

# DEVELOPMENT OF AN AUTOMATED RESIDENTIAL PROPERTY PRICE PREDICTING SYSTEM USING LINEAR REGRESSION MODEL

**Funke Florence Abiola**

Department Computer Science, The Federal Polytechnic, Idah, Kogi State

[funke.abiola.940@gmail.com](mailto:funke.abiola.940@gmail.com)

**ABSTRACT** One of the most important investments in the world today is real estate, particularly in places like Lagos, Abuja, Rivers, and the like, which are also popular places for people to live and work. Therefore, before you spend your hard-earned money on any property, it is crucial to know the current valuation of the residence. This paper's primary goal is to forecast the present market value of residential real estate in Nigeria. While doing so, consideration is given to several factors, such as the number of bedrooms and the availability of various amenities. The purpose of this forecast is to assist a customer in finding feasible solutions that better fit their needs. The cost of the different residences in question was predicted using the linear regression model. This concept also benefits the customer by doing away with the necessity to speak with a broker. The purpose of the study is to determine how the dependent variable—asking price—relates to the exploratory factors, which include neighbourhood, garage capacity, number of bedrooms and bathrooms, land parcel size, and home square footage. The developed model will allow for the prediction of housing market values based on a variety of property attributes using the linear regression method. This would allow both buyers and sellers to make an initial appraisal of any property in Nigeria. The data for this study was gathered from an online resource known as the UCL Repository. The data collected for this study will be used to specify the model, generate parameter estimates, check for model inadequacy, repeat until the model is adequate, validate the model, repeat until the model is valid, and then use the final model to estimate asking prices for properties in each Nigerian neighbourhood. The property asking price is the dependent variable, and five quantitative independent variables using linear regression: home square footage, lot square footage, number of bedrooms, number of bathrooms, year of construction, and number of cars accommodated in a garage will be the main focus of this work.

**KEYWORDS** Linear Regression Model, Real Estate, Data Mining, Housing Price Prediction

## I. INTRODUCTION

Nigeria's housing market is one of the most varied in the nation, with significant regional variations in costs. The size of the land parcel, the square footage of the house, and the facilities all have an impact on the cost of residential housing in the metropolitan areas of Kano, Lagos, Abuja, Port-Harcourt, and others. According to intuition, the size of a land lot will have a greater influence on the cost of residential property in the city centre, whereas the square footage of a home and features like fireplaces, garages, and swimming pools would have a greater impact on suburban area costs. The purpose of the study is to determine how the dependent variable—asking price—relates to the exploratory factors, which include neighbourhood, garage capacity, number of bedrooms and bathrooms, land parcel size, and home square footage. The developed model will make it possible to anticipate housing market values based on a variety of property attributes using the linear regression method. This would allow both buyers and sellers to make an initial appraisal of any property in Nigeria. The UCL Repository is an online resource from which the data for this study will be collected. Using the data gathered for this study, the model will be specified, parameter estimates will be created, the model's adequacy will be checked, it will be repeated until it is

adequate, it will be validated until it is valid, and the final model will be used to estimate the asking price of real estate in each neighbourhood in Nigeria. This study will primarily focus on five quantitative independent variables using linear regression—home square footage, lot square footage, number of bedrooms, number of bathrooms, year of construction, and number of cars accommodated in a garage—with the asking price of the property serving as the dependent variable. The accepted method for simulating the behavior of home prices is regression (Schulz and Werwatz, 2004). According to the model, a good has a wide range of qualities that come together to create bundles of qualities that the consumer values and that have an impact on utility (Ching and Chau, 2003). The price of the house is often estimated by the model based on its qualities. It calculates each housing attribute's marginal contribution to a home's price. It is important to remember that factors that influence home values in one location may not have the same effects in another. The hedonic technique has been used in numerous studies to estimate home prices, with varying findings from the researchers. Therefore, the results cannot be generalized (Abdulai and Owusu-Ansah, 2011). In response, (Sirmans et al., 2006) pointed out that the predicted coefficients for a few features from earlier research differ greatly depending on the location. Therefore, it seems sense to conclude that before any conclusions can be drawn, a separate investigation is necessary for each market or location. Very little research has been done using house price estimation in the Nigerian housing sector. In light of this, the necessity to use the method to pinpoint the primary factors influencing home values in the research region in order to direct real estate developers is what spurred this study. Additionally, the work will serve as a foundation for future research in this area. The housing market has a significant impact on economic activity. By increasing overall spending, employment, and home sales volume, housing building and renovation stimulate the economy. Additionally, they mimic the demand for related businesses like home durables. For the majority of households, when a home is the greatest single asset, fluctuations in home prices have an impact on the value of the asset portfolio (Leatham, 2011). Prospective homeowners, developers, investors, appraisers, tax assessors, and other real estate market players, such as mortgage lenders and insurers, rely on accurate house price estimates (Frew, 2007). Traditional house price forecasting relies on comparing expenses and sales prices in the absence of a recognized standard or certification system. Thus, having a house price prediction model accessible helps to bridge a substantial information gap and improves the efficiency of the real estate market (Calhoun 2008).

## II. RELATED WORKS

The general idea of this research project is provided by the related work on property price prediction that is covered in this part. Two significant issues in statistics are classification and regression. In each, a response variable  $y$  is predicted based on the values of a vector of predictor variables  $x$ . Let  $X$  stand for  $x$ 's domain and  $Y$  for  $y$ 's. The issue is known as regression if  $y$  is a continuous or discrete variable that takes real values, such as an automobile's weight or the frequency of accidents. Otherwise, the issue is known as categorization if  $Y$  is a finite set of unordered values (such as the kind of car or its nation of origin). The challenge, expressed mathematically, is to identify a function  $d(x)$  that converts every point in  $X$  to a point in  $Y$ . It is necessary to have a training sample of  $n$  observations  $L = \{(x_1, y_1), \dots, (x_n, y_n)\}$  in order to generate  $d(x)$ . The topic is referred to as supervised learning in computer science. Similar to this, a regression tree is a tree-structured solution where each partition's data is fitted to a constant or a comparatively basic regression model.

### REGRESSION MODEL

Regression problem-solving classifiers examine the connection between the numerical target (dependent) variables and the attribute- (independent) variables (van Wijk, 2008). Numerous methods that forecast a continuous numerical value can be used for this. Regression can handle both linear and nonlinear interactions and is appropriate for handling complicated prediction issues (Fayyad, Piatetsky-Shapiro & Smyth, 1996; van Wijk, 2008). Generally speaking, there are more potential outcomes from regression jobs than from classification using broad groups.

### TREE

Starting at the top, a tree symbolizes a sequence of splits. A tree is made up of nodes joined by edges. A unique data structure for data storage is the binary tree. The unique requirement for a binary tree is that each node can have no more than two children. Because insertion or deletion operations in a binary tree are as speedy as in a linked list and searches are as quick as in a sorted array, a binary tree combines the advantages of both linked lists and ordered arrays. With a root value and subtrees of children with a parent node represented as a collection of connected nodes, a tree is a commonly used abstract data type (ADT) or data structure in computer science that implements an ADT and simulates a hierarchical tree structure (Amit & Neelima 2008). With the restrictions that no reference is duplicated and none points to the root, a tree data structure can be defined recursively (locally) as a collection of nodes (beginning at a root node), where each node is a data structure made up of a value and a list of references to nodes (the "children") (Amit & Neelima 2008). An alternative definition of a tree is an ordered tree with a value assigned to each node that is abstracted as a whole (globally). Although a tree can be mathematically analyzed as a whole, it is typically represented and worked with separately by node when it is represented as a data structure (as opposed to a set of nodes and an adjacency list of edges between nodes, as one may represent a digraph, for example). When considering a tree in its whole, for instance, one can refer to "the parent node" of a specific node; yet, in a data structure, a given node typically just stores the list of its offspring. however, makes no mention of its parent (if any) (Amit & Neelima 2008). When adjusting the hyperparameters of the model (such as the number of hidden units in a neural network). By early terminating training when the error on the validation dataset rises—a symptom of overfitting to the training dataset—validation datasets can be used for regularization. The fact that the error of the validation dataset may change during training, resulting in several local minima, complicates this straightforward process in practice. Due to this complexity, numerous ad hoc methods have been developed to determine when overfitting has actually started (Jae Sung Lee & Eun Sung Lee, 2014). The test dataset, also known as a holdout dataset, is used to evaluate the final model fit on the training dataset. While the algorithms will be trained on our data, a separate test set is required to generalize the results (Witten, Frank & Hall, 2011).

### PREVIOUS WORKS ON PRICE PREDICTION

There are several studies that forecast real estate selling prices (Knight, 2002; Dubin, 1998; Peterson, 2008; Nguyen & Cripps, 2001). However, to our knowledge, no research on real estate prices utilizing text mining have been published. This implies that we must compare our findings to similar research that use various attribute variables, often known as dependent variables. These studies frequently employ broad real estate features to create predictions. This strategy is commonly known as hedonic price prediction (Peterson, 2008). Utilized as variables for attributes. We provide a summary of many works on price prediction using machine learning approaches in this section. We will use these as a contrast to our findings on asking price prediction because all of the studies that we identified concentrate on selling price prediction. The aforementioned research serve as the foundation for our previously stated hypothesis and offer a point of comparison for the outcomes of our investigation. A 1998 study by Dubin demonstrates a way to forecast real estate prices based on their attributes. The goal of this strategy was to demonstrate that it could be used to anticipate real estate prices by estimating the price values using maximum likelihood regression. Attribute variables were general real estate features. (Dubin, 1998) observes that the model had significant prediction errors and that adding spatial data allowed it to be improved. He points out that this significantly

improved RMSE, although he suggested using this strategy in addition to more conventional approaches. As digital technology advanced quickly, Nhuyen and Cripps released their article on real estate price forecast in 2001. Multiple Regression Analysis (MRA) was employed by the researchers in their study to predict prices. Their predictions were based on general housing features, much like Dubin (1998), and they employed training sets of three different sizes. Nhuyen and Cripps got price fluctuation results, which are the difference between asking and selling prices, and imply that MRA works better on smaller training sets. The researchers do not give conclusive results for their measurements. In addition, they discovered that the size of their training set had a significant impact on overall performance. Peterson (2008) conducted research that provided relevant comparisons for RMSE measurements. The survey included 46,467 residential homes from 1999 to 2005. In this study, Artificial Neural Networks (ANN) are employed as a nonlinear modelling technique and compared to a linear regression. According to earlier studies, both are used to forecast real estate prices utilizing attribute variables related to general house characteristics. Because training set size affects performance, Peterson trains using 10%, 25%, 50%, and 75% of his whole dataset. According to his experiment, ANN produces fewer mispricing mistakes than linear regression. He comes to the conclusion that nonlinear models, such as ANN, can prevent price errors in linear models, which can be rather serious. There are research on price prediction in various fields outside real estate. Ghani and Simmons (2004) experimented with both regression and classification to forecast online auction final prices. Using 1,300 training samples and 400 testing samples, attribute variables such as the seller's rating, the auction's minimum price, and the dates of the auction were used. Regression trees and linear regression were employed to forecast precise price values to the Nigerian.

### III. METHODOLOGY

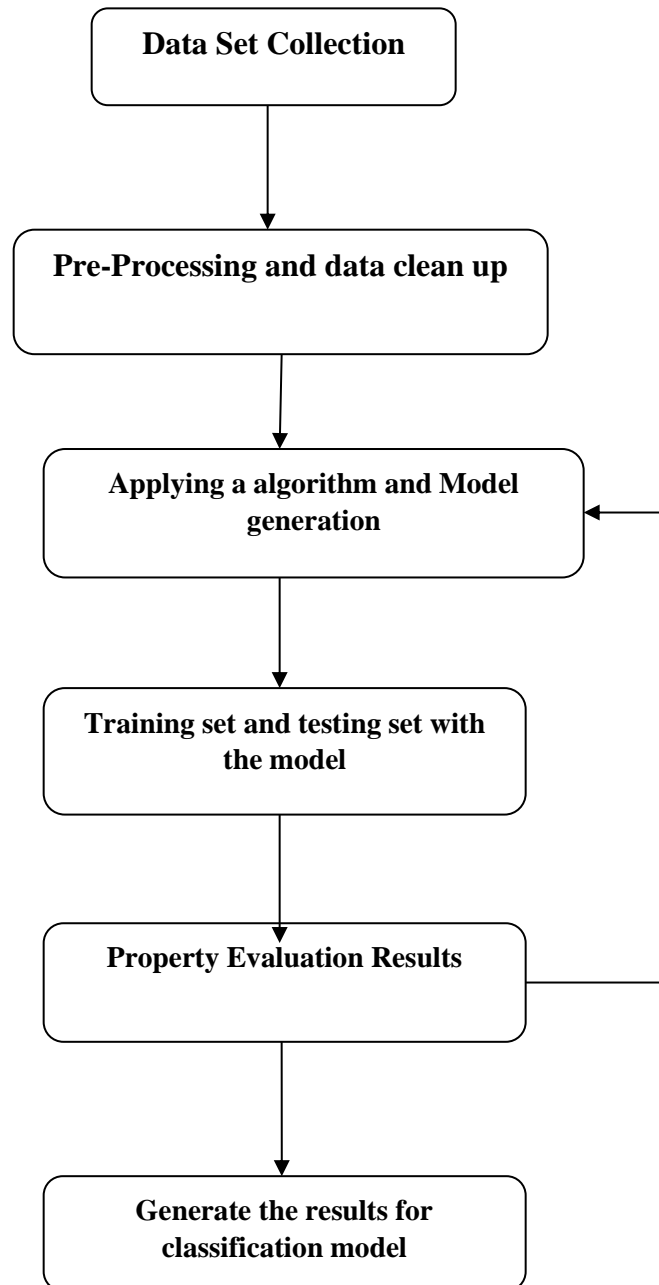
A regression tree and linear regression are built using binary recursive partitioning, an iterative process that splits the data into partitions or branches and then further divides each partition into smaller groups as the algorithm progresses along each branch. The Training Set, which comprises of pre-classified records required to determine the tree's structure, is originally separated in the same manner. Following that, the algorithm begins separating the data into the first two branches or divisions, utilizing each field's possible binary split. The method selects the split that minimizes the total of the squared deviations from the mean over the two separate divisions. Each newly generated branch is then subjected to the splitting rule. This operation is repeated until each node becomes a terminal node and fulfils the user-specified minimum node size. Even if a node has not reached the minimum size, it is considered a terminal node if the total of its squared deviations from the mean equals zero. Regression analysis examines the relationship between one or more independent variables and a dependent variable. Based on the cost and characteristics of recent home price transactions, the Estate Management and Valuation Department at Federal Polytechnic, Idah (F.P.I.) could be the source of data for this study. This study aims to forecast property prices using regression trees (SAP). Regression analysis is a statistical process used to determine the relationship between one or more independent variables and a dependent variable. When one of the independent variables is altered while the other independent variables stay constant, the regression approach allows you to understand how the average value of the dependent variable (also known as the criterion variable) evolves.

#### A. SYSTEM DESIGN

Analyzing the sample of data that provides a thorough features study of property price is made possible by data mining. This method's primary focus is on analyzing the standard dataset of attributes that is currently available, as well as pre-processing and getting it ready for additional analysis. At this point, the dataset is cleaned up of extraneous information that could affect predictions negatively or make computations more difficult, and the necessary filters are applied. Then, several models created for predicting property prices are based on the remaining significant features. The primary classification process now starts, and the suggested classification algorithms can be applied to the existing dataset. Following classification, each method's key performance metrics were computed. This is the work's prototype design.

**SUGGESTED ALGORITHM AND HARDWARE/SOFTWARE NEEDS**

The regression model is the algorithm I used to create the system that implements the predictions. The specifications and prerequisites that I utilized for the project are listed below: Windows OS, Netbeans, Weka, Microsoft Excel (.CSV file), and a computer system with a keyboard and vmouse. It had four gigabytes of RAM.



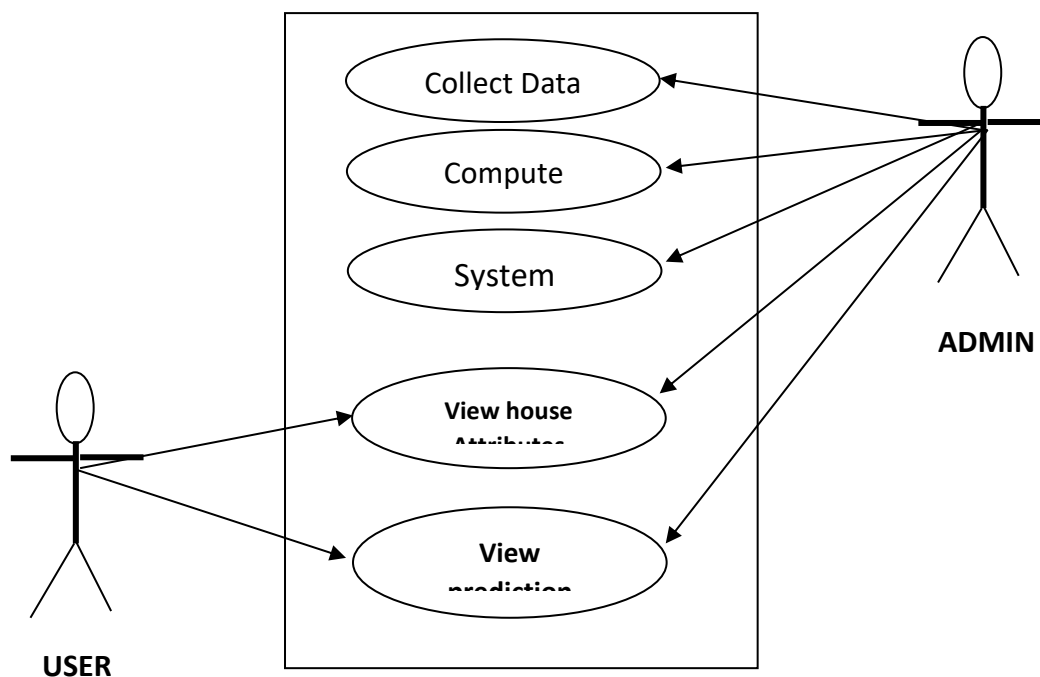
**Figure 1:** The Model Prediction of Property Price using Regression Tree

### CLASSIFICATION OPERATION

The data in this instance is prepared for the use of a classification algorithm. Classification is done on test data after the model has been created using training data. Following that, a number of the most significant metrics related to property prices are assessed and contrasted, including accuracy, precision, sensitivity, specificity, F-measure, and area under the ROC curve.

### USE CASE DESIGN

A use case is a technique used in system analysis to identify, outline, and categorize system requirements. The use case is a collection of possible interactions between people and systems in a given situation that are related to a certain goal. It consists of a set of components (such as classes and interfaces) that may be combined to provide a bigger impact than the sum of their individual parts. Similar to Figure 1 above, Figure 2 gives a brief summary of the system.



**Figure 2:** The Used Case Model Prediction of Property Price

### B. DATASETS

How to forecast the final feature, "Price," based on other attributes is the primary goal of the study. There are numerous properties with string descriptions during the data collection process, but since the algorithms we are studying require numerical values, some conversions are required. The table displays the conversion details. The basic principle is to assign an integer value to a string value of an attributes according how many records that "string value" classifies in whole data set. Attribute analysis can now be carried out on the dataset in order to achieve the goal of the Research work from it datasets.

### C. THE DESCRIPTION AND VALUE OF DATA FIELDS



ATTRIBUTE NAME DESCRIPTION VALUE 1 Bedroom(BR) The number of bedrooms. A numerical value 2 Country The country of Residential A string 3 ID The number of bedrooms A numerical value 4 Zip Code The number of bedrooms A numerical value 5 Location Geographical Location String 6 Kitchen(K) The number of kitchens A numerical value 7 Bathroom(T) The number of Bathroom A numerical value 8 Height (H) Height of the building A numerical value 9 State Location at state level String 10 accommodates The Number of Bedrooms A Numerical Value 11 Price(P) House price The numerical value of house price. 12 Room Type The type of rooms String 13 Property type The property type String Table

#### **D. FEATURE SET DESCRIPTION**

Numerous feature extraction techniques exist, such the Gain Ratio Attribute Evaluation method, ranker search, and others. Analysis can now be applied to the dataset to remove irrelevant characteristics.

#### **E. EXPERIMENTAL SETUP**

All of the predictions will be examined on a machine learning task where the sample dataset will be used for feature pricing. The experimental setup will be carried out based on the sample datasets and the property price prediction model, which was examined to maintain consistency in the outcome of the performed prediction. In order to predict property price each field will be analyze into training set and testing as well as verify the promising result.

#### **F. EVALUATION METRICS**

One method for summarizing a classification algorithm's performance is to create a confusion matrix. If your dataset has more than two classes or an uneven number of observations in each class, classification accuracy by itself may be deceptive. You can have a better understanding of what your classification model is doing correctly and what kinds of mistakes it is making by calculating a confusion matrix. Based on this approach, one may determine the confusion matrix for a machine learning problem. Following the analysis, you must apply the following technique:

- (a). What is the confusion matrix and why should you utilize it.
- (b). How can I generate a confusion matrix for a two-class classification task from scratch?
- (c). How to generate a confusion matrix with Weka, Python, and R.

### **IV. IMPLEMENTATION AND RESULTS**

This chapter analysis the concept of tree regression algorithm for developing a system in predicting the price of properties. Based on this requirement, analysis was carried to collect the records of recent transactions.

#### **A. IMPLEMENTATION**

The project work was implemented with sets of 13 features or attributes and was structured into the Excel, Weka tool and Netbean as to predicting the price which was implemented as statistical information. The results were based on inputting require filled as well as to providing the prediction solution. The below interface shows the input interface

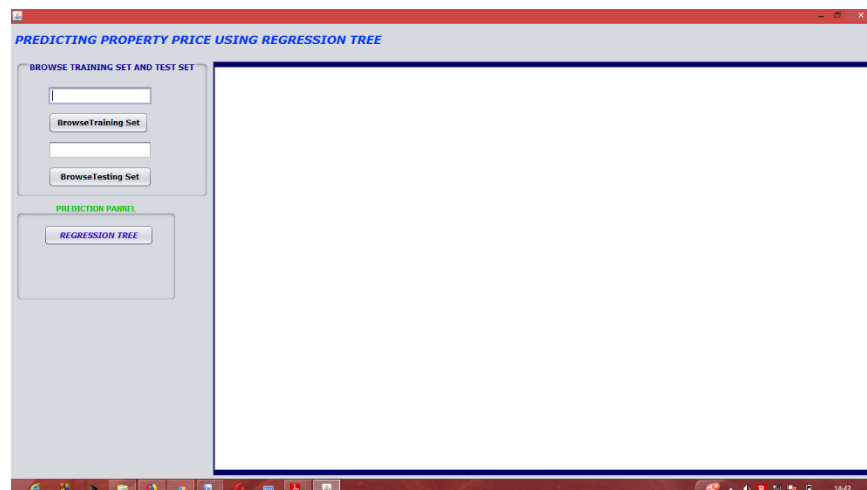


Figure. 1: Interface of Prediction

This dataset listed below data sets was gotten from referred sites from Estate Management and valuation department and it was used to perform this analysis, regression algorithm to build the model for predicting a promising result. Table 2 shows the practical example of this work.

ID	Zipcode	Location	County	State	Height of	Property	Room Type	Accommo	Bathroom	Bedrooms	kitchen	Price(Nmillion)
1	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	2	1	1	1	80
2	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	2	1	1	1	70
3	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	1	1	1	1	75
4	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	2	1	1	1	75
5	10002	Abuja	Nigeria	F.C.T	3	Apartment	Entire	4	1	2	1	250
6	10002	Abuja	Nigeria	F.C.T	3	Apartment	Entire	6	1	2	1	499
7	10002	Abuja	Nigeria	F.C.T	3	Apartment	Entire	1	1	1	1	65
8	10002	Abuja	Nigeria	F.C.T	3	Apartment	Entire	2	1	2	1	130
9	10002	Abuja	Nigeria	F.C.T	3	Apartment	Entire	6	1	2	1	249
10	10002	Abuja	Nigeria	F.C.T	3	Apartment	Entire	2	1	1	1	122
11	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	2	1	1	1	99
12	10002	Abuja	Nigeria	F.C.T	3	Apartment	Entire	3	1	1	1	150
13	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	2	1	1	1	95
14	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	1	1	1	1	90
15	10002	Abuja	Nigeria	F.C.T	3	Apartment	Private	2	1	1	1	90

Figure. 2: Interface of the Dataset Collected

Figure 2 shows the general concept of this project work, the analysis was done based on the sample of housing data set display above, the dataset was divided into 70% training and 30% testing for predicting property price, Weka was used to prepare the processing set up by training the model using the classifier algorithm to classify the datasets. Below is the data interface for the data sets after it has been classified into training and datasets.



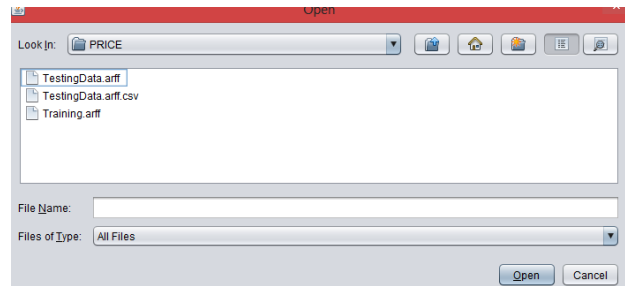


Figure 3: Classified Datasets

## B. DETAILS OF EXPERIMENTATION

The evaluation was done based on the housing attributes and the price point which was structured in order to prediction their property based on their price and numeric class equivalent of number of persons a room will accommodate with respect to model accuracy of regression comparison algorithm. The following is the output of running program of our proposed method.

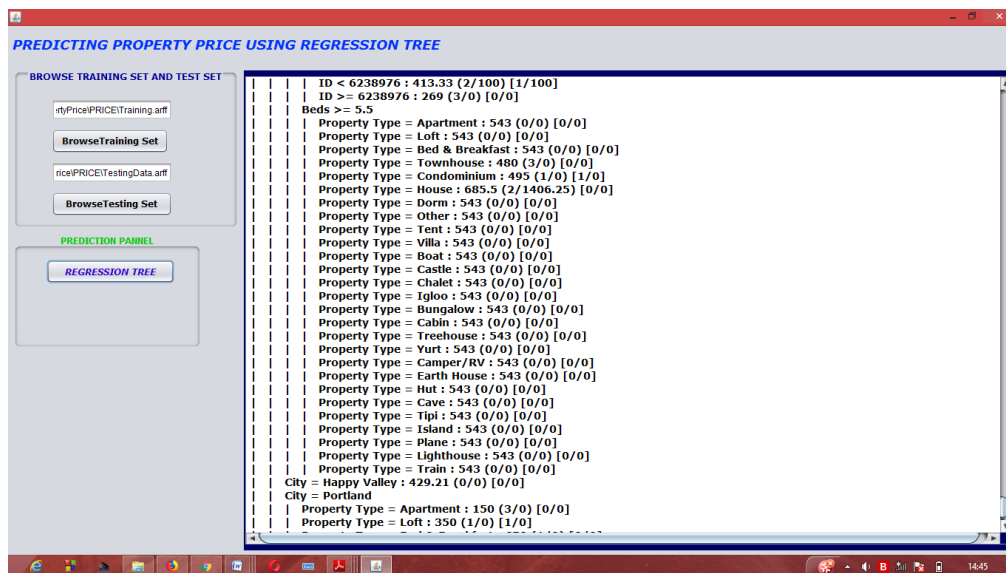


Figure 4: Output of the Predicted Result

Based on this analysis it shows that information gain from the sample of the dataset collected, the result show that RT (regression tree) accuracy measure is 50.5% based on this sample analysis was done with a series of step for RT where able to be used to predict the values of numeric price class are of variances in their results, the above present the analysis of property prediction price.

## V. CONCLUSION

This study has resulted in the development of a model that offers consumers a fresh perspective on future home price projections. Several regression approaches were tested and compared before settling on a prediction method-based regression analysis. We employed linear prior mean functions in our model to ensure that future price projections are more logical. We designed a strategy to use as much data as feasible in our prediction engine. This study has resulted in the development of a model that offers consumers a fresh perspective on future home price projections. Several regression approaches were tested and compared before settling on a prediction method-based regression analysis. We employed linear prior mean functions in our model to ensure that future price projections are more logical. We designed a strategy to

use as much data as feasible in our prediction engine. Take into account the model's performance. The strategy of developing a model for producing predictions has proven successful and may be used to various problem sets with regional differences in the prediction model. In this study, we found that the kind of home, its size and age, number of bathrooms, condition, location, distance, availability of neighbourhood services, and the year the property was sold all had a substantial impact on house price. However, the number of rooms is not particularly large. The finding on the number of rooms was not surprising. In Abuja, it is common to see tenement buildings with a larger number of rooms that are frequently occupied by different families; thus, a tenement house with, say, 14 rooms and occupied by multiple families may be the same size (in terms of total floor area) as a detached bungalow with only 5 rooms and more amenities and occupied by one family. This is assumed to be because the samples included in the study contain a range of various types of properties, including single-family homes with fewer rooms and multifamily homes with more rooms, even if both have roughly identical total floor space.

This shows that some communities in the research region have low living standards and poor housing conditions. It is believed that if separate regressions had been undertaken for the various property types, the problem may have been remedied. However, the current study was unable to do this due to a lack of recorded transaction data required for analysis. We conclude that the physical condition of the property, the number of bathrooms, the availability of utilities, the house's location, and the kind of home are the most important criteria that buyers consider when selecting whether to buy a home in Abuja. I suggest that going forward, Abuja's residential real estate developers should take these aspects into account while creating new properties. When determining the price of homes, participants in the real estate market should also consider these elements. In the future, it is recommended that adequate data should be acquired for the different house types so that each can be tested separately. In the future, we ought to investigate whether the data inferred from a house's satellite image may be utilized to calculate its cost. The list of upcoming research and enhancements that will enhance the results' applicability for commercial use follows. Including information regarding a home's condition at the time of the transaction is another enhancement.

## REFERENCES

- [1] Limsombunchai, V., Gan, C., and Lee, M. (2004). 'House price prediction: hedonic price models vs. artificial neural network', *American Journal of Applied Sciences*, 1(3), 193 – 201..
- [2] Mohamad, J. B. (2012). Assessment of Property Values in Thin Market Using Rank Transformation Regression and Multiple Regression Analysis. Master of Science, Universiti Teknologi Malaysia.
- [3] Ojetunde, I. (2013). Revisiting the Interaction between the Nigerian Residential Property Market and the Macroeconomy. Peer Review paper presented at FIG Working Week 2013, 6-10 May, Abuja, Nigeria.
- [4] Ojetunde, I., Popoola, N. I., and Kemiki, O. A. (2012). 'On the Interaction between the Nigerian Residential Property Market and the Macroeconomy', *Journal of Geography, Environment and Planning (JOGEPP)*, 7(2). Retrieved from: <http://ssrn.com/abstract=2105998>
- [5] Owusu-Ansah, A. (2012). Examination of the Determinants of Housing Values in Urban Ghana and Implications for Policy Makers. African Real Estate Society Conference. 24 -27 October 2012. Accra, Ghana.
- [6] Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., and Case, N. (2003). 'Real estate appraisal: a review of valuation methods', *Journal of Property Investment and Finance*, 21(4), pp.383-401.
- [7] Rosen, S. (1974). 'Hedonic prices and implicit markets: product differentiation in pure competition', *The Journal of Political Economy*, 82(1), pp.34-55.
- [8] Schulz, R., and Werwatz, A. (2004). 'A state space model for berlin house prices: estimation and economic interpretation', *The Journal of Real Estate Finance and Economics*, 28(1), pp.37-57.
- [9] Selim, S. (2008). 'Determinants of house prices in turkey: a hedonic regression model', *Doğuş Üniversitesi Dergisi*, 9(1), pp.65 – 76.

- [10] Shapiro, E., Mackmin, D., and Sams, G. (2013). *Modern Methods of Valuation*. Routledge, Abingdon.
- [11] Sirmans, G. S., MacDonald, L., Macpherson, D. A., and Zietz, E. N. (2006). 'The value of housing characteristics: a meta-analysis', *The Journal of Real Estate Finance and Economics*, 33(3), pp.215-240.
- [12] Sirmans, S. G., and Macpherson, D. A. (2003). 'The state of affordable housing', *Journal of Real Estate Literature*, 11(2), pp.131-156.
- [13] Abigail Bola Adetunji, Oluwatobi Noah Akande, Funmilola Alaba Ajala, Ololade Oyewo, Yetunde Faith Akande, Gbenle Oluwadara, House Price Prediction using Random Forest Machine Learning Technique, *Procedia Computer Science*, Volume 199, 2022, Pages 806-813, ISSN 1877-0509, <https://doi.org/10.1016/j.procs.2022.01.100>.
- [14] Hang, T. (2010). 'The impact of neighborhood types on the prices of residential properties', *Sunway Academic Journal*, 7, pp.77-88.
- [15] Watkins, C. A. (2001). 'The definition and identification of housing submarkets', *Environment and Planning A*, 33(12), pp.2235-2254.
- [16] Wezel, M. V., Kagie, M., and Potharst, R. (2005). Boosting the Accuracy of Hedonic Pricing Models. *Econometric Institute Research Papers*. No. EI 2005-50.
- [17] Ye, L (2013). The impact of neighborhood types on the prices of residential properties *Econometric Institute Research Papers*. 2 (1): 1-13

