Email: editor@ijerst.com or editor.ijerst@gmail.com

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

The Use of Modified Extreme Boosting in Machine Learning for the Prediction of Home Prices

[1] Dr.Taj Uddin [2] Mohammed Abdul Maaz, [3] Shaik Arshad Ali, [4] Abdul Samad Faizan,
[1]Associate Professor,Department of CSE-AIML, Lords Institute of Engineering & Technology.
[234] Student Department of CSE-AIML, Lords Institute of Engineering & Technology.

ABSTRACT

Machine learning's impact on commonplace spoken commands and predictions has grown in recent years. In its place, it offers an improved customer service system and a safer automated driving experience. The evidence suggests that ML is a promising technology with the potential to revolutionize many different markets. The home Price Index is a common tool used to measure changes in home values (HPI). A person's home price cannot be reliably predicted using only the HPI because of the strong relationship between property prices and other factors like population, area, and location. While some research has been able to accurately forecast home values using traditional machine learning methods, these studies almost never compare and contrast various models and completely exclude the more intricate but less well-known ones. The adaptive and probabilistic model selection procedure of Modified Extreme Gradient Boosting led us to suggest it as our model for this investigation. This procedure includes creating features, training and optimizing hyperparameters, interpreting models, and finally, selecting and evaluating models. House price indices, which are often used to bolster housing policy efforts and provide cost estimates. Using machine learning techniques, this research builds models to predict future changes in house values. Location, square footage, home price, and modified extreme gradient boosting are all relevant terms.

1. Introduction

An important component of any thriving economy is the real estate market, of which the housing market is a part. Many recently hired individuals have made it their career goal to become homeowners, a goal that is highly valued in many areas of the globe due to the status symbol nature of that possession. Regardless, investors seek out the real estate market for the opportunities it presents, rather than the commodities it really is [1]. If the economy is growing, the real estate market—and the housing market in particular—must be booming as well. Buying a property is a dream of many young professionals throughout the world because of the social prestige it conveys. Reason being, having a house is a prestige symbol. However, investors are interested in the housing market because, contrary to popular belief, they see real estate not as a commodity but as an opportunity for financial gain. Most people who purchase a home or invest in real estate do so with the hope of eventually seeing a profit from the property's value increasing. In general, the proportion of the population that owns their home decreases as house prices rise. Countries with high homeownership rates, especially those whose economies are still growing, have traditionally been the primary subjects of studies [3]. Because housing costs have a major impact on the market's sustainability over the long run, it is critical that individuals have access to affordable housing that satisfies their fundamental necessities. Whether or whether buying a home is a smart long-term financial plan is heavily dependent on how affordable it is. Real estate market volatility is much lower compared to that of the stock market, interest rates, and currency exchange markets. Real estate has been one of the most lucrative investment industries in recent years, especially in the last fifteen, and price swings in homes have a major influence on this industry. One of the most hotly debated subjects in real estate as of late has been the process used to determine property prices. This issue has encouraged several players to speculate about the future of property prices. These include residential investors, real estate investment trusts, individual investors, and officials from different government organizations. These people have used a wide variety of strategies to accomplish their goal. Urban populations have grown exponentially since the start of the Industrial Revolution, leading to a critical shortage of housing. This is a result of the unprecedented rate of urbanization that is happening right now on every continent. Many aspects of this issue have became clearer as time has progressed. While Germany was still building its infrastructure in the 1980s, this book delves at the massive housing crises that hit the country's major cities. At this juncture, people all throughout the world began to accept the reality of the housing crisis, which was caused by economic disparity, and the need for more social housing. Cities in emerging nations with a younger urban population are still in their formative years, and these

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

characteristics have been recognized as issues there as well. For instance, many studies have focused on the problem of subpar housing in India. Abhay said that, while the housing scarcity issue is stimulating the construction supply, low-quality housing is common across the city (2001-2011) when providing statistics on the number of dwellings created in Karnataka over the last decade. This was brought out by Abhay while discussing the amount of residential development projects in Delhi over the analyzed decade [8]. The scarcity of housing alternatives is exacerbated, according to many, by the disproportionately large number of older dwellings. Estimating a home's worth is difficult due to the interconnected nature of the property's physical features, the surrounding area, and its location. Specifically, compared to more traditional methods for solving this issue, the house price projection model has gotten much less attention in the current literature. There is general agreement that this is a critical issue that requires immediate attention, yet it continues happening nevertheless. The emergence of Big Data has led to machine learning's prominence as a prediction approach [9]. Because of this, it is now feasible to generate better predictions about future property values based on the property's characteristics alone, without using past data. Some studies have looked into this and discovered that machine learning works, but most have just compared model performances without considering how to combine them optimally [10]. In order to predict home prices, this study employs an algorithmic tweak called extreme gradient boosting. Therefore, the ultimate goal of this study is to have a better understanding of regression approaches in machine learning. The supplied datasets must be processed in order to get best performance. To get a good estimate, we need to identify which of a house's distinctive qualities are most significant before we can remove the less important ones, as these qualities make up the house's worth [11]. Because these modifications wouldn't be available to all properties, the data is biased; without them, house prices wouldn't fluctuate as much. Using the Modified Extreme Gradient Boosting technique, the primary goal of this research is to forecast the home price. It is used for price prediction based on factors such as area type, location, BHK, and more. We compare the accuracy and performance of the Modified XGBoost with a number of other methods.

2. Literature Survey

In their study, Park et al. [4] examined the best accurate ML models for forecasting property prices. They did this by looking at 5359 row dwellings in Virginia. To solve a classification problem, they used the RF technique; to solve a regression problem, they used the nave Bayesian procedure. Conclusions drawn from the research indicate that the RIPPER algorithm considerably improved price prediction. Banerjee et al. [12] looked at several machine learning methods to forecast whether real estate values would rise or fall in the future. Despite having the best accuracy rate, the RF technique was shown to have the highest overfitting. The SVM method, on the other hand, was the most reliable as it did not alter during the research. Several machine learning algorithms were examined by Kok et al. [13] for their potential use in real estate evaluations. The RF technique and the GBR method were among these approaches. In terms of overall performance, the results showed that the XGBM algorithm was the best. If Ceh et al. [14] wanted to know which method would provide better price predictions, they contrasted the hedonic pricing model with the RF algorithm. A total of 7,407 residences in the Ljubljana region were polled by the writers between 2008 and 2013. It was evident from the results that the RF model performed better in terms of prediction (Slovenia). 5. Hu et al. [15] examined the accuracy of predicting the cost of renting a property in Shenzhen (China) using supervised learning algorithms. The study was conducted by the authors using several algorithms, including k-NN, ETR, SVR, RF, and multilayer perceptron neural network (MLP-NN). Based on the results, it seems that the RF and ETR algorithms were more reliable in their operations. An updated methodology for assessing the geographical proximity of London and better performance estimates for property prices was developed by S. Lu et al. [16]. Neither of these geographical characteristics can be measured in a linear fashion since they are non-linear. Predicting the future value of real estate in Saudi Arabia using an ANN-based model is the objective of Elham Alzain et al. [17]. The information was collected from Aqar in four major cities in Saudi Arabia: Riyadh, Jeddah, Dammam, and Al-Khobar. The results showed a high level of agreement between the predicted and experimental values. Using ANN for this yields an 80% success rate. P. Durganjali et al. [18] used classification algorithms to forecast houses' resale prices. Leaner regression, Decision Tree, K-Means, and Random Forest are among the categorization techniques used in this research to forecast a property's selling price. Many factors, including the

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

current economic climate, a home's location, and the home's physical characteristics, contribute to its final selling price. In this paper, we use these methods, identify the best model for predicting better outcomes using RMSE as the performance matrix across various datasets. To predict future home prices, Sifei Lu et al. [19] developed a hybrid regression method. This research puts the creative feature engineering technique to the test using a small dataset and data characteristics. A few of the most recent submissions to the "House Price: Advanced Regression 6 Methods" Kaggle competition used the proposed technique as its cornerstone. Given the readers' budgetary constraints, the objective of the essay is to provide clients with appropriate price estimates. In their study, Fletcher et al. [20] (2000) look at the question of whether aggregate or disaggregated data is better for hedonic analysis in home price forecasting. It turns out that the hedonic pricing coefficients of a number of attributes change drastically with respect to age, location, and property type. Nevertheless, it is thought that this might be adequately replicated using an aggregate approach. It is also possible to use the hedonic pricing model to assess the separate external effects of things like environmental variables on home prices. As an example, the hedonic pricing model has been used in several studies to assess the impact of noise and air pollution on property values. Chapter 2.1: The Issue Some simple and basic houses may be predicted using the present method. The accuracy of other algorithms may also be predicted using this method. It makes a price prediction based on only one parameter. The present training phase is not beneficial for the system in its current state. They can't foresee a lavish residence. Optimal algorithm suggestions are not possible [21]. Predictions cannot be made using certain criteria. At iteration t, the objective function that has to be minimized is the loss function with regularisation:

Real value (label) known from the training data-set

$$\mathcal{L}^{(t)} = \sum_{i=1}^{n} l(y_i, \hat{y}_i^{(t-1)} + f_t(\mathbf{x}_i)) + \Omega(f_t)$$

Can be seen as f(x + Δx) where x = $\hat{y}_i^{(t-1)}$

The XGBoost objective "cannot be maximized using typical optimization methods in Euclidean space," as said, as it is clearly a function of functions. For the function f(x simplest), the following is the basic linear approximation:

$$f(x) \approx f(a) + f'(a)\boxed{(x-a)}$$

$$\Delta x = f_t(\mathbf{x}_i)$$

A function of x is the sole valid choice for representing the initial function: Only in terms of x is the initial function expressible. Applying Taylor's theorem, one may find the simplest function of x at a given point and transform it into f(x). A function of the currently-selected tree (step t) was the objective function f(x), where x denoted the sum of the t CART trees, before the Taylor approximation was applied. Here, x represents the extra student that has to be added in step t, the loss function is represented by the equation f(x), and the expected result from step t-1 is an. By recasting the goal (loss) function as a direct function of the new learner added at each iteration, optimization in Euclidean space may be achieved using the previously given knowledge [22]. This paves the way for optimization in a space (t) that is defined by the geometry of the Earth. Predictions produced in stages (t-1) and (x-a) stand as the extra learner that has to be added in step (t) to eagerly lower the goal, as previously stated. You can't eagerly lower the goal without doing this. The second-order Taylor approximation is used when:

$$f(x) \approx f(a) + f'(a)(x-a) + \frac{1}{2} f''(a)(x-a)^2$$

$$\mathcal{L}^{(t)} \simeq \sum_{i=1}^{n} [l(y_i, \hat{y}^{(t-1)}) + g_i f_t(\mathbf{x}_i) + \frac{1}{2} h_i f_t^2(\mathbf{x}_i)] + \Omega(f_t)$$

XGBoost objective using second-order Taylor approximation-

Where:

$$g_i = \partial_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)}) \text{ and } h_i = \partial^2_{\hat{y}^{(t-1)}} l(y_i, \hat{y}^{(t-1)})$$

The loss function's first and second order gradient statistics-

At least, if we eliminate the constant components, we are left with the simplified objective to minimise at step t as follows:

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

$$\tilde{\mathcal{L}}^{(t)} = \sum_{i=1}^{n}[g_i f_t(\mathbf{x}_i) + \frac{1}{2}h_i f_t^2(\mathbf{x}_i)] + \Omega(f_t)$$

XGBoost simplified objective Our next objective is to identify a learner that minimises the loss function at iteration t since the aforementioned is a sum of straightforward quadratic functions of a single variable that can be minimised using well-known methods.

$$argmin_x\ Gx + \frac{1}{2}Hx^2 = -\frac{G}{H},\ H > 0 \quad \min_x\ Gx + \frac{1}{2}Hx^2 = -\frac{1}{2}\frac{G^2}{H}$$

Minimizing a simple quadratic function Notice how the following scoring function resembles the "basic quadratic function solution" above when using the authors' method to "assess the quality of a tree structure q":

$$\tilde{\mathcal{L}}^{(t)}(q) = -\frac{1}{2}\sum_{j=1}^{T}\frac{\overset{\text{instances mapped to leaf}_j}{\boxed{(\sum_{i\in I_j} g_i)^2}}}{\sum_{i\in I_j} h_i + \lambda} + \gamma T.$$

The tree learner structure q scoring function

$$y\ln(p) + (1-y)\ln(1-p) \text{ where } p = \frac{1}{(1+e^{-x})}$$

Classes are classified binaryly using the Cross Entropy loss function, where p is the probability score and y is the real label; both variables may take on values between 0 and 1 [23]. The output x of the model is the sum of all the CART tree learners. The first and second derivatives of the gradient and hessian with respect to x need to be determined if we are to minimize the log loss objective function. The formula for the hessian is (p-*) (1-p), while the formula for the gradient is (p-y), as you may have learned from this Stats Stack Exchange topic.

3. Proposed System

Predicting the property price with this technique is a safe and trustworthy method. It is able to forecast the value of a home in a desirable geographic area. For improved precision, it employs Extreme Gradient Boosting. It is simpler and more accurate to make predictions using this approach, which employs linear functions, likelihood, etc.

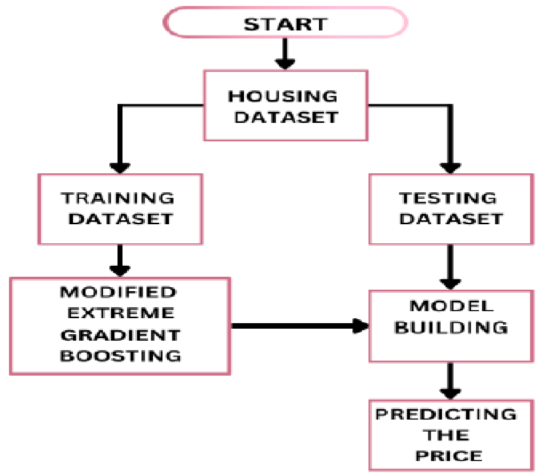

Figure 1. Flow Diagram

We pre-process the data by removing noise, detecting anomalies, and filling in missing values, among other things, when we get it from the database. Several methods are used to train the model on the pre-processed data, and then testing data is used to test the model's accuracy. Make use of a tweaked version of the extreme boosting technique to construct the model. Launch the program by linking the model to flask. Collect data entered by website users and then extract relevant keywords from that data. Make a prediction based on the provided data.

3.1. System Architecture

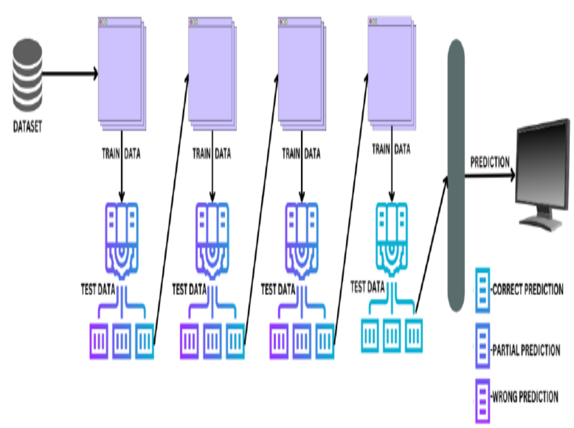

Figure 2. Proposed architecture

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

Algorithm 1. Modified Extreme Boosting

Initialization:

STEP 1: Given training information from the instance space S = (x1, y1),..., (m, ym) where xi EX and y; Y = -1,+1.

STEP 2: start the distribution off. $D_1(i)$ =.

Using the algorithm for t = 1,..., T, do

Utilizing the distribution Dt, train a weak learner ht: X R.

Zt is a normalisation factor chosen to create the distribution of Dt+1, and its formula is

Dt+1(i) = Dt(i)e-tyiht(xi) Zt.

f(x) = oatht(x) and H(x) = sign(f(x)) are the final scores.

3.2.1 Extreme Boosting Algorithm with Revised Steps A. Overarching Criteria quiet: The value is initially set to zero. You need to enter 0 to print ongoing messages and 1 to enter quiet mode. The default is booster: GBTree. It is necessary to specify the booster to be used: both GBTree and GBLinear, which are tree-based and linear, respectively. the XGBoost algorithm sets the buffer automatically; the user has to do nothing to change it. num feature: The value is automatically set by the XGBoost Algorithm; no human intervention is necessary. Parameters for the Booster B As a default, ETA is set to 0.3. When making an update, be sure to indicate the step size shrinkage that was used to prevent overfitting. The weights of newly added features are instantly available after each round of boosting. The boosting method is made more conservative by eta by lowering the feature weights. There is a range of 0 to 1. As the eta value decreases, the model's resistance to overfitting increases. The GAMMA variable has a default value of 0. The least loss reduction needed to establish another division must be indicated on a leaf node of the tree. The bigger the algorithm becomes, the more cautious it will be. The range is from zero to. The bigger the gamma, the more cautious the algorithm becomes. By default, the MAX DEPTH parameter is set to six. You need to tell me how deep a tree can go. 1 in order to finish the range. The minimum child weight is 1 by default. It is necessary to specify the absolute minimum instance weight (hessian) that a kid must have. After the tree partitioning process, a leaf node will emerge if its combined instance weight is smaller than the minimal child weight. More divisions will be abandoned by the building procedure. relates to the minimal number of instances needed for each node when using linear regression mode. The bigger the algorithm becomes, the more cautious it will be. The range is from zero to. Maximal Deletion: The default value is 0. The biggest interval that may be used to predict the mass of each tree. Setting the value to 0 removes all constraints. Changing it to a positive value may make the updating process more careful. This parameter isn't usually used, however logistic regression might benefit from it. particularly in cases when there is a dramatic disparity between the sexes. Changing it to a value between 1 and 10 could make control updating easier. The range is from zero to.

Table 1. Algorithm used and its function

ID	Model	Function
1	Linear Regression	sklearn.linear_model.LinearRegression
2	Random Forest Regressor	sklearn.ensemble.RandomForestRegressor
3	Gradient Boosting Regressor	sklearn.ensemble.GradientBoostingRegressor
4	Ridge Regression	sklearn.linear_model.Ridge
5	Lasso Regression	sklearn.linear_model.Lasso
6	Ada Boosting Regression	sklearn.ensemble.AdaBoostRegressor
7	Decision Tree Regression	sklearn.tree.DecisionTreeRegressor
8	Modified Extreme Boosting	XGBRegressor()

4. Result and Discussion The figure 3 shows the data and parameter present in the dataset. It also shows the columns and the rows in the csv file.

International Journal of Engineering Research and Science & Technology
ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

```
In [57]: housing_clean.sort_values(["price"], ascending=False)
```

Out[57]:

	area_type	availability	location	bath	balcony	price	bhk	sqft	price_per_sqft	sqft_per_bhk
12443	Plot Area	Ready To Move	Other	8	4	2600.0	4	4350.0	59770.114943	1087.5
6421	Plot Area	Soon to be Vacated	Bommenahalli	3	2	2250.0	4	2940.0	76530.612245	735.0
8398	Super built-up Area	Ready To Move	Bannerghatta Road	4	5	1400.0	5	2500.0	56000.000000	500.0
9535	Plot Area	Ready To Move	Indira Nagar	5	4	1250.0	4	2400.0	52083.333333	600.0
1299	Plot Area	Ready To Move	Chamrajpet	7	1	1200.0	9	4050.0	29629.629630	450.0
...	...	...	...	...	...	...	...	...	...	...
5410	Super built-up Area	Ready To Move	Attibele	1	1	10.0	1	400.0	2500.000000	400.0
11091	Built-up Area	Ready To Move	Attibele	1	1	10.0	1	410.0	2439.024390	410.0
7482	Super built-up Area	Ready To Move	Other	2	1	10.0	1	470.0	2127.659574	470.0
12579	Super built-up Area	Ready To Move	Chandapura	1	1	10.0	1	410.0	2439.024390	410.0
8594	Built-up Area	Ready To Move	Chandapura	1	1	9.0	1	450.0	2000.000000	450.0

12339 rows × 10 columns

Figure 3. Dataset Description

Figure 4 displays a graph including information from the dataset. Visual representations of data ratios include bar graphs and histograms. You can see the count of each region type in the dataset represented by the bar graph. The first histogram shows the density and count of sqft_per_bhk, which is the parameter's weight. Displayed in the second histogram is the price per square foot. The dataset's square footage is shown in the third histogram. The purchase price per square foot is shown in the fourth histogram. Dataset prices are shown via the last histogram.

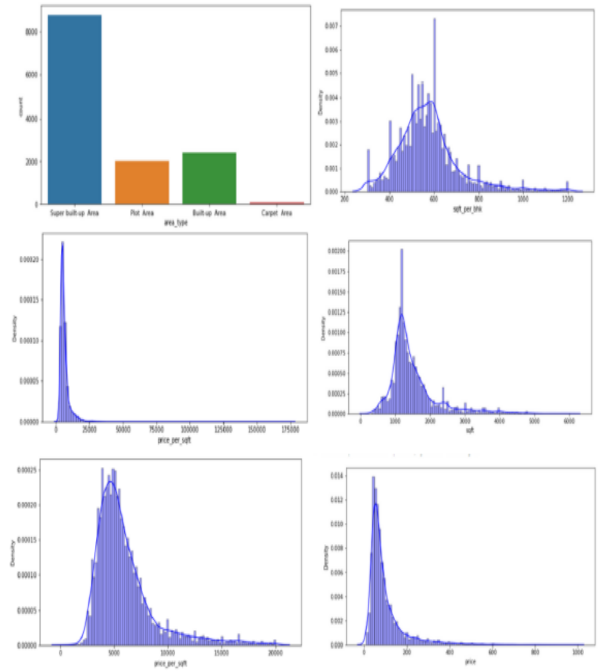

Figure 4. Data Graph

The figure 5 shows the various algorithm used and its accuracy. The algorithm used for Linear Regression, Ridge Regression, AdaBoost Regression, Decision Tree Regression, Random Forest Regression, Gradient Boosting Regression and Modified Extreme Boosting Algorithms are used to compare the accuracy with Modified XGBoost.

```
In [83]: lin_reg = LinearRegression()
         lin_reg.fit(X_train, y_train)
         lin_reg.score(X_test, y_test)

Out[83]: 0.7902192001533112

In [84]: ridge_reg = Ridge(alpha = 0.1)
         ridge_reg.fit(X_train, y_train)
         ridge_reg.score(X_test, y_test)

Out[84]: 0.7901996166492569

In [88]: ab_reg = AdaBoostRegressor(loss = "linear")
         ab_reg.fit(X_train, y_train)
         ab_reg.score(X_test, y_test)

Out[88]: 0.6924090517277011

In [89]: gb_reg = GradientBoostingRegressor(max_depth = 7, max_featu
         gb_reg.fit(X_train, y_train)
         gb_reg.score(X_test, y_test)

Out[89]: 0.6086057413153001

In [86]: dt_reg = DecisionTreeRegressor()
         dt_reg.fit(X_train, y_train)
         dt_reg.score(X_test, y_test)

Out[86]: 0.7447872870304995

In [87]: rf_reg = RandomForestRegressor()
         rf_reg.fit(X_train, y_train)
         rf_reg.score(X_test, y_test)

Out[87]: 0.8106514648362296

In [92]: xgb_reg = XGBRegressor()
         xgb_reg.fit(X, y)
         xgb_reg.score(X, y)

Out[92]: 0.9131660566803912
```

Figure 5. Accuracy

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

Table 2. Accuracy table

S. No.	Algorithm	Accuracy
1	Linear Regression	0.78021
2	Ridge Regression	0.79019
3	Lasso Regression	0.76216
4	Decision Tree (DT)	0.74478
5	Random Forest (RF)	0.81065
6	AdaBoost (AB)	0.69240
7	Gradient Boosting Tree (GB)	0.60806
8	Modified Extreme Boosting	0.82912

The figure 6 shows the comparison graph between the algorithms to show the performance and efficiency of the algorithms and it clearly shows that Modified XGBoost is more efficient than any other algorithms

```
In [1]: import matplotlib.pyplot as plt


        algorithms = ['Linear','Ridge','Lasso','DT','RF','AB','GB','XGBoost']
        accuracy = [0.79021, 0.79019, 0.76216, 0.74478, 0.81065, 0.69240, 0.608060, 0.91912]
        plt.plot(algorithms, accuracy)
        plt.xlabel("Algorithm")
        plt.ylabel("Accuracy")
        plt.title("Accuracy of Various Algorithms")
        plt.ylim(0, 1)
        plt.show()
```

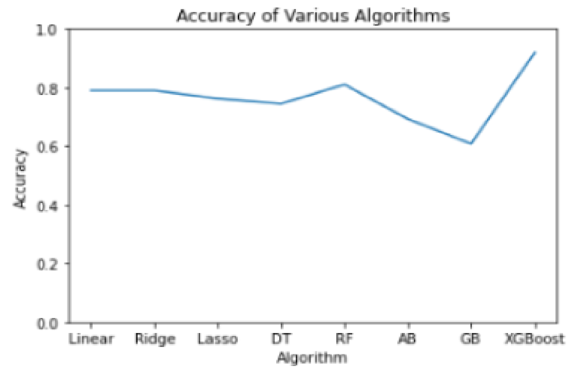

Figure 6. Accuracy graph

This figure 7 shows that front-end of the application which is running on the server 127.0.01.5000 using flask. It contains the details such as location, area, availability, square footage, BHK, bathrooms. The input needs to be entered according the user.

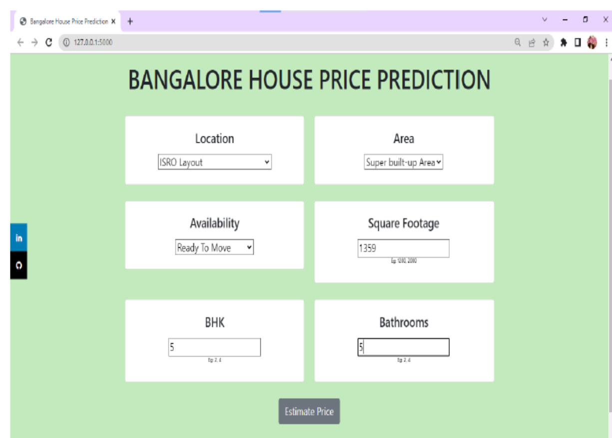

Figure 7. Running website on flask

International Journal of Engineering Research and Science & Technology

ISSN 2319-5991 www.ijerst.com

Vol. 21, Issue 2, 2025

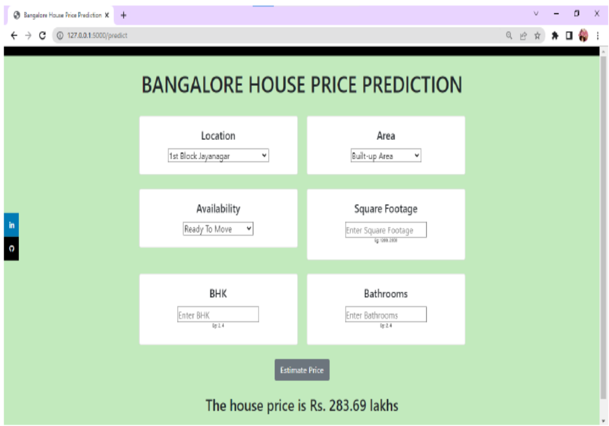

Figure 8. House price prediction

The figure 8 shows the prediction of house price for the given input.

5. Conclusion

The real estate industry may play a role in helping the economy grow by creating new employment. Owners and recipients of property are entangled in this situation. That is why accurate predictions of future property values are crucial. Many people keep a close eye on house value movements since they are a solid indicator of the state of the economy. One useful technique for controlling property usage and budgeting is a model that can predict future housing expenses. Predicting the future worth of real estate has several uses, including helping politicians establish fair prices and empowering owners and brokers to make educated judgments. In this study, we evaluate Bayesian Regression and other popular regression algorithms for their ability to forecast Bangalore housing prices. Given the large number of attributes and strong connection in the publicly accessible data, the results showed promise. Consequently, the local data need more attributes, preferably those that are highly correlated with property price. Nevertheless, XGBoost yielded the most favorable outcomes. Research shows that when compared to other prediction algorithms, Modified XGBoost performs better. Incorporating social media pricing data, Google Maps photos, user assessments of the property's features, and economic information might enhance ML projections in the future.

References

[1] Gallent, Nick, Phoebe Stirling, and Iqbal Hamiduddin. (2022). Pandemic mobility, second homes and housing market change in a rural amenity area during COVID-19–The Brecon Beacons National Park, Wales. Progress in Planning, 100731.

[2] Gajalakshmi, R. K., et al. (2020). An Optimized ASM based Routing Algorithm for Cognitive Radio Networks. International Conference on System, Computation, Automation and Networking (ICSCAN), IEEE.

[3] Padmini, S. V. (2022). Socio-economic conditions of slum dwellers in Karnataka-a case study in Tumkur district. Archers & Elevators Publishing House.

[4] Raj, S. Gokul, N. Srinath, and T. Ananth Kumar (2019). Real-Time Trespasser Detection Using GPS based UAV. IEEE International Conference on Innovations in Communication, Computing and Instrumentation (ICCI).

[5] Tokunaga, Robert S. (2010). Following you home from school: A critical review and synthesis of research on cyberbullying victimization. Computers in Human Behavior, 26(3): 277-287.

[6] Arunmozhiselvi, S., et al. (2022). A Systematic Approach to Agricultural Drones Using a Machine Learning Model. Machine Learning Approaches and Applications in Applied Intelligence for Healthcare Data Analytics. CRC Press, Pages 41-60.

[7] Nguyen, Quang Hung, et al. (2021). Influence of data splitting on performance of machine learning models in prediction of shear strength of soil. Mathematical Problems in Engineering, 2021: 1-15.

[8] Rajmohan, R., et al. (2022). G-Sep: A Deep Learning Algorithm for Detection of Long-Term Sepsis Using Bidirectional Gated Recurrent Unit.

[9] Ford, Elizabeth, et al. (2019). Identifying undetected dementia in UK primary care patients: a retrospective case-control study comparing machine-learning and standard epidemiological approaches. BMC Medical Informatics and Decision Making, 19(1): 1-9.

[10] Muthukumarasamy, Sugumaran, et al. (2020). Machine learning in healthcare diagnosis. Blockchain and Machine Learning for E-Healthcare Sys., 343.

ISSN 2319-5991 www.ijerst.com
Vol. 21, Issue 2, 2025

[11] Mair, Carolyn, et al. (2000). An investigation of machine learning based prediction systems. Journal of systems and software, 53(1): 23-29.

[12] Saravanan, V., et al. (2023). Design of deep learning model for radio resource allocation in 5G for massive IoT device. Sustainable Energy Technologies and Assessments, 56: 103054.

[13] Čeh, Marjan, et al. (2018). Estimating the performance of random forest versus multiple regression for predicting prices of the apartments. ISPRS International Journal of Geo-Information, 7(5): 168.

[14] Hong, Yao, Jianhong Kang, and Ceji Fu. (2022). Rapid prediction of mine tunnel fire smoke movement with machine learning and supercomputing techniques. Fire Safety Journal, 127: 103492.

[15] Lu, Binbin, et al. (2014). Geographically weighted regression with a non-Euclidean distance metric: a case study using hedonic house price data. International Journal of Geographical Information Science, 28(4): 660-681.

[16] Alzain, Elham, et al. (2022). Application of Artificial Intelligence for Predicting Real Estate Prices: The Case of Saudi Arabia. Electronics, 11(21): 3448.

[17] Durganjali, P., and M. Vani Pujitha. (2019). House resale price prediction using classification algorithms. International Conference on Smart Structures and Systems (ICSSS), IEEE.

[18] Lu, Sifei, et al. (2017). A hybrid regression technique for house prices prediction. International Conference on Industrial Engineering and Engineering Management (IEEM), IEEE.

[19] Prasanalakshmi, B. (2022). Deep Regression hybridized Neural Network in human stress detection. International Conference on Smart Technologies and Systems for Next Generation Computing (ICSTSN), IEEE.

[20] Fletcher, Mike, Jean Mangan, and Emily Raeburn. (2004). Comparing hedonic models for estimating and forecasting house prices. Property Management, (3): 189-200.

[21] Hallegatte, Stéphane, et al. (2012). Investment decision making under deep uncertainty-application to climate change. World Bank Policy Research Working Paper – 6193.

[22] Sun, Shiliang, et al. (2019). A survey of optimization methods from a machine learning perspective. IEEE Transactions on Cybernetics, 50(8): 3668-3681.

[23] Coenen, Lize, et al. (2022). Machine learning methods for short-term probability of default: A comparison of classification, regression and ranking methods. Journal of the Operational Research Society, 73(1): 191-206.