International Journal of Creative and Open Research in Engineering and Management
ISSN: **3108-1754 (Online)**
Volume 02 Issue 04 April-2026 | Impact Factor: 3.5

House Price Prediction

Mayank Tyagi[1], **Varun Sharma**[2]
[1,2]Department of Computer Science, Student of Master of Computer Applications, Jagan Institute of Management Studies (JIMS), New Delhi, India

Article Info

Article history:

Received: month dd, yyyy
evised: month dd, yyyy
ccepted: month dd, yyyy
ublished: month dd, yyyy

Keywords:

Libraries
Exploratory Data Analysis
E.D.A)
Data Preprocessing/Feature
ngineering
Feature Selection Scikit-Learn
R^2-score Algorithm

ABSTRACT

Machine Learning (ML) is a rapidly developing branch of Artificial Intelligence (A.I.) that helps computers learn from the past data and improve their performance without explicit programming. Machine learning has gained significant attention in recent years due to its ability to analyze large amounts of data and provide intelligent solutions for complex problems. The most significant application of machine learning is house price prediction, which is helpful in understanding the prices of houses based on different factors. The present study utilizes the machine learning technique to make accurate predictions in house price prediction using the given data set of houses, which consists of 1,460 records and 81 features. Data preprocessing is the initial step in which the data is made suitable for further analysis. The data preprocessing steps include missing value handling, data cleaning, and removal of outliers. After preprocessing the data, Exploratory Data Analysis (EDA) is performed to understand the data, identify the patterns, and analyze the relationships between different features and the target variable. The dataset is split into training and testing data sets using the train-test split. Various machine learning algorithms were implemented to build prediction models. The algorithms used were Linear Regression, Decision Tree Regressor, Random Forest Regressor, and K-Nearest Neighbors. The R^2 score was used to evaluate the models' performances. The results show that the Random Forest Regressor performs better than the other algorithms since it has the highest R^2 score. This proves that the machine learning algorithms can be used to make house price predictions. This study proves that machine learning algorithms can be used in predicting house prices. The process is simple and easy to follow even for a beginner who wants to learn the process of building a machine learning model.

1. INTRODUCTION

One of the first definitions of machine learning was provided by Arthur Samuel in 1959. He stated, "Machine learning is the field of study that gives computers the ability to learn without being explicitly programmed." Later, Tom M. Mitchell proposed a more technical definition of machine learning.

According to him, "a computer program is said to learn from experience E with respect to some task T and performance measure P, if its performance on T, as measured by P, improves with experience E."

Machine learning (ML) is a subfield of artificial intelligence (AI), which allows computer systems to learn from data, detect patterns, and make decisions with almost no human involvement. Instead of receiving step-by-step instructions for each action through explicit programming, machine learning algorithms continuously enhance their ability to perform a particular task using statistical methods when provided with more information. In essence, machine learning revolves around feeding a model training data for example (data in the form of images, numbers, texts, etc.) and their respective outcomes (supervised) or not (unsupervised). Then, the machine learning algorithm uncovers patterns within the data. Rather than instructing a spam filter to recognize specific terms associated with spam, the program undergoes extensive training by examining tens of thousands of emails and determining whether a message is legitimate or spam based on the characteristics of the message. As the model is presented with new information, it learns how to update its internal parameters to minimize errors in prediction. The capability to generalize from past experiences to future ones is the defining characteristic of successful machine learning.

Machine Learning is a technique used in computer science to enable computers to learn from data and make the required predictions without being explicitly programmed. This technique enables computers to analyze the data and find the relationship between the variables. These learned patterns help the system make decisions or predictions when the computer is provided with the new data. Many applications such as recommendation systems, spam email detection, image recognition, and speech recognition rely on machine learning techniques, and they provide the intelligent solutions to the users. Some real-world applications of ML include Netflix recommendations, Siri voice recognition, and fraud detection systems. There are three categories of machine learning, namely: supervised learning, which involves learning from labelled data; unsupervised learning, which involves discovering patterns from unlabelled data; and reinforcement learning, which involves learning to take actions by trial and error. Machine learning systems can solve complicated tasks in a way that would otherwise not be possible or very difficult to code manually.

One of the important real-world applications of machine learning is the prediction of house prices. It is a difficult task to predict the house prices because the prices of the houses vary based on many factors, including the location of the house, the size of the house, the number of rooms, the year the house was built, the conditions of the neighborhood, and the structural aspects of the house. ML models can analyze historical housing data and identify patterns that influence property prices. Therefore, the prediction of the house prices is a useful application of machine learning as it is important to understand the value of the house, which is helpful for the buyers and sellers of the properties. And also can be very useful to explain it to beginners who are interested in Machine Learning as this dataset contains of many steps that a beginner should know while creating Machine Learning Model.

Prior to developing a prediction model, Exploratory Data Analysis (E.D.A) is conducted to understand the data, identify relationships between variables, and identify any missing information or unusual data in the dataset. Visualization of data and statistical analysis are conducted to identify the features that are highly influential in determining house prices. Once the data is analyzed and cleaned, machine learning algorithms are used to develop a model that can be used to forecast prices of houses based on selected features.

In this research paper, a machine learning technique is utilized to predict the prices of houses using a dataset containing the details of houses. There are a total of 1460 entries and 81 features, and the features represent different attributes of the houses, such as the lot area, the quality of the houses, the year the houses were built, the number of bedrooms, the garage area, and so on. This dataset is small which is perfect for a beginner to begin with their Machine Learning journey as to understand what's happening in each step. In this dataset, the target column is SalePrice that indicates the selling price of the houses.

2. RELATED WORK

There are various research works conducted in the field of house price prediction using machine learning techniques. These research mainly concentrate on analyzing housing data and applying various algorithms to improve accuracy in house price prediction.

The House Prediction Price Using Machine Learning proposed by Zhao et al. [1], proposed a house price prediction model using the multi-source data fusion approach. In this study, the researchers proposed the idea of integrating multiple types of feature such as property features, amenities, traffic conditions, and environmental factors, to improve the accuracy of the house price prediction model. The proposed house price prediction model was evaluated using actual real estate data with more than 28,000 housing transactions in Beijing, China. The results indicated that the combining multiple data sources can effectively capture the complex relationships between housing features and market conditions. The study concluded that the multi-source data fusion approach can improve the accuracy and reliability of the house price prediction model. In the research conducted by Chatwani et al. [2], the authors explained the importance of the prediction of house prices with the help of machine learning models. They explained that the prices of houses influenced by many factors, which include the location of the house, size of the house, number of rooms, and the economic conditions. In this research, the researcher used many machine learning algorithms to predict the house prices. These algorithms include regression, decision trees, k-nearest neighbors, random forest, and support vector machine. With the help of machine learning techniques, the historical market data can be analyze and provide useful predictions for the real estate market.

Another research conducted by Rahman et al. [3], they concentrated on the advanced machine learning algorithms for predicting house prices in Kuala Lumpur. The researchers used a dataset that had thousands of records and their attributes such as location, number of rooms, size of the houses, and the distance to facilities qnd many more. The researcher used several prediction models, such as Multiple Linear Regression, Ridge Regression, LightGBM, and XGBoost, and found that the best prediction model, based on the lowest prediction error and highest R-squared value, was the XGBoost model which indicates a more accurate prediction model. In the study conducted by Kalidass et al. [4] published in IRJET journal, the focus was to predict house prices using ensemble machine learning algorithms like Random Forest and Gradient Boosting. In this study, the researchers used housing datasets with features like the number of bedrooms, living area, lot size, etc. After preprocessing the data and training the models, it was concluded that the Random Forest and Gradient Boosting algorithms were giving lower errors and better accuracy compared to the results obtained using traditional methods for predicting house prices. The study concluded that ensemble machine learning models can effectively capture complex relationships between housing features and property prices.

In a research by Xiangjun Yang [5], various machine learning algorithms, i.e., Linear Regression, Random Forest, and Decision Trees, were compared for their ability to predict house prices. In the research, a housing dataset from Kaggle, containing various features of properties, was used. Various preprocessing techniques were used for the dataset such as handling missing values, normalisation, and feature selection were applied before model training. And the results showed that the Random Forest algorithm works better than other regression algorithms because of its ability to deal with complex datasets effectively and it can capture non-linear relationships. Rawool et al. [6] developed a model for the prediction of house prices using various machine learning algorithms, namely Linear Regression, Random Forest, and Decision Tree. They used various attributes of houses, i.e., the number of rooms, location, and other structural attributes to predict the prices of houses. According to the result, the Random Forest algorithm gave better results as compared to other algorithms.

Patel et al. [7] used different machine learning techniques such as multiple linear regression and decision tree regression were used to predict house prices. The model was trained and tested using real estate data, and the results showed that multiple linear regression performed better compared to decision tree regression in predicting property values. The system can help users estimate house prices based on various features such as location, size, and amenities. This research demonstrates that machine learning can be effectively used for house price prediction. In the future, more features and advanced models can be added to improve prediction accuracy and make the system more reliable. And another research by Luna et al. [8], multiple machine learning algorithms such as Linear Regression, Random

Forest, and Gradient Boosting were used to predict house prices using the Melbourne housing dataset. The dataset was analyzed and preprocessed before applying the models to understand the relationships between different features. The performance of the models was evaluated using metrics such as MAE, RMSE, and R^2 score. The results showed that the Gradient Boosting model performed better than the other models, providing higher accuracy and lower error values. This study highlights the effectiveness of machine learning techniques in predicting house prices. In future work, combining multiple models can further improve prediction performance.

3. METHODOLOGY

The flowchart (figure 1) shows the entire process flow of creating a model for predicting the cost of houses through machine learning techniques. In the first stage, the dataset is loaded, which includes data related to different attributes of houses. Once the loading of the dataset is done, the next stage involves the pre-processing of data by dealing with missing data, converting the categoric values, and preparing the dataset for further analysis. The pre-processing stage is followed by EDA, where the dataset is explored to get insights about it. This phase involves exploring different relations between attributes and the target attribute, which will help us detect some anomalies in the dataset. Once we get insights about the dataset, the next phase involves the cleaning of the dataset, including removing the noise from it.

Following the data cleaning process, the next step is the selection of appropriate features that are most important for predicting house prices. The final step in the process involves splitting the dataset into training and test datasets using the train-test split method, where the model is trained on one part of the data and evaluated on unseen data.The next stage involves applying various machine learning models to the split dataset. These machine learning models' performances will be assessed through suitable evaluation measures. In the decision phase, it will be decided whether the model chosen is producing accurate results or not. In case it is producing poor results, tuning and re-evaluation of the model will take place.

After getting an acceptable model, the final stage will involve deploying the model to predict future house prices. In the end, the process terminates by choosing the best-performing model.

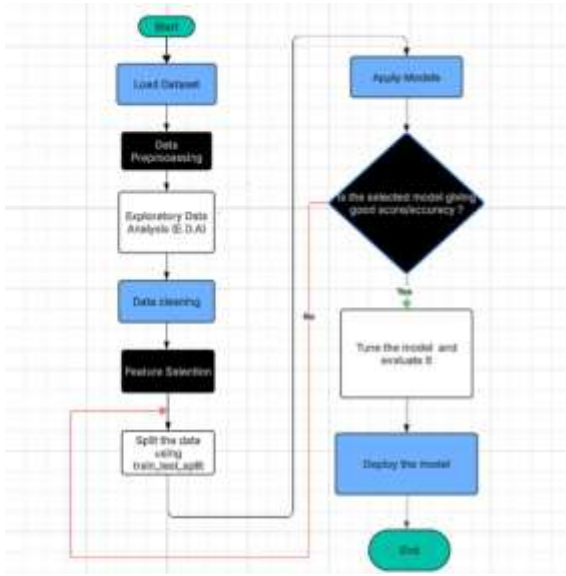

Figure 1: How to build a model in Machine Learning

To build a Machine Learning model, we mostly follow these steps:

1. Data collection: Before applying Machine Learning technqiues, first we need data and we can collect the data from public datasets like Kaggle. It is the most popular dataset site where every dataset is available in form of csv files. And then dataset is uploaded using Python libraries that is Pandas. Numpy library is used for numerical operations. Since, the size of dataset consisted of 81 columns, so I had inserted a link below where you can click the link and directly goes to the dataset: Dataset is available here

2. Data Preprocessing/Feature Engineering: Data preprocessing helps to improve the quality of the dataset and ensures that the models can learn effectively from the data. Find the NULL values in each column by using “isnull().sum()” function. This will give you NULL values present in each feature (column). In this dataset, we had filled these missing values by using “fillna([]median())” function. Column which are almost fully empty can be dropped out of the dataframe (in which we stored our dataset as variable). We can also handle missing values using imputation or interpolation techniques. For feature Engineering, apply Normalization, Standardization techniques to column consists of numerical values (Numerical column) so that they can be come in a range. And apply One Hot Encoding, Label Encoding to catgeorical column. In this dataset, we had used both “MinMaxScaler” and “StandardScaler” to normalzie numerical feature and for categorical feature, we had used “one-hot encoding” using the “get_dummies()” method was used to convert categories into binary variables.

3. Exploratory Data Analysis (E.D.A): Exploratory Data Analysis is carried out to discover any inconsistencies within the dataset and reveal any patterns within the data set that can assist us in improving the effectiveness of the data set. The methods employed include statistical analysis, correlation analysis, and data visualization. EDA can be conducted using python libraries, namely "Matplotlib" and "Seaborn". In our house price prediction model, we have used graphs like distplot, boxplot, and heatmap. The purpose of conducting EDA is to gain insight into how the characteristics of houses affect the prices of the houses. This analysis is also useful in discovering how the data is distributed and revealing skewness in the attributes. Exploratory Data Analysis facilitates the detection of outliers and the relationship between the independent variables and the dependent variable.

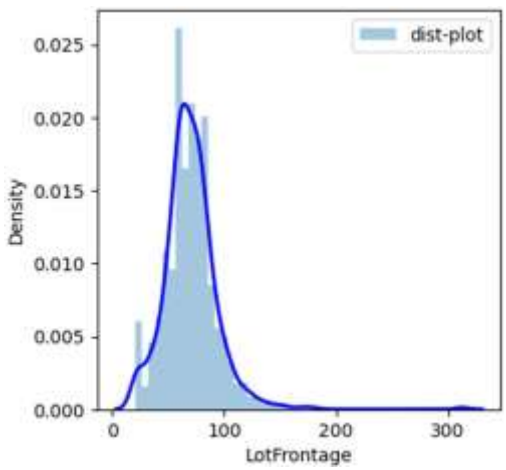

Figure 2(a): Before filling the NaN values for the feature 'LotFrontage'.

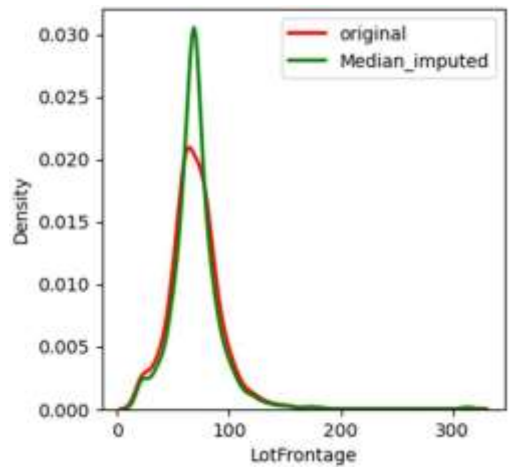

Figure 2(b): After filling the NaN values for the feature 'LotFrontage'.

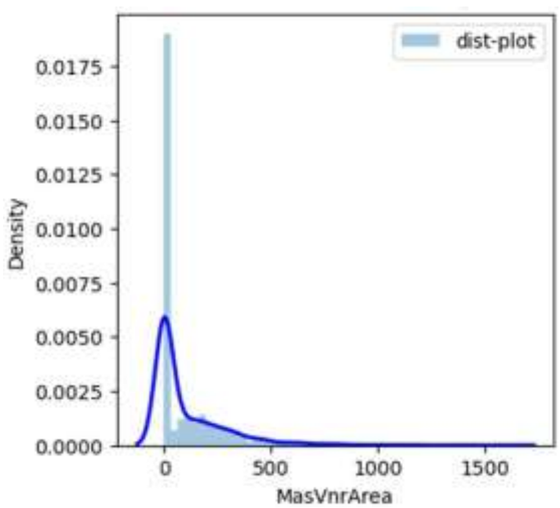

Figure 3(a): Before filling the NaN values for the feature 'MasVnrArea'.

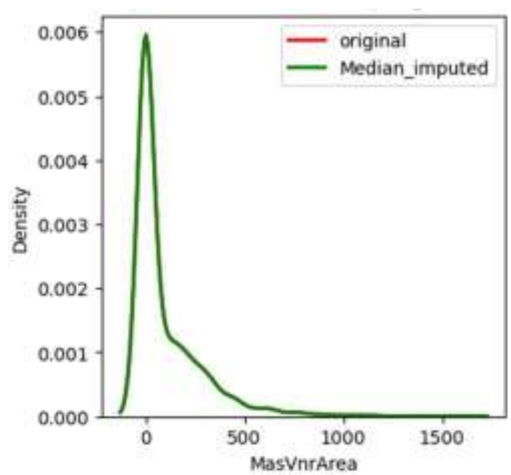

Figure 3(b): After filling the NaN values for the feature 'MasVnrArea'.

4. Removing Outliers: This step typically falls under data cleaning and data preprocessing. The reason why outliers have been removed is because they are data points that exhibit a behavior which is not the same as that shown by the rest of the data points and which tend to impair the performance of the model. The way through which outliers were removed in our analysis was through the use of a popular method called Inter Quartile Range (IQR). IQR is a very powerful statistical tool used for the detection and removal of outliers from a particular set of data. This is achieved by finding the distance between Q1 and Q3, and any data point beyond that is considered an outlier. Removing such extreme values helps in reducing noise and improving the quality of the dataset. This step also helps in making the model more stable and improves the accuracy of predictions. As a result, the model becomes more reliable when applied to new data.

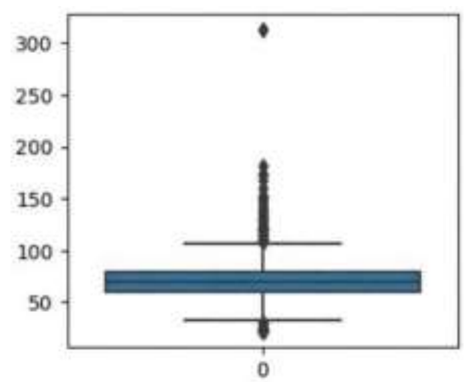

Figure 4(a): Outliers present in the feature 'LotFrontage'.

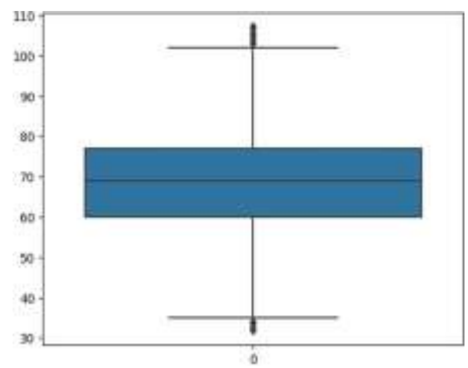

Figure 4(b): Outliers removal using the IQR method from the feature 'LotFrontage'.

5. Feature Selection: Feature selection is the step where we select the most relevant features and create new meaningful ones to enhance model accuracy and reduce complexity. Good feature selection often improves the model perform ance rather than changing the algorithms. Selecting relevant features helps reduce unnecessary data and imrpoves the efficiency of the machine learning algortihms. Feature selection helps in reducing overfitting is through discarding any irrelevant or less important factors from the data set. It improves the training speed of the model as fewer features require less computational power. Techniques such as correlations, importance measures provided by the models, and the application of the variance threshold technique.Appropriate feature selection helps to ensure that only the relevant features are selected, which improves the generalization ability of the model. By using appropriate features, the prediction model can provide more accurate and reliable results. A Correlation “Heatmap” was used to visualize the relationships between columns present in the dataset. By analyzing the heatmap, it becomes easier to undertsand which variables have impact on house prices. But from

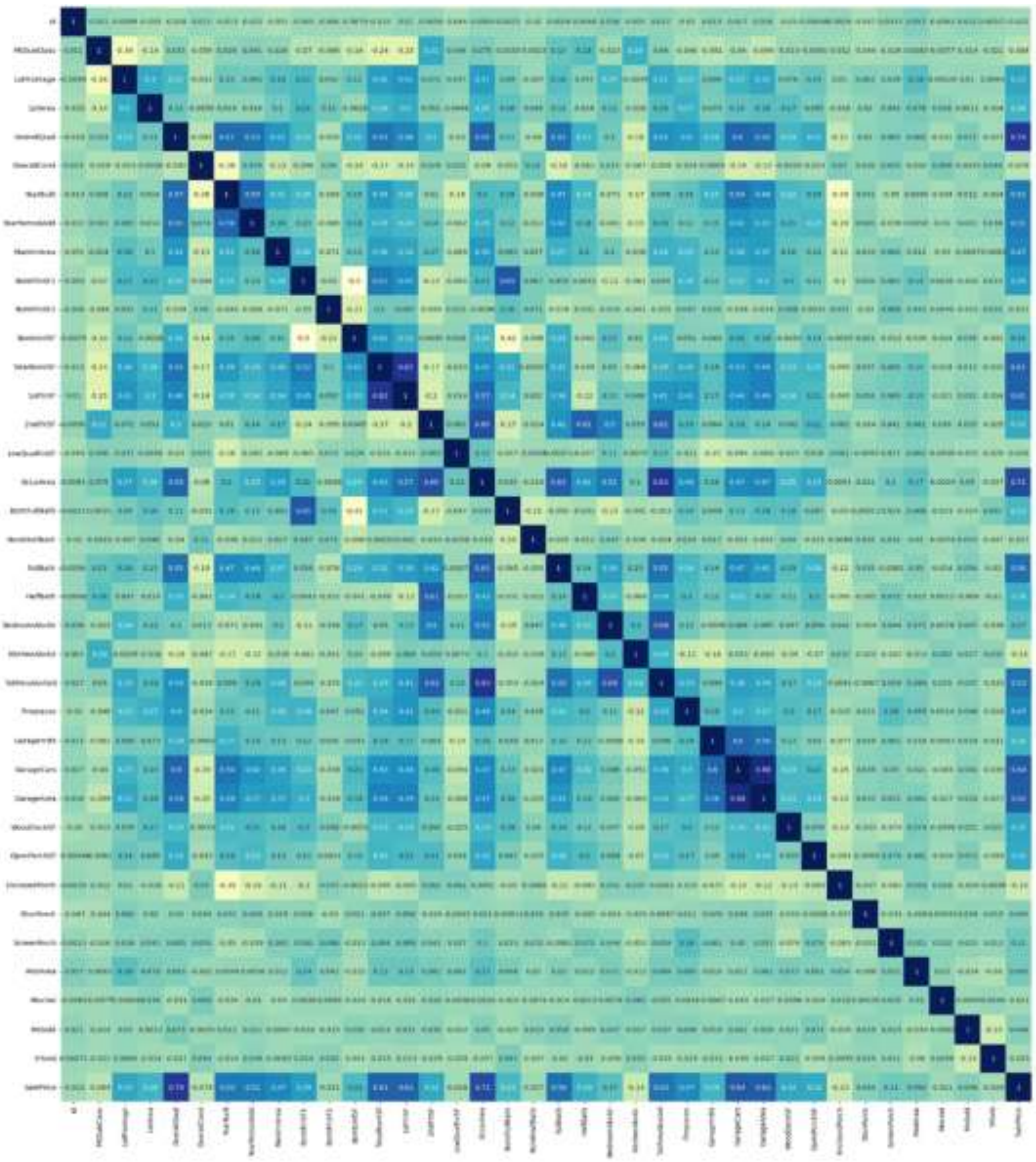

Figure 5: A Heatmap to check the relationship of each column with Target column

6. Correlation Metrices, we concluded that there are none of the feature which is not making a correlation with the Target variable (SalePrice), and hence we cannot say that removal of any feature will make a good impact on the R^2-score of the model. But after the training the Machine Learning Model, if we are unable to achieve a good R^2-score, then we could try to remove some of the features by checking their correlation with the Target variable again which could make an impact on the R^2-score of the trained model.

7. Select a Machine Learning Algorithm: Before selecting a machine learning algorithm to build ML model, we need to divide the data using "train_test_split()" function available in "scikit-learn" library. This function divides the dataset so that the model learns from one part and is evaluated on unseen data. Splitting of the dataset will help evaluating how effective the model will perform when presented with new data, preventing overfitting. A fixed random state can also be used in the train_test_split() function to achieve consistent results. Proper data splitting ensures that the model does not memorize the data but learns general patterns from it. There should be a balance in the distribution of data in both the training and testing datasets for better evaluation. This process is very important in validating machine learning algorithms reliability and effectiveness. In simple words, the training dataset is used to train the ML model while the testing dataset is used to evaluate the performance of the trained models. How much data needs to be divided in training and testing dataset can be specified in parameter present in this function. For our House price prediction model, we had used 75% of dataset for training and 25% for testing it. In this research paper, we used several machine learning models to predict the house price. And also, we compared them to check the performance of each algorithm. The models we used while creating house price prediction model were Linear Regression, Decision Tree Regressor, Random Forest Regressor and K-nearest neighbors (KNN).

8. Evaluate the model: The above-mentioned algorithms were implemented using the scikit-learn library, which was also used earlier for splitting the dataset. Each model was trained and evaluated using the R^2-score as the performance metric. The obtained scores are as follows: Linear Regression: 0.786 Decision Tree Regressor: 0.736 Random Forest
Regressor: 0.888 KNN: 0.716
These results show that the Random Forest Regressor performed better than all other algorithms used in this study. This is because it can handle complex relationships and reduce overfitting by combining multiple decision trees. Although the R^2-scores can be further improved by using a larger dataset or better feature engineering, the current results are acceptable for this dataset size. Additionally, other evaluation metrics such as Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) can also be used to measure prediction error and evaluate model performance more effectively. These metrics provide a better understanding of how close the predicted values are to the actual values.

- MSE: It measures how much your predictions are wrong with the help of the formula MSE = $(1/n)\sum(\text{yactual} - \text{ypredicted})$
- RSME: Just square root of MSE i.e. RSME = $\sqrt{\text{MSE}}$.

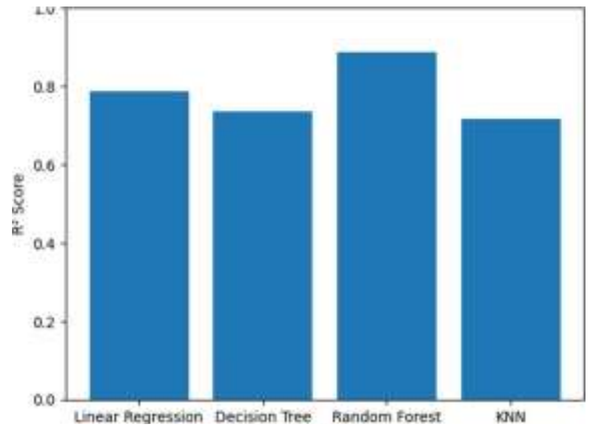

Figure 6: Performance comparison of Machine Learning Algorithms for House Price Prediction

Deploy the model: The house price prediction model is now fully prepared for deployment. We can deploy it via platforms such as Streamlit or via cloud-based services. The deployment of our model will enable us to use it in practical applications where users will input house details and receive the price predictions right away. It will also aid in making our model more available by presenting it with an easy-to-use interface. With the help of platforms like Streamlit, we can develop web applications effortlessly.

4. RESULTS AND DISCUSSION

In this research paper, various machine learning models were used to predict the house prices based on the provided housing dataset. These models were Linear Regression, Decision Tree Regressor, Random Forest Regressor, and K-Nearest Neighbors. The performance of the models is based on the R^2-score since it was a Regression problem, which is used to check how accurately the predicted values correlate with the actual house prices. The results obtained from the models used in this research are in the Figure 7. The comparison of different models also shows that model performance depends on how well the algorithm can capture patterns present in the dataset. Furthermore, it is evident that appropriate data processing and selection of the features led to improved results of the predictions.
Additionally, the difference between the R^2 scores for various algorithms demonstrates the necessity of selecting the correct regression model.

Based on the results obtained from the models, it is observed that the Random Forest Regressor has the highest R^2-score of 0.888, which means that this model has performed better compared to the other models used in this research. This is because this model is able to handle the complex relationships that exist among the features of the houses.

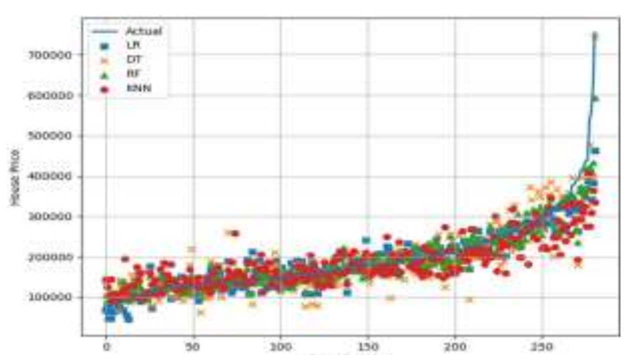

Figure 7: Comparison of Actual vs Predicted House Prices

On the other hand, Linear Regression had a moderate performance, while Decision Tree and KNN had lower prediction score values compared to Random Forest. This indicates that ensemble learning algorithms such as Random Forest can produce better results for house price prediction. And finally, the results show that house price prediction can be analyzed by machine learning algorithms and produce reliable prediction results. This information allows comprehending the advantages and disadvantages of each algorithm examined in this paper.

5. CONCLUSION

In this research study, a machine learning method was employed to predict house prices based on a housing dataset. We performed Exploratory Data Analysis (EDA) to understand the dataset, as well as the relationship between different housing features and the target variable. Feature selection and preprocessing were also employed to prepare the dataset for machine learning models. Different machine learning models, such as Linear Regression, Decision Tree Regressor, Random Forest Regressor, and K-Nearest Neighbors (KNN), were implemented to develop the house price prediction model. The performance of the implemented models was also evaluated using the R^2 score.

Among the implemented models, the Random Forest Regressor performed better with a high R^2 score of 0.888, indicating that the model was more effective in predicting house prices than the other models. The results of the current study show that machine learning models can effectively analyze the housing dataset to provide accurate predictions for house prices.

Although the model achieved satisfactory results, there are several ways in which the study can be improved in the future. First, the prediction accuracy can be improved by using big housing dataset. More complex machine learning and deep learning models can also be used to improve the prediction performance.

Feature engineering techniques can also be used to generate new features, which can be significant in determining the price of houses. Other complex algorithms like XGBoost, Gradient Boosting, and Neural Networks can also be used to compare their performance with already implemented models.

Another way in which the study can be improved is by adding more features like the presence of schools, transportation facilities, and the quality of the neighborhood, which can be significant in determining the price of houses. Future work can also be focused on improving the model by adjusting different parameters of the algorithms to achieve better prediction accuracy.

ACKNOWLEDGMENTS

The authors would like to express their sincere gratitude to the mentor Dr. Deepshika Aggarwal for their valuable guidance and continuous support throughout the completion of this research work. Their suggestions and feedback helped in improving the quality of the study.

FUNDING INFORMATION

The authors state that no funding was involved in this research work.

AUTHORS CONTRIBUTIONS STATEMENT

This journal uses the Contributor Roles Taxonomy (CRediT) to recognize individual author contributions, reduce authorship disputes, and facilitate collaboration.

Name of Author	C	M	So	Va	Fo	I	R	D	O	E	Vi	Su	P	Fu
Mayank Tyagi	☐	☐	☐	☐	☐	☐	☐	☐	☐		☐	☐	☐	
Varun Sharma	☐	☐			☐	☐	☐			☐	☐	☐	☐	

C : **C**onceptualization
M : **M**ethodology
So : **So**ftware
Va : **Va**lidation
Fo : **Fo**rmal analysis
I : **I**nvestigation
R : **R**esources
D : **D**ata Curation
O : Writing - **O**riginal Draft
E : Writing - Review & **E**diting
Vi : **Vi**sualization
Su : **Su**pervision
P : **P**roject administration
Fu : **Fu**nding acquisition

CONFLICT OF INTEREST STATEMENT

Authors state no conflict of interest.

INFORMED CONSENT

Not applicable as this study does not involve human participants.

ETHICAL APPROVAL

Not applicable as this study does not involve human or animal objects.

DATA AVAILABILITY

The data that support the findings of this study are publicly available on Kaggle at: Dataset is available here

REFERENCES

Book

- A. Géron, *"End-to-End Machine Learning Project"* in Hands-on Machine Learning with Scikit-Learn, Keras & Tensorflow, Nicole Tache, 2nd ed. Sebastopol, CA, USA: O'Reilly Media, 2019, pp. 37-79.
- Andreas C. Müller and Sarah Guido, *"Model Evaluation and Improvement"* in Introduction to Machine Learning with Python, Dawn Schanafelt, 1st ed. Sebastopol, CA, USA: O'Reilly Media, 2016,pp. 251–302.

Research Paper

[1] Yaping Zhao, Jichang Zhao, and Edmund Y. Lam, "House Prediction Price: A Multi-Source Data Fusion Perspective", *Big Data Mining and Analytics*, vol. 7, no. 3, pp. 604-618, Sept. 2024, doi: 10.26599/BDMA.2024.9020019 .

[2] Smit Chatwani, Mayank Chotaliya, and Vinaya Sawant, "House Price Prediction Review", *International Journal of Computer Applications (0975-8887)*, vol. 184, no. 7, pp. 36-39, April 2022.

[3] Shuzlina Abdul-Rahman, Nor Hamizah Zulkifley, Ismail Ibrahim, and Sofianita Mutalib, "Advanced Machine Learning Algorithms for House Price Prediction: Case Study in Kuala Lumpur", *International Journal of Advanced Computer Science and Applications*, vol. 12, No.12, pp. 737-743, 2021.

[4] Prof. J. Kalidass, T. Dharshalini, R. Nivetha, and AP. Subasri, "House Prediction Price using Machine Learning", *International Reasearch Journal of Engineering and Technology (IRJET)*, vol. 11, issue 04, pp. 1351–1357, April 2024.

[5] Xiangjun Yang, "Research on House Price Prediction based on Machine Learning", *ITM Web of Conferences*, artcile no. 02018, pp. 02-08, 2025, doi: https://doi.org/10.1051/itmconf/20257002018.

[6] A. G. Rawool, D. V. Rogye, S. G. Rane, and Dr. Vinayaka A. Bharadi, "House Price Prediction Using Machine Learning", *Iconic Research and Engineering Journals*, vol. 4, issue 11, pp. 29–33, May 2021.

[7] Deepanshu Patel, Parmeshwar Nayak, Shubham Gupta, and Jayanth C, "House Price Prediction Using Machine Learning", *International Journal For Research in Applied Science and Engineering Technology*, vol. 11, issue 06, pp. 1870-1873, June 2003, doi: https://doi.org/10.22214/ijraset.2023.53841 .

[8] Lale EL Mouna, Hassan Silkan, Youssef Haynf, Mohamedade Farouk Nann, and Stéphane C. K. Tekouabou, "A Comparative Study of Urban House Price Prediction using Machine Learning Algorithms", *E3S Web of Conferences,* artcile no. 03001, pp. 02–06, 2023, doi: https://doi.org/10.1051/e3sconf/202341803001 .

[9] Fatbardha Maloku, Besnik Maloku, and Akansha Agarwal Dinesh Kumar, "House Price Prediction Using Machine Learning and Artificial Intelliegence", *Journal of Artificial Intelligence & Cloud Computing*, vol. 3(4), pp. 02–09, August 2024, doi: 10.47363/JAICC/2024(3)357 .

[10] Sindiya Jasmin.C, and Ms. P Sathya, "House Price Prediction", *International Journal of Advanced Research in Education and Technology (IJARETY)*, vol. 11, issue 03, pp. 851–856, June 2024, doi: 10.15680/IJARETY.2024.1103004 .

[11] Matthew C. Okoronkwo, Ugochukwu E. Orji, Chikodili H. Ugwuishiwu, Caroline N. Asogwa, Emenike C. Ugwuagbo, and Bande S. Ponsak, "Ensemble Machine Learning-Based Real Estate Price Prediction with Explainable Artificial Intelligence Methods for Determinant Analysis", *International ournal of Engineering Trends and Technology*, vol. 73, issue 10, pp. 80-91, Oct. 2025, doi: https://doi.org/10.14445/22315381/IJETT-V73I10P106 .