

# International Journal of Information Technology, Research and Applications (IJITRA)

Renju K, Freni S, (2024). An Ensemble Approach for Predicting the Price of Residential Property, 3(2), 27-38.

ISSN: 2583 5343

DOI: 10.59461/ijitra.v3i2.99

The online version of this article can be found at:  
<https://www.ijitra.com/index.php/ijitra/issue/archive>

Published by:  
PRISMA Publications

IJITRA is an Open Access publication. It may be read, copied, and distributed free of charge according to the conditions of the Creative Commons Attribution 4.0 International license.

**International Journal of Information Technology, Research and Applications (IJITRA)** is a journal that publishes articles which contribute new theoretical results in all the areas of Computer Science, Communication Network and Information Technology. Research paper and articles on Big Data, Machine Learning, IOT, Blockchain, Network Security, Optical Integrated Circuits, and Artificial Intelligence are in prime position.



<https://www.prismapublications.com/>

Journal homepage: <https://ijitra.com>

# An Ensemble Approach for Predicting the Price of Residential Property

Renju K<sup>1</sup>, Freni S<sup>2</sup>

<sup>1,2</sup> Department of Computer Science, Mount Carmel College, Autonomous,  
Bengaluru, Karnataka, India, 560052

## Article Info

### Article history:

Received April. 01, 2024  
Accepted June 04, 2024  
Published June. 15, 2024

### Keywords:

Web Scraping  
Ensemble Strategy  
Random Forest  
XGBoost  
Support Vector Regression

## ABSTRACT

Today, determining the rent for a property is crucial given that the cost of housing increases annually. Our future generation requires a straightforward method to forecast future property rent. Various factors influence the price of a house, including its physical condition, location, and size. This study utilizes web scraping techniques to collect data from pertinent websites for analytical and predictive purposes. Employing an ensemble strategy, the research predicts housing rents in Bangalore. Seven ensemble models of machine learning algorithms, such as Random Forest, XGBoost, Support Vector Regression (SVR), and Decision Trees, are integrated into the analysis. The objective was to determine the optimal model by evaluating their performance scores obtained from a comparative analysis.

This is an open access article under the [CC BY-SA](https://creativecommons.org/licenses/by-sa/4.0/) license.



## Corresponding Author:

Freni S  
Department of Computer Science,  
Mount Carmel College, Autonomous,  
Bengaluru, India  
E-mail: P.BAVITHRA.MATHARASI@mccblr.edu.in

## 1. INTRODUCTION

This research paper emphasizes on the dynamic real estate market of Bangalore, a city that has swiftly emerged as Asia's fastest-growing metropolis. Bangalore's remarkable growth trajectory has surpassed not only the economic growth rate of its own nation but has also solidified its reputation as the Silicon Valley of India, largely propelled by its booming IT sector [4]. This exponential expansion has been further fueled by robust social infrastructure, esteemed educational institutions, and rapid advancements in physical infrastructure, presenting a myriad of challenges and opportunities for both tenants and landlords.

Despite the escalating demand for rental properties, the existing body of literature has largely overlooked the development of predictive models for house rent prices, despite the acknowledged urgency in addressing this issue. Machine learning has become a highly effective prediction tool, especially with the rise of Big Data, enabling more accurate forecasts of property prices based solely on property characteristics, rather than relying on past data. While some studies have showcased the effectiveness of machine learning, the majority have focused solely on individual algorithmic models, neglecting the potential benefits of combining multiple algorithms into ensembles.

In this study, we embark on a comprehensive exploration of machine learning techniques, with a particular emphasis on ensemble learning, to craft an accurate predictive model for house rent prices in Bangalore. Our evaluation and comparison of various algorithms and methodologies seek to identify the most effective approach for meeting the urgent need for precise rent price prediction in Bangalore's thriving real estate market. By leveraging features such as area name, square footage, number of bedrooms, and area type, Our research endeavors to offer invaluable insights and tools not only to stakeholders in the real estate sector but also to newcomers to Bangalore. By doing so, we aim to facilitate informed decision-making in this dynamic environment, empowering individuals to navigate the complexities of the property market with confidence [8].

## 2. LITERATURE REVIEW

The projection of house rent prices in Bangalore has become increasingly vital in the backdrop of its rapid urbanization and burgeoning real estate market[9]. Machine learning, particularly ensemble techniques, has emerged as a promising approach for predicting house rent prices, leveraging the strengths of multiple algorithms to enhance predictive accuracy. This literature survey aims to explore existing research endeavors focused on predicting house rent prices in Bangalore using ensemble methods. By reviewing previous studies and methodologies, this survey seeks to identify trends, challenges, and advancements in the field, paving the way for further research and development in this critical domain of real estate analytics[10].

Vivek Singh Rana et.al [11] have explored various regression techniques for predicting house prices. This study has highlighted the importance of factors such as physical condition, location, amenities, and interest rates in determining house prices. Different algorithms, including Support Vector Regression, XGBoost, Decision Tree Regression, and Random Forest, have been compared to identify the most effective method for accurately predicting the values of houses.

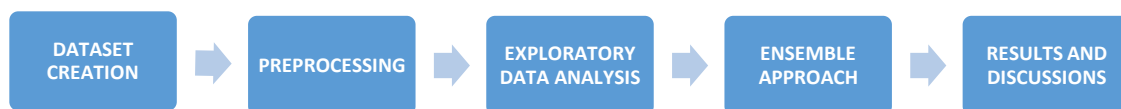
Sanit Kumar et.al [12] provides an overview of existing research on web scraping tools and techniques, emphasizing the importance of extracting valuable data efficiently from the internet. It discusses the challenges of dealing with unstructured web data and highlights the significance of using tools like BeautifulSoup, requests, Selenium, and Pandas for effective web scraping. The review also mentions the role of PyCharm IDE in creating Python programs for web scraping and the use of Entity-Relationship Diagrams to represent data structures.

J. Avanijaa et.al [13] shows the prediction of house prices using the XGBoost regression algorithm. It discusses the importance of factors such as location, neighborhood, and amenities in predicting house prices accurately. The study emphasizes the significance of data preprocessing in developing the prediction model. Additionally, the paper highlights the use of ensemble learning with base learners to achieve a single prediction.

## 3. METHODOLOGY

Explaining research in chronological order, including research design, research technique (as algorithms, pseudocode, or otherwise), how to test, and data gathering [4][5]. The summary of the research course should be accompanied by references, so that the explanation can be accepted scientifically [6]. Figures 1 show Overview of Prediction Modeling.

The research follows a systematic approach consisting of five key steps, as illustrated in Figure 1.[3].



**Figure 1.** Experiment Architecture

### a. Dataset Creation

This dataset was acquired to anticipate house prices in Bangalore by employing a technique known as web scraping. This involves automatically retrieving information from websites. To facilitate this process, we utilized a Python library called BeautifulSoup[15].

BeautifulSoup is a Python tool that assists in parsing HTML and XML documents. It parses the HTML structure of a webpage and enables us to search for particular HTML tags, attributes, or text strings within it. Once identified, we can extract the desired data from these elements.

In this research, we utilized BeautifulSoup to navigate through the pages of NoBroker.com, a property listing website. We targeted HTML elements containing information such as BHK, area size, latitude, longitude, total size in square feet, deposit amount (in Rs), rental amount (in Rs), furnishing status, property age, availability status, and information regarding immediate possession[12]. By automating the process with BeautifulSoup, we were able to efficiently gather a substantial amount of property data from NoBroker.com. This dataset serves as the basis for our investigation into predicting house prices in Bangalore. This dataset

comprises of 11,524 records and incorporating 10 features initially[2]. Refer to Table 2 for the dataset's features and their respective descriptions.[6]

**TABLE 1 Features and their datatype**

Features	Description	Data Type
BHK	Number of bedrooms	Numerical
Area_name	Area of the house	Non-numerical
Area_size	Size of the property area	Numerical
Area_type	Type of area (e.g., Urban, Rural)	Non-numerical
Latitude	Geographic latitude coordinates.	Numerical
Longitude	Geographic longitude coordinates.	Numerical
Rent	Rental price of the property	Numerical
Furnishing	Furnishing status of the property.	Non-numerical
Property age	Age of the property	Numerical
Availability status	Status of property availability.	Non-numerical

#### b. Preprocessing

Initially, non-numerical features such as Area\_name, Area\_type, and ventilation underwent transformation into numerical features utilizing techniques like One Hot Encoding and Label Encoding available in the scikit-learn library. Moreover, less correlated features such as latitude, longitude, and availability were removed from the dataset, resulting in a reduction from 11 features to 9 features, as depicted in Table 3.

Addressing missing values represented another critical step in the preprocessing phase. Empty cells within the dataset were replaced with the mean of the respective column, employing the SimpleImputer function from the scikit-learn library. This strategic approach ensured the dataset's completeness and suitability for further analysis [16].

Additionally, other preprocessing techniques were employed to enhance the quality of the dataset was acquired to enhance the performance of subsequent modeling steps. These techniques may include data normalization or scaling to ensure all features are on a comparable scale, outlier detection and removal to mitigate the impact of extreme values, feature engineering to create new informative features, and dimensionality reduction techniques such as principal component analysis (PCA) to reduce the complexity of the dataset while retaining important information. Each of these preprocessing techniques plays a crucial role in preparing the dataset for accurate and reliable predictive modeling.

**TABLE 2. Features after pre – processing**

Features
1. BHK
2. Area_name
3. Area_size
4. Area_type
5. Deposit
6. Rent
7. Furnishing
8. Property age

### c. Exploratory Data Analysis

Before deploying any model, it's crucial to validate the accuracy and appropriateness of the dataset for analysis. To accomplish this, we conducted a thorough Exploratory Data Analysis (EDA), delving into the dataset's features, attributes, and their relationships.

Indicating the presence and strength of correlation between these variables. The correlation coefficient, a numerical indicator ranging from +1 to -1, elucidates the extent of association between two variables: a positive coefficient denotes a positive relationship, a negative coefficient signifies a negative relationship, and a coefficient of 0 indicates independence between variables. Moreover, we constructed a relationship matrix, as shown in Figure 3.5, to delve deeper into the interrelationships among six variables, thereby providing a comprehensive understanding of their associations[18], Figure 3.5



Figure 2 Correlation Matrix

### d. Ensemble Approach

After completing the dataset analysis, we refined and adjusted it based on our insights. Our dataset involves 6 pairs of four algorithms: XG Boost, SVR, Decision Tree Regression, and Random Forest[19]. We plan to utilize an ensemble approach across these pairs.

Ensemble learning, a methodology in machine learning, involves combining predictions from several individual models to improve overall performance. A prevalent technique within ensemble learning is averaging, where predictions from each model are averaged to generate the ultimate prediction. Each model within the ensemble operates autonomously to make predictions, which are subsequently averaged to derive the final ensemble prediction. This approach serves to mitigate the influence of individual model biases and errors, thereby enhancing predictive accuracy.[20].

In our study, we will implement ensemble approach across the 6 pairs of algorithms. This strategy aims to harness the strengths of different algorithms and mitigate their weaknesses and reliability of our predictions. Ensemble approach across the following 6 pairs of algorithms:

- i. XG Boost and Decision Tree Regression.
- ii. Decision Tree Regression and Random Forest.
- iii. Random Forest and XG Boost.
- iv. SVR and XG Boost.
- v. SVR and Decision Tree Regression.
- vi. SVR and Random Forest.

### i. XG Boost And Decision Tree Regression.

An ensemble technique known as "averaging" is utilized to merge predictions from two distinct models: XGBoost and Decision Tree Regression[21]. Through this approach, the individual predictions are combined by averaging them, thereby enhancing the overall prediction's reliability. This technique helps counteract potential biases and errors inherent in each individual model, leading to a more robust and accurate prediction.

Mathematically, the ensemble prediction  $\hat{Y}_{ensemble}$  is computed as follows:

$$\hat{Y}_{ensemble} = \frac{1}{2} (\hat{Y}_{XGB} + \hat{Y}_{DT})$$

Where:

- $\hat{Y}_{ensemble}$  represents the ensemble prediction.
- $\hat{Y}_{XGB}$  denotes the prediction from the XGBoost model.
- $\hat{Y}_{DT}$  denotes the prediction from the Decision Tree Regression model.

(1)

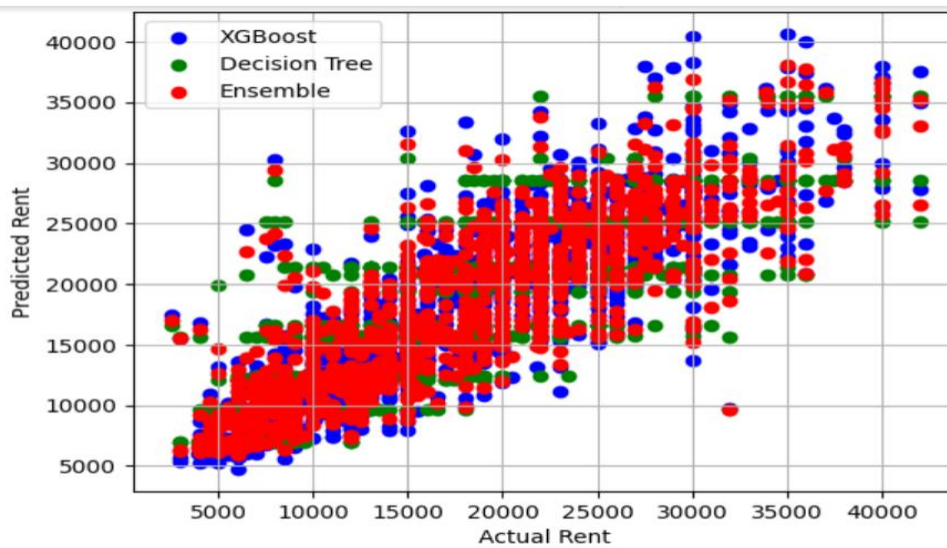


Figure 3

- Each data point in the plot corresponds to one observation from the test set.
- The x-axis represents the actual Rent values.
- The y-axis represents the predicted Rent values.
- Predictions made by the XGBoost model are denoted by blue points.
- Predictions made by the Decision Tree model are denoted by green points.
- Predictions made by the ensemble model are denoted by red points.
- A gray dashed line signifies the perfect prediction line, where actual and predicted values are identical.

### ii. Decision Tree Regression and Random Forest

A technique known as Stacking Ensemble is utilized in this predictive model. Stacking involves the combination of predictions from multiple base models through a two-stage process. Firstly, diverse base models are trained on the dataset. Then, a meta-model is trained using the predictions made by these base models as input features. The meta-model learns to make the final prediction based on the combined outputs of the base models.

**Mathematical Representation:** Mathematically, the ensemble prediction  $\hat{Y}_{ensemble}$  is calculated

using the predictions  $\hat{Y}_{base_1}, \hat{Y}_{base_2}, \dots, \hat{Y}_{base_n}$  from the base models:

$$\hat{Y}_{ensemble} = f(\hat{Y}_{base_1}, \hat{Y}_{base_2}, \dots, \hat{Y}_{base_n})$$

(2)



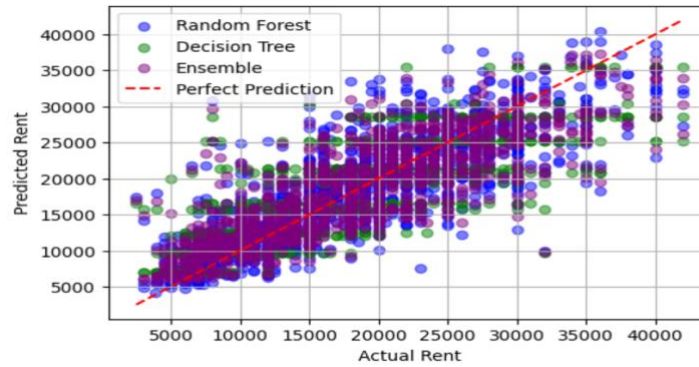


Figure 4

- Each data point on the graph corresponds to an observation from the test dataset.
- The x-axis represents actual rent values, while the y-axis displays predicted rent values.
- Blue points indicate predictions made by the RandomForest model.
- Green points depict predictions from the DecisionTree model[21].
- Purple points represent predictions generated by the ensemble model (averaging RandomForest and DecisionTree predictions)[22].

### iii. Random Forest and XG Boost

This model is independently trained on various subsets of the training data. For instance, employing BaggingRegressor with RandomForestRegressor as the base estimator involves training multiple RandomForestRegressor models, each on a distinct subset of the training data. Similarly, utilizing another BaggingRegressor with XGBRegressor as the base estimator entails training multiple XGBoost models on diverse data subsets.[23].

After training, predictions from these individual models are combined. For instance, predictions from Bagging RandomForestRegressor and Bagging XGBRegressor are averaged to create combined predictions. Refer figure 3.8. This ensemble approach aims to leverage the diverse perspectives of multiple models to produce more accurate and robust predictions compared to any single model alone.

Mathematically,

- $M$  as the number of base models (Random Forest and XGBoost models).
- $N$  as the number of instances of each base model (i.e., the number of subsets on which each base model is trained).
- $f_i(x)$  as the prediction of the  $i^{th}$  base model for input  $x$ .
- $\frac{1}{N}$  as the weighting factor for each base model's prediction (since they're averaged).

With these notations, the ensemble model's prediction  $\hat{y}(x)$  for input  $x$  can be expressed as:

$$\hat{y}(x) = \frac{1}{M} \sum_{i=1}^M \left( \frac{1}{N} \sum_{j=1}^N f_i(x) \right)$$

(3)

Combined Bagging RandomForestRegressor and XGBRegressor: Actual vs Predicted Rent

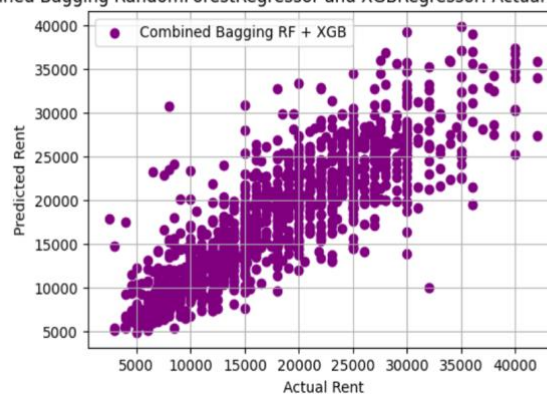


Figure 5

#### iv. SVR and XG Boost

This model demonstrates ensemble learning through a technique called stacking. Two different types of models are trained: XGBoost and SVR. XGBoost is a gradient boosting algorithm, while SVR is a kernelized Support Vector Regressor.

In stacking, the predictions of these base models (XGBoost and SVR) are combined using another model called the final estimator. Here, an XGBoost model is chosen as the final estimator.

The stacking model is trained on the scaled training data. It learns to combine the predictions from the base models to produce an ensemble prediction (fig 3.9)

Mathematically,

- $f_{XGBoost}(x)$  as the prediction of the XGBoost model for input  $x$ .
- $f_{SVR}(x)$  as the prediction of the SVR model for input  $x$ .
- $\hat{Y}_{stacking}(x)$  as the prediction of the stacking model for input  $x$ .

The ensemble model's prediction  $\hat{Y}_{stacking}(x)$  for input  $x$  can be expressed as:

$$\hat{Y}_{stacking}(x) = f_{final\_XGBoost}([f_{XGBoost}(x), f_{SVR}(x)])$$

Here,  $f_{final\_XGBoost}$  is the final XGBoost model which combines the predictions of the base models  $f_{XGBoost}(x)$  and  $f_{SVR}(x)$ .

(4)

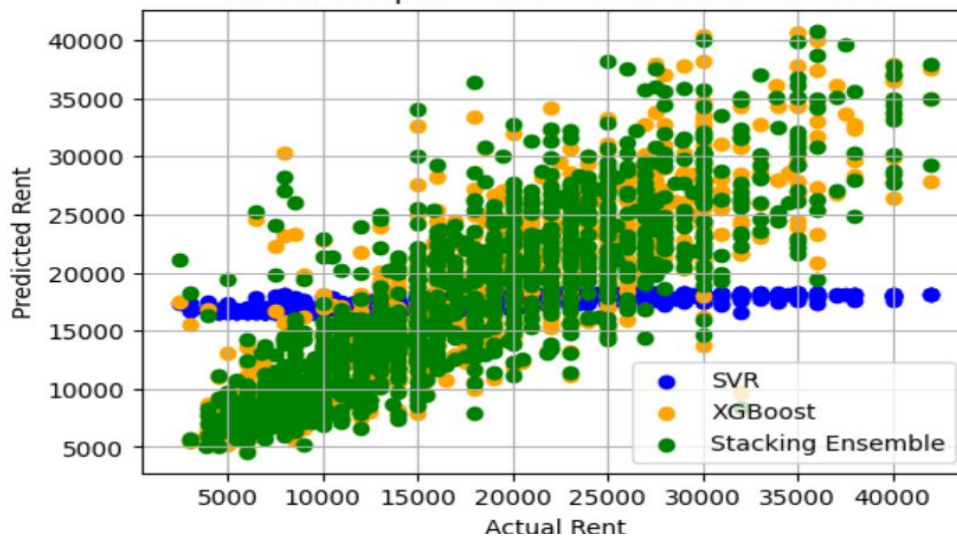


Figure 6

#### v. SVR and Decision Tree Regression

In this model, Ensemble is achieved by averaging predictions from SVR and Decision Tree models. This ensemble technique aims to leverage diverse modeling approaches to potentially improve prediction accuracy.

However, this ensemble model might not be good for prediction due to the following reasons:

- Both SVR and Decision Tree models may capture different aspects of the data. When combined, their predictions might not necessarily complement each other.
- Averaging predictions assumes that both models are equally reliable, which might not always be the case. If one model significantly outperforms the other, averaging could dilute the strengths of the better-performing model.
- Ensemble techniques like averaging are effective when individual models have diverse behaviors or when some models tend to overfit while others underfit. In this case, the models might have similar behavior or limitations, leading to limited improvement through ensemble averaging.(fig 3.10)



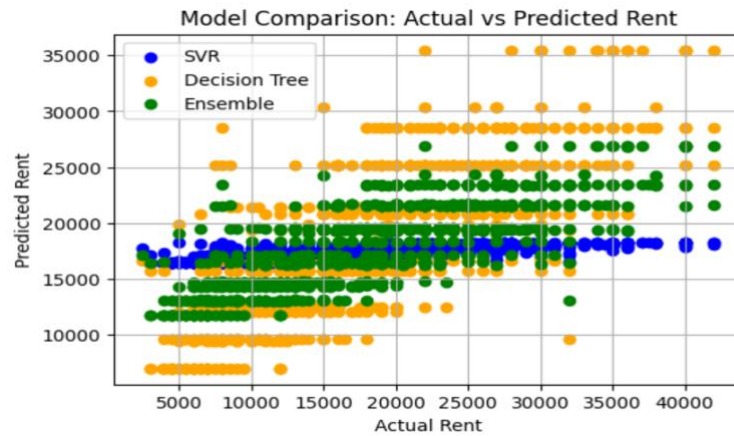


Figure 7

#### vi. SVR and Random Forest

This model implements ensemble learning techniques, specifically AdaBoost, Bagging, and Stacking, for regression tasks using scikit-learn in Python.

- **AdaBoost (Boosting):** AdaBoostRegressor is employed with RandomForestRegressor as the base estimator. AdaBoost iteratively corrects the mistakes of the weak models by adjusting the weights of training instances. It combines the predictions of multiple weak learners to build a strong model[13].
- **Bagging:** BaggingRegressor is used with Support Vector Regressor (SVR) as the base estimator. Bagging (Bootstrap Aggregating) builds multiple models (often of the same type) from different subsamples of the training dataset. It reduces variance and helps to avoid overfitting.
- **Stacking:** StackingRegressor is utilized, which combines multiple regression models (SVR and RandomForestRegressor) using another regressor (RidgeCV in this case) as the final estimator. Stacking combines the predictions of base estimators, often referred to as level-0 models, to build a final model (level-1 model).

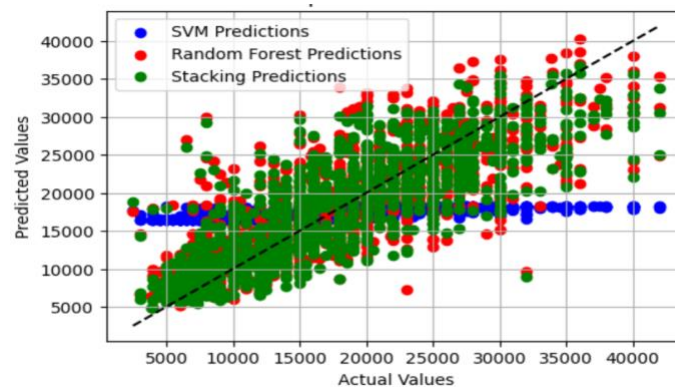


Figure 8

#### vii. RandomForestRegressor, GradientBoostingRegressor and DecisionTreeRegressor

In the analysis, various combinations of machine learning algorithms including XGBoost (XGB), Support Vector Regressor (SVR), Decision Tree, and Random Forest were experimented with. Among these, the combination of XGB and Decision Tree yielded the highest accuracy of 75.9%.

To further enhance the predictive performance, an ensemble stacking technique was employed. The process involved combining the best-performing algorithms identified earlier, namely RandomForestRegressor, GradientBoostingRegressor, VotingRegressor, BaggingRegressor, and DecisionTreeRegressor, in a stacking ensemble[24].

In this approach, the algorithms that demonstrated the highest accuracy, XGB and Decision Tree, were chosen for inclusion in the stacking ensemble. Subsequently, StackingRegressor from scikit-learn was utilized as the

ensemble technique. This method enables the combination of multiple regression models by employing a meta-regressor to predict the target variable.

Each base model (RandomForestRegressor, GradientBoostingRegressor, VotingRegressor, BaggingRegressor, DecisionTreeRegressor) was independently trained on the training dataset. The predictions made by these base models on the training dataset were then aggregated to serve as features (meta-features) for training the meta-regressor[25] figure 3.12

It was discovered that the stacking ensemble achieved an increased accuracy of 80.4%, surpassing the previous best accuracy of 75.9%. This approach leveraged the strengths of multiple well-performing algorithms to further enhance the accuracy of house rent price prediction by capturing a broader range of patterns in the data[14].

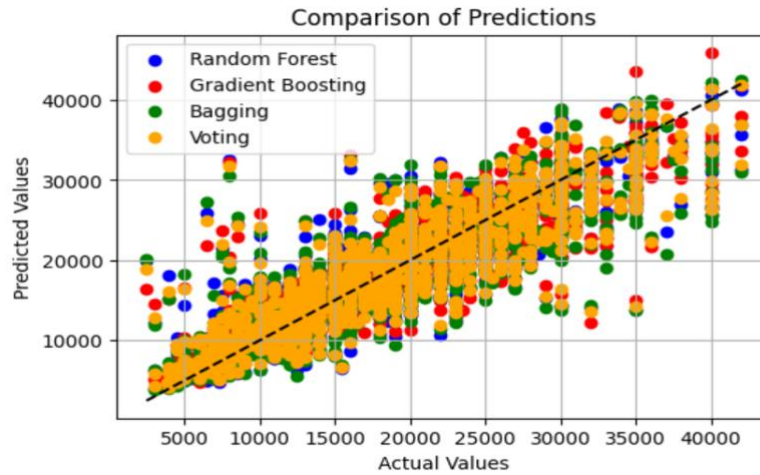


Figure 9

**4. RESULT AND DISCUSSIONS**

The experimental assessment focused on evaluating the efficacy of diverse ensemble models by employing RMSE and accuracy percentage as performance metrics.

In previous research, an in-depth analysis was conducted on individual algorithms including XG Boost, Decision Tree Regression, Random Forest, and Support Vector Regressor (SVR) to discern the most effective approach. Refer Table 4 [3].

However, in this study, an innovative strategy was adopted to further enhance prediction accuracy. Specifically, an ensemble approach was employed which amalgamated these individual models to leverage their complementary strengths and mitigate weaknesses (Refer Table 5). This novel approach sought to capitalize on the diverse perspectives and modeling techniques offered by each algorithm, ultimately aiming to outperform the performance of any single model alone.

**TABLE 3 Individual Experimental Analysis**

<i>Models</i>	<i>Test Accuracy</i>	<i>Training Accuracy</i>	<i>RSME</i>	<i>MSE</i>	<i>MAE</i>
Decision Tree Regression	0.29	0.99	119.56	14294.38	39.67
Random Forest	0.61	0.94	88.68	7864.09	32.09
SVR	0.31	0.29	118.14	13956.22	41.29
XGBoost	0.63	0.90	85.69	7342.85	32.30

**TABLE 4 Ensemble Experimental Analysis**

Ensemble models	RMSE	Accuracy in Percentage
XG Boost and Decision Tree Regression	5025.52	75.8%
Random Forest and XG Boost	4492.94	75.5%
SVR and XG Boost	4708.94	74.4%
Decision Tree Regression and Random Forest	4350.94	75.1%
SVR and Decision Tree Regression	5632.57	69.3%
SVR and Random Forest	4626.22	74.6%
RandomForestRegressor, GradientBoostingRegressor and DecisionTreeRegressor	3398.38	80.3%

The results suggest that certain ensemble combinations, such as RandomForestRegressor, GradientBoostingRegressor, and DecisionTreeRegressor, outperformed others in terms of both RMSE and accuracy. However, the performance varied depending on the specific combination of models used. These findings underscore the importance of carefully selecting ensemble models to achieve optimal performance in regression tasks.

## 5. CONCLUSION

In conclusion, our study underscores the significance of ensemble techniques in predicting residential property rents accurately. Through a meticulous exploration of diverse machine learning algorithms and ensemble methodologies, we have contributed to advancing predictive model in the realm of real estate. Notably, the ensemble approach, particularly leveraging combinations of RandomForestRegressor, GradientBoostingRegressor, and DecisionTreeRegressor, showcased superior performance, achieving an accuracy of 80.3%. Our findings highlight the necessity of tailored algorithmic combinations to address the complexities of the property market effectively. Moving forward, further research could delve into refining ensemble techniques and incorporating additional data sources to enhance predictive accuracy and robustness in forecasting residential property rents.



## REFERENCES

- [1] Z. Cai and Y. Zhao, "House Rent Analysis with Linear Regression Model-A Case Study of Six Cities in India," 2023. [Online]. Available: <https://www.magicbricks.com>
- [2] S. Kumar, J. Thakur, D. Ekka, and I. Sahu, "Web Scraping Using Python," *International Journal of Advances in Engineering and Management (IJAEM)*, vol. 4, p. 235, 2022, doi: 10.35629/5252-0409235237.
- [3] V. S. Rana, J. Mondal, A. Sharma, and I. Kashyap, "House Price Prediction Using Optimal Regression Techniques," in *Proceedings - IEEE 2020 2nd International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 203–208. doi: 10.1109/ICACCCN51052.2020.9362864.
- [4] P. Zhao chan, "A Case Study of Six Cities in India," 2023. [Online]. Available: <https://www.magicbricks.com>

- [5] J. L. Alfaro-Navarro, E. L. Cano, E. Alfaro-Cortes, N. Garcia, M. Gamez, and B. Larraz, "A Fully Automated Adjustment of Ensemble Methods in Machine Learning for Modeling Complex Real Estate Systems," *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/5287263.
- [6] D. Kumar Rao, N. Panini Challa, N. Kumar Rao, B. Naseeba, and S. Chakrvarthi, "WEB SCRAPING (IMDB) USING PYTHON", [Online]. Available: <https://www.researchgate.net/publication/368575445>
- [7] G. Uzut, S. Buyrukoglu, and R. Asst, "Prediction of real estate prices with data mining algorithms PREDICTION OF REAL ESTATE PRICES WITH DATA MINING ALGORITHMS Ömmu Gülsüm UZUT Selim BUYRUKOĞLU," 2020.
- [8] "Housing cost estimation of banglore".
- [9] M. Satish, "Bangalore House Price Prediction," 2021. [Online]. Available: <https://www.researchgate.net/publication/354403038>
- [10] K. Abebe, "Article ID: IJARET\_12\_04\_029 Bangalore City)," *International Journal of Advanced Research in Engineering and Technology (IJARET)*, vol. 12, no. 4, pp. 270–279, 2021, doi: 10.34218/IJARET.12.4.2021.029.
- [11] V. S. Rana, J. Mondal, A. Sharma, and I. Kashyap, "House Price Prediction Using Optimal Regression Techniques," in *Proceedings - IEEE 2020 2nd International Conference on Advances in Computing, Communication Control and Networking, ICACCCN 2020*, Institute of Electrical and Electronics Engineers Inc., Dec. 2020, pp. 203–208. doi: 10.1109/ICACCCN51052.2020.9362864.
- [12] S. Kumar, J. Thakur, D. Ekka, and I. Sahu, "Web Scraping Using Python," *International Journal of Advances in Engineering and Management (IJAEM)*, vol. 4, p. 235, 2022, doi: 10.35629/5252-0409235237.
- [13] P. Kora, J. Avanija, G. Sunitha, K. Reddy Madhavi, R. Hitesh, and S. Vittal E A Associate, "Prediction of House Price Using XGBoost Regression Algorithm," 2021. [Online]. Available: <https://www.researchgate.net/publication/350810698>
- [14] N. Ragapriya, T. A. Kumar, R. Parthiban, P. Divya, S. Jayalakshmi, and D. R. Raman, "Machine Learning Based House Price Prediction Using Modified Extreme Boosting," *Asian Journal of Applied Science and Technology*, vol. 07, no. 01, pp. 41–54, 2023, doi: 10.38177/ajast.2023.7105.
- [15] D. Kumar Rao, N. Panini Challa, N. Kumar Rao, B. Naseeba, and S. Chakrvarthi, "WEB SCRAPING (IMDB) USING PYTHON", [Online].
- [16] "BANGALORE RESIDENCE PRICE FORECASTER," *International Research Journal of Modernization in Engineering Technology and Science*, Jul. 2023, doi: 10.56726/irjmets43077.
- [17] R. Gupta Professor and N. N. S Professor, *Machine Learning based Predicting House Prices using Regression Techniques*. 2020.
- [18] "House rent estimation of banglore".
- [19] W. K. O. Ho, B. S. Tang, and S. W. Wong, "Predicting property prices with machine learning algorithms," *Journal of Property Research*, vol. 38, no. 1, pp. 48–70, 2021, doi: 10.1080/09599916.2020.1832558.
- [20] J. L. Alfaro-Navarro, E. L. Cano, E. Alfaro-Cortes, N. Garcia, M. Gamez, and B. Larraz, "A Fully Automated Adjustment of Ensemble Methods in Machine Learning for Modeling Complex Real Estate Systems," *Complexity*, vol. 2020, 2020, doi: 10.1155/2020/5287263.
- [21] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

- [22] Vignesh M, Vijay V, Krishna S, Sathyamoorthy K, and Student, "HOUSE PRICE PREDICTION USING MACHINE LEARNING BY RANDOM FOREST ALGORITHM," *www.irjmets.com @International Research Journal of Modernization in Engineering*, 1374, [Online]. Available: [www.irjmets.com](http://www.irjmets.com)
- [23] N. Ragapriya, T. A. Kumar, R. Parthiban, P. Divya, S. Jayalakshmi, and D. R. Raman, "Machine Learning Based House Price Prediction Using Modified Extreme Boosting," *Asian Journal of Applied Science and Technology*, vol. 07, no. 01, pp. 41–54, 2023, doi: 10.38177/ajast.2023.7105.
- [24] Z. Cai dan, "House Rent Analysis with Linear Regression" 2023. [Online].
- [25] B. Charbuty and A. Abdulazeez, "Classification Based on Decision Tree Algorithm for Machine Learning," *Journal of Applied Science and Technology Trends*, vol. 2, no. 01, pp. 20–28, Mar. 2021, doi: 10.38094/jastt20165.

### BIOGRAPHIES OF AUTHORS .

	<p><b>Renju K</b> Received a master's degree in computer application and master of philosophy. She has worked for 17 years as Assistant Professor at Mount Carmel College Autonomous, Bengaluru, Karnataka, India. She holds the position of Head of the Department, Department of Computer Science at Mount Carmel College Autonomous, Bengaluru, Karnataka, India. Her research interests include deep learning, machine learning. Published many papers in various national and international conferences and Journals. She can be contacted at email: <a href="mailto:renju.k@mccbbl.edu.in">renju.k@mccbbl.edu.in</a>.</p>
	<p><b>Freni S</b>, Research Scholar, Department of Computer Science, Mount Carmel College Autonomous, Bengaluru, Karnataka, India. Received a B.Sc. (Computer Science) from Vel Tech College, Chennai, Tamil Nadu, India. Her research interests include machine learning, data mining, cloud computing, big data analytics. She can be contacted at email: <a href="mailto:frenimartin28@gmail.com">frenimartin28@gmail.com</a></p>