Multilayer Perceptron Neural Network Classification of HIV Status of Children

T. O. James[1]; S. U. Gulumbe[2]; A. Danbaba[3]; M. T. O. Ibrahim[4]

[1,2,5]Department of Mathematics,
Kebbi State University of Science and Technology, Aliero, Nigeria.
e-mail:tolujam@gmail.com[1]whafsat@gmail.com[2]belloml2015@gmail.com[5]

[2,3]Department of Mathematics,
Usmanu Danfodiyo University,
Sokoto, Nigeria.
e-mail: ugulumbe@gmail.com[2]; danbabako@gmail.com[3]

[4]Al-Hikmah University,
Ilorin, Nigeria.
e-mail: Ibrahimtaofeekk@yahoo.com

***Abstract*—Classification plays an essential task in medical decision, especially in HIV status. The classification model used in medical diagnosis is of paramount importance, it becomes a major problem if a classifier records a higher percentage of misclassification. Classifying HIV status wrongly can cause anxiety; further cost of testing and even untimely death. Therefore, accurate classification of HIV is very imperative. In this paper, a multilayer perceptron neural network with a logistic function is trained using resilient back-propagation with back-tracking algorithm and a cross entropy error function is used in classification of Child HIV status. Three artificial neural network based models were developed, the trained artificial neural networks are (4 – 5 – 1), (4 – 4 – 1) and (4 – 3 – 1). The models were trained using data obtained from antenatal clinic (ANC). Performance of the models in the held-out sample were assessed using the Akaike Information Criterion (AIC), Hannan-Quinn Information Criterion (HQIC), Schwartz Information Criterion(SIC). The results obtained show that the neural network (4-3-1) model at decay constant set at λ=0.001 was the best neural network model for classifying HIV status of children with an accuracy of 94.3%.The classifiers' performance was built on the confusion matrix and the sensitivity and specificity of this neural network models was 96% and 94%.The ROC curve and the Area under the ROC curve (AROC) was also utilized, the area lies between 0.7 to 0.9, which indicate that the model shows accuracy in classifying between HIV positive children and HIV negative Children.**

Keywords-*Artificial neural network, Classification, child, HIV status.*

.

I. Introduction

Classification and prediction of medical datasets poses real challenges in Medical Data Mining [1] but there are so many traditional statistical classification techniques like discriminant analysis, logistic regression, decision trees, and neural networks, etc. Despite, the availability of several classification methods most of them does not offer adequate results, but studies available in recent years have reported that the artificial neural network approach supersede them and is an alternative tool for classification even in most of the medical problems [2][3]. The following properties buttress this. They do not require any prior assumptions regarding the distribution of the training data, nonlinear, parallel processing and their adaptability and learning process [4][5][6].

Classification is a two-step process says Han & Kamber [7] which are learning and classification. They opined that classifier is built describing a predetermined set of classes in learning. The classification algorithm builds the classifier by analyzing a training set and their associated class labels. The learned model or classifier is represented in the form of classification rules. Manaswini and Ranjit [8] reveal that in medical applications, classification is desired to differentiate a pattern of low frequency from a pattern of high frequency. While, Zhang [9] says the goal of classification is to assign an object to a predefined group or class, based on observed attributes related to the object.

HIV/AIDS continues to be a major global health priority. Globally, 3.2 million children under 15 were living with HIV in 2013, comprising 9.1% of all people living with

HIV. From the estimated 3.2 million children living with HIV, 91% live in sub-Saharan Africa, 6% live in Asia and the Pacific and the remaining 3% are situated in the rest of the world. 240 000 children worldwide acquired HIV in 2013: one new infection every two minutes [10]. An estimated 1.5 million of the 115 million annual births in low- and middle-income countries are born to HIV-infected mothers [11]. Therefore, this paper describes the application of neural networks in classification of HIV status of children.

Neural Network

Neural networks are powerful mathematical models suitable for almost all data mining tasks, with special emphasis on classification and estimation problems [12]. It is inspired by the human brain and its ability to adapt on the basis of the inflow of new information and it is a massively parallel processor made up of simple processing units, which has a natural propensity for storing experimental knowledge, it can develop nonlinear statistical models to deal with complex biological systems and making it available for use and very useful in neural classification [13][14][2][6][15].

Related Work

Some researchers have shown applicability of Multi Layered Perceptron (MLP) on classification. Some worked on the classification of the health status of HIV/AIDS patients while some on classification of AIDS versus HIV status patients and classification of HIV/AIDS infected and non-infected status of individuals utilizing different training algorithms and concluded that the MLP network trained using backpropagation algorithm produced the best performance of all [16][17][8]. Many more researchers have utilized the benefit of Artificial Neural Network knowledge in clinical outcome of classification, prediction and even prediction of prevalence of HIV/AIDS [18][19][20][21][22].

II. Research Methodology

In this study, the model was built by the multilayered perceptron (MLP) neural network using resilient back propagation (BP) algorithm, initially the activation of the neuron is computed as the weighted sum of it inputs and the weights multiply the normalized input information. Where input is denoted by Xi, and each weight wi, then the activation is equal to

$$\sum x_i w_{ij} \qquad (1)$$

And next this activation is transformed into a response by using a transfer function. In this study, a log-sigmoid activation function was used because of the nonlinear nature of the problem; it is given by

$$f(x) = \frac{1}{1+e^{-x}} \qquad (2)$$

The model developed to classify a HIV/AIDS status of children involves the three-layer back-propagation algorithm as shown in figure 1.0 modified from Khan, et al [23]. An input layer is used to represent a set of input variables. The dataset was divided in to two set, the first set of 86 cases (70% of data) for training set and 37 cases (30% of data) for testing the model and the study used the training sample to build the models. The choice of using this breakdown is to allow enough data for the training phase so that the trained neural network models could generalize best on the out-of-sample data.

The predictor was Mother's CD4 count (MCD4), Delivery mode and ART Drug which was applied as input data to the MLP network, it distributes the input to the hidden layer. Three hidden layer was used in this study. The outputs from the hidden layer then become the inputs to the output layer, which provides the network output. The decay or penalty parameters were varied at the values 0.001, 0.01 and 0.05 for each neural network model trained. Epochs for training set are repeated 200 times as a learning rate. In order to avoid the over fitting the network, the learning process was stopped when the total number of epoch repeats reached 300.

The trained resilient back propagation is then adopted as a model to classify the children as either HIV positive or negative. The neural network module in R was used for training the networks.

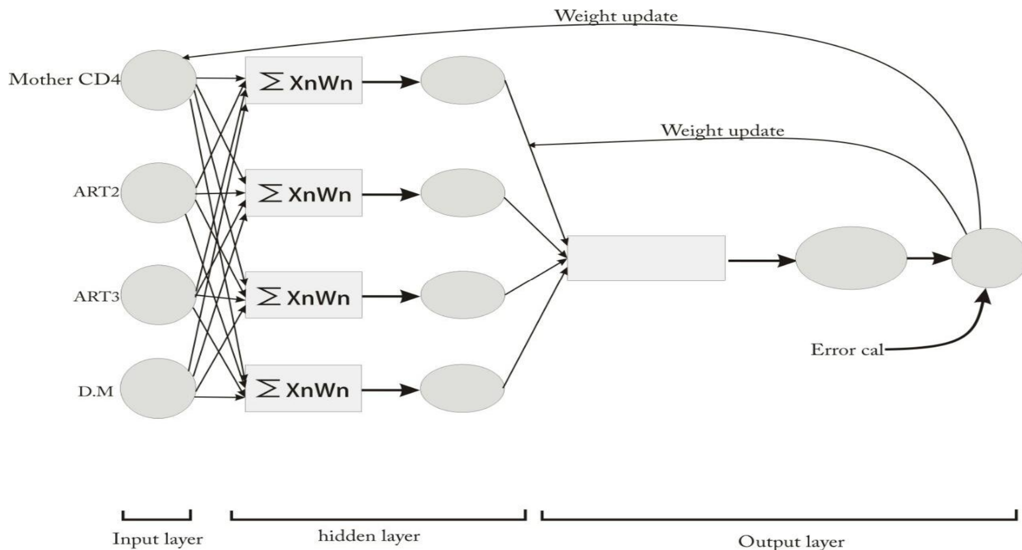

Figure 1: Schematic Training phase of the model

Model Evaluation Criteria

The performance of the resilient back-propagation neural network is evaluated by means of three statistics: the Akaike Information Criterion (AIC), mean Hannan-Quinn Information Criterion (HQIC), and Schwartz Information Criterion (SIC). They are expressed as follows:

Akaike Information Criterion (AIC)

$$AIC(k) = \log_e\left(\frac{SSE}{N}\right) + \frac{2k}{N} \quad (3)$$

Hannan-Quinn Information Criterion:

$$HQIC(k) = \log_e\left(\frac{SSE}{N}\right) + \frac{k}{N}\log_e(\log_e(N)) \quad (4)$$

Schwartz Information Criterion:

$$SIC(k) = \log_e\left(\frac{SSE}{N}\right) + \frac{k}{N}\log_e(N) \quad (5)$$

where K is the number of parameters, n is the number of observation and SSE is the sum of squares error due to the network.

Confusion Matrix

The confusion matrix was used for the output of series of experiments of classification models.

Correctly Classified Instances (Accuracy): The proportion of children that are correctly classified.

$$\text{Error Rate} = \frac{TP+TN}{TN+FP+FN+TP} \quad (6)$$

Incorrectly Classified Instances(Error Rate): The proportion of children that are incorrectly classified.

$$\text{Error Rate} = \frac{FP+TN}{TN+FP+FN+TP} \quad (7)$$

Incorrectly Classified Instances(Error Rate): The proportion of children that are incorrectly classified.

Proportion of correct classifications in percentage= $\frac{TP+TN}{TN+FP+FN+TP} \times 100\%$ (8)

$$Se = \frac{TP}{TP+FN} \quad (9)$$

$$Sp = \frac{TN}{TN+FP} \quad (10)$$

III. Results And Discussion

Table 1 presents the summary of the best fitted network models for each value of the decay λ=0.001,0.01, 0.05.

Table 1: Summary of the best fitted NETWORK 4-3-1 neural networks using MLP

Best fitted model	Decay(λ)	No of hidden nodes	F. entropy	AIC	HQIC	SIC
Fit 1	0.001	3	12.90545	-2.652222	-2.764039	-2.109983
Fit 2	0.01	3	16.10535	-2.475667	-2.587484	-1.933427
Fit 3	0.05	3	19.96646	-2.353322	-2.465139	-1.811082

Table 2: Confusion matrix for classification of Child HIV status

		Predicted Values		***Total***
		Negative	Positive	
Actual HIV Status	Negative	(TN) 68	(FP) 4	72
	Positive	(FN)2	(TP)48	50
Total		70	52	122

Key: TN = True Negative FP= False Positive FN= False Positive TP= True Positive

TN: The number of HIV negative child that are classified as negative.

FN: The number of HIV positive child that are classified as negative.

TP: The number of HIV positive child that are classified as positive.

FP: The number of HIV negative child that are classified as positive.

The confusion matrix in Table 2 shows that, out of the total 72 actually HIV negative children, only 68 children are classified as HIV negatives and the rest are misclassified as HIV positive. And out of the total 50 actually HIV positive children, only 48 children are classified as HIV positive and the rest are misclassified as HIV negative. This means the model has better performance in terms of correctly classifying HIV positives than HIV negatives.

Table 3: Results from classification of the NN model – Sensitivity, Specificity, Accuracy, Error Rate, False positive and False Negative

NN model	%
Sensitivity	96
Specificity	94
Accuracy	95
Error rate	5
False Positive	3
False Negative	7

Performance of Chosen Model

In order to check the performance of the chosen model, data were simulated and MSE for both ANN and logistic regression were generated to compare ANN performance to that of conventional logistic regression models. Table 4 summarizes the result that shows the MSE of both models. Artificial Neural Network (ANN) outperformed the logistic model and was the best models with the lowest mean square error (MSE).

Table: 4: The MSE of both ANN and Logistic regression

ANN	Logistic
0.0016	0.0305
0.0101	0.0239
0.0396	0.0400
0.0200	0.0298
0.0237	0.0282
0.0284	0.0317
0.0311	0.0371
0.0289	0.0347

Receiver Operating Characteristics Curve for HIV Status

ROC curves are a useful tool for comparing classification models and also show the trade-off between the TP rate and the FP rate for a given model [21]. This is done by plotting the true positive rate (sensitivity) against the false positive rate (1-specificity) at different cut-off points. And, the closer the ROC plot is to the upper left corner, the higher the overall accuracy of the models. From Figure 1.2 the area lies between 0.7 and 0.9 which indicates an excellent model.

ROC curve for Comparison of Logistic Regression (LR) and Artificial Neural Networks (ANN)

The area under ROC curve in Figure 1.3 was utilized in this study to compare logistic regression and neural networks.

The area under ROC curve was 0.98 and 1.0 respectively. It showed that the area was not significantly different in the two models.

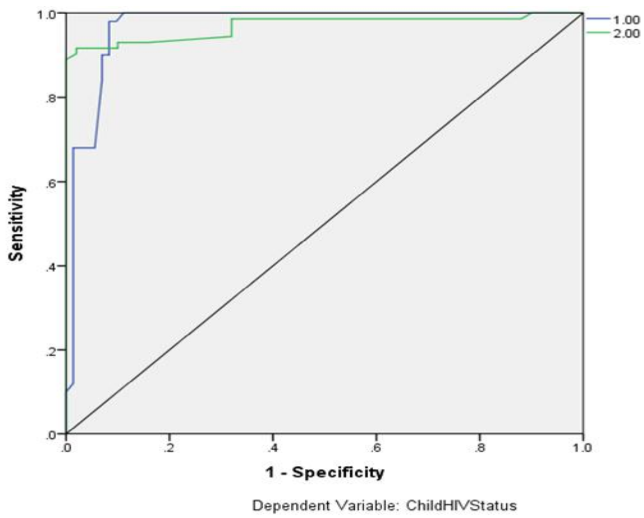

Figure 1.2: The Roc Plot of Child HIV Status

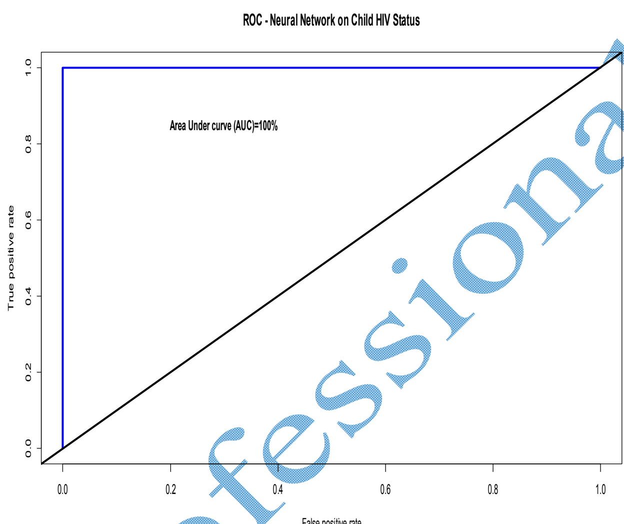

Figure 1.3: Receiver-operating characteristic (ROC) curves for Logistic regression (LR) and Artificial neural networks (ANN).

Discussion
Three different models were fitted using three hidden neurons and classification was done for both the training set (In-Sample) and testing/validation set (Out-sample). The decay or penalty parameters were varied at the values 0.001, 0.01 and 0.05 for each neural network model trained. They are NETWORK 4-5-1, NETWORK 4-4-1 and NETWORK 4-3-1. Based on the simultaneous agreement of the three information criteria, AIC (-2.652222), HQIC (-2.764039) and SIC (-2.109983) in Table 1.0. The network model NETWORK 4-3-1 was selected as the best architecture by MLP NETWORK model for further study.

The output of series of experiments of classification models were analyzed and evaluated in terms of the details of the confusion matrix of the model in Table 1.1. The accuracy, sensitivities, specificity, error rate, false positive and false negative of the neural network architecture for the training and testing dataset of this study were 95%, 96%, 94%, 5%, 3%, and 7% respectively as seen in table 1.3. The classification accuracy shows that 95% of the proportions of HIV status cases are correctly classified; sensitivity reflects that 96% of positive HIV cases are classified as positive and specificity measures 94% of HIV negative cases are classified as negative, while the error rate shows that 5% were misclassified and since the lower, the error rates the better the procedure therefore the classification procedure is good. [24]

The result shows that the MLPNN model shows accuracy in classifying between HIV positive children and HIV negative Children. The performances of the chosen models was checked by comparing the MSE of both ANNs and LRMs in Table 1.1, the performance of the best models was chosen based on the lowest value of mean square error and Neural network gave the best performance. This study for classification of HIV showed that the area under ROC curve in figure 1.2 lies within 0.7 to 0.9 which indicate excellent and very good model. They were also tested using receiver operating characteristic (ROC) curve, in logistic regression and neural networks the area under ROC curve was 0.98 and 1.0 respectively has shown in figure 1.3 and then logistic regression and neural networks was similar in classification subjects but the neural network perform better than the conventional logistic regression.

IV. Conclusion

In this paper, multilayer perceptron networks models were used for classification of HIV/AIDS data to distinguish between HIV positive and negative Children. The classification accuracy shows that 95% of the proportions of HIV status cases are correctly classified, while from the Confusion Matrix, sensitivity reflects that 96% of positive HIV cases are classified as positive and specificity measures 94% of HIV negative cases are classified as negative This model can be used as an alternative for classification problems in HIV status of children. Hospitals and various agencies working on prevention of AIDS/HIV should use this type of research outputs in classifying and predicting the HIV status of children.

Acknowledgment

The authors are grateful to anonymous reviewers for their valuable comments on the original draft of this manuscript.

References

[1] Raghavendra S. and Indiramma M. (2015): Classification and Prediction Model using Hybrid Technique for Medical Datasets. International Journal of Computer Applications 127(5):20-25

[2] Baxt W.G (1995). Application of Artificial Neural Networks to Clinical Medicine. Lancet 346: 1135–1138.

[3] Sargent D.J (2001). Comparison of Artificial Neural Networks with other Statistical Approaches: results from medical data sets. Cancer; 91, 1636–1642

[4] Huang, W. Y. and Lippmann, R. P. (1987) "Comparisons Between Neural net and Conventional Classifiers," in IEEE 1st Int. Conf. Neural Networks, San Diego, CA, pp. 485–493.

[5] Denton, J. W., Hung, M. S., & Osyk, B. A. (1990). A neural network approach to the classification problem. Expert Systems with Applications, 1(4), 417–424.

[6] Haykin, 1999 Haykin, S. (1999). Neural Networks: A Comprehensive Foundation. 2nd ed. Prentice Hall PTR Upper Saddle River, NJ, USA

[7] Han and Kamber, M (2006). Data Mining: Concepts and techniques. (2nded.) San Francis Co: Morgan Kaufmann publishers

[8] Manaswini, P and Ranjit, K.S (2011) Multilayer Perceptron Network in HIV/AIDS Application. International Journal of Computer Application in Engineering Science. Vol 1, Issue 1, 41-48.

[9] Zhang, G. P. (2000). "Neural Networks for Classification: A Survey." IEEE Transactions on Systems, Man and Cybernetics-Part C: Applications and Reviews, vol. 30, pp. 451–462.

[10] UNAIDS (2014) Questions and Answers, Q&A II: Basic facts about the HIV/AIDS 1944–1946, Annals of the History of Computing, Vol. 2, No. 4, pp.349–361.

[11] Newell M. L, Coovadia H, Cortina-Borja M, Rollins N, Gaillard P, Dabis F, (2004). Mortality of infected and uninfected infants born to HIV-infected mothers in Africa: a pooled analysis. Ghent International AIDS Society (IAS) Working Group on HIV Infection in Women and Children. 364(9441):1236-43.

[12] Refaat, M. (2007) Data Preparation for Data Mining Using SAS.Morgan Kaufmann Publishers, San Francisco.

[13] Rumelhart, D. E, Hinton,G.E. and McClelland, J.L.,(eds) (1986) Learning internal representations by error propagation. In Parallel Distributed Processing: explorations in the microstructure of cognition, vol. 1: foundations: 318-362. MIT Press, Cambridge.

[14] Barnard, E. (1992). Optimization for Training Neural Nets, IEEE Trans. Neural Networks, vol. 3, pp. 232–240.

[15] Gupta, J. N. D. & Sexton, R. S., (2013). Comparing backpropagation with a genetic algorithm for neural network training. Omega, 27, 679–684.

[16] Kwak, N. K. and Lee, C. (1997). A Neural Network Application to Classification of Health Status of HIV/AIDS Patient. Journal of Medical Systems, 21(2), 87-97.

[17] Lee, C. and Park, J. (2001). Assessment of HIV/AIDS-related Health Performance using an Artificial Neural Network. Journal. In f. Manage, 38, 231–238.

[18] Pradhan M. and Sahu R.K. (2011). Multilayer Perceptron Network in HIV/AIDS Application. International Journal of Computer Applications in Engineering Sciences I(I).

[19] Taryn, N.H (2006). Predicting HIV Status Using Neural Networks and Demographic Factors. Unpublished Master's Thesis. Faculty of Engineering, University of the Witwatersr and, Johannesburg.

[20] Sibanda and Pretorius, 2011, Sibanda, W. and Pretorius, P (2011) Novel Application of Multi-Layer Perceptrons Neural Networks to Model HIV in South Africa using Seroprevalence Data from Antenatal Clinics. International Journal of Computer Applications 35 (5), 26-31.

[21] Elias, L. (2011). HIV Status Predictive Modeling Using Data Mining Technology. A thesis of Degree of Master of Science in Health Informatics in Addis Ababa University School of Graduate Studies School of Information Science and School of Public Health.

[22] Baridam, B and Irozuru C (2012). The Prediction of Prevalence and Spread of HIV/AIDS using Artificial Neural Network-The case of Rivers State in the Niger Delta, International Journal of Computer Application 44(2), 42-45.

[23] Khan, Z.H., Alin, T, S and Hussain, M.A (2011), Price Prediction of Share Market using Artificial Neural Network (ANN). International Journal of Computer Applications 22(2), 8887 – 8975.

[24] James, T.O., Onwuka,G.I, Babayemi, I, Etuku I and Gulumbe, S.U (2012) Comparison Classifier of Condensed KNN and K-Nearest Neighbourhood Error rate method. The Computing Science and Technology International Journal 2(3), 44-46