American Journal of Epidemiology & Public Health

Research Article

A Comparison of Multinomial Logistic Regression and Artificial Neural Network Classification Techniques Applied to TB/HIV Data

Samuel Ominyi Ejeh[1]*, Olatayo Oluwasegun Alabi[1], Opeyemi Oyekola Ogungbola[1], Olawumi Olumayowa Olatunde[2] and Zainab Olabisi Dere[3]

[1]*Department of Statistics, Federal University of Technology Akure, Akure, Ondo State, Nigeria*

[2]*National Bureau of Statistics, Plot 762, Independence Avenue, Central Business District, Abuja, Nigeria*

[3]*Deptartment of Mathematics, Florida State University, USA*

***Address for Correspondence:** Samuel Ominyi Ejeh, Department of Statistics, Federal University of Technology Akure, Akure, Ondo State, Nigeria, E-mail: bosammie@gmail.com

Submitted: 21 February 2022; **Approved:** 31 March 2022; **Published:** 07 April 2022

Cite this article: Ejeh SO, Alabi OO, Ogungbola OO, Olatunde OO, Dere ZO. A Comparison of Multinomial Logistic Regression and Artificial Neural Network Classification Techniques Applied to TB/HIV Data. American J Epidemiol Public Health. 2022 Apr 07;6(1): 014-018. doi: 10.37871/ajeph.id53

ABSTRACT

Multinomial Logistic Regression (MLR) is the extension for the (binary) logistic regression when the categorical dependent outcome has more than two levels. Artificial Neural Networks (ANN) are a special type of machine learning algorithms that are modeled after the human brain. A Multilayer Perceptron (MLP) is a feedforward neural network with one or more hidden layers. An MLP uses backpropagation as a supervised learning technique. This research compares the performance of Multinomial Logistic Regression (MLR) models and ANN (Multilayer Perceptron, MLP) models using TB/HIV co-infected patients' data obtained from the Directly Observed Treatment Short Course (DOTs) Clinic of Nigerian Institute of Medical Research (NIMR) who had been registered between 2012 and 2021. The dependent (target) variable are the categorical data; Completed and Cure, Died and Defaulted. The referent group in the MLR is the Completed and Cured group and two models was estimated for the Died relative to Completed and Cured and Defaulted relative to Completed and Cured. The research compared the percentage of correct classification from MLR and ANN models, it was established that ANN model classifies the dataset better with an overall correct percentage classification of 94.4% than the MLR with an overall correct classification of 92.7%.

Keywords: Multinomial logistic regression model; Artificial neural networks; Multilayer perceptron TB/HIV infection; Absolute lymphocyte count; Body mass index

INTRODUCTION

Tuberculosis (TB) is and will continue to be a global public health issue, as well as one of the world's top 10 leading causes of death. In Nigeria, TB continues to be a major public health issue. Since the development of AIDS, Tuberculosis (TB) and HIV have been tightly associated, and tuberculosis is the most common opportunistic illness afflicting HIV-positive individuals, as well as the leading cause of mortality in AIDS patients.

Health professionals at the Nigerian Institute of Medical Research's (NIMR) Directly Observed Treatment Short Course (DOTs) center are continually presented with three possible outcomes (Completed and Cure, Died, or Defaulted) for TB/HIV co-infected patients who enrolled in their facilities. They are constantly looking for techniques to precisely classify patients with certain medical variables and treatment outcomes in order to efficiently monitor patients who are more likely to die or default before completing their treatment. Multinomial Logistic Regression (MLR) and Artificial Neural Network (ANN) are two of the most common statistical tools for tackling classification problems nowadays.

Multinomial Logistic Regression Analysis (MLR)

The multinomial logistic regression model is an extension of the binary one, implying that the outcomes are multi-leveled rather than dichotomous. When discussing a discrete-outcome regression model with at least three responses, the measurement scale should be considered. The date's output variable in this study is on a nominal scale.

When the dependent variables are categorical with more than two levels, multinomial logistic regression is used. There will be multiple regression equations if the dependent variable has more than two values. In fact, the number of regression equations is one fewer than the number of possible outcomes.

Artificial Neural Networks (ANN)

Artificial Neural Network is considered nonlinear statistical data modeling tools where the complex relationships between inputs and outputs are modeled. Artificial Neural Networks (ANNs) are processing models inspired by the human brain that, due to their input-output mapping capacity, have been widely used to tackle regression and classification problems. That is, just as our nervous system's neurons can learn from previous data, the ANN may learn from the data and deliver responses in the form of predictions or classifications. There are two important types of Artificial Neural Networks –:

Feedforward Neural Network - Information flows solely in one way in feedforward ANNs. That is, data flows from the input layer to the concealed layer, then to the output layer. In this neural network, there are no feedback loops at all. This form of neural network is commonly employed in supervised learning applications like classification and image recognition. They're useful when the data are not always in a logical order.

Feedback Neural Network - The feedback loops are an element of the feedback ANNs. Recurrent neural networks are a form of neural network that is mostly used for memory retention. These networks are best used in situations where data are sequential or time dependent. An ANN with four input neurons, a schematic representation of a generic totally connected feedforward *ANN* is depicted in figure 1.

AIM AND OBJECTIVES

The aim of this study is to compare the performance of the Multinomial Logistic regression and Artificial Neural Network models in Classification of Tuberculosis/HIV co-infected patients.

The specific objectives are to:

i. Formulate Multinomial Logistics Regression and Artificial Neural Network models for Tuberculosis / HIV co-infected patients.

ii. Determine the significance of Tuberculosis preventive therapy regimens among TB/HIV co-infected patients.

iii. Evaluate the performances of Multinomial Logistics Regression and Artificial Neural Network model in term of Accuracy, Sensitivity, Specificity, and the percentage of correct classification rate.

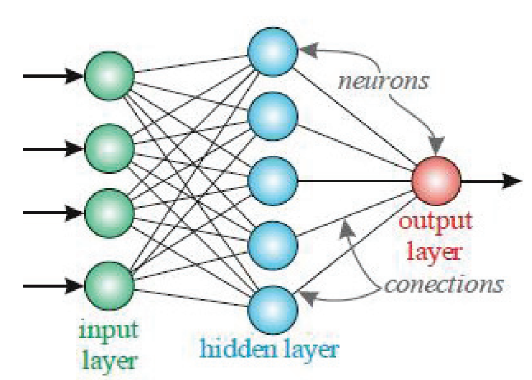

Figure 1: Schematic representation of a generic totally connected feed forward ANN.

DOI: https://dx.doi.org/10.37871/ajeph.id52

ISSN: 2644-0032

This study precisely identifying individuals with certain medical variability into treatment outcomes, which will aid in the effective monitoring of patients who are more likely to die or default before the treatment is finished. It also aims to assess the value of the treatment regimens used at the Directly Observed Treatment Short Course (DOTs) facility for TB/HIV co-infected patients in Lagos, Nigeria. Patients who registered at the Lagos Directly Observed Treatment Short Course (DOTS) center were studied. DOTS is a clinic run by the Nigerian Institute of Medical Research (NIMR), a parastatal under the Federal Ministry of Health that has treated over 5000 patients with tuberculosis in the last decade (2012-2021) [1]. This researcher work will be limited to performance evaluation of the Multinomial Logistic Regression and that of Artificial Neural Network in accurately classifying the treatment outcome of patients in the center thereby providing healthcare worker the opportunity to pay more attention to patients with certain medical test result variability which usually could determine treatment outcome.

Artificial Neural Networks (ANN) was frequently used as a modelling tool in the analysis of complex problems. Their study compared the estimates for total duration during log skidding operations stations in Turkey's Eastern Black Sea region (Giresun Forest District Directorates) with those of Multiple Regression Analysis (MRA) [2]. This study used a multinomial logistic regression model to examine the perceptions of these actors on the consequences of the pandemic at the local level and to discover the main elements that influenced their evaluation. The findings revealed a widespread concern about issues such as employment, job stability, and household debt. When assessing the vulnerability of some groups, particularly women and the elderly, to the consequences of the crisis and their role as citizens, the variables of age and sex were significant [3]. The findings of an experimental comparison study of Logistic Regression Analysis (LRA) and Artificial Neural Network (ANN) for predicting the academic success of prospective mathematics instructors when they join graduate school. Our model was trained and tested using 372 student profiles. Logistic Regression Analysis can be used to assess the model's strength (LRA). For ANN, pupils had a greater average correct success percentage than for LRA. The Back-Propagation Neural Network (BPNN, or a typical type of ANN) had a successful prediction rate of 93.02 percent, whereas the LRA had a prediction rate of 90.75 percent [4]. The researcher was curious about the relationship between different birth control methods and the intervention, as well as age, race, IPV/Reproductive Coercion (RC) experiences, and relationship status. The study proposed a way for determining if intervention, IPV, RC, and women's characteristics were linked to contraceptive methods; it also proposed a method for classifying different groups of women based on 1-year longitudinal patterns of women's recognition of abusive behaviors [5]. The characteristics that influence the types of domestic violence against women were summarized and assessed using a multinomial logistic regression model. Furthermore, twelve independent variables were employed, with the irrelevant variables excluded from the data set using the chi-square test of independence [6]. Using multinomial logistic regression, a study was conducted to investigate the association between children's work status and their demographic variables. Maximum likelihood estimate was used to test the validity of the model in the study, and the model was found to be significant [7]. Reported the findings of a study that compared two models for classifying pupils based on their academic performance: Multinomial Logistic Regression (MLR) and Artificial Neural Network (ANN). The average Classification Correct Rate was used to determine each model's predicted accuracy (CCR). The Multinomial Logistic Regression had a lower average classification correct rate than the Artificial Neural Network [8]. A Multinomial Logistic Regression model was used to analyze categorical data in this study. Whether nominal or ordinal, the model works with a single nominal/ordinal answer variable with more than two categories. This paradigm has been used to analyze data in a variety of settings, including health, social, behavioral, and educational settings [9].

The programming language used in this manuscript is Statistical Packages for Social Sciences (SPSS). The evaluation model used in the manuscript is split test (also known as Train/Test split). The data was split in two, 70% for training and 30% for test.

METHODOLOGY

Multinomial Logistic Regression (MLR)

In brief, multinomial logistic regression model is an expanded version of the two-category model (binary model) for *J* category. Thereby, *J*-1 multinomial logistic regression models occur [10]. It is allowed that the response probabilities depend on nonlinear transformations of the linear function of equation:

$$x_i^{'}\beta_i = \sum_{k=0}^{K} \beta_{jk} x_{ik} \tag{1}$$

where K is the number of the predictors, i represents *ith* individual, x represents independent variable and j expresses the category dependent variable.

The multinomial logistic model can be viewed as an extension of the binary logit model. For example, in case of three categories (*J*-3) the models can be written as below:

$$P_{i1} = P\left(y_i = 1|x_i\right) = \frac{1}{1+\exp\left(x_i^{'}\beta_2\right)+\exp\left(x_i^{'}\beta_3\right)} \tag{2}$$

$$P_{i2} = P\left(y_i = 2|x_i\right) = \frac{\exp\left(x_i^{'}\beta_2\right)}{1+\exp\left(x_i^{'}\beta_2\right)+\exp\left(x_i^{'}\beta_3\right)} \tag{3}$$

$$P_{i3} = P\left(y_i = 3|x_i\right) = \frac{\exp\left(x_i^{'}\beta_3\right)}{1+\exp\left(x_i^{'}\beta_2\right)+\exp\left(x_i^{'}\beta_3\right)} \tag{4}$$

In here, β_2 and β_3 denote the covariate effects specific to the second and third response categories with the first category as the reference. Besides, a reference category (baseline category) is determined at first to compare and analyze. At this stage, the researcher can select the reference category (j) optionally for instance, if there are categories such as 1, 2, and 3 in a dependent variable, 1 can be selected as the reference category. In this way, two different logistic models can be obtained for comparison of 1-2 and 1-3 [11].

On the other hand, the equation (2) for P_{i1} can be derived from the constraint that the three probabilities sum to 1.

$$P_{i1} = 1-(P_{i2}+P_{i3}) \tag{5}$$

The sum of the probabilities of categories of dependent variable

DOI: https://dx.doi.org/10.37871/ajeph.id52

ISSN: 2644-0032

should be equal to 1 as in binary logit model. For instance, if the dependent variable has a category three-level structure, the sum of the probabilities for each category will be equal to 1 as follows [12];

$$[P_{i1} = P(y_i = 1|x_i)] + [P_{i2} = P(y_i = 2|x_i)] + [P_{i3} = P(y_i = 3|x_i)] = 1 \tag{6}$$

In general, the probabilities of a dependent variable with j categories can be expressed in multinomial logit as below:

$$P_{ij} = P(y_i = j|x_i) = \frac{\exp(x_i'\beta_3)}{1 + \sum_{j=2}^{J} \exp(x_i'\beta_j)} \tag{7}$$

If J is selected as the baseline category, the probability of the dependent variable to lie within the baseline category is defined as given in Equation (7)

$$P_{i1} = P(y_i = 1|x_i) = \frac{1}{1 + \sum_{j=1}^{J-1} \exp(x_i'\beta_j)} \quad j = 1,2,\ldots,J-1 \tag{8}$$

Furthermore, the probability that it lies within the baseline category can be computed with the help of other probabilities as given in Equation (8), if and only if the other probabilities are known [13].

$$P_j = P(y = J) = 1 - [P(y = 1) + P(y = 2) + \ldots + P(y = J-1)] \tag{9}$$

In a multinomial logistic regression model, the logit transformation is obtained by taking the logarithms of the odds ratios after selecting the baseline category. For the four-category example, when 0 is selected as the baseline category, the logarithms of odds ratios can be obtained as given in Equation (10), Equation (11), and Equation (12) [14].

$$In\left[\frac{P(y = 1|x_1)}{P(y = 0|x_1)}\right] = \beta_1 + \beta_{11}x_1 \tag{10}$$

$$In\left[\frac{P(y = 1|x_2)}{P(y = 0|x_1)}\right] = \beta_2 + \beta_{21}x_1 \tag{11}$$

$$In\left[\frac{P(y = 1|x_3)}{P(y = 0|x_1)}\right] = \beta_3 + \beta_{31}x_1 \tag{12}$$

As its seen, the baseline category is taken as "$y = 0$" in all the three odds ratios. The notation of the model can be generalized as in Equation (13) with all these given [13].

$$In\left[\frac{P_j}{P_J}\right] = In\left[\frac{P(y = j)}{P(y = J)}\right] = \left(\sum_{k=1}^{K} \beta_{jk}x_k\right) \quad j = 1,\ldots J-1 \tag{13}$$

As Equation (11) indicates, multinomial logistic regression model can be transformed into binary logit model for $J = 2$. The multinomial logit model is estimated using maximum likelihood with the log-likelihood function for a sample of n-observations given by

$$1 + \sum_{j=1}^{J-1} \exp(x_i'\beta_j) \tag{14}$$

$$InL = \sum_{i=1}^{n} \sum_{j=1}^{J} d_{ij} log(P_{ij}) \tag{15}$$

Where d_{ij} is a dummy variable that takes a value 1 if observation i takes the jth category and 0 otherwise because P_{ij} is a nonlinear function of parameters of the regression model.

Artificial neural network

Artificial Neural Networks (ANNs) are processing models inspired by how the brain functions, and they are one of the most widely used data analysis tools in science. An ANN is made up of a group of neurons and the connections that connect them, and it is distinguished by its high processing capability despite the simplicity of the neurons that make up the ANN [15].

Multilayer neural networks

A feedforward neural network having one or more hidden layers is referred to as a multilayer perceptron. A source neuron input layer, at least one middle or hidden layer of computational neurons, and a source neuron output layer make up the network. On a layer-by-layer basis, the input signals propagate forward. With no necessity for linear separability, a multilayer neural network is far more adaptable than a single neuron. Training takes longer and is less obvious. MLP is the most frequent type of multilayer perceptron (Multi-Level Perceptron). Network of Backpropagation (alluding to a common method of training these networks; other training methods could conceivably be used). A pictorial diagram of Backpropagation training cycle is shown in figure 2.

RESULTS

Multinomial Logistic Regression Analysis (MLR)

For the Multinomial Logistics analysis, since our output category is three (Completed and Cure, Died and Defaulted), the Completed and Cured is treated as the referent group and parameter estimation is obtained for Died relative to Completed and Cure group then also for Defaulted relative to Completed and Cure group. A descriptive

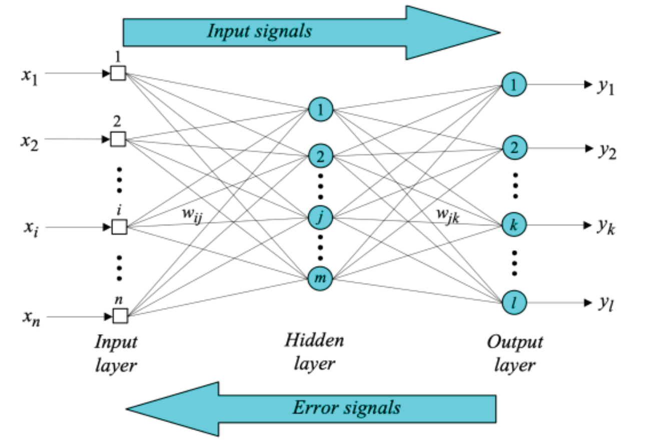

Figure 2: Back-propagation training cycle.

DOI: https://dx.doi.org/10.37871/ajeph.id52

ISSN: 2644-0032

Statistics of the 466 participants in this study is presented in table 1, while Description, Code and Values for the Variables used for the study is presented in table 2. Other results for MLR analysis like goodness of fit, Logistics Regression (LR) chi-square and Pseudo R^2 Model fitting information and the Classification result are presented in subsequent sections.

Goodness of fit test: The goodness-of-fit test based on deviance is a likelihood-ratio test between the fitted model and the saturated one. Measures of goodness of fit typically summarize the discrepancy between observed values and the values expected under the model in question.

In table 3, we modelled the effects of treatment regimens as to its significant to tuberculous preventive therapy, using the Chi-square (Pearson test). Since (p = 0.996) > (α = 0.05), we concluded that the effect of the treatment regimens does not have significant to TB preventive therapy for TB/HIV co-infected adults. There is no observed difference in the observed values and the expected values in the model which is why the result is not significant. Pseudo R^2 presented in the table 4 above are examined also, this result indicates the proportion of variation being explained by the model. From table 4, it is observed that independent variable defined 11.5% of the variance of the dependent variable according to Cox and Shell R^2 value, 16.5% according to Nagelkerke R^2 values and 10.3% according to McFadden R^2 value.

Table 1: Descriptive statistics of the 466 participants in this study.

Variable	Mean	Standard Deviation	Coefficient of Variation
Age	37.6137	12.7125	33.7975
Body Mass Index	0.3495	0.1245	35.6223
CD4 Count	202.5073	168.5325	83.2229
Creatinine	100.4553	68.4812	68.1708
Hemoglobin	52.7834	27.5678	52.2282

Table 2: Description, code and values for the variables used for the study.

Variable	Description of Variable	Codes or values used for variable levels
Age	Age	year(s)
Gender	Patient's sex	0 = Male, 1 = Female
Status	Marital status	0 = Single, 1 = Married 2 = Divorce, 3 = Widow
Type of Tb	Types of Tuberculosis Disease	Pulmonary = 0 Extra-pulmonary = 1
Weight	Weight	Kg
Art Status	HIV positive on Anti-retroviral drug (ARV) Status	HIV positive on ART = 0 HIV positive non-ART = 1
Treatment	Treatment Outcomes	Completed and cured =0 Died = 1 Defaulted = 2
Hemoglob	Hemoglobin	g/dL
Creatinine	Creatinine level	Mg/Dl
Cd4count	CD4 cells level measurement	Cell/mm^3
Glucose	Glucose Level	Mg/dL

Table 3: Chi-square value, degree of freedom and significance level computed.

Methods	Chi-Square	df	Sig.
Pearson	796.997	906	0.996
Deviance	231.289	906	1.000

Table 4: Various pseudo R^2 values computed.

Measurement	R^2 values
Cox and Snell	0.115
Nagelkerke	0.165
McFadden	0.103

Artificial Neural Network (ANN)

Here we present the results for ANN (Multilayer Perceptron) analysis which includes Case Processing Summary, Model summary and Independent Variable Importance.

The summary of ANN (Multilayer Perceptron) is presented in table 5. The sum of squares Error for Training and Testing (22.767 and 6.197) respectively is relatively low since the percentage of incorrect prediction of 7.9% in Training and 5.6% in testing shows that the model achieved 92.1% accuracy in training and 94.4% in testing respectively.

Gradient descent: Gradient descent takes little, predictable strides towards the local minima and when the inclinations are small it can require some time to converge. Momentum then again considers the past slopes and speeds up union by pushing over valleys quicker and staying away from local minima.

Batch gradient descent, also called vanilla gradient descent, calculates the error for each example within the training dataset, but only after all training examples have been evaluated does the model get updated. This whole process is like a cycle and it's called a training epoch. We present the Training summary in the following table 6.

The independent Variable Importance shows the contribution of each independent variable to predicting the Outcome (dependent). The Importance of independent variables and its normalized values is presented in table 7. The Normalized importance shows that creatinine level (100%) follows by Glucose level (81.4%), Type of Tuberculosis (36.9%) and CD4 count (25.5%) has major contribution to predicting the output of the dependent variables.

A simple bar chart showing the distribution of various variable contribution in the model is also presented in figure 3 for better understanding.

ANN's multilayer network relationship

The relationship network is presented in figure 4, the effects of the different dependent variable are presented to the network input layer. The network propagates the input pattern back and forward the hidden layer until the output pattern is generated. It was observed that parameter coefficients changed 224 times during the interaction in "treatment" and Perceptron process to improve the model. The weights are modified as the error is propagated until the desired output is achieved.

Relative Operating Characteristics Curve (ROC)

The ROC curve shows at different Classification threshold, the Sensitivity (the proportion of true positive i.e., the proportion of

DOI: https://dx.doi.org/10.37871/ajeph.id52

ISSN: 2644-0032

Table 5: Parameter estimation.

Treatment Outcome		B	Std. Error[a]	Wald	df	Sig	Exp(B)	95% confidence interval for Exp(B)	
								Lower bound	Upper bound
Died	Intercept	-6.404	1.979	10.476	1	.001			
	Age	.031	.024	1.748	1	.186	1.032	.985	1.081
	Marital Status	-.066	.198	.111	1	.739	.936	.635	1.380
	Weight	-.016	.016	1.030	1	.310	.984	.953	1.015
	Gender	.492	.507	.941	1	.332	1.635	.605	4.417
	Type of TB	1.337	.465	8.246	1	.004	3.806	1.529	9.479
	ART Status	.083	.484	.029	1	.864	1.086	.420	2.807
	Hemoglob	.003	.008	.168	1	.682	1.003	.987	1.020
	Creatinine	.007	.002	7.623	1	.006	1.007	1.002	1.011
	CD4Count	-.002	.001	1.423	1	.233	.998	.996	1.001
	Glucose	.000	.003	.000	1	.997	1.000	.994	1.006
Defaulted	Intercept	-9.342	2.742	11.611	1	.001			
	Age	.039	.030	1.737	1	.187	1.040	.981	1.103
	Marital Status	.160	.238	.449	1	.503	1.173	.735	1.872
	Weight	.017	.023	.577	1	.448	1.018	.973	1.065
	Gender	-.208	.639	.106	1	.745	.812	.232	2.841
	Type of TB	2.710	.620	19.098	1	.000	15.032	4.458	50.688
	ART Status	.383	.650	.347	1	.556	1.467	.410	5.242
	Hemoglob	-.007	.011	.349	1	.555	.993	.972	1.015
	Creatinine	.006	.003	4.236	1	.040	1.006	1.000	1.012
	CD4Count	-.004	.003	1.908	1	.167	.996	.991	1.002
	Glucose	-.021	.007	10.242	1	.001	.979	.967	.992

a. The reference category is: Completed and Cure

Table 6: Model summary for ANN.

Model Summary		
Training	Sum of Squares Error	22.767
	Percentage of Incorrect Predictions	7.9%
	Stopping Rule Used	1 consecutive step(s) with no decrease in error[a]
	Training Time	0:00:00.35
Testing	Sum of Squares Error	6.197
	Percentage of Incorrect Predictions	5.6%

a. Error computations are based on the testing sample.

Table 7: Normalized importance in percentages.

	Importance	Normalized Importance
Age	.056	17.0%
Gender	.036	10.9%
Marital Status	.014	4.2%
Type of TB	.121	36.9%
ART Status	.012	3.6%
Hemoglob	.023	7.1%
Creatinine	.327	100.0%
CD4Count	.083	25.5%
Glucose	.266	81.4%
Weight	.062	19.0%

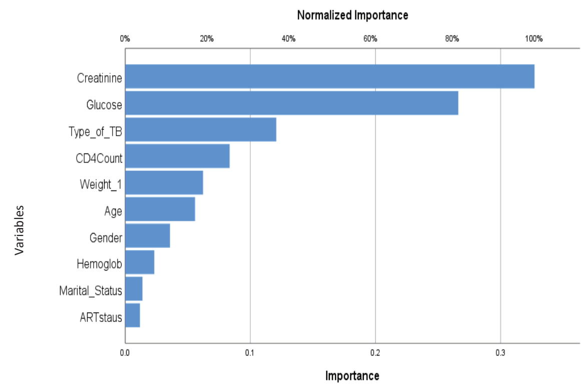

Figure 3: Normalized importance.

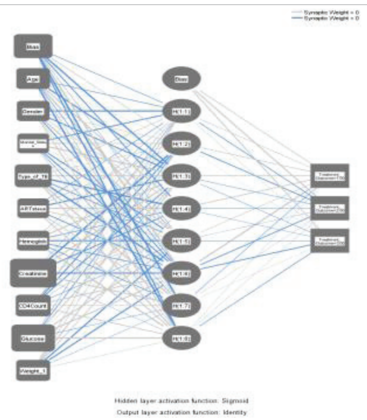

Figure 4: Network relationship between the dependent and independent variable.

DOI: https://dx.doi.org/10.37871/ajeph.id52

 ISSN: 2644-0032

Completed and Cured that are correctly classified as Completed and Cure) and Specificity (the proportion of true negative i.e., the proportion of not Completed and Cured that are correctly classified as not Completed and Cured).

The ROC can be better explained by the Area Values under the Curve (AUC) covered by each group under treatment outcome as given in table 8. Like we noted earlier in this paper, a model whose predictions are 100% wrong has an AUC of 0.0 and one whose predictions are 100% correct has an AUC of 1.0 as shown in figure 5. The AUC values of 0.817 for Completed and cures is considered acceptable, and this is similar for the rest group. The higher the AUC, the better the performance of the model at classifying variable.

Comparison between MLR and ANN classification: Finally, here we can objective view the classification task done by the Multinomial Logistic Regression and the Multilayer perceptron of the Artificial Neural Network.

Multinomial logistic regression classification: Percentage of Correct Classification which is used to determine which category were best predicted by the model. Completed and Cured was correctly predicted by the model 99.3% of the time i.e. [429 of the 432 patients who are Completed and Cured were predicted to do so by the model: $\frac{429}{429+1+2} = 0.993$. This is 99.3% correct prediction rate.

The model particularly did not very good job at predicting DIED i.e. [1out of 20 patients who died was predicted correctly by the model as $\frac{1}{20} = 0.05$. This is (5%) correct prediction rate]. Same goes for Defaulted i.e. [2 our of 14 patients who defaulted was predicted correctly by the model as $\frac{2}{14} = 0.1428$. This is about 14.3 correct prediction rate]. Generally, the overall correct prediction percentage of 92.7% was excellent.

Artificial neural network (multilayer perceptron) classification: Here we presented the classification output of the ANN (Multilayer Perceptron). Table 9 shows the classification statistics used to determine which group memberships were best predicted by the model. COMPLETED AND CURED was correctly predicted by the model 100% of the time both in training and Testing but the Model was unable to classify both the DIED and DEFAULTED insufficient data, however it achieved 94.4% over all prediction accuracy which is an improvement over the Multinomial Logistic regression (92.7%) which was excellent.

DISCUSSION AND CONCLUSION

In this paper, we find out that the observed TB/HIV preventive treatment regimens do not differ significantly from the expected value. That is the effect of the preventive therapy regimen, and the main disease treatment regimen are not statistically different. We further move into the classification task by treating the Completed and Cured as the referent group and therefore estimated a model for Died relative to Completed and Cured and a model for Death relative to Completed and Cured in Multinomial Logistics Regression analysis reported our findings in Table 9. Therefore, since the parameter estimates are relative to the referent group, the standard interpretation of the Multinomial Logistic Regression model is that for a unit change in the predictor variable, the logit of outcome of the likelihood ratio test (Table 9) relative to the referent group is expected to change by its respective parameter estimate (which is in log-odds units) given the variables in the model are held constant. For ANN, the independent Variable Importance shows the contribution of each independent variable to predicting the Outcome (dependent). The Normalized importance shows that creatinine level (100%) followed

Table 8: Values of Area Under the Curve (AUC) for treatment outcomes.

Area Under the Curve (AUC)		Area
Treatment Outcome	Completed and Cured	0.817
	Died	0.742
	Defaulted	0.896

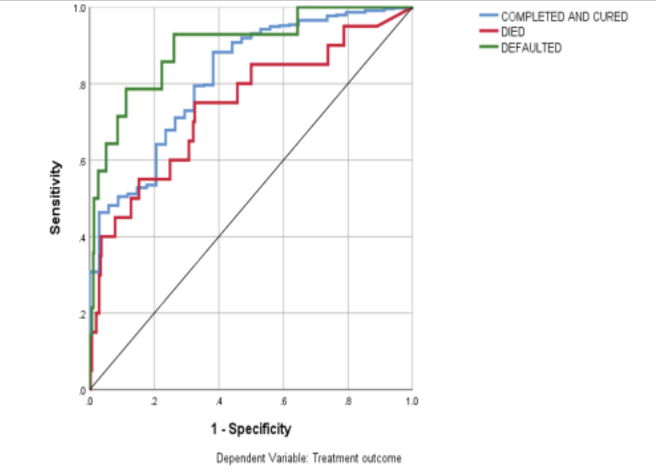

Figure 5: Relative operating characteristics curve.

Table 9: Multinomial logistics regression and ANN classification result.

Multinomial Logistics Regression (MLR) classification result

Observed	Predicted			
	Completed and Cured	Died	Defaulted	Percentage of Correct Classification
Completed and Cured	429	1	2	99.3%
Died	19	1	0	5.0%
Defaulted	12	0	2	14.3%
Overall Percentage	98.7%	0.4%	0.9%	92.7%

ANN'S Multilayer Perceptron Classification result

Sample	Observed	Predicted			
		Completed and Cured	Died	Defaulted	Percentage Of Correct Classification
Training	Completed and Cured	315	0	0	100.0%
	Died	18	0	0	0.0%
	Defaulted	9	0	0	0.0%
	Overall Percent	100.0%	0.0%	0.0%	92.1%
Testing	Completed and Cured	117	0	0	100.0%
	Died	2	0	0	0.0%
	Defaulted	5	0	0	0.0%
	Overall Percent	100.0%	0.0%	0.0%	94.4%

Dependent variable: Treatment outcome.s

DOI: https://dx.doi.org/10.37871/ajeph.id52

ISSN: 2644-0032

by Glucose level (81.4%), Type of Tuberculosis (36.9%) and CD4 count (25.5%) has major contribution to predicting the output of the dependent variables.

We observe the following in Multinomial Logistic regression Classification, First, Completed and Cured were correctly predicted by the model 99.3% of the time. Secondly, the model did not do a very good job at predicting Died and Defaulted (at a rate of 5.0% and 14.3% respectively) but the overall prediction percentage of 92.7% was better. For ANN Completed and Cured was correctly predicted by the model 100% of the time both in training and Testing but the Model was unable to classify the Died and Defaulted, however, it achieved a 94.4% overall prediction ability.

CONCLUSION

This paper is based on many participants from Lagos residents in Nigeria, where the prevalence of TB infection and HIV are very high. In this study, the MLR model and the ANN (MLP) model have been compared using TB/HIV co-infected data. The TB/HIV preventive treatment therapies and the main diseases treatment regimen was examined too, and it shows there is not enough statistical evidence to show they differ.

After the classification task done by both the Multinomial Logistic Regression and the Artificial Neural Network (MLP), it was observed that Multilayer Perceptron of ANN perform better with an overall 94.4 % correct classification than the Multinomial Logistic Regression with an overall 92.7% correct classification.

RECOMMENDATIONS

In the light of the analysis in this paper, The Artificial Neural Network (ANN) model have a more realistic interpretation and provides more informative results as compared to MLR model for the available data.

Therefore,

a) We suggest that using the Multinomial Logistic Regression model may not be the optimum approach for TB/HIV patient data. The Artificial Neural Network provide an alternative method to fit such data.

b) Determining the effect of the treatment regimens may be additional values to research.

The results from this model could then be compared in the future. In addition, further study can be carried out to evaluate the effects of practical cases such as large censoring.

REFERENCES

1. Ogungbola OO, Akomolafe AA, Musa AZ. Accelerated failure time model with application to data on tuberculosis/HIV co-infected patients in Nigeria. American J Epidemiol Public Health. 2018;2(1):021-026. https://bit.ly/3NMGVP7

2. Caliskan E, Sevin Y. A comparative study of artificial neural networks and multiple regression analysis for modelling skidding time. Applied Ecology and Environmental Research Journal. 2018;17(2):1741-1756. https://bit.ly/3uQEOkw

3. Benito U, Hanns F, Claudio E, Marisela F. Multinomial logistic regression to estimate and predict the perceptions of individuals and companies in the face of the COVID-19 pandemic in the Ñuble region, Chile. Sustainability. 2020;12: 9553. doi: 10.3390/su12229553.

4. Elif B. Using neural network and logistic regression analysis to predict prospective mathematics teachers academic success upon entering graduate education. Education Sciences: Theory and Practice. 2016;16:943-964. doi: 10.12738/estp.2016.3.0214.

5. Gao Q. Multinomial logistic regression and group-based trajectory modeling for longitudinal data of contraceptive methods and recognition of abusive behaviors among women seeking family planning clinical care. Thesis submitted to the Department of Biostatistics Graduate School of Public Health. 2019. https://bit.ly/3jbosxs

6. Erkan A, Noyan A. Examination by multinomial logistic regression model of the factors affecting the types of domestic violence against women: A case of Turkey. International Journal of Scientific & Technology Research. 2016;5(11):67-74. https://bit.ly/3DK7NKJ

7. Erkan A. Using multinomial logistic regression to examine the relationship between children's work status and demographic characteristics. Research Journal of Politics, Economics and Management. 2016;4(1):77-93. https://bit.ly/3KiHm1a

8. Asogwa OC, Oladugba AV. On the comparison of Artificial Neural Network (ANN) and Multinomial Logistic Regression (MLR). West African Journal of Industrial and Academic Research. 2015;13(1):3-9. https://bit.ly/3DIXuqi

9. Abdalla ME. An application on multinomial logistic regression model. Pakistan Journal of Statistics and Operation Research. 2012;8(2):271-291. doi: 10.18187/pjsor.v8i2.234.

10. Freese J, Long JS. Regression models for categorical dependent variables using stata. College Station: Stata Press; 2006.

11. Hosmer DN, Lemeshow S. Applied logistic regression. New York: Wiley; 1989.

12. Hosmer DN, Lemeshow S. Applied logistic regression. 2nd Ed. New York: Wiley; 2000.

13. Wilson RL, Sharda R. Bankruptcy prediction using neural networks. Decision Support Systems. 1994;11:545-557. doi: 10.1016/0167-9236(94)90024-8

14. Kleinbaum DG, Klein M. Logistic regression: A self-learning text. 3rd Ed. Statistics in the health sciences. New York: Springer; 2010. https://bit.ly/3JcvCvN

15. Haykin SS. Neural networks: A comprehensive foundation, New Jersey, USA: Prentice Hall; 1999.

DOI: https://dx.doi.org/10.37871/ajeph.id52