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

Samuel Ominyi Ejeh*, Olatayo Oluwasegun Alabi¹, Opeyemi Oyekola Ogungbola¹, Olawumi Olumayowa Olatunde² and Zainab Olabisi Dere³

¹Department of Statistics, Federal University of Technology Akure, Akure, Ondo State, Nigeria
²National Bureau of Statistics, Plot 762, Independence Avenue, Central Business District, Abuja, Nigeria
³Department 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

Copyright: © 2022 Ejeh SO, et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.
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 for TB/HIV co-infected patients 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.

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.

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 and recurrent 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.

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

\[
x_j \beta_j = \sum_{k=0}^{K} \beta_{jk} x_{jk}
\]

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:

\[
P_{i1} = P(y_i = 1|x_i) = \frac{1}{1 + \exp(x_i \beta_2) + \exp(x_i \beta_3)}
\]

\[
P_{i2} = P(y_i = 2|x_i) = \frac{\exp(x_i \beta_2)}{1 + \exp(x_i \beta_2) + \exp(x_i \beta_3)}
\]

\[
P_{i3} = P(y_i = 3|x_i) = \frac{\exp(x_i \beta_3)}{1 + \exp(x_i \beta_2) + \exp(x_i \beta_3)}
\]

In here, \(\beta_2\) and \(\beta_3\) 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.

\[
P_{i1} = 1 - (P_{i2} + P_{i3})
\]

The sum of the probabilities of categories of dependent variable
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]:

\[ P_{y1} = P(x_1 = 1|y_1) + P(x_2 = 2|y_1) + P(x_3 = 3|y_1) = 1 \]

(6)

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

\[ P_{ij} = P(y_i = j|x_i) = \frac{\exp(x_i^\prime \beta_j)}{1 + \sum_{j=1}^{J-1} \exp(x_i^\prime \beta_j)} \]

(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)

\[ P_{iJ} = P(y_i = J|x_i) = \frac{1}{1 + \sum_{j=1}^{J-1} \exp(x_i^\prime \beta_j)} \quad j = 1, 2, ..., J-1 \]

(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].

\[ P_{ij} = P(y = J) = 1 - \{P(y = 1) + P(y = 2) + \ldots + P(y = J - 1)\} \]

(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].

\[ \ln \left[ \frac{P(y = 1|x_1)}{P(y = 0|x_1)} \right] = \beta_1 + \beta_1 x_1 \]

(10)

\[ \ln \left[ \frac{P(y = 1|x_2)}{P(y = 0|x_1)} \right] = \beta_2 + \beta_2 x_1 \]

(11)

\[ \ln \left[ \frac{P(y = 1|x_3)}{P(y = 0|x_1)} \right] = \beta_3 + \beta_3 x_1 \]

(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].

\[ \ln \left[ \frac{P(y = j)}{P(y = J)} \right] = \ln \left[ \frac{P(y = j)}{P(y = J)} \right] = \left( \sum_{k=1}^{K} \beta_{jk} x_k \right) j = 1, \ldots, J-1 \]

(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

\[ 1 + \sum_{j=1}^{J-1} \exp(x_i^\prime \beta_j) \]

(14)

\[ \ln L = \sum_{i=1}^{n} \sum_{j=1}^{J} d_{ij} \log(P_{ij}) \]

(15)

Where \( d_{ij} \) is a dummy variable that takes a value 1 if observation \( i \) takes the \( j \)th 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
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 R² 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 \( p = 0.996 > \alpha = 0.05 \), 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 R² 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 R² value, 16.5% according to Nagelkerke R² values and 10.3% according to McFadden R² value.

**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 updated 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...
Table 5: Parameter estimation.

<table>
<thead>
<tr>
<th>Treatment Outcome</th>
<th>B</th>
<th>Std. Error*</th>
<th>Wald</th>
<th>df</th>
<th>Sig</th>
<th>Exp(B)</th>
<th>95% confidence interval for Exp(B)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Lower bound</td>
</tr>
<tr>
<td>Died</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-6.404</td>
<td>1.979</td>
<td>10.476</td>
<td>1</td>
<td>.001</td>
<td>1.032</td>
<td>.985</td>
</tr>
<tr>
<td>Age</td>
<td>.031</td>
<td>.034</td>
<td>0.111</td>
<td>1</td>
<td>.739</td>
<td>.963</td>
<td>.635</td>
</tr>
<tr>
<td>Marital Status</td>
<td>-.066</td>
<td>.198</td>
<td>0.111</td>
<td>1</td>
<td>.739</td>
<td>.963</td>
<td>.635</td>
</tr>
<tr>
<td>Weight</td>
<td>-.016</td>
<td>.016</td>
<td>1.030</td>
<td>1</td>
<td>.310</td>
<td>.984</td>
<td>.953</td>
</tr>
<tr>
<td>Gender</td>
<td>.492</td>
<td>.507</td>
<td>1.268</td>
<td>1</td>
<td>.268</td>
<td>1.003</td>
<td>.967</td>
</tr>
<tr>
<td>Type of TB</td>
<td>1.337</td>
<td>.465</td>
<td>8.246</td>
<td>1</td>
<td>.004</td>
<td>3.806</td>
<td>1.529</td>
</tr>
<tr>
<td>ART Status</td>
<td>.083</td>
<td>.029</td>
<td>1.748</td>
<td>1</td>
<td>.186</td>
<td>1.086</td>
<td>.420</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>.003</td>
<td>.008</td>
<td>0.168</td>
<td>1</td>
<td>.704</td>
<td>.990</td>
<td>.923</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.007</td>
<td>.002</td>
<td>7.623</td>
<td>1</td>
<td>.007</td>
<td>1.007</td>
<td>1.002</td>
</tr>
<tr>
<td>CD4Count</td>
<td>-.002</td>
<td>.001</td>
<td>1.423</td>
<td>1</td>
<td>.233</td>
<td>.998</td>
<td>.996</td>
</tr>
<tr>
<td>Glucose</td>
<td>.000</td>
<td>.003</td>
<td>0.000</td>
<td>1</td>
<td>.997</td>
<td>1.000</td>
<td>.994</td>
</tr>
<tr>
<td>Defaulted</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Intercept</td>
<td>-9.342</td>
<td>2.742</td>
<td>11.611</td>
<td>1</td>
<td>.001</td>
<td>1.040</td>
<td>.981</td>
</tr>
<tr>
<td>Age</td>
<td>.039</td>
<td>.030</td>
<td>1.737</td>
<td>1</td>
<td>.187</td>
<td>1.040</td>
<td>.981</td>
</tr>
<tr>
<td>Marital Status</td>
<td>.160</td>
<td>.238</td>
<td>4.499</td>
<td>1</td>
<td>.036</td>
<td>1.173</td>
<td>.735</td>
</tr>
<tr>
<td>Weight</td>
<td>.017</td>
<td>.023</td>
<td>0.577</td>
<td>1</td>
<td>.448</td>
<td>1.018</td>
<td>.973</td>
</tr>
<tr>
<td>Gender</td>
<td>-.208</td>
<td>.639</td>
<td>1.066</td>
<td>1</td>
<td>.745</td>
<td>.812</td>
<td>.232</td>
</tr>
<tr>
<td>Type of TB</td>
<td>2.710</td>
<td>.200</td>
<td>19.098</td>
<td>1</td>
<td>.000</td>
<td>15.032</td>
<td>4.458</td>
</tr>
<tr>
<td>ART Status</td>
<td>.383</td>
<td>.150</td>
<td>.347</td>
<td>1</td>
<td>.556</td>
<td>1.467</td>
<td>.410</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>-.007</td>
<td>.011</td>
<td>.349</td>
<td>1</td>
<td>.555</td>
<td>.993</td>
<td>.972</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.006</td>
<td>.003</td>
<td>4.236</td>
<td>1</td>
<td>.040</td>
<td>1.006</td>
<td>1.000</td>
</tr>
<tr>
<td>CD4Count</td>
<td>-.004</td>
<td>.003</td>
<td>1.908</td>
<td>1</td>
<td>.167</td>
<td>.996</td>
<td>.991</td>
</tr>
<tr>
<td>Glucose</td>
<td>-.021</td>
<td>.007</td>
<td>10.242</td>
<td>1</td>
<td>.001</td>
<td>.979</td>
<td>.967</td>
</tr>
</tbody>
</table>

a. The reference category is: Completed and Cure

Table 6: Model summary for ANN.

<table>
<thead>
<tr>
<th></th>
<th>Sum of Squares</th>
<th>Percentage of Incorrect Predictions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>22.767</td>
<td>7.9%</td>
</tr>
<tr>
<td>Stopping Rule Used</td>
<td>1 consecutive step(s) with no decrease in error*</td>
<td></td>
</tr>
<tr>
<td>Testing</td>
<td>6.197</td>
<td>5.6%</td>
</tr>
</tbody>
</table>

a. Error computations are based on the testing sample.

Table 7: Normalized importance in percentages.

<table>
<thead>
<tr>
<th>Independent Variable Importance</th>
<th>Importance</th>
<th>Normalized Importance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age</td>
<td>.056</td>
<td>17.0%</td>
</tr>
<tr>
<td>Gender</td>
<td>.036</td>
<td>10.9%</td>
</tr>
<tr>
<td>Marital Status</td>
<td>.014</td>
<td>4.2%</td>
</tr>
<tr>
<td>Type of TB</td>
<td>.121</td>
<td>36.9%</td>
</tr>
<tr>
<td>ART Status</td>
<td>.012</td>
<td>3.6%</td>
</tr>
<tr>
<td>Hemoglobin</td>
<td>.023</td>
<td>7.1%</td>
</tr>
<tr>
<td>Creatinine</td>
<td>.327</td>
<td>100.0%</td>
</tr>
<tr>
<td>CD4Count</td>
<td>.083</td>
<td>25.5%</td>
</tr>
<tr>
<td>Glucose</td>
<td>.266</td>
<td>81.4%</td>
</tr>
<tr>
<td>Weight</td>
<td>.062</td>
<td>19.0%</td>
</tr>
</tbody>
</table>
Completed and Cured that are correctly classified as Completed and Cured) 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 Cured 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: 
\[
\frac{429}{429 + 1 + 2} = 0.993. \text{ This is 99.3% correct prediction rate.}
\]

The model particularly did not very good job at predicting DIED i.e. [1 out 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 COMPLETED AND CURED was correctly predicted by the model 99.3% of the time i.e. [429 of the 432 patients 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 were best predicted by the model.

Percentage 98.7% 0.4% 0.9% 92.7%

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 Logistic 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>
<thead>
<tr>
<th>Area Under the Curve (AUC)</th>
<th>Area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Completed and Cured</td>
<td>0.817</td>
</tr>
<tr>
<td>Died</td>
<td>0.742</td>
</tr>
<tr>
<td>Defaulted</td>
<td>0.896</td>
</tr>
</tbody>
</table>

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

<table>
<thead>
<tr>
<th>Dependent Variable: Treatment outcome.s</th>
<th>Predicted</th>
<th>Percentage Of Correct Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Observed</td>
<td>Completed and Cured</td>
<td>Died</td>
</tr>
<tr>
<td>Training</td>
<td>315</td>
<td>0</td>
</tr>
<tr>
<td>Died</td>
<td>18</td>
<td>0</td>
</tr>
<tr>
<td>Defaulted</td>
<td>9</td>
<td>0</td>
</tr>
<tr>
<td>Overall Percent</td>
<td>100.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Table 9: Multinomial logistics regression and ANN classification result.

<table>
<thead>
<tr>
<th>Sample</th>
<th>Observed</th>
<th>Predicted</th>
<th>Percentage Of Correct Classification</th>
</tr>
</thead>
<tbody>
<tr>
<td>Training</td>
<td>Completed and Cured</td>
<td>117</td>
<td>0</td>
</tr>
<tr>
<td>Died</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Defaulted</td>
<td>5</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Overall Percent</td>
<td>100.0%</td>
<td>0.0%</td>
<td>0.0%</td>
</tr>
</tbody>
</table>

Table 10: Multilayer Perceptron Classification result.
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


