Real Estate Price Prediction System Using Hybrid Lasso Algorithm

Dr. N. Moganarangan[1], Kishorre Raj. M[2], Ashathulla. J[3], Devasanjay.T[4], Shalinthraj Adhithyan.R[5]

[1] Professor, Sri ManakulaVinayagar Engineering College, Puducherry
[2,3,4,5] Student, Sri ManakulaVinayagar Engineering College, Puducherry

House Price prediction system is a website where the user/ buyer can be able to predict the accurate cost of real estate. To analyse the pertinent characteristics and the best models for predicting the price of homes, a literature review is done.. The model will then use the user's data, and the user will be able to view the predicted price of the property they are selling or looking to buy in a matter of seconds. The feature selection process is done with a type of regression called LASSO regression. A kind of linear regression that makes advantage of shrinkage is called LASSO regression. So we are proposing a system that uses Hybrid LASSO Regression algorithm to do the Feature Selection. Hybrid LASSO Regression is hybrid of both LASSO and Ridge Regression. When models exhibit significant levels of multi-collinearity or when you wish to automate specific steps in the model selection process, such as variable selection and parameter elimination, this specific sort of regression is ideally suited. This would greatly help academics and housing developers identify the most important factors that influence home values and recognize the most effective machine learning model to follow when conducting the field investigation.

I. INTRODUCTION

The field of machine learning (ML) iscrucial to modern industry and research.Through the use of algorithms and neural network models, computer system performance is graduallyincreased. Using sample data, often known as "training data," machine learning algorithms automatically create a mathematical model that they then use to make judgments without being explicitly trained to do so.

Both individuals and real estate companies buy and sell homes. Individuals buy for habitation or as an investment, while companies buy to operate. In either case, we think that everyone deserves to receive exactly what they pay for. Housing market overvaluation and undervaluation have always been problems, and there aren't enougheffective detection methods. A primary pass is given by broad indicators like house price-to-rent ratios. However, a thorough analysis and decision-making are required to resolve this matter. When it comes to this, machine learning is useful. By feeding an ML model thousands of data points to learn from, a solution that is both powerful enough to properly predict pricing and flexible enough to meet everyone's demands may be created.

Utilizing these machine learning techniques to create ML models that can then benefit consumers is the main goal of this article. A buyer's first goal is to find their ideal home, complete with all the facilities they require. They search for these homes and real estate with a budget in mind, but there is no

assurance that they will pay a fair price for the item and avoid being overcharged. Similar to this, a seller searches for a specific figure that they can use to tag the property's worth. This figure cannot simply be an educated guess; much research is required to arrive at a home's assessment. Additionally, there's a chance that the product will be underpriced. If the price is anticipated for these users, it might be possible for them to purchase lands for the prices they deserve, neither more nor less.

LIMITATIONS OF PREVAILING METHODOLOGIES

There has been a sizable amount of research in the field of housing price prediction, but very little of it has produced any practical solutions. There isn't much proof that a corporation has set up a functional house price predictor. For now, very few digital solutions exist for such a huge market and most of the methods used by people and companies are as follows:

Buyers/Customers:
People frequently go online to research trends and other relevant information when they first consider buying a home or other real estate. People do this to search for a home with all the amenities they require. People take note of the cost associated with these houses as they work. However, the typical person lacks in-depth knowledge and precise information regarding what the real price should be. They may be misinformed as a result because they think the prices listed online are real. When looking for a property, contacting different estate agents is the second thought that comes to mind. The issue with this is that these agents should only be paid a small portion of the price just to look for a home and establish a price for you. Because they have no other choices, people frequently believe this price tag without question. There may be instances where agents and sellers engage in a covert transaction, leading to the customer being unawarely sold an overpriced home.

Seller/Agencies:
When someone considers selling their home, they compare it to the hundreds of thousands of other properties that are listed all over the world. Calculating the price by comparing it to several estates takes a lot of effort and runs the danger of being inaccurately priced. Large real estate organizations need to sell a variety of items, and they must designate individuals to handle each product. This also relies on a human to forecast a price tag, thus there is a chance for inaccuracy. These designated individuals must also be paid. But you can save a lot of time, money, and accuracy by having a computer perform this work for you by crunching the large numbers

II. LITERATURE REVIEW

In the twenty-first century, real estate has evolved into something much more than a basic requirement. not just for those wishing to purchase real estate, but also for the businesses that market these properties. Real estate experts claim that in addition to being a man's basic need, property now also stands for a person's wealth and status. Real estate investments often appear to be profitable because the values of the properties do not drop off quickly. Many different types of household investors, bankers, policymakers, and others may be impacted by changes in the real estate price. The real estate industry appears to be a desirable choice for investment. Predicting real estate value is a key economic indicator, which implies that every organisation in the real estate industry today is working successfully to gain a

competitive advantage over other competitors. It was suggested to employ artificial intelligence and machine learning techniques to create an algorithm that can anticipate house prices based on specific input features in order to make the process easier for a typical human being while still producing the best results.

This algorithm's commercial use is that classified websites can use it to anticipate the prices of new properties that are going to be posted by using a few input variables and predicting the right and justified price, avoiding client price inputs and preventing error from entering the system. using the Jupiter/Google Colab IDE. An open-source web application called Jupiter IDE enables us to share and create documents with LiveCode, visualisations, equations, and text that provides context. It includes tools for data cleansing, data translation, numerical value simulation, statistical modelling, data visualisation, and machine learning. created a mechanism that would assist individuals in learning about the approximate real estate pricing. Users can input their specifications to receive prices for the desired homes in accordance with those needs. The sample house plan is made available to users as a resource for houses. Following the removal of the original data set's missing samples, 400 samples are used as training data and 52 samples as test data. Value of the training data for housing. According to the results, the Random Forest Regressor and Decision Tree Regressor both offered the best accuracy. There is no significant difference between any of the feature selection groups, strong or weak, across the board. It is a positive indication that model over-fitting can be reduced when only the purchase prices are used to predict the selling prices. In the group of very weak features, there is also a noticeable drop in accuracy. Instead of running the calculations sequentially, we might use many processors and run them in parallel, which might cut down on preparation and prediction times. Include Even more features in the model so that we may provide clients the option to choose another location to create the high-temperature maps rather than entering their choices manually. a 100-house data set with a number of parameters was employed. The machine was trained on half of the data set, and it was tested on the other half.

The results are truly accurate and experimented with some of the most basic machine learning methods, including multiple linear regression, decision tree classifier, and decision tree regression. With the use of this work, users may forecast both the availability and cost of homes in the city. employed artificial intelligence algorithms to forecast home price changes. The step-by-step process for analysing the data set has been given. These feature sets were then used as the input for four algorithms, which produced a CSV file with forecasted home prices. stated that a combination of these models must be used. A high bias (under-fit) is produced by a linear model, whereas a large variance is produced by a high complexity model (over-fit). The results of this study can be applied to the annual adjustment of the land guideline value, which could increase the amount of money that the State Government receives from this transaction. says that it is evident that lasso regression and ridge regression are doing better with more accuracy % and with fewer error values after executing this experiment with different machine learning algorithms. These algorithms perform well when this experiment is compared to the label and the achieved outcome.

III. PROPOSED WORK

With the help of the many features that the user inputs, this system aims to calculate the cost of a house. These features are provided to the ML model, which then makes a prediction based on how they affect the label. This will be accomplished by initially looking for an adequate dataset that meets the requirements of both the user and the developer. Additionally, the dataset will undergo a procedure

known as data cleaning after it has been finalised, during which all unnecessary data will be removed and the raw data will be converted into a.csv file. Additionally, the data will undergo data preprocessing, during which any missing data will be resolved and, if necessary, label encoding will be performed. Additionally, this will go through data transformation and be turned into a NumPy array before being sent to train the model. The model will be trained using a variety of machine learning methods, and when they are utilised, their error rates will be extracted. As a result, a final algorithm and model that can produce correct predictions will be developed. Users and businesses can log in and then fill out a form about the many characteristics of their property that they wish to anticipate the price of.

The form will be submitted as well after carefully choosing the attributes. In a matter of seconds, the user will be able to see the estimated price of the property they input thanks to the data that was sent to the model by the user. Users and businesses can log in and then fill out a form about the many characteristics of their property that they wish to anticipate the price of. The form will be submitted as well after carefully choosing the attributes. The user's data will then be sent to the model, where they can view the estimated price of the property they input in a matter of seconds.

ALGORITHM USED

Hybrid LASSO Regression Algorithm - It is nothing but a Hybrid model that uses penalties from both LASSO regression and Ridge Regression. The technique combines both the lassoand ridge regression methods by learning from their shortcomings to improve the regularization of statistical models and that is to provide accurate results using highly correlated features.
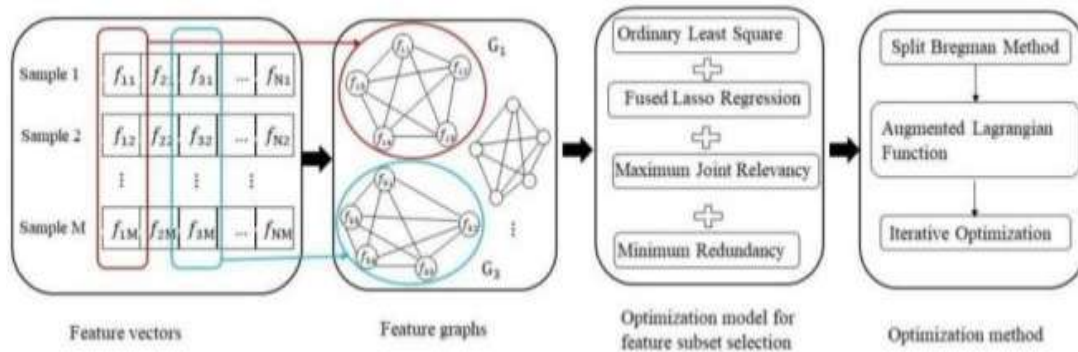
ADVANTAGES OF ALGORITHM:

➢ Robust Tool for approximating a target function.
➢ Speed up the training process.
➢ Fast response time.

STATE CHART DIAGRAM
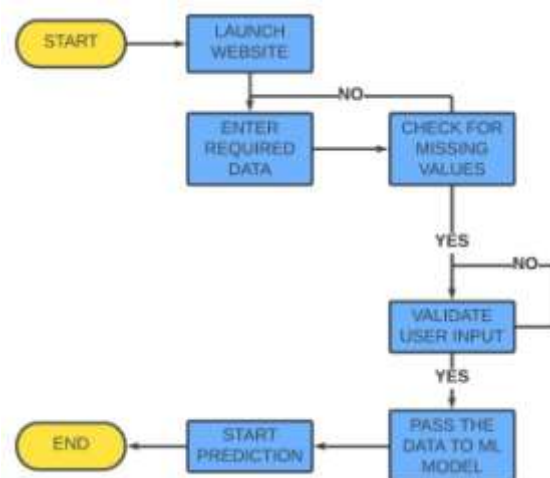
ARCHITECTURE DIAGRAM

ADVANTAGES OF THE PROPOSED SYSTEM:

- ➢ Improved traceability
- ➢ Provides a tangible improvement of final classification performance
- ➢ Quick and Efficient to use
- ➢ Excellent empirical performance
- ➢ Relatively simple and computationally inexpensive method
- ➢ Simple, fast and less complex.
- ➢ It is a fast and easy procedure to perform or imprecise data.

IV. IMPLEMENTATION

Dataset:

We have used one datasetin this paper where we used Hybrid LASSO machine learning algorithm are applied to the datasetforpredicting prices.The dataset that is used here is a dataset based on Bangalore. There are 8 data attributes that are used in the dataset.

Attributes	Description
area_type	Whether the area is well developed or developing.
availability	If the property is ready to be sold
location	Where the property is located
size	Number of bedrooms
total_sqft	Total Size of the property
bath	Number of Bathrooms
balcony	Number of Balconies
price	Price of property

Table 1: Attributes and label in thedataset

Data Cleaning:

The Bangalore dataset contains a large number of missing values. The machine-learning model would behave incorrectly if the data were used without any action. So, it is best to eliminate the data that has missing attributes in order to prevent the machine-learning model from failing. Any missing values in the future will be addressed as soon as the data enters the pipeline by inserting the simple imputer function as well.

Data Preprocessing:

The dataset has been divided into two sets, a training set and a testing set. 80% of the dataset is in the training set, while 20% is in the testing set. We wanted to ensure that the splitting would divide the values in the columns, which included just two distinct values, equally.

Feature Selection:

The different attributes listed in the above table were used to train our machine-learning model. However, only the attributes "total sqft, location, size, and bath" are the features that are utilised to determine the price.The aforementioned characteristics are independent characteristics that are valued when one contemplates purchasing real estate.

V. CONCLUSION

Buying your own house is what every human wish for. Using this proposed model, we want people to buy houses and real estate at their rightful prices and want to ensure that they don't get tricked by sketchy agents who just are after their money. Additionally, this model will also help Big companies by giving accurate predictions for them to set the pricing and save them from a lot of hassle and save a lot of precious time and money. Correct real estate prices are the essence of the market and we want to ensure that by using this model. The algorithm we use (i.e) Hybrid LASSO Regression algorithm will be potent enough for Predicting the prices of different houses with various features and will be able to handle large sums of data. The system is quite userfriendly and time-saving too.

REFERENCES

1. Thuraiya Mohd, Suray Masrom, Noraini Johari, "Machine Learning Housing Price Prediction in Petaling Jaya, Selangor, Malaysia", International Journal of Recent Technology and Engineering (IJRTE), Volume-8, Issue-2S11, 2019.
2. G. Naga Satish, Ch. V.Raghavendran, M.D.Sugnana Rao, Ch.Srinivasulu , "House Price Prediction Using Machine Learning" , International Journal of Innovative Technology and Exploring Engineering (IJITEE), Volume-8 Issue-9, 2019.
3. Kuvalekar, Alisha and Manchewar, Shivani and Mahadik, Sidhika and Jawale, Shila, House-price Forecasting Using Machine Learning (April 8, 2020). Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST) 2020.
4. Neelam Shinde, Kiran Gawande , "Valuation Of House Prices Using Predictive Techniques", International Journal of Advances in Electronics and Computer Science, Volume-5, Issue-6, 2018.
5. Jingyi Mu, Fang Wu,andAihuaZhang,"Housing Value Forecasting Based on Machine Learning Methods", Hindawi Publishing Corporation Abstract and Applied Analysis, Volume 2014.

6. Sayan Putatunda, "PropTech for Proactive Pricing of Houses in Classified Advertisements in the Indian Real Estate Market".

7. Atharva Chouthai, Mohammed Athar Rangila ,Sanved Amate, Prayag Adhikari, Vijay Kukre , "House Price Prediction Using Machine Learning" , International Research Journal of Engineering and Technology(IRJET), Vol:06 Issue: 03, 2019.

8. B.Balakumar,P.Raviraj, S.Essakkiammal,"Predicting Housing Prices using Machine Learning Techniques".

9. Akshay Babu, Dr. Anjana S Chandran , "Literature Review on Real Estate Value Prediction Using Machine Learning" , International Journal of Computer Science and Mobile Applications, Vol: 7 Issue: 3, 2019.

10. Mr. RushikeshNaikare, Mr. Girish Gahandule, Mr.AkashDumbre, Mr. Kaushal Agrawal, Prof. Chaitanya Manka , "House Planning and Price Prediction System using Machine Learning",International Engineering Research Journal, Vol:3 Issue: 3, 2019.

11. Aswin Sivam Ravikumar, Thibaut Lust, "Real Estate Price Prediction Using Machine Learning", 2016.

12. Bindu Sivasankar, Arun P. Ashok, Gouri Madhu, FousiyaS , "House Price Prediction", International Journal of Computer Science and Engineering(IJCSE), Vol: 8 Issue: 7, 2020.

13. M.Thamarai, S.P.Malarvizhi, "House Price Prediction Modeling Using Machine Learning", International Journal of Information Engineering and Electronic Business(DJIEEB),VoL12, No.2, pp. 15-20,2020.DOI: 10.5815/ijieeb.2020.02.03

14. Thuraiya Mohd, Suraya Masrom, Noraini Johari, "Machine Learning Housing Price Prediction in Petaling Jaya, Selangor, Malaysia", International Journal of Recent Technology and Engineering (IJRTE), Volume-8, Issue- 2S11, 2019.

15. G. Naga Satish, Ch. V. Raghavendran, M.D.Sugnana Rao, Ch.Srinivasulu , "House Price Prediction Using Machine Learning" , International Journal of Innovative Technology and Exploring Engineering(IJITEE), Volume-8 Issue-9, 2019.

16. Kuvalekar, Alisha and Manchewar, Shivani and Mahadik,Sidhika and Jawale, Shila, House Price Forecasting Using Machine Learning (April 8, 2020). Proceedings of the 3rd International Conference on Advances in Science & Technology (ICAST), 2020.