Awareness of Artificial Intelligence Technologies for Residential Property Valuation in Lagos Metropolis, Nigeria

Amos Olaolu ADEWUSI and Oluwakemi Faith GBADEBO
Department of Estate Management,
Federal University of Technology, Akure, Nigeria
e-mail: aoadewusi@futa.edu.ng

DOI: 10.56201/rjpst.v6.no3.2023.pg204.225

Abstract

Property valuation is a core aspect of estate surveying and valuation practices, providing crucial estimates for real estate stakeholders including developers, institutional lenders, insurance companies among others. Traditional valuation methods often fall short in accurately estimating property values, leading advanced economies to promote the use of Artificial Intelligence Technologies (AIT). While significant research has been conducted on the awareness and adoption of AIT in developed countries, there is limited research on this topic in developing countries. This study aims to assess the level of awareness of AIT among professionals involved in property valuation in Lagos, Nigeria. Data for the study were collected from estate firms across Lagos metropolis and analyzed using descriptive statistics and Chi-square tests. Sixteen types of AI technologies were identified, and six training algorithms were evaluated. The results reveal a general lack of awareness of AIT among estate valuers practicing in Lagos metropolis. Specifically, for CatBoost, XGBoost, LGBM, and Random Forest, there is no significant relationship between years of professional qualification or educational qualification and awareness levels of AIT, as indicated by Pearson Chi-Square tests with p-values > 0.05. In addition, symmetric measures (Phi and Cramer's V values) indicate weak to very weak associations between years of professional experience, educational qualifications, and awareness levels of the selected AI technologies. This highlights a significant gap in the training of valuers in the study area. The findings offer valuable insights for property professionals, real estate investors, and policymakers, suggesting a need for enhanced training and awareness programs to bridge this gap.

Keywords: *Artificial Intelligence (AI), Property value, Property valuation*

1.0 Introduction

Property valuation is a critical aspect of estate surveying and valuation practices. According to the Royal Institution of Chartered Surveyors (RICS, 2006), valuation is defined as the opinion of a professional about capital value on a defined basis at a specified time. Over the years, property valuation practices have significantly evolved (Yomralioglu & Nisanci, 2004). The global trend toward modernization affects every sector, including real estate, particularly property valuation. Both traditional and advanced valuation methods have faced various criticisms. Traditional methods, in particular, are often considered subjective, inaccurate, unreliable, and costly (Pagourtzi et al., 2003; Abidoye, 2017; Zurada et al., 2006; Adewusi, 2022). Factors such as educational background, experience, culture, and technological exposure influence the choice of valuation methods in a given area (Abidoye, 2018).

However, advancements in technology and data processing have led to the integration of Artificial Intelligence (AI) technologies in property valuation. Machine learning algorithms like Artificial Neural Networks (ANN), Fuzzy Logic (FL), K-Nearest Neighbour (KNN), Expert Systems (ES), Support Vector Machines (SVM), Random Forest (RF), Gradient Boosting Machines (GBM), Catboost, XGBoost, and LightGBM are increasingly used to determine valuation estimates especially in the advanced countries The primary goal of incorporating these technologies is to achieve consistent, clear, easily understood, applicable, and internationally acceptable valuation figures, thus fostering a sustainable valuation practice that meets the complex requirements of the 21st century.

Despite the potential benefits, the application of AI technologies in property valuation in Nigeria remains minimal (Abidoye & Chan, 2018; Abidoye, 2019). In contrast, these technologies are widely used in advanced countries across Asia, Europe, and other regions (Guan et al., 2008; Amri, 2012; Tabales, 2013; Sarip & Hafez, 2015; Yildrim, 2019). Assessing the knowledge, awareness, adoption, and usage of AI technologies among practicing valuers is essential to enhance property valuation practices.

Against this background, this paper assesses the level of awareness of these technologies for valuation purpose among Estate Valuers with a view to enhancing their awareness for better valuation precision.

2.0 Literature Review

2.1 Artificial Intelligence Technology

AI technique is an interactive computer-based system specifically developed to assist decision makers in resolving complex and ill-defined situations. These systems utilize a blend of models, analytical approaches, and information retrieval to aid in the creation and assessment of suitable alternatives (Raul, 1999). Information technology advancements have led to the development of computerized systems aimed at enhancing decision-making efficacy (Arnott & Pervan, 2005). AI technology strive to create an interactive platform where computerized systems may automate the structured aspects of an issue, while individuals focus on the complicated and unstructured portions of the decision-making process (Silver, 1991).

A brief description of some of the technologies are discussed as follows:

2.1.1 Artificial Neural Network

Artificial neural networks are computational systems that utilize parallel connections between microprocessors, arranged in layers, to imitate the organization of neuronal networks in the brain. This is in contrast to traditional computers, which have microprocessors arranged in series. (Mora-Esperanza, 2004). The inception of artificial neural networks can be traced back to McCulloch and Pitts (1943). They were the pioneers in attempting to simulate the human brain neuron by showing that neural network can successfully perform arithmetic logical tasks. Subsequently, Hu (1964) applied this approach to weather forecasting.

There are three main components of the ANN model. They are;

The input data layer;

The hidden layer (sometimes referred to as the black box); and

The output measure layer, that includes the value predicted.

Khalafallah (2008) employed neural networks-based models to forecast the performance of the housing market throughout testing and validation procedures. The analysis identified a prediction error within the range of -2% to +2%. Abidoye and Chan (2017) conducted a study on modelling the prices of residential properties utilizing artificial neural networks in the Lagos property market. The data adopted for this study was based on sales transactions data involving property features which were retrieved from practicing estate firms in Lagos. Evidence indicates that the ANN model possesses a strong capacity for prediction, suggesting that it is appropriate and dependable for property appraisal.

Kang et al., (2020) developed a regression model artificial neural network and genetic algorithm to forecast property value in Seoul. The study employed apartment sales data between 2013 and 2017 in order to determine the sales prices as well as compare the forecasting accuracy of the models developed. Results shows that both models have a high level of precision.

2.1.2 CATBOOST (Categorical Boosting)

CatBoost, which simply means categorical boosting, is an algorithm that is based on decision trees and gradient boosting like extreme gradient boost, but with even better performance. This algorithm was developed by Dorogush, Ershov and Gulin (2018). In the study by Ibraheem et al., (2020), the performance of Catboost classifier was compared to that of another machine learning algorithm. The catboost method beats other classifier presented in the study based on feature of both data sets. The study however suggests, that Catboost be used for better forecast.

Oyedeji, Oyediran and Majekodunmi (2022) carried out a comparative assessment of the performance of Random Forest, Support vector machine, and Catboost models for forecasting property price. The three selected forecasting models demonstrated a precision and accuracy rate of over 80% in classifying residential properties.

Wang and Zhao (2022) focus their study on formulating a feasible method of predicting house prices. A datasets containing features and house prices of King County in the US were used. Eight models including Catboost, LightGBM, and XGBoost serve as candidate models. The findings from the study show that Catboost performs the best among all the models and can be used for house price prediction.

2.1.3 Extreme Gradient Boosting (XGBoost)

The Extreme Gradient Boosting algorithm was proposed by Friedman (2001). However, Chen et al., (2016) propounded XGBoost that combines the GBDT and Random Forest methods. XGBoost employs a more regularized model formalization to prevent over-fitting o data, resulting in improved performance. According to Punmoose and Ajit (2016), XGBoost is a boosted tree technique that uses the gradient boosting principle. The XGBoost algorithm is a scalable ensemble of decision trees that uses gradient boosting. Similar to gradient boosting, XGBoost optimizes a loss function to provide an incremental extension of the objective function.

The following is the procedure of XGBoost, according to Chen and Guestrin (2016); Gomez-Rios, Luengo, and Herrera (2017):

i. Feature Selection: Data cleaning, extraction of data characteristics, and data elements selection based on feature of relevance scores are the specific processes of feature selection using the XGBoost.
ii. Modelling Instruction: With default parameters, the model is trained using the selected characteristics.
    a. Parameter Optimization: The goal of parameter optimization is to reduce the differences between expected and actual values.

Jha, Babceanu, Pandey, and Jha (2021) conducted a study of house market price determination problem via the use of different algorithms in Florida. Using publicly available datasets, XGBoost, Catboost, random Forest, Lasso, and Voting regressor are being used to predict houses prices. According to the study, XGBoost algorithm surpasses other algorithms in terms of prediction model performance, coefficient determination, mean square error, mean absolute error, and computing time.

Siregar et al., (2022) examined housing value forecast employing a hybrid approach that combined genetic algorithm with extreme gradient boosting. The approach proposed underwent evaluation based on root mean square error, computing time, and the number of removed characteristics. The proposed approach was compared with XGBoosting. The results indicates that the proposed technique yields a root mean square error value of 0.129 which is smaller than the value of 0.133 obtained by using only XGBoosting.

Weng (2022) research on the house price forecast based on random forest, adaptive boosting, gradient boosting, and extreme gradient boosting. The study found that gradient boosting and extreme gradient have more accurate prediction results compared with other algorithms.

2.1.4 Random Forest Regression

The Random Forest algorithm is a widely recognized ensemble learning method based on trees and utilizes a bagging-type ensemble approach (Punmoose & Ajit, 2016). It was developed by Breiman in 2001, which was a better version of the bagging method that was developed in 1996 by Breiman. The Random Forest is a unique form of ensemble regression tree that utilizes majority voting or average of predictions from each of its constituent trees to make predictions (Antipov & Pokryshevskaya, 2012). According to Breiman and Cutler (2005), Random Forest is known for its exceptional level of accuracy, and it utilizes the decision tree algorithm to classify samples.

The accuracy of the Random Forest model in projecting residential property prices has been employed in a number of countries. According to Wang and Wu (2018), who conducted a study using housing assessment price data from Airlington County, Virginia, USA in 2015. They found that RF is more accurate than linear regression. Mohd et al., (2019) conducted an estimation of property value in Malaysia by considering set of variables such as the number of bedrooms, floor level, building age, and floor space. Their study evaluates and contrasts the results of RF, Decision tree, Ridge regression, Linear regression, and Lasso. The analysis revealed that RF is the most precise overall, as assessed by RMSE.

In South Korea, Hong, Choi, and Kim (2019) conducted a study on house price valuation using RF technique for residential property appraisal. The study evaluated the attributes of housing price predictor based on the RF approach with the hedonic price model. The study used apartment transaction data from 2006 to 2017. Findings shows that RF technique could be useful addition to hedonic models since it better portrays the complexity and nonlinearity of real-world housing markets. Adetunji et al. (2021) conducted a study on housing price prediction using the Random Forest (RF) approach. The study utilized datasets including 506 entries and 14 home attributes to assess the effectiveness of the RF model. An analysis of the forecasted and observed prices demonstrated that the RF model exhibited a satisfactory level of accuracy, with a margin of error of ±0.5 when comparing the anticipated and real values.

2.1.5 Light Gradient Boosting Model (LGBM)

Another GBDT technique that provides automatic categorical feature encoding is light gradient boosting model. According to LightGBM'S online documentation (Microsoft Corporation, 2020), the program employs a technique described by Fisher in his article "On grouping for maximal homogeneity". Prokhorenkova et al., (2017) refer to the LightGBM online documentation as well as the LightGBM source code. LGBM creates a histogram of a categorical feature's values, then sorted it using gradient statistics.

Sibindi et al., (2022) examined the utilization of a hybrid light gradient boosting machine and extreme gradient boosting for accurately predicting property values. The work aims to create a hybrid model by combining LGBM and XGBoost, with the goal of reducing overfitting by minimizing variance and enhancing accuracy. The hybrid LGBM and XGBoost model gives accurate price prediction.

2.1.6 Support Vector Machine

This is a supervised classification algorithm that seeks to find a hyperplane that can divide data into different classes. The separation is determined by maximizing the distance between the hyperplane and the boundary that determines the extent of each class. As stated by Masias et al. (2016), the Support Vector Machine algorithm utilizes a kernel function to transform the training data into a higher-dimensional space. This transformation allows for the identification of a more effective hyperplane separator. The SVM model has the capability to mitigate bias or issues in projecting property price (Sarip & Hafez, 2015).

In the study of Li et al., (2009), using SVR approach, five property value determinants in China were used to predict property prices. The study use data from 1998 to 2008. According to the traditional evaluation criteria, such as mean absolute error (MAE), mean percentage error (MAPE), and root mean square error (RMSE), the authors concluded that the SVR model is an outstanding technique for predicting property prices.

Rafieo and Adeli (2016) adopt Support Vector Machine to investigate the feasibility of a property developer proceeding with a new development or halting construction at the project's inception.

This determination is made by predicting future housing prices. The study employed data from 350 residential housing units constructed in Tehran, Iran between 1993 to 2008. The authors utilized model that was trained using 26 housing characteristics. Their results demonstrated that SVR is a suitable approach technique for predicting house prices. Winky, Ho, and Siu (2021) evaluated property values in Hong Kong by employing machine learning techniques, namely Support Vector Machine, Random Forest, and Gradient Boosting Machine. The evaluation was based on a dataset of more than 40,000 home sales over a span of 18 years. Three performance metrics, namely Mean Square Error (MSE), Root Mean Squared Error (RMSE), and Mean Absolute Percentage Error (MAPE), were used to compare the predictive power of the algorithms. The study determined that Support Vector Machine (SVM) is a valuable method for data fitting due to its ability to generate predictions that are reasonably accurate in a relatively short amount of time.

2.1.7 Fuzzy Logic

Fuzzy logic utilizes a mathematical framework to represent and analyze human reasoning (Ozcelik, 2023). With this technology, computers analyze data that are not defined numerically in a prudent way and makes them computable. As stated by Del Giudice, De Paola and Cantisani (2017), fuzzy logic was introduced by Zadeh in the 1960s as a method capable of modelling the improbability between normal spoken and written language. As a result of this, fuzzy logic includes certain qualitative linguistic functions.

The important attributes that differentiate fuzzy logic compared to logical is its inability to arrive at a definite judgement. According to Pagourtzi et al., (2003), classical logic evaluates a phenomenon as either present or absent whereas fuzzy logic evaluates it as being either some or not. Modeling using Fuzzy logic involves of three primary steps:

i. Fuzzification, which convert numerical inputs into verbal expressions and assigns membership degrees in fuzzy sets;
ii. Fuzzy inference and rule extraction, which utilizes “if-then” statements of membership functions to establish connections between fuzzy sets; and
iii. Defuzzification, which is to converts the fuzzy output value to precise numerical values (Ulvi, 2018; Ozcelik, 2023).

Hui, Lau, and Lo (2009) investigated the process of making real estate investment decisions in Hong Kong using Fuzzy logic approach. An alternative approach to risk management for investors is examined, using a portfolio of indicators instead of relying on a single specific indicator or index commonly used by professionals. The result obtained from this framework is subsequently compared to the property price index. Results of this study demonstrate that the composition of the housing-indicator portfolio provides a value that accurately reflects the intricacies of both the real estate market and investors’ expectations.

Del Giudice, Paola, and Cantisani (2017) examined the assessment of real estate investments utilizing Fuzzy Logic. A fuzzy logic system was employed to assess the conditions of real estate market using imprecise and ambiguous data. The results demonstrated that by effectively implementing fuzzy logic, operators and investors can enhance the quality of their investment decisions and mitigate risks associated with unclear inputs. To establish transparency in the real estate transactions in Pune city, in India, Kamire, Chaphalkar, and Sandbhor (2021) conducted research on Real property value prediction capability using Fuzzy Logic and ANFIS. The authors attempt to compare their output by evaluating the accuracy of the developed model and analyzing the properties of the available datasets.

2.1.8 K-Nearest Neighbor

The K-nearest neighbor (K-NN) algorithm is normally used for classification tasks. The K-nearest neighbor technique utilizes variable similarity to predict the value of any new data points. There are three primary phases in the stage of K-NN. These are: determining the distance between the point to be estimated and each training point; selection of the K closet points based on distance; and estimating the new point by averaging the selected data points. Hassan (2019) was of the opinion that K-NN regression method cannot be defined as a traditional model because of dependency individual samples in dataset. This method involves predicting a new sample by taking the average of the values of the K-nearest neighbors (Kuhn & Johnson, 2013).

Hassan (2019) examined property value by employing artificial neural networks, hedonic regression, and closest neighbors’ regression techniques. Both K-NN (nearest neighbor regression) and artificial neural networks offer flexible and nonlinear fitting capabilities. The classical hedonic approach and its nonlinear variants were used to analyze a mixed dataset and contrasted using several performance indicators. According to the results, The K-NN regression approach yields a satisfactory outcome. Singh et al. (2017) created KNN, random forest, and Naive Bayes classifiers for the purpose of distinguishing between text, numeric, and alphanumeric data types. The model was developed and trained to store known classes of training datasets in order to acquire patterns for making predictions. The varying number of features in the dataset did not have a substantial impact on the performance of the Random Forest (RF) algorithm, however the success metrics of the K-Nearest Neighbors (KNN) and Naive Bayes (NB) algorithms fluctuated, leading to a low accuracy rate. Stanley et al., (2021) constructed K-NN and decision tree models to forecast the numerical values of the quantity of a one-pound table at a specific interest rate and duration in years. The model is beneficial for investors, accountants, data experts, surveyors, and valuers who have an interest in financial analysis and its practical uses. A cross-validation test was conducted using the predicted R-squared test to identify overfitting and assess the performance of the model on the testing dataset. The K-nearest neighbors (K-NN) and decision tree approaches were trained and tested, with accuracy rates of 96.76% and 99.86% respectively.

2.2 Level of Awareness of Artificial Intelligence Technology

AI technology has been employed in a variety of study areas during the last decade, and its versatility has attracted the usage of the algorithms for variety of application in which property valuation is included. Researchers have begun to utilize AI technology in property valuation in recent years, but its application in the real estate practice is yet to be explored especially in the developed countries.

Mugunthany and Muhammad (2008) examined the existence of automated valuation model among valuation firms in Malaysia. The study focused on 246 valuation firms in Malaysia with the aim of evaluating their awareness on the existence of automated valuation models. The data were analyzed based on 85 questions in a questionnaire survey. In order to consolidate the research findings, interview sessions were held with 11 valuation firms' head office and the analysis was conducted using Nvivo 2.0 software. The study demonstrated that there are significant differences between valuation firms which are aware of the existence of automated valuation models compared with those which were not. Further findings revealed that, valuation firms' head offices are more informed about the existence of automated valuation model compared to their branches or smaller valuation firm. The paper concluded that there is low level of awareness of automated valuation models among valuation firms.

A study conducted by Azmi, Nanawi, Latif, and Ling (2013) examined the level of knowledge that property valuers have regarding computer aided valuation systems. Data collection mostly involved the distribution of questionnaires to practicing Estate Surveyors and Valuers. The findings revealed that the participants had a moderate level of awareness regarding the system and were not prepared for its deployment. Valuers are still adhering to the conventional method of property appraisal.

Abidoye (2017) discovered that there is a low level of understanding and utilization of sophisticated property valuation techniques, as seen by the cases in Hong Kong and Nigeria. This suggests that practical Valuers have not yet adopted the advanced property valuation methodologies, despite the fact that these approaches are widely used. The need of using modern property valuation methods to achieve a globally sustainable property valuation practice cannot be emphasized enough.

3.0 Methodology

3.1 Data Collection and Source

To achieve the study's goal, the data for the study were collected using the combination of both structured and online questionnaires survey from One Hundred and Ninety-four (194) practicing estate surveying and valuation firms operating in Lagos, Nigeria. This was done to assess the level of awareness of AI technology available for residential property appraisal in the study area. Valuers were requested to specify their level of awareness with the AI technology.

The data acquired for the study was analyzed using descriptive statistics, specifically weighted mean scores. The weighted mean score is determined using equation:
WMS = $(5n_5+4n_4+3n_3+2n_2+1n_1)/ N$………………EQ(i)

The Chi-Square test was conducted to establish the statistical association and independence among the research variable (Verma, 2013). The study employed the chi-square test to ascertain the statistical association between valuers and their educational background, as well as their years of professional experience.

This was computed using the equation: $x^2 = \sum_{i=0}^{n} \frac{(fo-fe)}{fe}$

3.2 Results and Discussion

The research examined the opinion of the practicing Estate Surveying and Valuation firms in Lagos on the level of awareness of AI technologies used in the determination of residential property values in Lagos, Nigeria. The various opinions of the participants have been presented, examined, and ranked to enhance the significance and understanding of the research.

Table 1 summarizes the awareness levels of various AI technologies and training algorithms among respondents. Each AI technology and algorithm is evaluated on a scale ranging from "Very High" to "Very Low" awareness. The mean awareness score, standard deviation, and ranking are provided to facilitate a comprehensive understanding. As shown in Table 1, Artificial Neural Networks (ANN) are the most recognized among the AI technologies options. However, a significant portion of respondents (67.5%) still exhibit low to very low awareness. The high standard deviation indicates considerable variability in the awareness levels. Similarly, Case-Based Reasoning (CBR) and Expert Systems are equally recognized, with a majority of respondents demonstrating low to very low awareness (76.2% for CBR and 73.5% for Expert Systems). The standard deviations suggest a moderate spread in the awareness levels. Both K-Nearest Neighbors (KNN) and Memory-Based Reasoning show similar awareness levels, predominantly low to very low (74.9%). The standard deviation indicates moderate variability in awareness. Support Vector Machines (SVM) are among the least known DSS options, with a substantial majority (82.5%) reporting low to very low awareness. The relatively low standard deviation suggests less variability in responses. The Adaptive Neuro-Fuzzy Inference System (ANFIS) is also among the least known methods, with 82.2% of respondents indicating low to very low awareness. The variability in awareness is slightly higher than for SVM. Radial Basis Function (RBF) Networks and M5P Trees exhibit very low awareness, with most respondents (84.8% for both RBF and M5P Trees) indicating low to very low awareness. The standard deviations suggest a relatively consistent lack of awareness. These algorithms are tied for the second rank.

Awareness levels are predominantly low to very low for the majority of respondents (76.2% for Back Propagation, 81.5% for Fast Parallel Conjugate Gradient (FPCG), and 81.5% for Levenberg-Marquardt), with moderate variability. One Step Secant is less known, with a majority of respondents (83.4%) indicating low to very low awareness. The standard deviation shows a relatively consistent lack of awareness. This analysis indicates that while certain AI technologies and training algorithms like ANN are more recognized, there remains a significant portion of the respondent population with low awareness levels across various systems. The variability in awareness levels highlights the need for increased educational efforts to enhance familiarity with these important tools.

Judging by the analysis and corrugating with the mathematical expression of weighted mean score as stated by Akere and Gidado (2003) and Kadir (2005), the practicing Estate Surveying and Valuation firms in Lagos are rarely aware of the AI technology applicable in the determination of residential property value in Lagos.

Table 1: Level of Awareness of AI technology used in the determination of residential property values in Lagos, Nigeria

	Very High	High	Somewhat Aware	Low	Very Low	Mean	Std. Dev	Rank	Av. Mean/ Decision
Decision Support Systems									
Artificial Neural Network	4(2.6%)	7(4.6%)	38(25.2%)	47(31.1%)	55(36.4%)	2.0596	1.02132	1st	
Case-Based Reasoning (CBR)	-	10(6.6%)	26(17.2%)	59(39.1%)	56(37.1%)	1.9338	.89940	2nd	
Expert Systems	-	9(6.0%)	31(20.5%)	52(34.4%)	59(39.1%)	1.9338	.91410	2nd	
K Nearest Neighbour	-	10(6.6%)	28(18.5%)	51(33.8%)	62(41.1%)	1.9073	.92629	4th	
Memory Based Reasoning	-	10(6.6%)	28(18.5%)	51(33.8%)	62(41.1%)	1.9073	.92629	4th	
LGBM	-	-	30(19.9%)	72(47.7%)	49(32.5%)	1.8742	.71465	6th	1.7289
Fuzzy Logic	2(1.3%)	2(1.3%)	32(21.1%)	52(34.4%)	63(41.7%)	1.8609	.88724	7th	arely Aware
Genetic Algorithm	-	3(2.0%)	32(21.2%)	56(37.1%)	60(39.7%)	1.8543	.81973	8th	of the
Random Forest	-	-	21(13.9%)	84(55.6%)	46(30.5%)	1.8344	.64736	9th	Methods
CatBoost	-	-	12(7.9%)	101(66.9%)	38(25.2%)	1.8278	.55090	10th	
Multi-Layer Perceptrons (MLP)	-	10(6.6%)	26(17.2%	59(39.1%)	56(37.1%)	1.8079	.86960	11th	
XGBoost	-	-	8(5.3%)	101(66.9%)	42(27.8%)	1.7748	.53132	12th	
Adaptive Neuro-fuzzy interference system	-	2(1.3%)	25(16.6%)	54(35.8%)	70(46.4%)	1.7285	.78259	13th	
Support Vector Machine (SVM)	-	3(2.0%)	23(15.2%)	53(35.1%)	72(47.4%)	1.7152	.79480	14th	
Radial Basis Function	-	-	23(15.2%)	59(39.1%)	69(45.7%)	1.6954	.72105	15th	
M5p Trees	-	-	23(15.2%)	56(37.1%)	72(47.7%)	1.6755	.72617	16th	
Training Algorithm									
Gradient Descent	-	3(2.0%)	28(18.5%)	60(39.7%)	60(39.7%)	1.8278	.79801	1st	
Fletcher-Powell Conjugate Gradient (FPCG)	-	3(2.0%)	25(16.6%)	56(37.1%)	67(44.4%)	1.7616	.79757	2nd	
Back Propagation	-	3(2.0%)	30(19.9%)	46(30.5%)	72(47.7%)	1.7616	.83832	2nd	
Levenberg Marquardt Back Propagation	-	-	28(18.5%)	59(39.1%)	64(42.4%)	1.7616	.74573	2nd	1.7528 Rarely Aware the Methods
One Step Secant	-	-	25(16.6%)	58(38.4%)	68(45.0%)	1.7152	.73373	5th	
Conjugate Gradient	-	-	25(16.6%)	54(35.8%)	72(47.7%)	1.6887	.74104	6th	

Furthermore, Using Chi-Square analytical method, the cross tabulation of the educational and the level of awareness of AI technology are further analyzed in tables 2;

Table 2 presents the awareness levels of various AI technology including CatBoost, XGBoost, LGBM, and Random Forest, categorized by educational qualifications (H.N.D., B.Sc, M.Sc, PhD, and Others). Chi-square tests were conducted to examine the relationship between educational qualifications and awareness levels. The table includes observed counts, expected counts, and the results of the chi-square tests. For CatBoost, the Pearson Chi-Square value of 8.478 with a p-value > 0.05 indicates no significant relationship between educational qualification and awareness. Symmetric measures (Phi and Cramer's V) suggest a weak association. Similarly, XGBoost shows a Pearson Chi-Square value of 7.684 with a p-value > 0.05, indicating no significant relationship, with weak association measures.

Regarding LGBM, the Pearson Chi-Square value is 2.634 with a p-value > 0.05, suggesting no significant relationship and a very weak association according to symmetric measures. For Random Forest, the Pearson Chi-Square value is 12.134 with a p-value > 0.05, indicating no significant relationship. Symmetric measures suggest a weak to moderate association in this case. In summary, the analysis of CatBoost, XGBoost, LGBM, and Random Forest indicates no significant relationship between educational qualification and awareness levels for these DSS methods based on the Pearson Chi-Square tests (all p-values > 0.05).

Table 2: Cross Tab between Educational Qualification and Awareness of Artificial Intelligence Technology

			CatBoost			**Total**
			Very Low	**Low**	**Undecided**	
Educational Qualification	H.N.D.	Count	6	22	1	29
		Expected Count	7.3	19.4	2.3	29.0
	B. Sc	Count	15	20	3	38
		Expected Count	9.6	25.4	3.0	38.0
	M.Sc	Count	15	47	7	69
		Expected Count	17.4	46.2	5.5	69.0
	PhD	Count	1	10	1	12
		Expected Count	3.0	8.0	1.0	12.0
	Others	Count	1	2	0	3
		Expected Count	.8	2.0	.2	3.0
Total		Count	38	101	12	151
		Expected Count	38.0	101.0	12.0	151.0

Chi-Square Tests

	Value	**df**	**Asymp. Sig.(2-sided)**
Pearson Chi-Square	8.478[a]	8	.388
Likelihood Ratio	8.966	8	.345
Linear-by-Linear Association	1.033	1	.310
N of Valid Cases	151		

Symmetric Measures

Nominal by Nominal	Phi	.237	.388
	Cramer's V	.168	.388
N of Valid Cases	151		

Educational Qualification * XGBoost

Educational Qualification	H.N.D.	Count	7	18	4	29
		Expected Count	8.1	19.4	1.5	29.0
	B. Sc	Count	12	25	1	38
		Expected Count	10.6	25.4	2.0	38.0
	M.Sc	Count	19	48	2	69
		Expected Count	19.2	46.2	3.7	69.0
	PhD	Count	4	7	1	12
		Expected Count	3.3	8.0	.6	12.0
	Others	Count	0	3	0	3
		Expected Count	.8	2.0	.2	3.0

Pearson Chi-Square	7.684[a]	8	.465
Likelihood Ratio	7.622	8	.471
Linear-by-Linear Association	.327	1	.568
N of Valid Cases	151		

Nominal by Nominal	Phi	.226	.465
	Cramer's V	.160	.465

	Total	Count	42	101	8	151				
		Expected Count	42.0	101.0	8.0	151.0				
Educational Qualification * LGBM										
Educational Qualification	H.N.D	Count	9	15	5	29				
		Expected Count	9.4	13.8	5.8	29.0	Pearson Chi-Square	2.634[a]	8	.955
	B. Sc	Count	13	18	7	38	Likelihood Ratio	2.995	8	.935
		Expected Count	12.3	18.1	7.5	38.0	Linear-by-Linear Association	.000	1	.982
	M.Sc	Count	22	32	15	69	N of Valid Cases	151		
		Expected Count	22.4	32.9	13.7	69.0				
	PhD	Count	3	6	3	12				
		Expected Count	3.9	5.7	2.4	12.0				
	Others	Count	2	1	0	3	Nominal by Nominal	Phi	.132	.955
		Expected Count	1.0	1.4	.6	3.0		Cramer's V	.093	.955
	Total	Count	49	72	30	151	N of Valid Cases	151		
		Expected Count	49.0	72.0	30.0	151.0				
Educational Qualification * Random Forest										
Educational Qualification	H.N.D	Count	10	17	2	29				
		Expected Count	8.8	16.1	4.0	29.0				
	B. Sc	Count	10	25	3	38	Pearson Chi-Square	12.134[a]	8	.145
		Expected Count	11.6	21.1	5.3	38.0	Likelihood Ratio	10.776	8	.215
	M.Sc	Count	23	35	11	69	Linear-by-Linear Association	2.130	1	.144
		Expected Count	21.0	38.4	9.6	69.0	N of Valid Cases	151		
	PhD	Count	2	5	5	12				
		Expected Count	3.7	6.7	1.7	12.0				
	Others	Count	1	2	0	3				
		Expected Count	.9	1.7	.4	3.0				
	Total	Count	46	84	21	151				
		Expected Count	46.0	84.0	21.0	151.0				

Table 3 displays the awareness levels of various AI technologies such as CatBoost, XGBoost, LGBM, and Random Forest, categorized by the number of years of professional qualification. Chi-square tests were employed to assess the relationship between years of professional qualification and awareness levels. The table includes observed counts, expected counts, and the results of the chi-square tests. For CatBoost, the Pearson Chi-Square value is 10.742 with a p-value > 0.05, indicating no significant relationship, although the p-value approaches the 0.05 threshold, suggesting a potential weak association. Symmetric measures (Phi and Cramer's V) indicate a weak to moderate association. Similarly, for XGBoost, the Pearson Chi-Square value is 4.595 with a p-value > 0.05, indicating no significant relationship and a weak association according to symmetric measures. For Random Forest, the Pearson Chi-Square value is 1.787 with a p-value > 0.05, suggesting no significant relationship and a very weak association based on symmetric measures. Regarding LGBM, the Pearson Chi-Square value is 4.116 with a p-value > 0.05, indicating no significant relationship and a weak association according to symmetric measures.

In a specific term, for CatBoost, XGBoost, LGBM, and Random Forest, there is no significant relationship between years of professional qualification and awareness levels based on the Pearson Chi-Square tests (all p-values > 0.05). Symmetric measures (Phi and Cramer's V) generally indicate weak to very weak associations between years of professional qualification and awareness levels for all AI technologies analyzed.

Table 3: Cross Tab of Number of Years of Professional Qualification and Awareness of Artificial Intelligence Technology

			CatBoost			Total	Chi-Square Tests			
			Very Low	Low	Undecided			Value	df	Asymp. Sig. (2-sided)
Number of Years of Professional Qualification	1-10 Years	Count	21	41	3	65	Pearson Chi-Square	10.742[a]	6	.097
		Expected Count	16.4	43.5	5.2	65.0	Likelihood Ratio	11.136	6	.084
	1-20 Years	Count	4	27	5	36	Linear-by-Linear Association	.154	1	.694
		Expected Count	9.1	24.1	2.9	36.0	N of Valid Cases	151		
	21 - 30 Years	Count	5	22	3	30				
		Expected Count	7.5	20.1	2.4	30.0	**Symmetric Measures**			
	31 Years & Above	Count	8	11	1	20		Value		Approx. Sig
		Expected Count	5.0	13.4	1.6	20.0	Phil	2.67		.097
Total		Count	38	101	12	151	Cramer's V	.189		.097
		Expected Count	38.0	101.0	12.0	151.0	No of Valid Cases	151		
Number of Years of Professional Experience * XGBoost										
Number of Years of Professional Qualification	1-10 Years	Count	18	44	3	65				
		Expected Count	18.1	43.5	3.4	65.0				
	1-20 Years	Count	11	21	4	36	Pearson Chi-Square	4.595[a]	6	.597
		Expected Count	10.0	24.1	1.9	36.0	Likelihood Ratio	5.155	6	.524
	21 - 30 Years	Count	7	22	1	30	Linear-by-Linear Association	.071	1	.790
		Expected Count	8.3	20.1	1.6	30.0	N of Valid Cases	151		
	31 Years & Above	Count	6	14	0	20	Nominal by Nominal Phi	.174		.597
		Expected Count	5.6	13.4	1.1	20.0	Cramer's V	.123		.597
Total	Count		42	101	8	151	N of Valid Cases	151		

	Expected Count		42.0	101.0	8.0	151.0				
Number of Years of Professional Experience * Random Forest										
	1-10 Years	Count	18	39	8	65				
		Expected Count	19.8	36.2	9.0	65.0				
	1-20 Years	Count	13	18	5	36	Pearson Chi-Square	1.787[a]	6	.938
Number of Years of Professional Qualification		Expected Count	11.0	20.0	5.0	36.0	Likelihood Ratio	1.782	6	.939
	21 - 30 Years	Count	10	15	5	30	Linear-by-Linear Association	.046	1	.831
		Expected Count	9.1	16.7	4.2	30.0	N of Valid Cases	151		
	31 Years and aAbove	Count	5	12	3	20	Nominal by Nominal	Phi	.109	.938
		Expected Count	6.1	11.1	2.8	20.0		Cramer's V	.077	.938
Total	Count		46	84	21	151	N of Valid Cases	151		
	Expected Count		46.0	84.0	21.0	151.0				
Number of Years of Professional Experience * LGBM										
	1-10 Years	Count	22	28	15	65				
		Expected Count	21.1	31.0	12.9	65.0	Pearson Chi-Square	4.116[a]	6	.661
	1-20 Years	Count	13	18	5	36	Likelihood Ratio	4.126	6	.660
		Expected Count	11.7	17.2	7.2	36.0	Linear-by-Linear Association	.240	1	.624
Number of Years of Professional Qualification Total	21 - 30 Years	Count	9	17	4	30	N of Valid Cases	151		
		Expected Count	9.7	14.3	6.0	30.0				
	31 Years and aAbove	Count	5	9	6	20	Nominal by Nominal	Phi	.165	.661
		Expected Count	6.5	9.5	4.0	20.0		Cramer's V	.117	.661
	Count		49	72	30	151	N of Valid Cases	151		
	Expected Count		49.0	72.0	30.0	151.0				

5.0. Conclusion

This study aimed to assess the awareness of AI technologies among professionals involved in property valuation in Lagos, Nigeria. The objective was to suggest alternative methods for achieving more precise valuation. Data were collected from estate firms across Lagos metropolis and analyzed using descriptive statistics and Chi-square tests. The study identified sixteen types of AI technologies and evaluated six training algorithms based on respondents' understanding. Artificial Neural Network emerged as the most recognized, followed by Cased-Based Reasoning and Expert Systems. Support Vector Machine, Radial Basis Function, and M5p Trees were the least recognized. The findings highlighted a general lack of awareness among estate surveying and valuation firms in Lagos regarding AI technology's applicability in residential property valuation. Cross-tabulation of respondents' educational and professional qualifications with their awareness of AI technology indicated no significant association, with p-values ≥ 0.05. In conclusion, the study reveals a low level of awareness among Nigerian valuers concerning AI technologies, aligning with previous research by Abidoye and Chan (2018) indicating a gap in knowledge regarding advanced AI technologies widely used in real estate research. Future research should explore barriers to adopting AI technology in property valuation among valuers. These findings offer valuable insights for property professionals, real estate investors, and policymakers. A broader study covering various zones within the Nigerian property market could provide more conclusive results.

References

Abidoye, R. B. (2017). Towards property valuation accuracy: a comparison of hedonic pricing model and artificial neural network. Ph.D thesis submitted to the department of building and real estate, the Hong Kong Polytechnic University

Abidoye, R. B., & Chan, A. P. (2017). Valuers' receptiveness to the application of artificial intelligence in property valuation. *Pacific Rim Property Research Journal*, *23*(2), 175 193.

Abidoye, R. B., & Chan, A. P. (2018). Improving property valuation accuracy: A comparison of hedonic pricing model and artificial neural network. *Pacific Rim Property Research Journal*, *24*(1), 71-83.

Adetunji, A. B., Akande, O. N., Ajala, F. A., Oyewo, O., Akande, Y. F., & Oluwadara, G. (2022). House price prediction using random forest machine learning technique. *Procedia Computer Science*, *199*, 806-813.

Adewusi, A.O. (2022), Modelling Residential Property Prices Using Artificial Intelligence Technologies in Lagos Metropolis, Nigeria. World Journal of Innovation and Mogern Technology, 5(1), 20- 43.

Amri, S., & Tularam, G. A. (2012). Performance of multiple linear regression and nonlinear neural networks and fuzzy logic techniques in modeling house prices. *Journal of Mathematics and Statistics*, *8*(4), 419-434.

Antipov, E. A., & Pokryshevskaya, E. B. (2012). Mass appraisal of residential apartments: An application of Random Forest for valuation and a CART-based approach for model diagnostics. *Expert systems with applications*, *39*(2), 1772-1778.

Armugam, M., Najib, R. M., & David, M. J. (2007). The existence of automated valuation models (AVMS) among valuation firms in Malaysia. *The Surveyors*, *1*, 1-14.

Arnott, D., Pervan, G., & Dodson, G. (2005). Who pays for decision support systems research? Review, directions, and issues. *Communications of the Association for Information Systems*, *16*(1), 17.

Azmi, A. S. M., Nawawi, A. H., Latif, S. N. F. A., & Ling, N. L. F. J. (2013). Investigating the Learning Process amongst Young Practicing Property Valuers. In *International Real Estate Research Symposium (IRERS2014), INSPEN, Kuala Lumpur:(yet to publish)*.

Chen, T., & Guestrin, C. (2016, August). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining* (pp. 785-794).

Chi Man Hui, E., Muk Fai Lau, O., & Kak Keung Lo, T. (2009). Deciphering real estate investment decisions through fuzzy logic systems. *Property Management*, *27*(3), 163-177.

Choi, S. W., Ko, T., Hong, K. J., & Kim, K. H. (2019). Machine learning-based prediction of Korean triage and acuity scale level in emergency department patients. *Healthcare informatics research*, *25*(4), 305.

Del Giudice, V., De Paola, P., & Cantisani, G. B. (2017). Valuation of real estate investments through Fuzzy Logic. *Buildings*, *7*(1), 26.

Dorogush, A. V., Ershov, V., & Gulin, A. (2018). CatBoost: gradient boosting with categorical features support. *arXiv preprint arXiv:1810.11363*.

French, N. (2004). The valuation of specialized property: A review of valuation methods. *Journal of Property Investment & Finance*, *22*(6), 533-541.

Gómez-Ríos, A., Luengo, J., & Herrera, F. (2017). A study on the noise label influence in boosting algorithms: AdaBoost, GBM and XGBoost. In *Hybrid Artificial Intelligent Systems: 12th International Conference, HAIS 2017, La Rioja, Spain, June 21-23, 2017, Proceedings 12* (pp. 268-280). Springer International Publishing.

Guan, J., Zurada, J., & Levitan, A. (2008). An adaptive neuro-fuzzy inference system-based approach to real estate property assessment. *Journal of Real Estate Research*, *30*(4), 395 422.

Hassan, Y. (2019). Property value assessment using artificial neural networks, hedonic regression and nearest neighbors regression methods. *Selçuk Üniversitesi Mühendislik, Bilim ve Teknoloji Dergisi*, *7*(2), 387-404.

Hu, G. C. (1964). Variable-factors theory of bearing capacity. *Journal of the Soil Mechanics and Foundations Division*, *90*(4), 85-95.

Ho, W. K., Tang, B. S., & Wong, S. W. (2021). Predicting property prices with machine learning algorithms. *Journal of Property Research*, *38*(1), 48-70.

Ibrahim, A., Muhammed, M. M., Sowole, S. O., Raheem, R., & Abdulaziz, R. O. (2020). Performance of CatBoost classifier and other machine learning methods.

Kamire, A., Chaphalkar, N., & Sandbhor, S. (2021). Real Property Value Prediction Capability Using Fuzzy Logic and ANFIS. In *Proceedings of the International Conference on Smart Data Intelligence (ICSMDI 2021)*.

Kang, J. K., Cho, K. H., Kim, S. B., & Choi, N. C. (2020). Artificial neural network modeling for prediction of dynamic changes in solution from bioleaching by indigenous acidophilic bacteria. *Applied Sciences*, *10*(21), 7569.

Khalafallah, A. (2008). Neural network based model for predicting housing market performance. *Tsinghua Science and Technology*, *13*(S1), 325-328.

Kuhn, M., & Johnson, K. (2013). *Applied predictive modeling* (Vol. 26, p. 13). New York: Springer.

Kumar, S., Sharma, A., Reddy, B. K., Sachan, S., Jain, V., & Singh, J. (2022). An intelligent model based on integrated inverse document frequency and multinomial Naive Bayes for current affairs news categorization. *International journal of system assurance engineering and management*, *13*(3), 1341-1355.

Masías, V. H., Valle, M. A., Crespo, F., Crespo, R., Vargas, A., & Laengle, S. (2016). Property valuation using machine learning algorithms: A study in a Metropolitan-Area of Chile. In *Selection at the AMSE Conferences* (p. 97).

Mohd, T., Jamil, N. S., Johari, N., Abdullah, L., & Masrom, S. (2019). An overview of real estate modelling techniques for house price prediction. In *Charting a Sustainable Future of ASEAN in Business and Social Sciences: Proceedings of the 3rd International Conference on the Future of ASEAN (ICoFA) 2019—Volume 1* (pp. 321-338). Springer Singapore.

Mora-Esperanza, J. G. (2004). Artificial intelligence applied to real estate valuation: An example for the appraisal of Madrid. *Catastro*, *1*, 255-265.

Odunfa, V., Fateye, T., & Adewusi, A. (2021). Application of artificial intelligence (ai) approach to African real estate market analysis opportunities and challenges. Advances in Multidisciplinary & Scientific Research Journal Publication, 29, 121-132.

Oyedeji, O. J., Oyeniran, O. M., & Majekodunmi, S. O. (2022). Comparative Assessment of Radom Forest, SVC and Cat Boost Performances as Property Price Forecasting Models. *Mathematical Statistician and Engineering Applications*, *71*(4), 1283-1289.

Ozcelik (2023) land valuation with fuzzy logic in artificial intelligence application.

Pagourtzi, E., Assimakopoulos, V., Hatzichristos, T., & French, N. (2003). Real estate appraisal: a review of valuation methods. *Journal of property investment & finance*, *21*(4), 383-401.

Pitts, W. (1943). The linear theory of neuron networks: The dynamic problem. *The bulletin of mathematical biophysics*, *5*, 23-31.

Prokhorenkova, L., Gusev, G., Vorobev, A., Dorogush, A. V., & Gulin, A. (2018). CatBoost: unbiased boosting with categorical features. *Advances in neural information processing systems*, *31*.

Sarip, A. G., & Hafez, M. B. (2015). Fuzzy logic application for house price prediction. *International Journal of Property Sciences (E-ISSN: 2229-8568)*, *5*(1).

Sibindi, R., Mwangi, R. W., & Waititu, A. G. (2023). A boosting ensemble learning based hybrid light gradient boosting machine and extreme gradient boosting model for predicting house prices. *Engineering Reports*, *5*(4), e12599.

Silver, M. S. (1991). *Systems that support decision makers: description and analysis*. John Wiley & Sons, Inc.

Siregar, M. U., Hardjita, P. W., Asdin, F. A., Wardani, D., Wijayanto, A., Yunitasari, Y., & Anshari, M. (2022, November). Housing Price Prediction Using a Hybrid Genetic Algorithm with Extreme Gradient Boosting. In *Proceedings of the 2022 International Conference on Computer, Control, Informatics and Its Applications* (pp. 296-300).

Tabale, J. M. N., Carmona, F. J. R., & y Ocerin, J. M. C. (2016). Artificial Intelligence (AI) techniques to analyze the determinants attributes in housing prices. *Inteligencia Artificial*, *19*(58), 23-38.

Verma, S. (2016). Deciding admission criteria for master of computer applications program in India using Chi-Square test. In *Proceedings of the Second international conference on Information and Communication Technology for Competitive Strategies* (pp. 1-4).

Wang, C., & Wu, H. (2018). A new machine learning approach to house price estimation. *New Trends in Mathematical Sciences*, *6*(4).

Wang, Z., Zhao, Y., & Zhang, F. (2022). Simulating the spatial heterogeneity of housing prices in Wuhan, China, by regionally geographically weighted regression. *ISPRS International Journal of Geo-Information*, *11*(2), 129.Friedman (2001).

Weng, W. (2022). Research on the House Price Forecast Based on machine learning algorithm. *BCP Business & Management*, *32*, 134-147.

Yıldırım, H. (2019). Property value assessment using artificial neural networks, hedonic regression and nearest neighbors' regression methods. *Selçuk Üniversitesi Mühendislik, Bilim ve Teknoloji Dergisi*, *7*(2), 387-404.

Yomralioglu, T., & Nisanci, R. (2004). Nominal asset land valuation technique by GIS. In *Proceedings of the FIG Working Week* (pp. 1-9).

Ziweritin, S., Ibiam, I. A., Oyeniran, T. A., & Oko, G. E. (2021). KNN and Decision Tree Model to Predict Values in Amount of One Pound Table. *International Journal on Computer Science and Engineering*, *9*(7), 17-21.

Zurada, J. M., Levitan, A. S., & Guan, J. (2006). Non-conventional approaches to property value assessment. *Journal of Applied Business Research (JABR)*, *22*(3).