine learning. Int. J. Res. Appl. Sci. Eng. Technol, 9, 686-692.

[6]. Zulkifley, Nor Hamizah, et al. "House Price Prediction using a Machine Learning Model: A Survey of Literature." International Journal of Modern Education & Computer Science 12.6 (2020).

[7]. Quang Truong, Minh Nguyen, Hy Dang, Bo Mei,Housing Price Prediction via Improved Machine Learning Techniques,Procedia Computer Science,Volume 174,2020,Pages 433-442,ISSN 1877-0509,https://doi.org/10.1016/j.procs.2020.06.111.

[8]. Kuvalekar, Alisha, et al. "House price forecasting using machine learning." Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST). 2020.

[9]. Rawool, Anand G., et al. "House price prediction using machine learning." Int. J. Res. Appl. Sci. Eng. Technol 9 (2021): 686-692.

[10]. Chavan, P. A., Shejole, P., Sharma, R., Shaikh, S., Shelke, D., & Sharma, P. HOUSE PRICE PREDICITION USING ML.
[11]. Dabreo, Smith, et al. "Real estate price prediction." Int J Eng Res Technol (IJERT) 10.04 (2021).
[12]. Ja'afar, NurShahirah, JunainahMohamad, and Suriatini Ismail. "Machine learning for property price prediction and price valuation: a systematic literature review." Planning Malaysia 19 (2021).
[13]. Zhang, Qingqi. "Housing price prediction based on multiple linear regression." Scientific Programming 2021 (2021): 1-9.