Resilient Back-Propagation Algorithm in the Prediction of Mother to Child Transmission of HIV

T. O. James¹, S. U. Gulumbe², A. Danbaba²

¹Department of Mathematics, Kebbi State University of Science and Technology, Aliero, Nigeria
²Department of Mathematics, Usmanu Danfodiyo University, Sokoto, Nigeria
Email: tolujam@gmail.com

Abstract
Prediction of a child HIV status poses real challenges in medical research. Even though there are different statistical techniques and machine learning algorithms that have been used to predict models like HIV for the clinical data with binary outcome variables, yet neural network techniques are major participants for prediction purposes. HIV is the primary cause of mortality among women of reproductive age globally and is a key contributor to maternal, infant and child morbidity and mortality. In this paper, resilient back propagation algorithm is used for training the Neural Network and Multilayer Feed forward network to predict the mother to child transmission of HIV status.

Subject Areas
Mathematical Analysis, Mathematical Statistics

Keywords
Artificial Neural Network (ANN), Resilient Back Propagation Algorithm (RBP), HIV Prediction

1. Introduction
The transmission of HIV from mother to child is responsible for over 90% of infections among children under the age of 15 [1]. AIDS is beginning to converse years of steady growth in child survival. Most children living with HIV become infected through mother-to-child transmission. Since the first reported case of HIV-1 transmission in children in 1983, the global pandemic has had a serious impact on the health and survival of children. Maternal mortality was still very
high in Nigeria with 400,000 children in Nigeria living with HIV [2]. 22% of all new children that have HIV infections globally during 2013 were in Nigeria (51,000) [3]. Therefore, HIV prediction is a critical issue among children.

1.1. Vertical Transmission

Vertical transmission is the transmission of the HIV virus from mother to the child, which is the major source of pediatric infection of human immunodeficiency virus one (HIV-1). This is when a Human Immune deficiency Virus (HIV) positive woman passes the virus to her baby, which is spread when blood, semen, or another body fluid from an infected person enters the body of an uninfected person either through sex, sharing of syringes, needles etc. or from an infected mother to her baby at birth [4]. It is also sometimes called perinatal transmission, or maternal transmission. Transmission can occur as in Figure 1 during pregnancy (in utero), around the time of labor and delivery at birth (intrapartum), or breastfeeding (postnatally) [5]. The effects are intense. Therefore, the aim of this paper is to use the resilient back-propagation neural networks to predict mother-to-child transmission (MTCT) of HIV.

1.2. Artificial Neural Network

Artificial neural networks (ANN) are computational bright systems proven to mathematically copy the computational operations of the human brain. A neural network consists of a set of connected cells: the neurons [6], which are made of basic units. It consists of three layers, they are; the input layer, the hidden layer, and the output layer. The independent variables are introduced into the network system by the input layer, transited to the hidden layer through the input neurons linked to the hidden neurons. Figure 2 shows a graphical presentation of neuron. That is, neuron is a real function of the input vector \( y_i, \ldots, y_k \). \( F \) is a function, the output of the hidden layer is imbedded in the summing junction where standardization of the data is processed, where each input is multiplied by weights \( w_{kj} \) along its path and the weighted inputs are then summed and biased by adding a value unto the weighted input. The output of the summation is sent into a function named transfer function and is used to map the processed data to the output data. This function classically falls into one of three types: Linear (or map), Threshold and Sigmoid. The output of the function block is fed to the output neuron and obtained as Equation (1) [7]. Various studies like [8] and [9] used the artificial neural network (ANN) in the prediction of HIV/AIDS.

1.3. The Resilient Back Propagation Algorithm

In the basic Back Propagation (BP) algorithm the weights are adjusted in the steepest descent direction (negative of the gradient). However, efficient as the back-propagation may be, it still suffers from the trap of local minimum or a slow convergence rate and often yields suboptimal solutions rather than global
minimum convergence. Therefore a resilient back propagation method has been established to overcome the fiasco of back propagation [10] [11]. This is based on the developed modification of traditional back propagation algorithm that modifies the weights of a network in order to find a local minimum of the error function. The lower, the error rates the better the procedure [12]. It is a local adaptive learning technique that removes the dangerous influence of the size of the partial derivative on the weight step and converges very quickly and uses simply the sign of the derivative which is the gradient change the biases/weights of the network, instead of the magnitude of the gradient itself. Hence, the resilient algorithm provides faster local adaptation [10] [13]. That is, when Sigmoid transfer function is used the gradient can have a very small magnitude, causing small changes in the weights and biases, even though the weights and biases are far from their optimal values. The value of the learning rate and the momentum properties doesn’t affect it. Therefore, the gradient of the error function is calculated with respect to the weights in order to find a root. Instead of the magnitude of the partial derivatives only their sign is used to update the weights. This gives an equal influence of the learning rate over the whole network [11]. The weights are adjusted by the following rule:

\[
\frac{\partial E}{\partial \omega} = W^{(i)} - \eta_k \cdot \text{Sign} \left( \frac{\partial E^{(i)}}{\partial \alpha_k^{(i)}} \right)
\]

as opposed to

\[
W^{(i+1)} = W^{(i)} - \eta \cdot \frac{\partial E(t)}{\partial \alpha_k^{(i)}}
\]

In traditional backpropagation, where \(t\) indexes the iteration steps and \(k\) the weights. For speedy convergence in shallow areas, the learning rate \(\eta_k\) will be
increased if the corresponding partial derivative keeps its sign since a changing sign indicates that the minimum is missed due to too large a learning rate.

Weight backtracking is a procedure of undoing the last iteration and adding a smaller value to the weight in the next step [11].

The performance of the resilient back-propagation neural network is evaluated in this study by two different criteria which were used to select the best model on the testing dataset. They are Akaike Information Criterion (AIC) and Bayesian information criterion (BIC). They are defined as follows:

Akaike Information Criterion (AIC)

\[
AIC = e^{\frac{2k}{n}} \times \frac{SSE}{n}
\]

Taking natural logarithm,

\[
\ln AIC = \left( \frac{2k}{n} \right) + \ln \left( \frac{SSE}{n} \right)
\]

Bayesian information criterion (BIC)

\[
BIC(k) = \log_e \left( \frac{SSE}{N} \right) + \frac{k}{N} \log_e (N)
\]

where \( k \) is the number of independent variable (including the intercept) and \( n \) is the number of observation.

2. Model Development

2.1. Data

The input data was collected from the ANC center in a famous Hospital in Sokoto, Nigeria. Mother’s CD4 count (MCD4), Delivery mode and ART Drug used and mode of delivery are the input while the Child HIV Status is the output.

2.2. The Prediction Phase

To produce the prediction model, as in any statistical model, the parameters (weights) of neural network model need to be estimated before the network can be used for prediction purposes. After training with acceptable error the weights are set into the network then the trained network is given the input data set of the mother of the child we want to predict. This is divided into three portions: training, validation and testing sets. A model is assumed good if the error of out-of-sample testing is the lowest compared with the other models. The weights multiply the input information. Where input is denoted by \( X_i \) and each weight \( w_j \), then the activation is equal to \( \sum x_i w_j \) and enter into the neurons of hidden layer, in our model we used three hidden layer which passes to the next neuron of the output layer. The trained network then predicts the child’s HIV status using the mother’s given input data set.

3. Interpretation of Results

The prediction of the child’s HIV status is obtained from the network. The results
Table 1. Network 4-5-1.

<table>
<thead>
<tr>
<th>Threshold used</th>
<th>Error</th>
<th>AIC</th>
<th>BIC</th>
<th>Reached threshold</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>16.217987</td>
<td>84.435974</td>
<td>157.340521</td>
<td>0.0009894</td>
<td>52172</td>
</tr>
<tr>
<td>0.01</td>
<td>16.421605</td>
<td>84.843211</td>
<td>157.747758</td>
<td>0.005393</td>
<td>18245</td>
</tr>
<tr>
<td>0.05</td>
<td>16.515972</td>
<td>85.031944</td>
<td>157.936491</td>
<td>0.0252263</td>
<td>15245</td>
</tr>
</tbody>
</table>

Table 2. Network 4-4-1.

<table>
<thead>
<tr>
<th>Threshold used</th>
<th>Error</th>
<th>AIC</th>
<th>BIC</th>
<th>Reached threshold</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>16.214673</td>
<td>74.429346</td>
<td>133.313788</td>
<td>0.0009987</td>
<td>93058</td>
</tr>
<tr>
<td>0.01</td>
<td>16.345452</td>
<td>74.690904</td>
<td>133.575346</td>
<td>0.00860059</td>
<td>23456</td>
</tr>
<tr>
<td>0.05</td>
<td>16.929938</td>
<td>75.859877</td>
<td>134.744319</td>
<td>0.043962</td>
<td>5193</td>
</tr>
</tbody>
</table>

Table 3. Network 4-3-1.

<table>
<thead>
<tr>
<th>Threshold used</th>
<th>Error</th>
<th>AIC</th>
<th>BIC</th>
<th>Reached threshold</th>
<th>Steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>0.001</td>
<td>16.222842</td>
<td>64.445684</td>
<td>109.31002</td>
<td>0.0006404</td>
<td>46074</td>
</tr>
<tr>
<td>0.01</td>
<td>16.309438</td>
<td>64.618875</td>
<td>109.483212</td>
<td>0.009976</td>
<td>18881</td>
</tr>
<tr>
<td>0.05</td>
<td>16.595952</td>
<td>65.191904</td>
<td>110.056241</td>
<td>0.042744</td>
<td>12794</td>
</tr>
</tbody>
</table>

Table 4. Summary of best fitted neural network models using fundamental variables as inputs.

<table>
<thead>
<tr>
<th>Best fitted model</th>
<th>Decay (λ)</th>
<th>No of hidden nodes</th>
<th>Error</th>
<th>AIC</th>
<th>BIC</th>
<th>Reached threshold</th>
</tr>
</thead>
<tbody>
<tr>
<td>Fit 1</td>
<td>0.001</td>
<td>3</td>
<td>16.222842</td>
<td>64.445684</td>
<td>109.31002</td>
<td>0.0006404</td>
</tr>
<tr>
<td>Fit 2</td>
<td>0.01</td>
<td>3</td>
<td>16.309438</td>
<td>64.618875</td>
<td>109.483212</td>
<td>0.009976</td>
</tr>
<tr>
<td>Fit 3</td>
<td>0.05</td>
<td>3</td>
<td>19.595952</td>
<td>65.191904</td>
<td>110.056241</td>
<td>0.042744</td>
</tr>
</tbody>
</table>

of the preliminary trainings of the networks are given in Tables 2-4. Prediction model based on the simultaneous agreement of the two information criteria, network model (4-3-1) in Table 4, that is, Fit 1 was selected as a tentative model for further study. The trained neural network is with hidden unit size of 3 for λ = 0.001 was used for prediction problems in HIV status of children as seen in Table 4. After obtaining the optimum structure for the network, the performance of the MLP network was determined. The performance analysis of the MLP network is based on accuracy. It produces a high accuracy of 95%. This high accuracy is obtained keeping all the other factors constant for the training algorithms. This study indicates the good predictive capabilities of MLP neural network and it confirms the work of [14] though their study shows 90% accuracy in their model.

Tables 1-3 are the various neural network models for prediction in a given network architecture. The corresponding network plot of the selected best model is seen in Figure 3.
Figure 3. Neural network architecture for prediction of child HIV status for Best fitted Network 4-3-1 model using “threshold 0.001” & resilient back-propagation with back-tracking algorithm.

4. Conclusion

This study used Resilient Backpropagation (RBP) algorithm in predicting mother to child transmission of HIV. The outcome of this study shows that if the physician has some demographic variable factors of a HIV positive pregnant mother, the status of the child can be predicted before been born. The prediction of HIV status of children is obtained from the network as seen in Table 4 and Figure 3 shows the graphs of the fitted and actual predict.

References


