IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

BANGALORE HOUSE PRICE PREDICATION

[1] Kunal Kumar, [2] Hasanji Mariyambibi Zakirhusen, [3] Shaikh Shifa, [4] Sagar Rakeshbhai Patel, [5] Yassir Farooqui

[1]Student, [2] Student, [3] Student, [4]Student, [5] Professor

[1]Computer Science Engineering,

[1]Parul University, Vadodara, Gujrat, India

Abstract—The estimation of changes in housing prices is often done through house price prediction Because location, area, and population are closely connected with housing price, more information is needed in addition to house price prediction to estimate the cost of a specific home. Many studies have used conventional machine learning techniques to predict home prices effectively, but they seldom address the shortcomings of particular models and often overlook the more sophisticated but less well- liked models. Therefore, this study will use both conventional and sophisticated machine learning techniques to examine the differences between several sophisticated models in order to examine the diverse effects of features on prediction methods. Additionally, a thorough validation of several model implementation strategies on regression will be provided in this study, with positive outcomes.

As the purpose of literature is to extract meaningful information from historical property market data. Historical property is assessed with machine learning techniques.

exchanges in India to find practical models for home sellers and purchasers. The large disparity between the most expensive homes' pricing is made clear. Furthermore, studies show that the Multiple Linear Regression method, which relies on mean squared error assessment, is competitive.

***Index Terms* - Bangalore House price prediction, Machine learning**

I. INTRODUCTION

Nowadays, machine learning (ML) is an essential component of research and industry. Through the use of algorithms and neural network models, computer system performance is gradually increased. Without being expressly programmed to generate decisions, machine learning algorithms use sample data, often known as training data, to automatically create a mathematical model that forms decisions.

Both individuals and real estate companies buy and sell homes; individuals buy for their own use or as investments, and the agencies acquire to operate a business. In any case, our opinion is that one should receive exactly what is paid for. The problem of overvaluation and undervaluation in housing markets has long existed, and appropriate detection methods are lacking. Broad metrics, such as price-to-rent ratios for homes or real estate, provide a primary pass. However, a thorough examination and sound judgment are required to make a decision on this matter. Using hundreds of thousands of data points to train an ML model, a solution that can fulfil the needs of all parties and be powerful enough to anticipate pricing reliably can be created.

This research aims to apply Machine Learning Techniques to create ML models that can benefit consumers. Buyers aim to find their ideal home with all the necessary features. Additionally, when searching for houses or real estate, buyers often have a specific price in mind, but there is no guarantee of receiving a fair deal. Sellers must conduct extensive research to determine the appropriate price for their property.

The value of a house is significant as an outcome of a wide range of options. as a result of this, estimating home costs poses unique challenges. Although many room units are dedicated to the current task, their presentation and applicability are limited because too long delays in information processing, a lack of real-world situations, and limited lodging options. This research aims to estimate the value of house advertising abuse defining methods.

The majority of the modern investigations have concentrated on the breakdown of housing cost predictions. This document provides the most recent expectation research on financial indicators to utilize. The report provides a summary of anticipated markets and present business sectors, making it easier to predict future market trends.

II. EXISTING SYSTEM

Multi Linear Regression Multiple Linear Regression. It shows the relationship between two or more explanatory variables and scalar response variable. Independent variable value is associated with dependent variable value Limitations The dependent variable y must be continuous. The independent variables can be of any type. The dependent variable is usually affected by the independent variables.

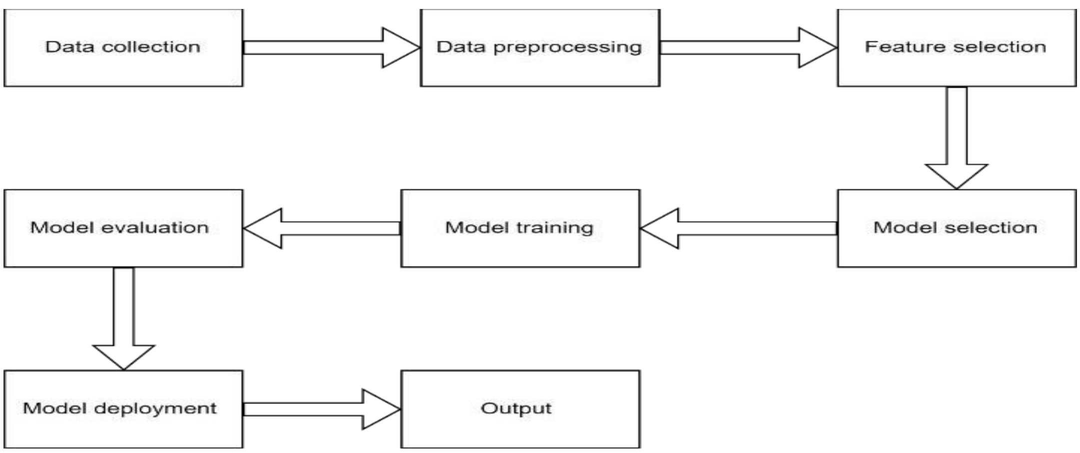
Limitations The dependent variable y must be continuous. The independent variables can be of any type. The dependent variable is usually affected by the independent variables.

III. PORPOSE SYSTEM

Machine Learning is an area of artificial intelligence that allows PC frameworks to learn and increase performance with the use of information. This study focuses on the development of algorithms for data prediction. Machine learning is applied to Perform a large number of computer jobs. It can also be utilized for computer-based predictions. Machine learning is sometimes used to create sophisticated models. Machine learning allows computers to learn on their own without human guidance. Machine learning is a valuable tool utilized globally. Machine learning includes collecting data and training computers to develop models using it.

SVM is a popular supervised learning technique used for both classification and regression applications. The SVM algorithm aims to establish a decision boundary that divides an n-dimensional space into classes, making it easier to assign fresh data points to the appropriate category. SVM Classifiers outperform Naïve Bayes algorithms in terms of accuracy and prediction speed.

Using a subset of training points during the decision phase reduces memory use. SVM performs well in high-dimensional spaces with unambiguous separations. Although SVMs do not generate probabilities directly, they can be converted to class probabilities using probability calibration methods. In the binary instance, probabilities are adjusted by Platt scaling, which involves logistic regression on SVM scores followed by cross-validation on training data. We will apply SVM machine learning to estimate property prices based on multiple factors, providing consumers with optimal and reliable results.



Research Flow Diagram

IV. LITERATURE SURVEY

1. Housing Price Prediction Using Machine Learning Algorithms: The Case of Melbourne City, Australia

Author: The Danh Phan,2018

House price Prediction is a crucial topic of land. The literature attempts to get useful knowledge from historical data of property markets. Machine learning techniques are applied to research historical property transaction sin Australia to get useful models for house buyers and sellers. Reveal dis the high discrepancy between house prices withinthe costliest and most affordable places within Melbourne city.

2. Predicting Sales Prices of the Houses Using Regression Methods of Machine Learning

Authors: Parasich Andrey Viktorovich; Parasich Viktor Aleksandrovich; Kaftannikov Igor Leopoldovich; Parasich Irina Vasilievna, 2018

This article we'll describe our solution for "House Prices: Advanced Regression Techniques" machine learning competition, which was persisted Kaggle platform. The goal is to predict house sale price by attributes like house area, year of building etc. In our solution, we use classic machine learning algorithms, and our original methods, which may bedescribed here. At the highest of the competition, we took 18th place among 2124 participants from whole world.

3. Real Estate Value Prediction Using Linear Regression

Authors: Nehal N Ghosalkar; Sudhir N Dhage, 2018

The real estate market may be a standout amongst the foremost focused regarding pricing and keeps fluctuating. Itis one among the prime fields to use the ideas of machine learning on the way to enhance and foresee the prices with high accuracy. There are three factors that influence the priceof a house which includes physical conditions, concepts and location. The current framework includes estimating the worth of homes with none expectations of market prices andprice increment.

4. Predicting Housing Market Trends Using Twitter Data

Authors: Marlon Velthorst; Cicek Güven, 2019

In this study, we attempt to predict the Dutch housingmarket trends using text mining and machine learning as anapplication of knowledge science methods in finance. Our main goal is to predict the short term upward or downwardtrend of the average house price in the Dutch market by using text data collected from Twitter. Twitter is widely usedalso and has been proven to be a helpful source of knowledge. However, Twitter, text mining (tokenization, bag-of-words, n-grams, weighted term frequencies) and machine learning (classification algorithms) have not been combined yet in order to predict the housing market trendsin short term.

5. House Price Prediction Using Machine Learning and Neural Networks

Authors: Ayush Varma; Abhijit Sarma; Sagar Doshi; Rohini Nair, 2018

Real estate is that the least transparent industry in our ecosystem. Housing prices keep changing day in and outing and sometimes are hyped instead of being supported valuation. Predicting housing prices with real factors is thatthe main crux of our scientific research. Here we aim to formour evaluations supported every basic parameter that's considered while determining the worth. We use various regression techniques during this pathway, and our results aren't sole determination of 1 technique rather it's the weighted mean of varied techniques to offer most accurateresults.

6. Forecasting house price index of China using dendritic neuron model

Authors: Ying Yu; Shuangbao Song; Tianle Zhou; Hanaki Yachi; Shangce Gao, 2016

The results of Chinese housing market continue to prosper or not is said to the event of China, and further it also has animpression on the planet finance. Thus, forecasting the house price level is extremely important and challenging. during this paper we propose an unsupervised learnable neuron 10 model (DNM) by including the nonlinear interactions between excitation and inhibition on dendrites.We use DNM to suit the House price level (HPI) data then forecast the trends of Chinese housing market. To verify theeffectiveness of the DNM, we use a standard statistical model (i.e., the exponential smoothing (ES) model) to form a performance comparison. Three quantitative statistical metrics including normalized mean square error, absolute percentage of error, and coefficient of correlation are wontto evaluate the forecasting performance of the 2 models. Experimental results demonstrate that the proposed DNM isbest than ES altogether of the three quantitative statistical metrics.

7. Prediction of real estate price variation based on economic parameters
   Authors: Li Li; Kai-Hsuan Chu, 2017

It is documented that a lot of economic parameters may more or less influence the important estate price variation. additionally, the banker and investor also are interesting to understand the important estate price future change. There had not appropriate model for including these factors for price prediction. Here, the influences of most macroeconomic parameters on land price variation are investigated before establishing the worth fluctuation prediction model. Here, back propagation neural network (BPN) and radial basis function neural network (RBF) two schemes are employed to determine the nonlinear model for real estate's price variation prediction of Taipei, Taiwan supported leading and simultaneous economic indices.

8. Predicting house sale price using fuzzy logic, Artificial Neural Network and K-Nearest Neighbor
   Authors: Muhammad Fahmi Mukhlishin; Ragil Saputra; Adi Wibowo, 2017

Determining the worth of land and residential are regularly determined at the earliest by the vendor, however determining the proper price within the sales process will affect the buyer's desire to elect and bid. Special characteristics in Indonesia, tax object value (NJOP) and site parameters are high influence to the worth. during this paper we proposed the prediction of land and house value using several methods. symbolic logic, Artificial Neural Network and K-Nearest Neighbor are compared during this paper to get the foremost appropriate method which will be used as a reference for determining the worth by the sellers. Google Maps is employed to represent the spatial data for prediction parameter.

9. Comprehensive Analysis of Housing Price Prediction in Pune Using Multi-Featured Random Forest **Approach Authors:** Rushab Sawant; Yashwant Jangid; Tushar Tiwari; Saurabh Jain; Ankita Gupta, 2018

The housing sector in India has been predicted to grow at 30-35% over subsequent decade. In terms of employment provided, it's second only to the agricultural sector. Housing is one among the main domain of land. Pune is emerging together of the main metropolitan cities of India and has many prestigious educational institutions and IT parks. This makes it a perfect place to shop for homes. Vagueness among the costs of homes makes it challenging for the customer to pick their dream house. 12 The interests of both buyer and seller should be satisfied in order that they are doing not overestimate or underestimate price.

10. Time-Aware Latent Hierarchical Model for Predicting House Prices
    Authors: Fei Tan; Chaoran Cheng; Zhi Wei, 2017

It is widely acknowledged that the worth of a home is the mixture of an outsized number of characteristics. House price prediction thus presents a singular set of challenges in practice. While an outsized body of works are dedicated to the present task, their performance and applications are limited by the shortage of while span of transaction data, the absence of real-world settings and therefore the insufficiency of housing features. to the present end, a time-aware latent hierarchical model is introduced to capture underlying spatiotemporal interactions behind the evolution of house prices. The hierarchical perspective obviates the necessity for historical transaction data of exactly same houses when temporal effects are considered.

V. METHODOLOGY

About Data Set: - We are using a dataset from Kaggle called 'Bangalore House Price Prediction.' There are **13320 rows** in this dataset, and each row represents a distinct entry. Apart from these, the dataset includes features for **Area_Type, Availability, Location, Size, Society, Total_Sqft, Bath, Balcony, and Price** offering a complete set of attributes for analysis the real estate market of Bangalore.

To analysis the data in this dataset, we use a variety of technologies, including:

- **Pandas:** A processing package that allows data to be filtered according to predefined criteria.
- **NumPy:** This tool is used to manipulate the dataset mathematically.
- **Matplotlib and Seaborn:** These libraries are essential for data visualization since they provide graphical representation of a data set and helps a more thorough comprehension of it.
- **Scikit-learn:** Machine learning algorithms are implemented using this set of tools to predict features based on the dataset.

```
data['bhk'] = data['size'].apply(lambda x:int(x.split()[0]))
data['bhk'].unique()
```

```
array([ 2,  4,  3,  6,  1,  8,  7,  5, 11,  9, 27, 10, 19, 16, 43, 14, 12,
       13, 18], dtype=int64)
```

1) **DATA CLEANING**

Managing Value Missing: Taking care of values that are missing is an essential part of data cleansing. The 'Society' and 'Balcony' features in our dataset have the most noticeable missing values.

```
data.isnull().sum().sort_values(ascending=False)
```

```
society         5502
balcony          609
bath              73
size              16
location           1
area_type          0
availability       0
total_sqft         0
price              0
dtype: int64
```

Choosing Features: In order to prepare our dataset for further modelling, we remove features that aren't needed, such **Society, Balcony, and Availability**. We choose not to include these elements since they can cause errors and make our modelling process more difficult.

```
data.drop(['availability','society','balcony'],axis=1,inplace=True)
```

Approving Missing Data: Using the **Fillan ()** method, we fill in any missing information for locations with **Sarjapur Road**. As the most common answer in the **Size** characteristic, we choose **2BHK** which guarantees a

```
data['location']=data['location'].fillna('Sarjapur Road')
```

more stable dataset.

```
data['size']=data['size'].fillna('2 BHK')
```

Managing the Last Missing Values: Using strategic imputation approaches, the remaining characteristics with missing values are included. For example, we use statistical measures like **Mean (), Median (), or mode ()** to fill in the missing data in the **Bath** feature.

```
data['bath']=data['bath'].fillna(data['bath'].median())
```

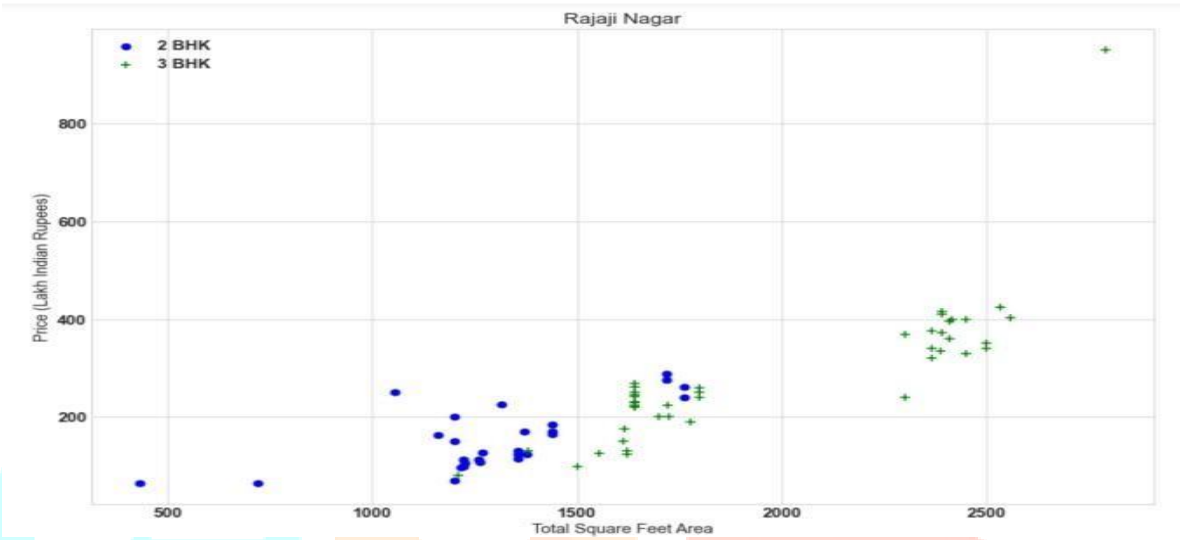
Engineering Features: We execute feature engineering in the **Size** feature by extracting the initial index, turning it into an integer, and adding a new column called **BHK.** Through the capture of an important component of the property, this improves the dataset.

New Functionality: price_per_sqft Add We improve the dataset's richness by adding values produced from the formula **price * 10000 / total_sqft,** and by introducing a new column named " **price_per_sqft** Add." This offers a more thorough understanding of the correlation between cost and total square footage.

```
data['price_per_sqft']=data['price'] * 100000 / data['total_sqft']
```

2) DATA VISULIZATION

Comparing Price Using Total_Sqft Footage: Our analysis compares the **price** of real estate to the overall area in **square feet**, with a particular emphasis on **2Bhk** and **3Bhk** layouts. According to our research,



residences in Rajnagar city featuring three-bedroom layouts are more highly valued.

Finding Outliers: We begin by characterizing the dataset's statistical features, defining its lowest and maximum values, in order to improve its integrity. After that, we set a cutoff point and remove any rows with fewer than 300 total square feet or bedrooms.

```
In [141]: # remover less than 300 sqft
          data=data[((data['total_sqft']/data['bhk']) >= 300)]
          data.describe()
```

Out[141]:

	total_sqft	bath	price	bhk	price_per_sqft
count	12530.000000	12530.000000	12530.000000	12530.000000	12530.000000
mean	1594.564544	2.559537	111.382401	2.650838	6303.979357
std	1261.271296	1.077938	152.077329	0.976678	4162.237981
min	300.000000	1.000000	8.440000	1.000000	267.829813
25%	1116.000000	2.000000	49.000000	2.000000	4210.526316
50%	1300.000000	2.000000	70.000000	3.000000	5294.117647
75%	1700.000000	3.000000	115.000000	3.000000	6916.666667
max	52272.000000	16.000000	3600.000000	16.000000	176470.588235

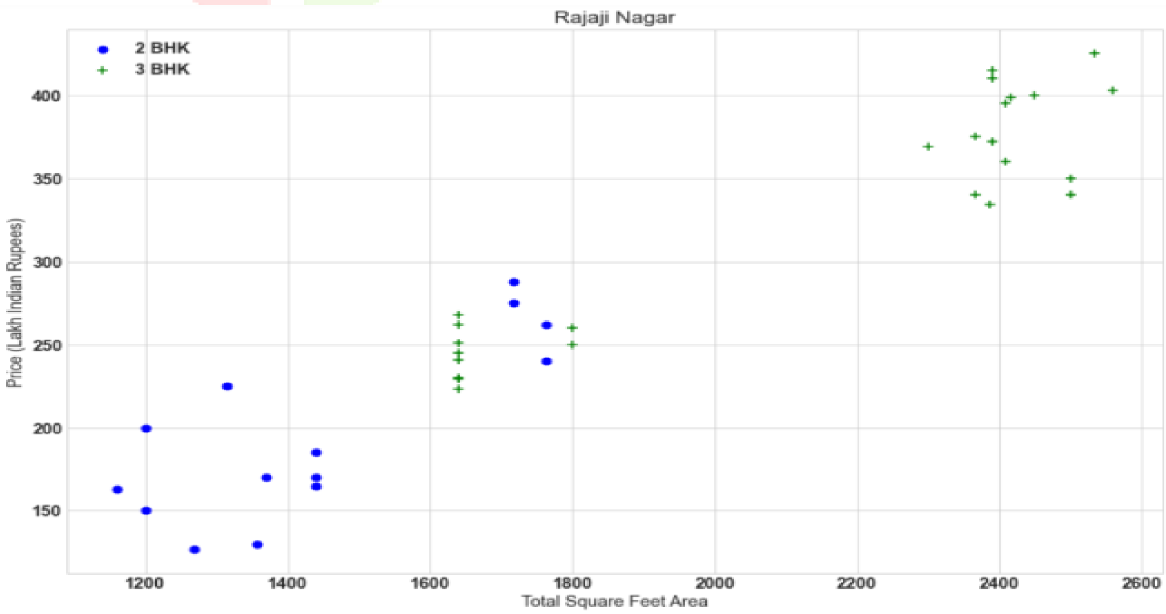
Total_Sqft Outlier Identification Method: To prevent extreme values from distorting our research, we employ a strong algorithm in the total_Sqft footage column to identify and eliminate outliers.

```
def remove_outlier_sqft(df):
    df_output=pd.DataFrame()
    for key,subdf in df.groupby('location'):
        m=np.mean(subdf.price_per_sqft)

        st=np.std(subdf.price_per_sqft)
        gen_df=subdf[(subdf.price_per_sqft > (m - st)) & (subdf.price_per_sqft <= (m + st))]
        df_output=pd.concat([df_output,gen_df], ignore_index=True)
    return df_output
data=remove_outlier_sqft(data)
data.describe()
```

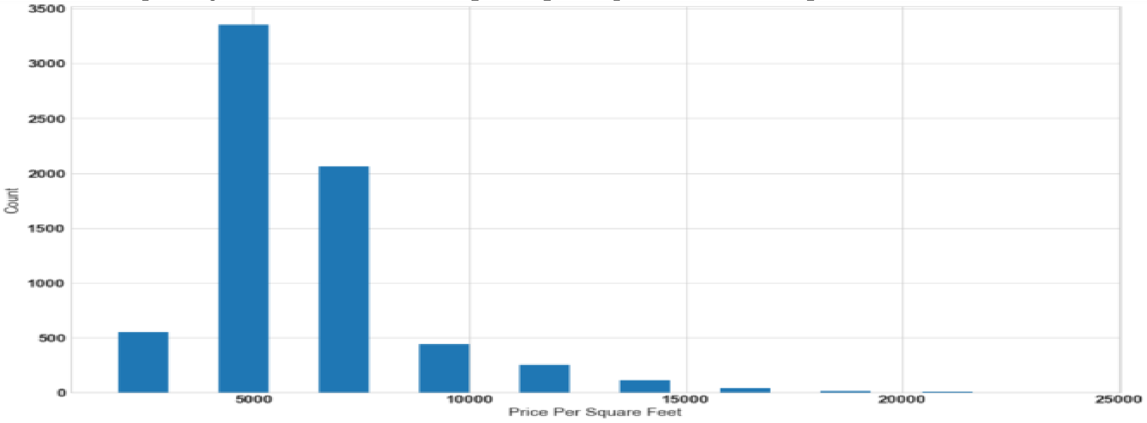
BHK Outlier Elimination: The **Bhk** column undergoes comparable outlier reduction procedures, which are implemented methodically to improve the dataset's dependability.

```
def bhk_outlier_remove(df):
    exclude_indices=np.array([])
    for location,location_df in df.groupby('location'):
        bhk_stats={}
        for bhk ,bhk_df in location_df.groupby('bhk'):
            bhk_stats[bhk]= {
                'mean':np.mean(bhk_df.price_per_sqft),
                'std':np.std(bhk_df.price_per_sqft),
                'count':bhk_df.shape[0]
            }
        for bhk ,bhk_df in location_df.groupby('bhk'):
            stats=bhk_stats.get(bhk-1)
            if stats and stats['count']>5:
                exclude_indices=np.append(exclude_indices,bhk_df[bhk_df.price_per_sqft<(stats['mean'])].index.values)

    return df.drop(exclude_indices,axis='index')
```

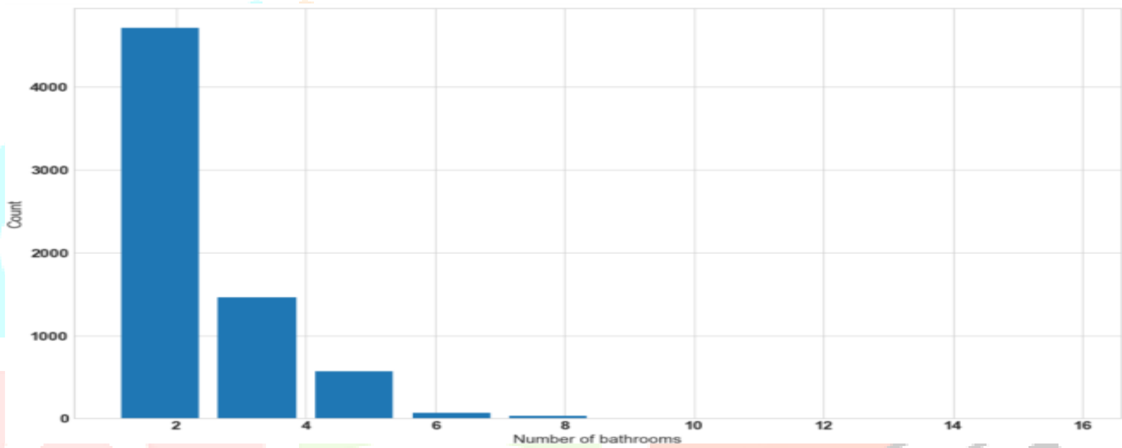
Visualization Following Outlier Elimination: After the thorough removal of anomalies, we visualize the dataset and compare **pricing** to the **total square foot** footage for location-based **2Bhk** and **3Bhk** setups.



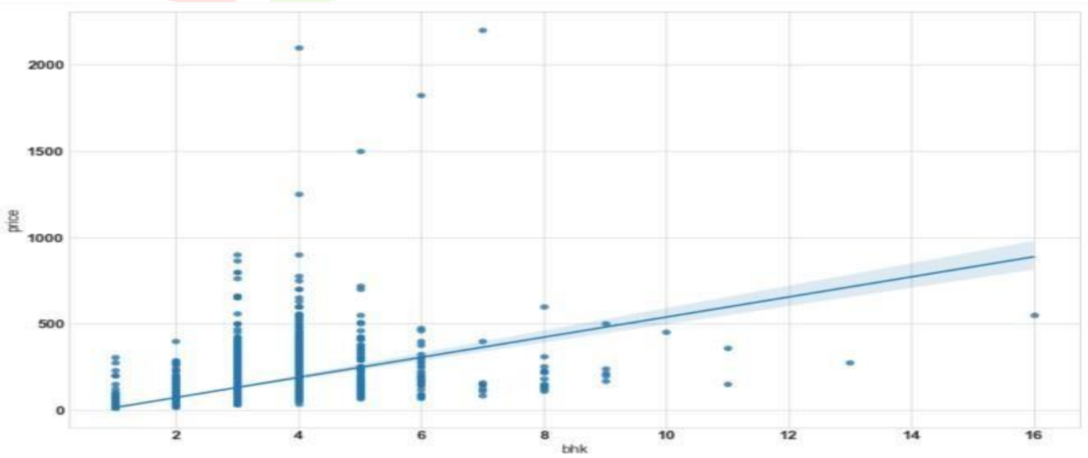
Analysis of Price Per Square Foot Frequency: In order to shed light on the pricing trends in our dataset, we examine the frequency distribution of the price per square foot in depth.



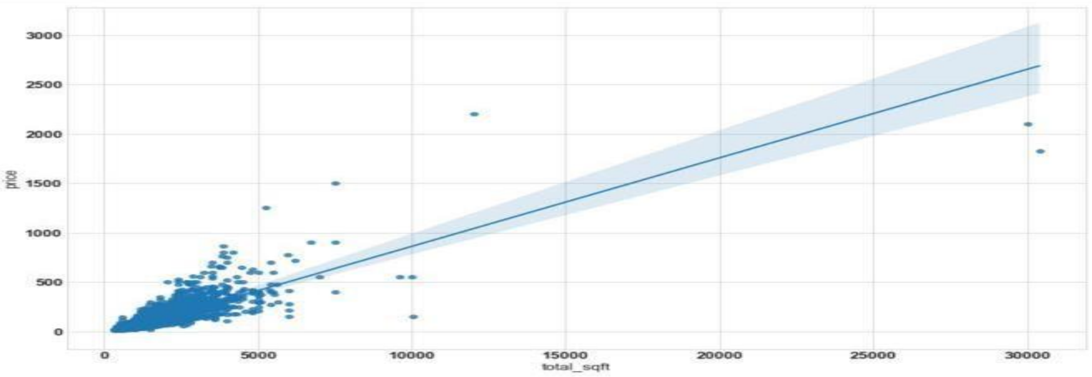
Analysis of Bathroom Frequency: An extensive analysis of the dataset provides insight into the most common bathroom designs by displaying the frequency distribution of the number of bathrooms.



Plot of Seaborn Regression: We create a regression figure using Seaborn to show the connection between **property values** and the number of **bedrooms (Bhk).** Understanding how the number of bedrooms affects **pricing** dynamics is made easier with the help of this plot.



We also demonstrate how **property values** and **total square** footage are related. This figure makes it easy to understand how **pricing dynamics** are affected by **total square foot.**



APPLY MODEL:

Our data is separated into two sections: one for training and the other for testing. We include 80% of the data for training and 20% for testing.

```
x_train,x_test,y_train,y_test=train_test_split(x,y,test_size=0.2,random_state=0)
```

```
print(x_train.shape)
print(x_test.shape)
```

```
(5921, 4)
(1481, 4)
```

then, in order to apply the machine learning model, we are building pipeline and passing the data through pipeline.

```
pipe.fit(x_train,y_train)
```

```
Pipeline(steps=[('columntransformer',
                 ColumnTransformer(remainder='passthrough',
                                   transformers=[('onehotencoder',
                                                  OneHotEncoder(sparse=False),
                                                  ['location'])])),
                ('standardscaler', StandardScaler()), ('ridge', Ridge())])
```

We convert location features to numeric values using one hot encoding before running the model.

```
column_trans=make_column_transformer((OneHotEncoder(sparse=False),['location']),
                                     remainder='passthrough')
```

Utilizing five distinct machine learning models, such as **linear regression, Lasso Regression, Ridge Regression, Elastic net Regression**. we obtain a linear regression result of 89.49%

whereas other model Score is, the Score of **ridges regression** is 89.48%, The Score of **Lasso Regression** is 89.48%, and the Score of **Elastic net Regression** is 78.22%.

EXAMPLE: -

Apply Ridge

```
In [71]: ridge=Ridge()

In [72]: pipe=make_pipeline(column_trans,scaler,ridge)

In [73]: pipe.fit(x_train,y_train)

Out[73]: Pipeline(steps=[('columntransformer',
                          ColumnTransformer(remainder='passthrough',
                                            transformers=[('onehotencoder',
                                                           OneHotEncoder(sparse=False),
                                                           ['location'])])),
                         ('standardscaler', StandardScaler()), ('ridge', Ridge())])

In [74]: y_pred_ridge=pipe.predict(x_test)

In [75]: r2_score(y_test,y_pred_ridge)

Out[75]: 0.8948686298478481
```

ALL MODEL ACCURECY

```
print("Linear Regression ",r2_score(y_test,y_pred_lr))
print("Lasso Regression ",r2_score(y_test,y_pred_lasso))
print("Ridge  Regression ",r2_score(y_test,y_pred_ridge))
print("Elasticnet Regression ",r2_score(y_test,y_pred_es))

Linear Regression  0.8949492042671273
Lasso Regression  0.8948686298478481
Ridge  Regression  0.8948686298478481
Elasticnet Regression  0.7822139133621335
```

Thus, while our data is being passed through a pipeline, linear regression is the optimal model for it.

OVERALL ACCURACY IN TABLE FORMAT:

Model Name	R2_Score
Linear Regression	89.49
Lasso Regression	89.48
Ridge Regression	89.48
Elastic net Regression	78.22

VI. FUTURE WORK

Include any recent pertinent information, such as crime statistics, school rankings, accessibility to public transportation, or impending infrastructure projects. Include temporal elements that might significantly affect house prices, such as patterns or seasonality. Make your model usable in the real world by deploying it as an API or web application that users may interact with. Provide a mechanism for ongoing model monitoring and updating in response to shifting market dynamics and consumer preferences.

VII. CONCLUSION

Predicting home values is a difficult procedure that involves taking a lot of variables into account. A few things that can affect house prices are the situation of the economy as a whole, the property's location and condition, and the accessibility of financing choices. Predictions can be made using a range of techniques and strategies, such as expert judgment, machine learning algorithms, and statistical modeling. But it's crucial to remember that any prediction is just an approximation and not a promise of future value. Making well-informed judgments about the purchase or sale of a property can be aided by carrying out extensive research and speaking with real estate experts.

VIII. ACKNOWLEDGEMENT

We sincerely appreciate the essential advice and assistance provided by Professor Yassir Farooqui. We are appreciative of the research possibilities provided by Parul University's Department of Computer Science and Engineering (CSE). We also thank our colleagues and research participants for their assistance and input.

REFERENCES

[1] Aminah Md Yusof and Shahida Ismail, Multiple Regressions in Analyzing House Price Variations. IBIMA Publishing Communications of the IBIMA Vol. 2012 (2012), Article ID 383101, 9 pages DOI: 10.5171/2012.383101.

[2] Babyak, M. A. What you see may not be what you get: A brief, nontechnical introduction to over fitting regression- type models. Psychosomatic Medicine, 66(3), 411-421.

[3] Atharva choogle, Priyanka khaire, Akshata gaud, Jinal Jain. House Price Forecasting using Data Mining Techniques International Journal of Advanced Research in Computer and Communication Engine g ISO 3297:2007 Certified Vol. 6, Issue 12, December 2017

[4] Darshan Sangani, Kelby Erickson, and Mohammad Al Hasan, Predicting Zillow Estimation Error Using Linear Regression and Gradient Boosting, IEEE 14th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), Page(s):530 – 534

[5] Model, Azme Bin Khamis, Nur Khalidah Khalilah Binti Kamarudin, Comparative Study on Estimate House Price Using Statistical and Neural Network, International journal of scientific and technology, research volume 3, ISSUE 12, December 2014, Page(s):126-131.

[6] Adyan Nur Alfiyatin, Hilman Taufiq, Ruth Ema Febrita, Wayan Firdaus Mahmudy, Modeling House Price Predictionusing Regression Analysis and Particle Swarm Optimization Case Study: Malang, East Java, Indonesia. (IJACSA) International Journal of Advanced Computer Science and Applications, Vol. 8, No. 10, 2017, Page(s):323-326.

[7] Nageswara Rao Moparthi, Dr. N. Geenthanjali “Design and implementation of hybrid phase-based ensemble technique for defect discovery using SDLC software metrics”,An International Conference by IEEE, PP. 268 – 274,2016

[8] Nihar Bhagat, Ankit Mohokar, Shreyash House Price Forecasting using Data Mining. International Journal of Computer Applications 152(2):23-26, October 2016.

[9] Valeria Fonti, Feature Selection using LASSO Research Paper in Business Analytics, VU Amsterdam, March 30, 2017.

[10] Peter B. Luh, Neural Network-Based Market Clearing Price Prediction and Confidence Interval Estimation with an Improved Extended Kalman Filter Method, IEEE Transactions on Power Systems 20(1):59 - 66, March 2005.

[11]. Visit Limsombunchai, Christopher Gan and MinsooLee, House Price Prediction: Hedonic Price Model vs. Artificial Neural Network, Lincoln University, Canterbury 8150, New Zealand, American Journal of Applied Sciences 1 (3): 193-201, 2004.

[12]. Dr. Nageswara Rao Moparthi, Ch Mukesh, Dr. P. Viday Saga, “Water Quality Monitoring System Using IoT”, An International Conference by IEEE, PP. 109 – 113 ,2018

[13]. Ahmed Khalafallah, Neural Network Based Model for Predicting Housing Market Performance, Tsinghua Science & Technology 13(S1):325-328, October 2008.

[14]. Steven C. Bourassa, Eva Cantoni, Martin Edward Ralph Hoesli, Spatial Dependence, Housing Submarkets and House Price Prediction the Journal of Real Estate Finance and Economics, 143-160, 2007.

[15]. Chris Brunsdon, A. Stewart Fotheringham and MartinE. Charlton, Geographically Weighted Regression: A Methodfor Exploring Spatial Nonstationarity, Geographical analysis, Volume 28, Issue 4, Pages: 281-375,1996.