Residential Tenants Classification:A Test of Performance of Five Selected Artificial Neural Networks training Algorithms

A.O Adewusi*(aoadewusi@futa.edu.ng) and J. A Oguntokun**

[1,2]Department of Estate Management, Federal University of Technology Akure, Nigeria

Abstract

Personal judgment of the property manager and traditional statistical methods are often used in reaching decisions on tenant selection, however, these methods have continued to be unreliable, inconsistent, and time-consuming. Artificial neural networks (ANNs) are known as powerful support tools in modeling unknown data relationships in decision making due to their abilities in pattern recognition of complex relationships, classification, prediction, forecasting, etc. While ANN has enjoyed continued application in many fields of endeavor, its application has been very limited in the fields of property management and investment.The current paper assesses the performance of ANN training algorithms in residential tenant classification with a view to choosing the best ANN training algorithm suitable for the classification of residential rental applicants in the Lagos Metropolis property market. Five ANN training algorithms are selected for analysis, namely, Levenberg-Marquardt (LM), Gradient Descent Backpropagation (GD), Resilient Back Propagation (RP), One Step Secant Back Propagation (OSS), and Gradients Descent with Momentum and Adaptive Rate Backpropagation (GDX). A total of 724 data samples of rental applications were obtained from the databases of the practicingproperty managers in the Nigerian property markets, the total samples were subdivided into 70%, 15%, and 15% for training, validation, and testing respectively. Test datasets (representing 15% of the total datasets) were used in evaluating the classification performance of the modeled (ANN training algorithms) The paper concludes that all the selected ANN training algorithms except GD produced good and efficient resultsin tenant classifications, however, GDX and OSS appear to be most suited for residential tenant classification in the Nigerian property market as they outperformed the other ANN training algorithms.The results provide decision inputs for professional real estate managers and cost–time saving frameworks for tenants selection.

Keywords: Artificial Neural Network, Training Algorithms, property investment

1. Introduction

An increase in the population of most cities in the world has attracted investment in the rental property markets (Olatoye,2005 and Viejo et al, 2019). One of the key issues in realizing the goal of investing in the rental market in choosing the right tenants with the ability to pay the initial and subsequent rents, able to care for the property, and maintain good relationships with others in the property. Property managers have been more meticulous in selecting the right tenants, however, with subjective approaches of using personal judgment and experiences, these approaches more often than not are unable to cope with the uncertainties associated with choosing the right tenants (Dabara et al, 2017).

Decision making using personal judgment is fraught with errors, inconsistency, time-consuming, and exposure to higher risks. The use of traditional methods of data analysis as decision support tools such as Discriminant Analysis (DA), Multiple Regression Analysis (MRA), Logistic Regression Analysis (LRA), etc have also been used in aiding classification decisions but have often proven inefficient because of the linearity assumptions underlying their proper use.

Adoption of tenant selection approaches with less reliance on human experience and personal judgment may undoubtedly produce good insight into choosing the right tenants(Gbadegesin and Ojo, 2013). Artificial Neural Networks (ANNs) have proven very useful in providing better performance than traditional statistical methods and subjective methods of screening rental applicants. ANNs have continued to enjoy wide applications in many fields such as medicine, engineering, finance, science, etc for classification, prediction, pattern recognition, forecasting, and so on. ANNs are computer-based systems with the ability to learn and detect patterns for predicting outputs (Lin et al, 2012 and Viejo et al, 2019).

Unlike traditional statistical methods of data analysis, ANNs can implicitly detect a non-linear relationship between the dependent and independent variables (Lau et al, 2019). ANN is a mathematical model that emulates the activity of the biological neural network in the human brain (Sharif et al, 2018 and Fonseka et al, 2018).

ANNs are trained, tested, and validated using different algorithms (Alizadeh et al, 2018) while ANNs are known to produce better performance in pattern classification, however, different ANNs training algorithms and architectures affect their classification performance (Lau et al, 2019). ANNs have different architectures such as feedforward, recurrent, and a host of hybrids, also, different training algorithms are in use to model ANN architectures. These architectures and training algorithms possess varying performances in classification and prediction (Wu and Ji, 2015).

As noted earlier that the application of ANN to property management related fields is limited when compared with other fields of disciplines, although, a few studies have begun to apply ANNs to real estate valuation in both the advanced and developing countries (Yacim and Boshoff, 2018). It has, however, been observed that the use of ANN to model tenant classification for risk evaluation is sparse especially in developing countries of which Nigeria is one.

Therefore, this paper examines the application of ANN to tenant selection using five different ANN training algorithms with a view to determining their respective classification performances in residential tenant selection.

The rest of the paper is structured as follows, the next section focuses on a brief review of the literature followed by the methodology, results and discussion, and conclusion.

2. Literature Review

As earlier noted, investment in rental property is driven by returns. The paramount stakeholder in ensuring regular returns from property investments is the quality of tenants in occupation. The source of income in rental investment is the tenant, as such, it is a major concern for the landlord or his representative to select tenants with the highest sense of caution when selecting tenants into the apartments (Fonseca et al 2018).

Tenant selection is the most important aspect of successful property management, as most mistakes can be corrected but putting a bad tenant in possession would not only affect the regularity of returns but could also result to endless frustrations (Olawande, 2011, Sano& Gbadegesin,2015 and

Dabara et al, 2017). Due to the importance of tenant selection in the rental submarket, research attention has been directed to finding objective solutions to the challenge of selecting tenants.

Yau and Davis(1994)investigated tenant selection using Manual Decision Support (MDS) for selecting tenants into a shopping mall. The paper concluded that manual decision support is capable of reducing the selection of bad tenants, however, this approach is largely hinged around human judgment. Aickelin and Dowsland (2002) examined the model for classification to solve mall layout and tenant selection using Enhanced Direct and Indirect Algorithm (EDIA), the result revealed that the use of algorithms will help in solving tenant selection problem.

Also, Furick, (2006) assessed the application of ANN to screen rental applicants in Florida, USA. The results revealed that the use of ANN helps in understanding tenant behaviors for the future forecast. Although the study addressed tenant behaviors for future forecast but did not examine the performance of different ANN training algorithms which isconsidered germane to ANN classification performance.

Olawande (2011) assessed criteria for tenant selection in Nigeria using Analytic Hierarchy Process (AHP), the study found that the income of the prospective tenants is central to being selected to occupy the vacant premises. Zu et al (2018) investigated the effect of tenant attributes on tenant classification by using the Extended Baye Model (EBM), the study found that the model is good in tracking off bad tenants.

Furthermore, Gbadegesin and Oletubo (2013) investigated approaches to tenant selection using Hierarchical Cluster Analysis, the study concluded that the use of multiple criteria could help in selecting good tenants. However, the study did not adopt any machine learning algorithms like ANN. Salleh et al (2014)examined the factors influencing tenants' ability to pay rent in Malaysia Public housing using descriptive and multiple regression analyses, the study concluded that ethnicity significantly influenced the ability to pay rent.Dabara et al (2017) investigated default factors in residential property in Osogbo Metropolis, Nigeria, the study revealed that careful selection of tenants, periodic inspection of property among others significantly affected the menace of rent default. Fonseca et al (2018) investigated the classification and selection of tenants in residential real estate using a constructivist approach based on mapping & Decision Expert (DEX). The study found that the integrated use of cognitive and DEX holds great potentials for broadening the understanding of property manager incorrectly classifying and select tenants into residential apartments.

From the foregoing, the previous studies are either analyzing factors that influence rent default or adopting selection approaches that feature subjectivity in tenant selection in one way or the others. However, none of the studies has investigated the application of ANN using different training algorithms, the only known study where ANN was used in screening renters is that of Furick (2006) which used a back-propagated algorithm to train the ANN.

Therefore, the current study intends to add to knowledge by comparing the classification performance of five different Neural Network training algorithms with a view to providing information on the best-suited algorithm for residential rental selection in the Nigerian property markets using 13 performance metrics. The models developed would no doubt provide a more objective decision input for tenant selection in the Nigerian property markets.

3Research methodology

3.1 Data Collection and Variable Description

A total of 724 data samples of rental applicants were obtained from the databases of the firms of Estate Surveyors and valuers which registered with the Nigerian Institution of Estate Surveyors and Valuer (NIESV). Estate Surveyors and valuers (ESV) are statutorily empowered by Decree NO. 24 of 1975 to determine the worth of interests and maintain such interest in land and buildings, the participatedESVvolunteered information on tenants selection from their firms' databases without disclosing 'sensitive' information of the rental applicants in question.

'Large dataset' is often required for ANN to learn very well, the total of 724 datasets collected for this study is relatively 'small'. ANNs are usually considered as tools that can help to analyze the cause-effect relationship in a complex system within a big data framework, however, health and behavioral sciences undergo a lot of challenges in gathering big data for analysis than any other discipline(Karim et al 2018).In most developing countries especially Nigeria, it is either that the data is unavailable or not in sufficient quantity as may be needed at a given time (Ogundari, 2017 and Bryant, 2017).

While it may be necessary to have 'big data' with which the ANNs could readily learn, in general, there is no threshold for the amount of data that one could consider as 'small' or 'big' (Qolomany et al, 2019). A review of previous researches has shown that small and big data have been used both to model ANN depending on the complexity of the problem at hand and the chosen algorithms (Amasaki and Lokan,2016, Karim et al, 2018).

However, to achieve good model learning, 507 data samples (70%) were set aside for training, 109 data samples (15%) were set aside for validation to prevent model overfitting, and 109 data samples (15%) for testing purpose.This subdivision used in the current study follows after the subdivision adopted by previous researches, for example, Fonseca et al (2019) adopted 80% and 20% for training and testing respectively, also Basaran et al (2019) used 75% and 25% for data training and testing respectively while Aliyu et al (2019)in a similar pattern used 85% and 15% for training and testing respectively.

A total of 14 characteristics of the selected and non-selected prospective tenants were extracted from the pre-tenancy assessment form (PTAF) usually given to any prospective tenants to determine his /her selection status.These 14 variables are used by the Estate Surveyors and Valuers (ESV) as screening criteria for selection, hence the adoption of 14 independent variables in this study as shown in table 1.
The selected prospective tenant is coded as 1 while the non-selected group was coded 0.

Table 1: Operationalization of variables

Variable	**Definition**	**Measurement**
Dependent Variable		
SLST	Selection Status	Dummy (1if tenant is selected, 0 if otherwise)
Independent Variables		
GEND	Gender	Dummy (1, if an applicant is male, 0 if otherwise)
MRTS	Marital Status	Dummy (1, if an applicant is married,

		0 if otherwise)
RELG	Religion	Nominal
ETHN	Ethnicity	Nominal
FAMS	Family size	Actual in number
AGE	Age	Actual in year
EDUC	Education	Ordinal
OCCP	Occupation	Nominal
SLRY	Salary	Scale (₦)
PPTY	Property type	Nominal
REFS	Reference source	Nominal
REWM	Relationship with the property manager	Nominal
DUOD	Duration of default	Scale (Monthly)
TENH	Tenant history	Dummy (1, if an applicant has a bad rental history, 0 if otherwise)

Table 1 shows the dependent and independent variables. In particular, the inclusion of ethnicity, gender, and religion as tenant selection criteria does not in any way meant to discriminate against any rental applicants. Ethnicity, gender, and religious characteristics are very sensitive in the developing countries especially in Nigeria where crises are usually skewed towards the lines of ethnicity, religion, and even gender (Olapade and Omodunbi, 2019 & Pogoson and Saleh, 2011). The response of the rental applicants to the content of PTAF reveals the background of such a rental applicant and this provides information for further actions to be taken by the property managers.

It is noteworthy that knowing the ethnicity, gender, and religious backgrounds of the rental applicants helps in appropriately sorting of rental applicants to ‘like’ and ‘unlike’ terms. In Nigeria, communal/ ethnic and religious crises are a common experience. For example, mixing tenants with known historical hostility in multitenant blocks of flats might jeopardize the interest of the property and lives of the occupants. Hence the need to understand the background of each of the rental applicants to achieve a peaceful co-existence in the property(Gbadegesin snd Ojo, 2013).

3.2. ANN Training Algorithms

A feed-forward Neural Network was used with five Neural Network training algorithms as listed in table 2;

Table 2: Category of ANN Training Algorithms

Main Function Type	Algorithms	Function Code
Backpropagation with Jacobian Derivatives	Levenberg-Marquardt	LM
Backpropagation with Gradient Derivatives	Gradient Descent Backpropagation	GD
	Resilient Back Propagation	RP

One Step Secant Back Propagation	OSS
Gradients Descent With Momentum And Adaptive Rate Backpropagation	GDX

ANN Feedforward Architecture Design

A typical feed-forward ANN consists of three layers: the input layer, hidden layer, and output layer. A feed-forward fully connected network is trained in a supervised manner (Al-Shayea, 2011). Following Lau et al (2019), the ANN model can simply be expressed as a mathematical function in equation 1;

$$\tilde{Y} = f(\tilde{X}, \tilde{W}) \qquad (1)$$

Where $\tilde{Y}$ and $\tilde{X}$ are the output and input vectors respectively, $\tilde{W}$ is a vector of weight parameters representing the connections within the ANN.

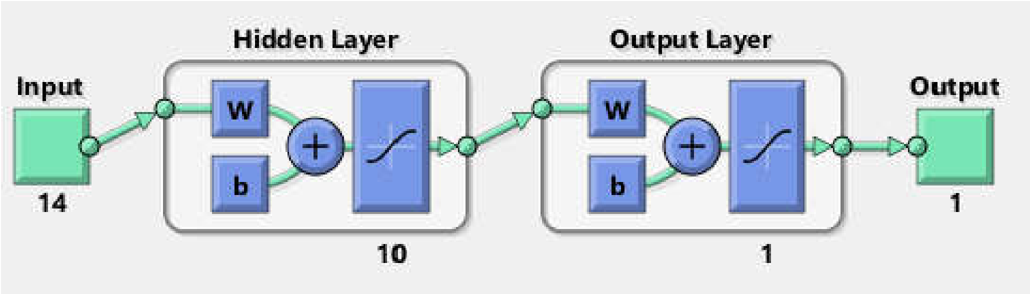
The input values were fed through the input layer into the hidden layer, the output value of j^{th} neuron y_j of the vector $\tilde{Y}$ are computed utilizingthe weighted sum of input element x and was contained in equation 2;

$$y_j = \theta\left(\sum_{i=1}^{N_i} w_{ij} x_i\right) \qquad (2)$$

θ is the activation function, N_i is the total number of j the connection lines to the jth neurons, and x_i is the output values from the previous layer of i^{th} neurons. The activation θ is used to transfer the value of the weighted sum of the input to the output layer.

A feed-forward neural network with 14 inputs, 10 hidden neurons, and 1 linear output neurons was created, using customized MATLAB version (2018a) as indicated in figure 1. MATLAB (2018a) possesses a unique Long Short-term memory (LSTM) networks for solving the connection problems.For model development, 507 datasets (representing 70%) were used as training data, 109 datasets (representing 15%) for model validation to prevent overfitting during training, and 109 datasets (representing 15%) were set aside as test data to evaluate the predictive power of the models.

The information in the input layer moves only ina forward direction, that is, from the input nodes, through the hidden nodes, and to the output nodes. The proposed Feedforward Neural Network is shown in figure 2; where w is the weight connection and b is bias.



Fig, 2: A two-layer feedforward ANN model indicating 14 inputs in the input layer, 10 number of neurons in the hidden layer, and 1 target/output used in creating the ANN model

3.3. ANN Model Performance Criteria

The dataset in this study contains 724 samples (73.3% for selected prospective tenants and 26.7% for non-selected prospective tenants) which is a case of an imbalanced dataset. It is pertinent to understand that an imbalanced dataset isa common characteristic of datasets in classification problems (Ayouche et al 2017). One of the biggest challenges in data mining is dealing with imbalanced datasets, this is often encountered in several real-world applications including credit card, fraud detection, customer retention, medical diagnostics, finance, and many other classification domains (Akosa, 2017).

An imbalanced dataset occurs when one or more classes (minority class) have a very low proportion in the dataset as compared to the other classes (majority class). Mostly in this kind of situation, the main interest is incorrectly classifying the minority class. Although, accuracy is widely used to measure the performance of classifiers, however, accuracy does not work well in an imbalanced dataset because the classifier will tend to be biased towards the majority class (selected tenant) and perform poorly in the minority class (non-selected tenants)(Brownlee, 2017).

Generally, analysts would want to balance both the false positive (FP) and false-negative (FN) rates, the performance measures that try to balance between the FP and FN rates are to be adopted in the reporting the performance of the classifier(Akosa, 2017). When dealing with class imbalance data, the final decision in model performance should consider a combination of different performance measures. Geometric means, discriminant power, F- measure, correlation coefficient, Cohen's Kappa, Younder's index, likelihood, precision, recall, and receiver operating characteristics can be used in determining the performance of any classifiers(Yaqub and AL-Amhedi, 2018).

In this regard, while standing on the findings of the previous researches, the current study evaluates the performance of ANN training algorithms using a blend of performance measures, these include precision, recall, F- measure, Geometric means, correlation coefficient, Cohen's Kappa, likelihood, mean squared error, training time, number of epoch and Area under the curve. The test data sample of 109 was used in evaluating the comparative performance of the developed models.

A brief description of the confusion matrix performance metrics adopted in this study is given as follows;

3.4. Combined Confusion Matrix Performance Measures.

Table 3; Confusion Matrix

		Predicted Value	
		Positive (P)	Negative (N)
Real Val	Positive (P)	True Positive (TP)	False Negative (FN)
	Negative (N)	False Positive (FP)	True Negative (TN)

From table 3,True Positive (TP). This is the correct classification of a good rental applicant as a good rental applicant. i.e good applicants classified as good, True Negative (TN) is a correct classification of a bad rental applicant as a bad rental applicant. i.e Bad applicants classified as Bad. False Positive (FP) refers to an incorrect classification of a bad rental applicant as a good rental applicant. i.e Bad applicants classified as good, while False Negative (FN) refers to the incorrect classification of a good rental applicant as a bad rental applicant i.e good applicants classified as bad.

Geometric Mean;

According to (Commert and Kocamaz, 2012),the distribution of data between classes may be unbalanced in many cases, in which case the geometric mean (GM) metric becomes extremely useful to achieve more objective results. GM maximizes the classification accuracy of the entire population based on the balance accuracy between the positive and negative classes. In other words, GM is only high if both classification accuracy is high (Tahir et al, 2019). The Geometric Mean (G-Mean) measures the balance between classification performances on both the majority and minority classes. A low G-Mean is an indication of poor performance in the classification of the positive cases even if the negative cases are correctly classified as such. This measure is important in the avoidance of overfitting the negative class and underfitting the positive class (Akosa, 2017). GM is expressed as;

$$G \quad mean = \sqrt{Sensitivity \times Specificity} \qquad (3)$$

F-Measure and Adjusted F-Measure

F-Measure is also called F1-score, and it represents the harmonic mean between recall and precision values. FM is used when the performance on both positive and negative classes are needed to be high, the value ranges from 0 to1, and high values of F-measure indicate high classification performance (Tahir et al, 2019). The F-Measure conveys the balance between precision and sensitivity. The measure is 0 when either the precision or the sensitivity is 0 (Akosa, 2017). It is calculated;

$$F \quad Measure = \frac{2 \times Sensitivity \times Precision}{Sensitivity + Precision} \qquad (4)$$

The Adjusted F-Measure (AGF) is an improvement over the F-Measure especially when the data is imbalanced. The AGF metric is introduced to use all elements of the confusion matrix and provide more weights to samples that are correctly classified in the minority class (Tahir et al, 2019). The AGF is calculated by first computing

$$F_2 = 5 \times \frac{Sensivity \times Precision}{(4 \times Sensitivity) + Precision} \quad (5)$$

After, the class labels of each case are switched such that positive cases become negative and vice versa. A new confusion matrix concerning the original labels is created and the quantity:

$$Inv\ F_{0.5} = \frac{5}{4} \times \frac{Sensitivity \times Precision}{(0.5^2 \times Sensitivity) + Precision} \quad (6)$$

The AGF is finally computed by taking the geometric mean of $F2$ and $InvF_{0.5}$ as

$$AGF = \sqrt{F_2 \times Inv\ F_{0.5}} \quad (7)$$

Where F2 and InvF2 are F measure and inversion of F measure respectively.

Matthew's CorrelationCoefficient

The Matthews correlation coefficient (MCC) is least influenced by imbalanced data, it represents the correlation between the observed and predicted classifications, and it is calculated directly from the confusion matrix (Akosa, 2017), The value ranges from -1 to +1. A coefficient of 1 shows that there is a perfect disagreement between the actual and the prediction, +1 when there is a perfect agreement,while 0 means no better than random and -1 the worst possible prediction(Hague et al, 2016 and Boughobel et al, 2017)

$$MCC = \frac{TP \times TN \times -FP \times FN}{\sqrt{(TP+FP)(TP+FN)(TN+FP)(TN+FN)}} \quad (8)$$

The area under the curve (AUC).

AUC provides an aggregate measure of performance across all possible thresholds, when AUC value is more than 0.5 and it is close to unity, the better the discriminating power of the model, it also means thatthe model is good enough to distinguish between the positive and negative classes (Narkhede, 2017).

4. RESULTS AND DISCUSSIONS

As earlier indicated, the study aims to examine, evaluate and compare the relative performance of the five selected Artificial neural networks training algorithms to determine the best performing among them using a blend of model performance measures which include, Mean Squared Error (MSE), training time (TT), number of the epoch, correlation coefficient, coefficient of determination, geometry mean, precision, recall, F-score, specificity, sensitivity, accuracy, and AUC. Tables4 and 5 show the frequency distribution of research variables and the model's Mean Square Error (MSE), training time, and the number of the epoch respectively. While figure 2 shows the performance of the model using the regression function on training, validation, and test.

Table 4: Frequency Distribution of Research Variables

S/no.	Variables		Frequency	Percentage
	Dependent variable	Not selected	193	26.7
1	Prospective rental applicant status	Selected	531	73.3
		Total	**724**	**100**
	Independent Variables	Female	241	33.3
2	Gender	Male	483	66.7
		Total	**724**	**100**
3	Marital	Marital	464	64.1
		Single	260	35.9
		Total	**724**	**100**
4	Religion	Others	6	0.82
		Muslim	133	18.377
		Christianity	585	80.8
		Total	**724**	**100**
5	Ethnicity	Hausa	6	0.82
		Yoruba	503	69.47
		Igbo	215	29.69
		Total	**724**	**100**
6	Family	7-10	57	7.9
		4-6	304	42.0
		1-3	363	50.1
		Total	**724**	**100**
7	Age	41 and above	191	26.4
		31-40	381	52.6
		18-30	152	21
		Total	**724**	**100**
8	Educational status	Ph.D	9	1.24
		MSc	79	10.91
		BSc	456	62.98
		O level	180	24.86
		Total	**724**	**100**
9	Occupation	Self-employed	282	38.9
		Private company	299	41.3
		Civil Servant	143	19.8
		Total	**724**	**100**
10	Income	400,000 and above	34	4.7
		200,00-350,000	216	29.8
		50,000-150,000	474	65.5
		Total	**724**	**100**
11	References	Community leader	16	2.2
		Religious leader	103	14.2
		Friend/ colleague	314	43.4
		Employer	291	40.2
		Total	**724**	**100**
12	Property type	Residential	724	100

13	Relationship with the property manager	None	458	63.3
		Same ethnic	63	8.7
		Church/mosque member	72	9.9
		Family	10	1.4
		Friend	121	16.7
		Total	**724**	**100**
14	Length of default	None	570	78.7
		>7 months	39	5.4
		4-6 months	27	3.7
		1-3 months	88	12.2
		Total	**724**	**100**
15	History of default	Bad History	163	22.5
		Good history	561	77.5
		Total	**724**	**100**

Table 4 shows the frequency distribution of the variables in the study, the number of non-selected prospective tenants is 193 representing 26.7% while the number of the selected prospective tenants is 531 representing 73.3%. This is a case of an imbalanced dataset which is usually common in the classification tasks. The frequency distributions of other variables are contained in table 4.

Table 5: Performance of ANN Training AlgorithmsUsing MSE, TT, and number of Epoch

S/NO	Algorithms	TT (seconds)	Epoch	MSE
1	LM	1	6	0.03
2	GD	2	1000	0.07
3	RP	0	6	0.06
4	OSS	0	14	0.06
5	GDX	0	102	0.07

Table 5 shows the mean squared error, the number of the epoch, and training time achieved by the models. The mean squared errors (MSE) capture the difference between the outputs and targetsZero error means no mistake indicating good performance (Ayouche et al, 2017). The closer to zero the error generated by the algorithm the better the efficiency of the algorithm. From table 5, LM generated the smallest MSE of 0.033 followed by RP and OSS with MSE values of 0.06 each, while GD and GDX both obtained MSE values of 0.07 each, in this regard, LM, RP and OSS recorded the lowest values of MSE which is consistent with the work of Wong et al(2018) which claim that LM, RP, and OSS are unique in generating low MSE.

Also, from table 5, the training time/ execution time for all the algorithms was evaluated, the time taken by RP, OSS, and GDX was almost zero seconds which makes them faster than other training algorithms under review. LM and GD trained between 1 and 2 seconds respectively. As it can be observed, the TT of all the algorithms is satisfactorily good. Although, the claim of Saji and Balachandran (2018) that LM is faster than other algorithms cannot be confirmed in the current study.

Moreover, the Training process ended in epoch 1000 for GD, GDX (102), OSS (14), RP (6), and LM (6). The study also detects positive relationships between TT and the number of epoch. The slowest ANN training algorithms ended with a larger epoch than other ANN training algorithms (GD and GDX).

Table 6:Statistical Analysis of the ANN training Algorithms using Coefficient of Determination, Correlation Coefficient and Regression Slope across training, validation and Testing Stages

S/No	Algorithms	Stages	R	R^2	b
		Training	0.90	0.86	0.78
1	LM	Validation	0.93	0.83	0.82
		Testing	0.82	0.72	0.68
		Overall	0.79	0.81	0.77
		Training	0.81	0.66	0.56
2	GD	Validation	0.73	0.53	0.56
		Testing	0.78	0.61	0.55
		Overall	0.83	0.59	0.56
		Training	0.88	0.77	0.77
3	RP	Validation	0.84	0.70	0.70
		Testing	0.80	0.69	0.68
		Overall	0.87	0.72	0.75
		Training	0.86	0.74	0.71
4	OSS	Validation	0.81	0.66	0.70
		Testing	0.85	0.72	0.70
		Overall	0.85	0.71	0.71
		Training	0.84	0.71	0.69
5	GDX	Validation	0.87	0.69	0.74
		Testing	0.79	0.67	0.66
		Overall	0.83	0.69	0.69

Table 6 indicates the coefficient of determination, correlation coefficient, and regression slopes across training, validation, and testing of the selected ANN models. The goodness of fit achieved by models is very key when evaluating model performance(Lau et al, 2019). To determine the respective goodness of fits and performances of all the selected models, coefficient of determination (R^2) and correlation coefficient (R) were used as shown in table 6. R^2 is the statistical measurement of the correlation between the predicted and target outputs which is often used in the linear regression to determine how well a line fits the data being observed especially when comparing models performance.

The higher the coefficient the better the goodness of fit and the performance of the model, for a good data fitting, the data should have the predicted output lying closely with the target output[34]. An R-value of 1 means a perfect relationship, 0 is a random relationship (Ayouche et al, 2017). As shown in table 6, LM, RP, OSS and GDX achieved good values of R^2 and R, for LM (0.81 and 0.90), RP (0.72 and 0.85), OSS (0.71 and 0.84) and GDX (0.69 and 0.83). However, the performance obtained by GD is somewhat not satisfactory as it obtained overall R^2 and R values of 0.59 and 0.77 respectively. Among the models, LM recorded the highest R^2 and R across the training, validation, and testing stages.

There is an indication of a good model without overfitting when the validation correlation coefficient R is close to the correlation coefficient value of the training stage (Viejo et al, 2019). All the models evaluated except GD passed this test of models without overfitting.

One important factor in any successful data mining is the issue of the model assessment technique. One of the commonly reported measures of the performance of an algorithm is the accuracy, this measures the overall efficiency of the algorithms but it could be misleading in the case of an imbalanced dataset thereby leading to a wrong decision(Brownlee, 2019). Accuracy tends to assign more weight to the majority class than the minority class. It is also noteworthy that the use of only one measure apart from accuracy could as well result in taking the wrong decision(Akosa, 2017).

Table 7: Performance of ANN using Combined Confusion Metrics

Measure	**LM**	**GD**	**RP**	**OSS**	**GDX**
Accuracy (ACC)	0.92	0.85	0.91	0.94	0.94
Sensitivity (SE)/ Recall	0.95	0.97	0.96	0.99	0.99
Specificity (SP)	0.83	0.63	0.78	0.79	0.81
Geometric Mean (GM)	0.89	0.78	0.87	0.89	0.89
Precision/Positive Predictive Value (PPV)	0.94	0.82	0.92	0.92	0.93
F-Measure (FM)	0.94	0.89	0.94	0.96	0.96
F_2- Measure (F_2M)	0.94	0.91	0.94	0.97	0.96
Adjusted F-Measure (AGF)	0.90	0.78	0.87	0.89	0.90
Mathew's Correlation Coefficient (MCC)	0.79	0.67	0.77	0.85	0.84
Random Accuracy (RA)	0.61	0.57	0.60	0.62	0.62
Kappa (KA)	0.79	0.65	0.77	0.84	0.84
Likelihood Ratio (LR(+)	5.72	2.60	4.43	4.83	5.07
Area Under Curve (AUC)	0.89	0.80	0.87	0.89	0.90

The balanced approach in model assessment is a combined approach of using more than one metric in evaluating the model performance.This approach is adopted in this study as shown in table 7.

The results of combined measures is shown in table 7, themodels exhibit varying performances as ranked by combined metrics. GDX recorded the highest predictiveACC (0.94),SE(0.99), GM(0.89),FM(0.96),AGF(0.90),TA(0.94),RA(0.89) and AUC(0.90).OSS also has the highest ACC (0.94),SE(0.99),GM(0.89),FM(0.96), F2(0.97), MCC(0.85), TA(0.94) and RA(0.84). Furthermore, LM also has the highest performance in SP(0.83), GM(0.89),PR(0.94), IF(0.85), AGF (0.09) andLR (+)(5.79).

The performance of RP is equally good but it is distantly following GDX, OSS, and LM in the overall ranking. while GD though somewhat high but ranks least in the performance table, this is similar to the conclusion reached in the earlier stated coefficient of determination and correlation coefficient in this study. The finding on GD is consistent with the studies of (Bache and Lichman, 2013 and Viejo, 2019) which claim that GD is traditionally lower in performance than other training algorithms. Also, from the analysis, GDX and OSS outperformed other ANN training algorithms as ranked by the combined confusion matrix measures. This finding is corroborated by the work of (Viego et al, 2019) which claims that GDX and OSS are among the foremost ANN training algorithms with higher classification efficiencies than other ANN training algorithms.

However, the paper does not find evidence of the claim made by Lau et al (2019) that LM and RP possessed better classification performance than GDX and OSS.

5 CONCLUSION

The study attempts to assess neural network training algorithms for residential tenant classification. ANN models were developed, trained, validated,and tested using five ANN training algorithms. The test datasets were used in evaluating the performance of the models. The performance of the networks was evaluated using training time, the number of the epoch, MSE, precision, recall, specificity, sensitivity, geometry mean, AUC, and coefficient of determination of actual outputs and target outputs. It could be concluded that all the ANN training algorithms produced good results which are considered adequate for residential tenant classification, however, GDX and OSS obtained the first position as ranked by the eight performance metrics, LM and RP also achieved very good performance in tenant classification. The study concludes that GDX and OSS are the most suitable for residential tenant classification in the Nigerian property market while GD should be adopted as being the least performing ANN training algorithm with caution.

The models developed in this study provide tenant selection decision supports for stakeholders in the field of property management such as property managers, investors, and other policymakers. The models also provide frameworks that reduce the long and cumbersome process in tenants selection exercise.

The finding of this study is limited to the datasets used and the location of the study. Further research efforts may be directed at examining other ANN architectures such as recurrent neural networks using more training algorithms and datasets.

References

Aickelin, U., and Dowsland, K.A., (2002). Enhanced direct and indirect genetic algorithm approaches for a mall layout and tenant selection problem. Journal of Heuristics, 8(5), 503-514.

Akosa, J., (2017), Predictive accuracy: a misleading performance measure for highly imbalanced data. In Proceedings of the SAS Global Forum (pp. 2-5).

Aliyu, I.N., Abdulrahaman, M.D., Nwaokolo, B.O., and Abdulkareem, S.A., (2019). An Effective Breast Cancer Prediction and Classification Using Artificial Neural Network. Journal of Engineering and Technology, 2(2), 37-48.

Alizadeh, M., Ngah, I., Hashim, M., Pradhan, B. and Pour, A.B., (2018). A hybrid analytic network process and artificial neural network (ANP-ANN) model for urban earthquake vulnerability assessment. Remote Sensing, 10(6), 975-976.

Al-shayea, Q.k., (2011). Artificial neural networks in medical diagnosis. International Journal of Computer Science Issues, 8(2). 150-154.

Amasaki, S. and Lokan, C., (2016); On applicability of fixed-size moving windows for ANN-based effort estimation. In 2016 Joint Conference of the International Workshop on Software Measurement and the International Conference on Software Process and Product Measurement (IWSM-MENSURA) (213-218). IEEE.

Ayouche, S., Aboulaich, R., and Ellaia, R., (2017). Partnership credit scoring classification problem: a neural network approach. International Journal of Applied Engineering Research, 12(5), 693-704.

Bache, K. and Lichman, M., (2013). UCI machine learning repository.

Başaran, E., Cömert, Z., Şengür, A., Budak, Ü. Çelik, Y. and Toğaçar, M., (2019). Chronic Tympanic Membrane Diagnosis based on Deep Convolutional Neural Network. In 2019 4th International Conference on Computer Science and Engineering (UBMK) (pp. 1-4). IEEE.

Boughorbel, S., Jarray, F. and El-Anbari, M., (2017). Optimal classifier for imbalanced data using Matthews Correlation Coefficient metric. PloS one, 12(6).

Bradley, A.P., (1997). The use of the area under the ROC curve in the evaluation of machine learning algorithms. Pattern recognition, 30(7), 1145-1159.

Brownlee, J., (2017). Long short-term memory networks with python. *Machine Learning Mastery. Retrieved from https://machinelearningmastery. com/deep-learning-with-python*.

Bryant, C., (2019). *Managing development in the Third World*. Routledge.

Cömert, Z. and Kocamaz, A.F., (2017). Comparison of machine learning techniques for fetal heart rate classification. *Acta Phys. Pol. A*, *132*(3), pp.451-454.

Dabara, D.I., Anthony, A.I., Olusegun, O.J., Eleojo, A.G. and Michael, A.O., (2017) Rent Default Factors in Residential Properties in Osogbo Metropolis Osun State, Nigeria. *International Journal of Business and Management Studies, 6(1);61-68*

Estabrooks A. and japkowicz, N., (2001), A mixture-of-experts frameworkfor learning from imbalanced data sets. In *international symposium on intelligent Data Analysis* (pp. 34-43). Springer, Berlin, Heidelberg.

Fonseca, M.B.B., Ferreira, F.A., Fang, W. and Jalali, M.S.,(2018). Classification and selection of tenants in residential real estate: a constructivist approach. International Journal of Strategic Property Management, 22(1), pp.1-11.

Fonseka, T.M., Bhat, V. and Kennedy, S.H., (2019). The utility of artificial intelligence in suicide risk prediction and the management of suicidal behaviors. Australian & New Zealand Journal of Psychiatry, 53(10), 954-964.

Furick, M. T. (2006). *Using Neural Networks to Develop a New Model to Screen Applicants for Apartment Rentals.* Doctoral dissertation, Nova Southeastern University, Graduate School of Computer and Information Sciences.

Gbadegesin, J. and Oletubo, A., (2013). Analysis of tenant selection criteria in an emerging rental market. Global Journal of Management and Business Research Interdisciplinary, 13(7),1-12.

Gbadegesin, J.T. and Ojo, O., (2013), Ethnic bias in tenant selection in metropolitan Ibadan private rental housing market. *Property management.6 (4), 29-45.*

Gerritsen,L., (2017). Predicting student performance with Neural Networks (Doctoral dissertation, Doctoral dissertation, Tilburg University.)

Haque, M.N., Noman, N., Berretta, R. and Moscato, P., (2016). Heterogeneous ensemble combination search using genetic algorithm for class imbalanced data classification. PloS one, 11(1), 12-27.

Karim, H., Niakan, S.R. and Safdari, R., (2018). Comparison of neural network training algorithms for classification of heart diseases. IAES International Journal of Artificial Intelligence, 7(4), 185-103.

Kaya, A., Keceli, A.S., Catal, C., Yalic, H.Y., Temucin, H. and Tekinerdogan, B., (2019). Analysis of transfer learning for deep neural network based plant classification models. Computers and electronics in agriculture, 158, 20-29.

Khor, R.C., Nguyen, A., O’Dwyer, J., Kothari, G., Sia, J., Chang, D., Ng, S.P., Duchesne, G.M. and Foroudi, F., (2019). Extracting tumour prognostic factors from a diverse electronic record dataset in genito-urinary oncology. International journal of medical informatics, 12(1), 53-57.

Lan, Z., Chen, M., Goodman, S., Gimpel, K., Sharma, P. and Soricut, R., (2019). Albert: A lite bert for self-supervised learning of language representations. *arXiv preprint arXiv:1909.11942*.

Lau, E.T., Sun, L. and Yang, Q., (2019). Modelling, prediction and classification of student academic performance using artificial neural networks. SN Applied Sciences, 1(9), 982-1000.

Lin, M.I.B., Groves, W.A., Freivalds, A., Lee, E.G. and Harper, M., (2012). Comparison of artificial neural network (ANN) and partial least squares (PLS) regression models for predicting respiratory ventilation: an exploratory study. European journal of applied physiology, 112(5),1603-1611.

Mustika, H.F., Syafiandini, A.F., Manik, L.P. and Rianto, Y., (2020). Evaluating Naïve Bayes Automated Classification for GBAORD. Computer Engineering and Applications Journal, 9(1), 29-37.

Narkhede, S., (2018). Understanding AUC-ROC Curve. *Towards Data Science*, *26*.

Ogundari, K., (2017). Categorizing households into different food security states in Nigeria: the socio-economic and demographic determinants. *Agricultural and Food Economics*, *5*(1), 8-19.

Olaopa, O.R. and Omodunbi, O., (2019). Politics of identity and crisis of nation building in Africa: the Nigerian experience. Journal of Nation-building & Policy Studies, 3(2), 45-65.

Olatoye, O., (2005). Borrowers' Perception of the Degree of Cumbersomeness of Lenders Requirements in Housing Financing in Southwestern Nigeria. In Conference Proceedings, Brisbane, Australia.

Olawande, O.A., (2011). Harnessing real estate investment through proper tenant selection in Nigeria. Property Management, 4(2), 15-36.

Pogoson, A.I. and Saleh, M.U., (2019). Gender and Nigeria's Internal Security Management. In Internal Security Management in Nigeria (pp. 633-647). Palgrave Macmillan, Singapore.

Qolomany, B., Al-Fuqaha, A., Gupta, A., Benhaddou, D., Alwajidi, S., Qadir, J. and Fong, A.C., (2019). Leveraging machine learning and big data for smart buildings: A comprehensive survey. *IEEE Access*, *7*, 90316-90356.

Saji, S.A. and Balachandran, K., (2015); Comparative Study of various training algorithms of Artificial Neural Networks on Diabetes dataset. International Journal on Recent and Innovation Trends in Computing and Communication, 3(2), 378-382.

Salleh, N.A., Johari, N. and Talib, Y., (2014). Identifying Variables Influencing Tenant Affordability to Pay Rent in Ipoh City Council Public Housing. In E3S Web of Conferences Vol. 3, EDP Sciences.

Sani, K.S. and Gbadegesin, J.T., (2015). A study of private rental housing market in Kaduna Metropolis, Nigeria. International Journal of Humanities and Social Science, 5(8), 173-183.

Sharif, M.S., Abbod, M., Krill, B., Amira, A. and Zaidi, H., (2011), Automatic PET volume analysis and classification based on ANN and BIC. In 2011 IEEE 15th International Symposium on Consumer Electronics (ISCE) (pp. 565-570). IEEE.

Sokolova, M., Japkowicz, N. and Szpakowicz, S., (2006). Beyond accuracy, F-score and ROC: a family of discriminant measures for performance evaluation. In Australasian joint conference on artificial intelligence (pp. 1015-1021). Springer, Berlin, Heidelberg.

Tahir, M.A.U.H., Asghar, S., Manzoor, A. and Noor, M.A., (2019). A classification model for class imbalance dataset using genetic programming. IEEE, 7, pp.71013-71037.

Tharwat, A., (2018). Classification assessment methods. *Applied Computing and Informatics*.

Viejo, C.G., Fuentes, S., Howell, K., Torrico, D.D. and Dunshea, F.R., (2019). Integration of non-invasive biometrics with sensory analysis techniques to assess acceptability of beer by consumers. Physiology & behavior, 2(2),139-147.

Wong, Y.J., Arumugasamy, S.K. and Jewaratnam, J., (2018). Performance comparison of feedforward neural network training algorithms in modeling for synthesis of polycaprolactone via biopolymerization. Clean Technologies and Environmental Policy, 20(9), 1971-1986.

Wu, G. and Chang, E.Y., (2003), Class-boundary alignment for imbalanced dataset learning. In *ICML 2003 workshop on learning from imbalanced data sets II, Washington, DC* (49-56).

Wu, Y. and Ji, Q., (2015). Discriminative deep face shape model for facial point detection. International Journal of Computer Vision, 113(1), 37-53.

Yacim, J.A. and Boshoff, D.G.B., (2018). Impact of artificial neural networks training algorithms on accurate prediction of property values. Journal of Real Estate Research, 40(3), 375-418.

Yan, Q., Xu, L., Shi, J. and Jia, J., (2013). Hierarchical saliency detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 1155-1162).

Yaqub, M.U. and Al-Ahmadi, M.S., (2016), Application of combined ARMA-neural network models to predict stock prices. In Proceedings of the 3rd Multidisciplinary International Social Networks Conference on Social Informatics (pp. 1-5).

Yau, C. and Davis, T., (1994). Using multi-criteria analysis for tenant selection. Decision Support Systems, 12(3), 233-244.

Zu, Q., Wu, T. and Wang, H., (2012). A multi-factor customer classification evaluation model. *Computing and Informatics*, *29*(4), 509-520.