

Artificial Neural Networks- A Review of Applications of Neural Networks in the Modeling of HIV Epidemic

Wilbert Sibanda

North West University

Vaal Triangle Campus

Van Eck Blvd, Vanderbijlpark, 1900, South Africa

Philip Pretorius

North-West University

Vaal Triangle Campus

Van Eck Blvd, Vanderbijlpark, 1900, South Africa

ABSTRACT

Neural networks have been applied successfully to a broad range of fields such as finance, data mining, medicine, engineering, geology, physics and biology. In finance, neural networks have been used for stock market prediction, credit rating, bankruptcy prediction and economic indicator forecasts. In medicine, neural networks have been used extensively in medical diagnosis, detection and evaluation of medical conditions and treatment cost estimation. Furthermore, neural networks have found application in data mining projects for the purposes of prediction, classification, knowledge discovery, response modeling and time series analysis. This review paper will present the application of neural networks to the study of HIV. HIV research falls into four broad areas namely, behavioral research, diagnostic research, vaccine research and biomedical research. Most of the research publications featured in this review paper emanate from the four broad HIV research areas and will be presented in three categories namely prediction, classification and function approximation.

General Terms

Neural networks, HIV, Modeling, Epidemic

Keywords

Multi-layer Perceptrons, neural networks, HIV/AIDS

1. INTRODUCTION

The artificial neural network (ANN) is set of processing units called neurons. The Ann can be used to approximate the relationship between input and output signals of the system (Rebizant et. al. 2011). ANNs use a mathematical or computational model for information processing. This paper explores the application of neural networks in the study of HIV/AIDS. This review is aimed at readers with little or no understanding of neural networks and is designed to act as a guide through the literature so that they may better appreciate this tool.

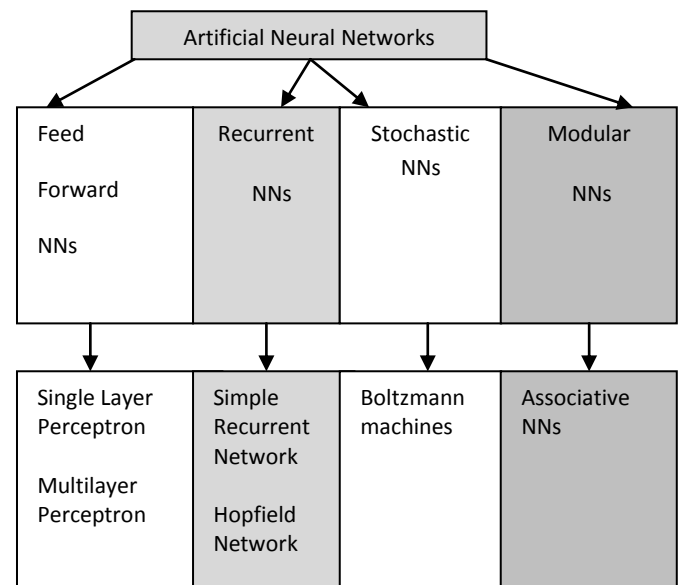


Fig. 1: A taxonomy of neural network architectures

This review is divided into several sections, starting with a brief introduction to the different types of neural networks (NNs) followed by a description of the most basic algorithm for training NNs known as backpropagation. This will be followed by a review of some of the recent applications of NNs to study HIV.

2. TYPES OF NEURAL NETWORKS

2.1. Feed-forward Neural Networks

These are the first and simplest type of ANNs. In these networks, the information moves in only one direction, forward from the input nodes, through the hidden nodes and to the output nodes. There are no cycles or loops in the network. Examples of feed-forward NNs are single layer perceptron (SLP) and multi-layer perceptron (MLP).

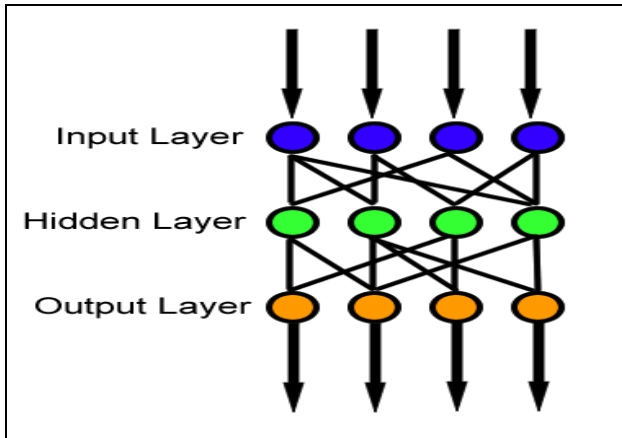


Fig. 2: A schematic representation of a Feed-forward NN.

2.1.1. Single Layer Perceptron (SLP)

An SLP network consists of a single layer of output nodes. Inputs are fed directly to the outputs via a series of weights. The sum of the products of the weights and inputs is calculated in each node, and if the value is above some threshold, the neuron fires and takes the activated value, otherwise it takes the deactivated value. Neurons with this kind of activation function are also called McCulloch-Pitts neurons or threshold neurons. These networks were described by Warren McCulloch and Walter Pitts in the 1940s (Graben and Wright, 2011).

2.1.2. Multi-Layer Perceptron (MLP)

MLPs consist of multiple layers of computational units, interconnected in a feed-forward way (Calcagno et. al. 2010). MLPs use a variety of learning techniques, the most popular being back-propagation, where the output values are compared with the correct answer to compute the value of some predefined error-function. The error is then fed back through the network. Using this information, the algorithm adjusts the weights of each connection in order to reduce the value of the error function by some small amount. A general method for non-linear optimization called gradient descent is applied to adjust the weights.

2.2. Recurrent Networks

Unlike feed-forward networks, recurrent neural (RN) networks are models with a bi-directional data flow. While a feed-forward network propagates data linearly from input to output, RN, also propagate data from later processing stages to earlier stages (Bitzer and Kiebel, 2012).

2.2.1. Simple Recurrent Network (SRN)

An SRN is a variation of the MLP. It is sometimes called an “Elman network” due to its invention by Jeff Elman. A three layer network is used, with the addition of a set of context units fixed with weights of one. At each step, the input is propagated in a standard feed-forward fashion, and then a back-propagation learning rule is applied. SRN can be used for sequence-prediction that is beyond the power of a standard MLP.

In a fully recurrent network, every neuron receives inputs from every other neuron in the network. These networks are not arranged in layers. Usually only a subset of the neurons receive external inputs in addition to the inputs from all the other neurons, and another disjunct subset of neurons report their output externally as well as sending it to all the neurons. These distinctive inputs and outputs perform the function of the input and output layers of a feed-forward or simple recurrent network, and also join all the other neurons in the recurrent processing.

2.2.2. Hopfield Network

The Hopfield is a recurrent neural network in which all connections are symmetrical. This network was invented by John Hopfield in 1982 (Dong et. al. 2011). The Hopfield guarantees that its dynamics will converge (Pajares et. al. 2010).

2.3. Stochastic Neural Networks

A stochastic neural network differs from a regular neural network in the fact that it introduces random variations into the network (Su et. al. 2011). Boltzmann machines is an example of a stochastic neural network.

2.4. Modular Neural Networks

A modular neural network is a neural network characterized by a series of independent neural networks moderated by some intermediary (Pandey, 2012). Each independent neural network serves as a module and operates on separate inputs to accomplish some subtask of the task the network intends to perform (Azam, 2000)). The intermediary takes the outputs of each module and processes them to produce the output of the network as a whole. The intermediary accepts the modules’ outputs but not respond to, nor otherwise signal, the modules. The modules do not interact with each other. One of the benefits of a modular neural network is the ability to reduce a large neural network to smaller, more manageable components (Azam, 2000). Examples of modular neural networks include committee of machines (CoM) and associative neural networks (ASNN).

2.4.1. Committee of machines (CoM)

CoM is a collection of different neural networks that together decide and vote on a given example, hoping that errors would cancel as there are several experts (Bettebghor et. al. 2011). This gives a much better result compared to other neural network models. CoM is similar to the general machine learning bagging method, except that the necessary variety of machines in the committee is obtained by training from different random starting weights rather than training on different randomly selected subsets of the training data (Bettebghor et. al. 2011).

2.4.2. Associative Neural Network (ASNN)

This is an extension of the CoM that goes beyond simple/weighted average of different models. ASNN

represents a combination of feed-forward neural networks and the k-nearest neighbour technique (kNN) (Sivaram et. al. 2011). It uses the correlation between ensemble response as a measure of distance amid the analyzed cases for the kNN. This corrects the bias of the neural network ensemble.

3. TRAINING A NEURAL NETWORK

A neuron that receives information from outside of the network is called an input layer. A neuron that contains the network's predictions or classifications is called the output layer. Neurons that are found between the input and output layers form the hidden layer. The transfer function is usually a log function, a sigmoid function, or a hyperbolic tangent and scales all of the information coming into the neuron. The sigmoid curve is often used as a transfer function because it introduces non-linearity into the network's calculations. The sigmoid function has the additional benefit of having an extremely simple derivative function, as required for back-propagation errors through a feed-forward neural network.

As values are sent from one layer to the next, a weight is assigned to each interconnecting line and is multiplied by the values. Each neuron on the hidden layer sums all inputs, and the combined input is modified by the function. The output value of the transfer function is generally passed directly to all neurons in the next layer, again with a weight assigned to each value. Values of the interconnecting weights predetermine the neural network's computation reaction to any arbitrary input pattern. As information is passed forward from the inputs towards the outputs, interconnecting weights are adjusted by a back-propagation algorithm during the learning phase so that known outputs will best match predicted outputs.

As stated above, back-propagation is the most common algorithm used to train NNs due to its ability to generalize well on a wide variety of problems. It is a feed-forward, fully interconnected, supervised network. It builds models that classify patterns or make predictions according to patterns of inputs and outputs it has learned.

Training of a NN is the procedure by which the values for the individual weights are determined such that the relationship the network is modeling is accurately resolved. By varying the weights through all possible values, and by plotting the errors in the three-dimensional space, an error surface is developed as shown in Fig. 3.

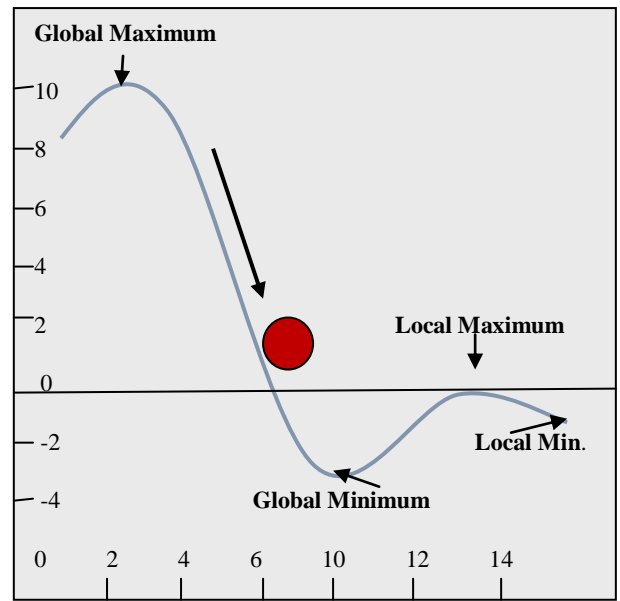


Fig. 3: A schematic representation of an error surface

The objective of training a neural network is to find the combination of weights which will result in the smallest error. In practice, it is not possible to plot such a surface due to the multitude of weights and hence it is more relevant to find the minimum point of the error. One possible technique is to use a procedure known as gradient descent. The back-propagation training algorithm uses this procedure to attempt to locate the global minimum of the error surface. The back-propagation algorithm is the most computationally straightforward algorithm for training the multilayer perceptron. Back-propagation has been shown to perform adequately in many applications. A great number of applications discussed in this paper used back-propagation to train NNs.

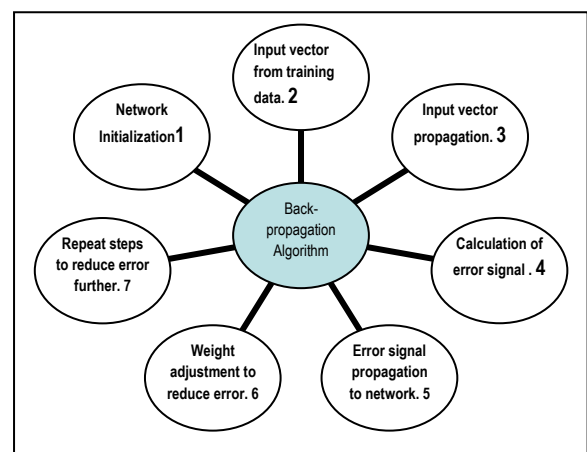


Fig. 4: A representation of the back-propagation algorithm

Fig. 4 shows a back-propagation algorithm also known as on-line training whereby the network weights are adapted after each pattern has been presented (Reuter et. al. 2010). The

alternative is known as batch training, where the summed error for all patterns is used to update the weights. In practice, many thousands of training iterations will be required before the network error reaches a satisfactory level as determined by the problem being addressed. Training should be stopped when the performance of the NN on the independent test data reaches a maximum.

The error surface in Fig. 3 contains more than one minimum. It is important that the training algorithm does not become trapped in a local minimum. The back-propagation algorithm contains two adjustable parameters, a learning rate and a momentum term, which can assist the training process in avoiding this. The learning rate determines the step size taken during the iterative gradient descent learning process. If this is too large then the network error will change erratically due to large weight changes, with the possibility of jumping over the global minima. Conversely, if the learning rate is too small then training will take a long time. The momentum term is used to assist the gradient descent process if it becomes stuck in a local minimum. By adding a proportion of the previous weight change to the current weight change it is possible that the weights can escape the local minimum.

4. GENERAL APPLICATIONS OF NEURAL NETWORKS

Neural networks have been applied to a wide variety of tasks, all of which can be categorized as either prediction, function approximation, or pattern classification. Prediction involves the forecasting of future trends in a time series of data given current and previous conditions. Function approximation is concerned with modeling the relationship between variables. Pattern classification involves classifying data into discrete classes.

HIV modeling involves using a variety of approaches. Choosing the most suitable approach depends on the complexity of the problem being addressed and the degree to which the problem is understood. Assuming a strong theoretical understanding of the problem is available then a full numerical model is perhaps the most desirable solution. However, in general, as the complexity of a problem increases the theoretical understanding decreases due to ill-defined interactions between systems. And thus statistical approaches are required. Recently, the use of neural networks has been shown to be an effective alternative to more traditional statistical techniques. Neural networks have been shown to be trained to approximate virtually any smooth, measurable function. Unlike other statistical techniques NNs make no prior assumptions concerning the data distribution. NNs can model highly non-linear functions and can be trained to accurately generalize when presented with new, previously unseen data. These features of NNs make them an attractive alternative to developing numerical models, and also when choosing between statistical approaches.

4.1. Function approximation and Prediction

Function approximation and prediction are very similar. Often, only one variable is modeled from the input data. To use a NN for prediction involves training the network to output the future value of a variable, given an input vector containing earlier observations.

The NN approximates highly non-linear functions and requires no prior knowledge of the nature of this relationship. This is one of the benefits NNs offer over conventional regression analysis. If the relationship between variables is non-linear then linear regression is clearly an inappropriate tool, although it may be possible to apply linear regression on a more local basis where the non-linearity can be dismissed. Non-linear regression is useful if the nature of the nonlinearity can be found and if the non-linearity is consistent over the entire range of measurements. Extremely non-linear relationships exist in the real world and it is inappropriate to attempt to understand these problems using traditional regression, attributing scatter to the presence of 'noise'. Under these circumstances NNs offer a vital tool.

4.2. Pattern classification

A classifier partitions the whole of measurement space, or the set of all possible input vectors, into disjoint subsets each representing one of the many target classes. Classifiers aim to minimize the probability of misclassification via Bayes theorem, by classifying a new example to the class that has the highest posterior probability. The posterior probability gives a measure of the likelihood of a particular measurement vector belonging to a particular class.

NNs can be used for classification by assigning output nodes to represent each class. Unlike function approximation and prediction, NNs used for classification will have more than one output node.

NNs have been shown to be superior to traditional classification approaches for several reasons. Firstly and foremost, this approach does not require any prior assumptions regarding the distribution of the training data. Classifications that use Bayes theorem rely on a Gaussian distribution of the data which is often not the case in practical applications. Another benefit of the NNs approach is that no decision regarding the relative importance of the various input measurements needs to be made. During training, the weights are adjusted to select the most discriminating input measurements.

5. USE OF NEURAL NETWORKS IN HIV MODELING

In this paper, a brief overview of applications of NNs in HIV modeling from prediction, function approximation and pattern classification will be presented. It is hoped that these papers illustrate the main principles of applying the NNs to real-world HIV modeling problems. Other papers will be mentioned for reference purposes.

5.1. Function approximation and Prediction

The multilayer perceptron has been applied within the field of HIV prediction.

Larder et al. (2008) describes the application of ANNs for decision support in medicine. The model takes treatment change episodes (TCEs) and HIV drug resistance mutations as basic input variables to train ANN models. In conclusion, the authors of this paper define the reliability of ANN predictions for HIV patients receiving routine clinical care, the utility of ANN models to identify effective treatments for patients failing therapy, strategies to increase the accuracy of ANN predictions and performance of ANN models in comparison to the rules-based methods currently in use.

Dechao Wang et. al. (2009) compared three computational methods for the prediction of virological response to combination HIV therapy. HIV treatment failure is commonly associated with drug resistance and the selection of a new regimen is often guided by genotypic resistance testing. The interpretation of complex genotypic data poses a major challenge and in that regard the authors developed artificial neural network (ANN) models that predict virological response to therapy from HIV genotype and other clinical information. The accuracy of ANN was compared with alternative modeling methodologies such as random forests (RF) and support vector machines (SVM). In conclusion, the researchers noted that RF and SVM models can produce predictions of virological response to HIV treatment that are comparable in accuracy to a committee of ANN models, however combining the predictions of different models further improved their accuracy.

Lamer et al. (2008) demonstrated the use of ANN trained using evolutionary computation to predict R5, X4, and R5X4 HIV-1 co-receptor usage. The results indicated identification of R5X4 viruses with predictive accuracy of 75.5%.

Yu-Dong Cai et al. (1998) studied the application of ANN method for predicting HIV protease cleavage sites in protein. The authors developed Kohonen's self-organisation model. Kohonen's self-organization neural network is a two-layer network. Output nodes are arranged regularly on a planar mapping grid. Each input node is connected to every output node via a variable connection weight. The self-organization model is well known for its low-dimensional topology-preserving mapping of high-dimensional patterns and stably evolving properties. Yu-Dong Cai et al, concluded that because of the neural network's superior ability in dealing with nonlinear problems such as predicting HIV protease cleavage sites (92.06%) in proteins it is quite reliable and accurate, and thus should be helpful in finding effective inhibitors of HIV protease.

Hatzakis et al. (2002) studied the application of neural networks to the modeling of mortality and morbidity during loss of HIV T-Cell homeostasis. Multiple factors may predict mortality of HIV patients, including initial response to therapy, viral factors and host immune parameters. Due to the complexity of this problem, the authors developed feed-forward back-propagation neural networks to optimally evaluate outcomes of therapy and predict morbidity and mortality. The neural networks featured 1 input, 1 hidden layer and 1 output layer over a sigmoid transfer function. Adaptive learning was used but not finally adopted since constant learning factor and momentum produced more accurate predictions. The maximum number of epochs was set at 100 000. The model stopped its training process at the epoch that the difference of two consecutive R^2 measurements satisfied the convergence criterion of 0.0001. The weights

were selected accordingly, and were initialized through a random number generator. The accuracy of the neural networks was validated using Cox regression modeling (RM). The results of this research indicated that although neural network and Cox modeling were successful in predicting mortality, the neural network was superior in assessing risk in the studied population. The authors concluded that neural network based modeling can be at least as accurate as regression modeling in predicting morbidity and mortality that occurs in late stages of HIV infection following loss of T-cell homeostasis. This technique may prove useful in deciding when to initiate therapy, evaluating the response to anti-retroviral treatment and eventually in the prediction of morbidity and mortality.

Hatzakis and Tsoukas, (2001) used neural networks for the assessment of HIV immunopathology. The study was primarily designed to assess the utility of neural networks for surrogate marker-based prediction modeling of HIV disease, evaluating the immune and virologic responses to HAART. A secondary objective of the study was to compare the performance of the NNs with multiple regression analysis-based modeling (MRM). Despite the documented ability of antiretroviral therapies to increase CD4 counts, the use of other surrogate markers in patient management is limited. The authors postulated that additional HIV-related surrogate markers might be used if effective analytical tools were available. The results demonstrated that the neural network was at least as accurate as a multi-regression model.

Hatzakis et al. (2005) applied neural networks for a longitudinal assessment of the electronic antiretroviral therapy to determine response to HIV treatment. The relationship between treatment outcomes with disease markers and other contributing factors is complex. The authors developed ANN models based on Jordan-Elman networks to longitudinally follow viral surrogate markers together with demographics, biochemical and laboratory data to describe the drug-virus-host interactions in over 4 000 HIV adult patients. Authors concluded that these tools can be used in real-time context of prospective, longitudinal clinical trials of newer antiretroviral drugs.

Herman et al. (1999) compared neural networks with five traditional methods for predicting creatinine clearance in patients with HIV infection. The 16 input variables were age, ideal body weight, actual body weight, body surface area, and the following blood chemistries: sodium, potassium, aspartate aminotransferase, red blood cell count, platelet count, white blood cell count, glucose, serum creatinine, blood urea nitrogen, and albumin. The only output variable was creatinine clearance. The ANN paradigm was a fully connected, three-layer, back-propagation algorithm with input, hidden, and output layers. The hyperbolic tangent function was used as the neuron transfer function. The selected neural architecture was trained for 1000 epochs. The authors concluded that average percentage prediction error, bias, and precision were greatly improved with the ANN over other equations. The approach is easy to use and apply to different kinds of problems and the flexibility of neural networks makes them a promising alternative to established methods.

Betechuoh et al. (2007) used neural networks in an inverse configuration for the adaptive control of HIV status of individuals. In this paper, a control mechanism to understand how demographic properties affect the risk of being HIV

positive is implemented. The research aims to understand whether HIV susceptibility can be controlled by modifying some of the demographic properties such as education. A feed-forward and inverse neural network comprising of 9 inputs and 1 output was constructed. A genetic algorithm was used to choose the optimal number of hidden units. The authors concluded that the proposed method is able to predict the educational level of individuals to an accuracy of 88% if the HIV status of individuals and other demographic characteristics are known. It is thus possible to understand how the educational level of individuals can be modified to control the prones of individuals to HIV contraction.

Ioannidis et al. (1997) proposed the use of neural networks to model complex immunogenetic associations on the progression of HIV infection. Complex immunogenetic associations of disease involving a large number of gene products are difficult to evaluate with traditional statistical methods and thus may require complex modeling. The authors evaluated the performance of feed-forward back-propagation neural networks in predicting rapid progression to acquired immunodeficiency syndrome (AIDS) for patients with HIV infection on the basis of major histocompatibility complex variables. Network network performance was compared with that of logistic regression. This research concluded that neural networks could be trained to recognize genetic patterns in conjunction with associated clinical outcomes. Their performance in modeling these complex associations in a training set was superior to logistic regression models. This was attributed to the ability of neural networks to model very complex data, when trained adequately.

Resino et. al. (2011) studied the development of an artificial neural network to predict significant fibrosis in HIV/Hepatitis C (HCV) coinfecting patients using clinical data derived from peripheral blood. Patients were randomly divided into an estimation group used to generate the ANN and a test group used to confirm its predictive power. The authors concluded that ANN technique is a helpful tool for guiding therapeutic decisions in the clinical practice of HIV/HCV coinfection.

Yoo et. al. (2010) researched the hierarchical kernel mixture models for the prediction of AIDS disease progression using HIV structural gp120 profiles. Changes to the glycosylation profile on HIV gp120 can influence viral pathogenesis and alter AIDS disease progression. The proposed new intelligent framework proved to be accurate and provided an important benchmark for predicting AIDS progression computationally. The model was trained using a novel HIV gp120 glycosylation structural profile to detect possible stages of AIDS disease progression for the target sequences of HIV positive individuals. The performance of the proposed model was compared to seven existing different machine learning models in terms of error-rate (MSE), accuracy, stability and complexity. The novel framework showed better predictive performance on the three stages of AIDS disease progression.

Zazzi et. al. (2011) studied the prediction of response to antiretroviral therapy by human experts and by the EuResist data-driven expert system. The EuResist expert system is a novel data-driven on-line system for computing the probability of 8-week success for any given pair of HIV-1 genotype and combination antiretroviral therapy regime plus optimal patient information. The objective of the study was to compare the EuResist system vs. human experts (EVE) for the ability to predict response.

Pasomsub et. al. (2010) researched the application of artificial neural networks for the phenotypic drug resistance prediction. Although phenotypic resistance testing provides more direct measurement of antiretroviral drug resistance than genotypic testing, it is costly and time consuming. However, genotypic resistance testing has the advantage of being simpler and more accessible. The study applied the artificial neural network (ANN) system to predict the HIV-1 resistance phenotype from the genotype. The results indicated that by using the ANN, within associated amino acid positions known to influence drug resistance for individual antiretroviral drugs, drug resistance was accurately predicted and generalized for individual HIV-1 subtypes.

Singh and Mars (2010) studied the application of support vector machines to forecast changes in CD4 count of HIV-1 positive patients. HIV infection can be effectively managed with antiretroviral (ARV) drugs, but close monitoring of the progression of the disease is vital. One of the better surrogate markers for the disease progression is the use of CD4 cell counts. Forecasting CD4 cell count helps clinicians with treatment management and resource allocation. The aim of this research was to investigate the application of machine learning to predict future CD4 count change. The model took as input the genome, current viral load and number of weeks from baseline CD4 count and predicted the range of CD4 count change. The model produced an accuracy of 83%.

Deeb and Jawabreh (2012) presented a quantitative structure-activity relationship study using artificial neural network (ANN) methodology to predict the inhibition constants of 127 symmetrical and unsymmetrical cyclic urea and cyclic cyanoguanidine derivatives containing different substituent groups. The results obtained by artificial neural networks gave advanced regression models with good prediction ability. Therefore, artificial neural networks provided improved models for heterogeneous data sets without splitting them into families.

Purwanto et. al. (2011) studied the application of adaptive neuro-fuzzy inference system for HIV/AIDS time-series prediction. Improving accuracy in time series prediction has always been a challenging task for researchers. Prediction of time series data in healthcare such as HIV/AIDS data assumes importance in healthcare management. Statistical techniques such as moving average (MA), weighted moving average (WMA) and autoregressive integrated moving average (ARIMA) models have limitations in handling the non-linear relationships among data. In general, for complex healthcare data, it may be difficult to obtain high prediction accuracy rates using the statistical or AI models and to solve this problem the authors proposed a hybrid model such as adaptive fuzzy inference system (ANFIS) to predict HIV/AIDS data. The results of this research indicated that the proposed model was superior to other models.

Goodarzi and Freitas (2010) studied the application of genetic function approximation (GFA) and artificial neural networks (ANN) to model the activities of a series of HIV reverse transcriptase inhibitor TIBO derivatives. The prediction results were found to be superior to those previously established. In this work, the researchers, applied multi-variate image analysis to quantitative structure-activity relationship (MIA-QSAR) was coupled to principal component analysis-adaptive neuro-fuzzy inference systems (PCA-ANFIS) was applied to the same set of compounds previously reported. The ANFIS procedure was capable of

accurately correlating the inputs (PCA scores) with the bioactivities. The result suggested that the present methodology may be useful to solve other QSAR problems, especially those involving non-linearities.

Nelwamondo et al. (2007) developed a comparison of neural networks and expectation maximization techniques to study missing data using industrial power plant, industrial winding process and HIV seroprevalence data. The authors compare two approaches to the problem of missing data estimation. The first technique is based on the current state-of-the-art approach to this problem, i.e. the use of maximum likelihood (ML) and expectation maximization (EM). The second approach is the use of a system based on auto-associative neural network and genetic algorithm (GA). The authors concluded that EM algorithm was more suitable and performed better in cases where there was little or no interdependency between the input variables, whereas the autoassociative neural network and GA combination was suitable when there was inherent nonlinear relationships between some of the given variables.

5.2. Pattern classification

Multilayer perceptrons have been applied to the classification of HIV/AIDS data to distinguish between HIV positive and negative subjects. Multilayer perceptrons have been shown to produce better classifications than discriminant analysis.

Pradhan and Sahu (2011) presented a new method of multilayered perceptron (MLP) network to classify HIV/AIDS infected and non-infected status of individuals. Seven features of a patient were used as input for training the neural networks, such as age, sex, weight, HB, CD4, CD8 and TB. In order to determine the applicability and best performance of the MLP network, three different training algorithms like backpropagation, Levenberg-Marquardt, and Bayesian rule algorithms were employed to train the MLP networks. The authors concluded that the MLP network trained using backpropagation algorithm produced the best performance with 89.80% accuracy as compared to Levenberg-Marquardt and Bayesian rule algorithms. The results also significantly demonstrated the suitability of the MLP network for calculating and specifying the HIV/AIDS positive/negative status of the patient.

Lee and Park, (2000) explored the application of neural networks to classify and predict the symptomatic status of HIV/AIDS patients. The purpose of the study was to apply an ANN to provide correct classification of AIDS versus HIV status patients. An ANN model was developed using publicly available HIV/AIDS data in the AIDS Cost and Services utilization survey (ACSUS). The authors concluded that an ANN model can facilitate planning, decision-making, and managerial control by providing hospital administration information.

Kwak and Lee, (1997) studied the application of neural networks to the classification and prediction of the health status of HIV/AIDS patients. Neural network modeling of HIV/AIDS issues involves the interaction of many diverse variables, whose relationships are often unclear and ill-defined. The study utilized the AIDS Cost and Services Utilization Survey (ACSUS), a longitudinal study of persons with HIV-related disease in which a combination of personal interviews and abstraction of medical records is used. The model developed was the three-layer back-propagation

algorithm neural network. Input pattern had nine variables: race-white, race-black, race-hispanic, exproute-IDU, exproute-IDU with specific reasons, totipngt (total number of inpatient nights), totamv (total number of ambulatory visits), toterv (total number of emergency room visits), and totobs (total observation days).

Tim and Marwala (2001) used computational intelligence methods for risk assessment of HIV. The design of the study consisted of two parts namely, use of neural networks trained using supervised learning on antenatal survey data to perform binary classification and use of trained neural networks to produce inferred risk probability using Bayesian classification methods to estimate class conditional densities. An autoassociative neural network was trained on complete datasets.

Pradhan and Sahu, (2011) studied the application of multilayered perceptron (MLP) network to classify HIV/AIDS infected and non-infected status of individuals. The inputs for the study were age, sex, weight, HB, CD4, CD8 and TB. In order to determine the applicability and best performance of the MLP network, three different training algorithms like Backpropagation, Levenberg-Marquardt, and Bayesian Rule algorithms were employed to train the MLP networks. This research concluded that the MLP network trained using backpropagation algorithm produced the best performance with nearly 90% accuracy as compared to Levenberg-Marquardt and Bayesian Rule algorithms. The study therefore demonstrated the suitability of the MLP network for calculating and specifying the HIV/AIDS positive/negative status of patients.

Goldbaum et. al (2011) applied machine learning classifiers (MLCs) to seek differences in visual fields (VFs) between normal eyes and eyes of HIV positive patients in order to find the effect of immunodeficiency on VFs and to compare the effectiveness of MLCs to commonly-used Statpac global indices in analysing standard automated perimetry (SAP). The authors concluded that eyes from both low and high CD4 HIV positive patients have VFs defects indicating retinal damage. Generalized learning classifier, SVM, and a Statpac classifier, MD, were found to be effective at detecting HIV eyes that had field defects, even when the defects were subtle.

Seyagh et. al (2011) studied the application of support vector machines, artificial neural networks and decision trees for anti-HIV activity prediction of organic compounds. Prediction of biological activity of molecules from their chemical structures is a challenge in drug discovery. In this regard pattern classification has gained attention as one of the methods that can be utilized in drug discovery. In this study, the authors used three classification models for anti-HIV activity, based on pattern recognition methods such as support vector machines, artificial neural networks and decision trees. All the methods gave good results in learning and prediction. The latter results indicated that the three models can be used as an alternative for classification problems in structure anti-HIV activity relationship.

Larder et. al. (2011) researched the potential utility of computational modeling as an HIV treatment selection tool by physicians. The HIV Resistance Database Initiative (RDI) based in the UK in conjunction with collaborating clinical centres around the world, has used antiretroviral treatment and response data from patients around the world to develop computational models that are highly predictive of virologic

response. The results of this study indicated that the majority of the physicians would use the proposed system if it were available, especially for highly treatment-experienced cases with challenging resistance profiles.

Kim et. al. (2010) explored an MLP-based feature subset selection for HIV-1 protease cleavage site analysis. In recent years several machine learning approaches have been applied to model the specificity of the human immunodeficiency virus type 1 (HIV-1) protease cleavage domain. The high dimensional domain dataset contains a small number of samples, which could misguide classification modelling and its interpretation. An appropriate feature selection can alleviate the problem by eliminating irrelevant and redundant features and thus improve prediction performance. In this regard the authors proposed a new feature subset selection method called FS-MLP that selects relevant features using multi-layered perceptron (MLP) learning. The method involved MLP learning with a training dataset and then feature subset selection using decompositional approach to analyze the trained MLP. The experimental results indicated that the FS-MLP is effective in analyzing multi-variate, non-linear and high dimensional datasets such as HIV-1 protease cleavage dataset. The authors concluded that the FS-MLP was a useful method for computational sequence analysis.

6. CONCLUSION

Artificial neural networks have been shown to be a useful tool for prediction, function approximation and classification. The real benefits of a modelling technique that can accurately reproduce any measurable relationship are enormous. The benefits of neural networks are particularly apparent in applications where a full theoretical model cannot be constructed, and especially when dealing with non-linear systems. The large number of challenges involved implementing, training and interpreting neural networks must be juxtaposed and weighed against the performance benefits of more traditional, and often inappropriate, techniques.

Twenty seven studies published between 1997 and 2012 met the minimum criteria for inclusion in our literature review. All the studies concluded that artificial neural networks can play a vital role in explaining the dynamics of HIV within communities.

7. ACKNOWLEDGEMENTS

Wilbert Sibanda acknowledges Doctoral funding from South African Centre for Epidemiological Modeling (SACEMA), Medical Research Council of South Africa (MRC) and the North West university (Vaal Triangle). Special thanks to Tlaleng Cathrine Sibanda for the perpetual encouragement. The authors would like to give a special word of gratitude to the National department of Health in South Africa for the 2006 and 2007 antenatal seroprevalence data.

8. REFERENCES

- [1] Azam 200. Biologically Inspired Modular Neural Networks. PhD Dissertation. Virginial Polytechnic Institute and State University.
- [2] Betechuoh B.L. et. al. 2007. Using Inverse Neural Networks for HIV Adaptive Control. *International Journal of Computational Intelligence Research*. vol.3, No. 1. 11-15.
- [3] Bettebghor D. et. al. 2011. Surrogate Modeling Approximation using a Mixture of Experts based on EM joint Estimation. *Structural and Multidisciplinary Optimization*. Vol. 43, No. 2, 243-259.
- [4] Bitzer S. and Kiebel J. 2012. Recognizing Recurrent Neural Networks (rRNN): Bayesian Inference for Recurrent Neural Networks. arXiv: 1201, 4339v1.
- [5] Calcagno G. and Staiano Antonino. 2010. A Multilayer Neural Network-based approach for the identification of responsiveness to interferon therapy in multiple sclerosis patients. *Information Sciences*. Vol. 180, Issue 21, 4153-4163.
- [6] Deeb O. and Jawabreh M. 2012. Exploring QSARs for Inhibitory Activity of Cyclic Urea and Nonpeptide-Cyclic Cyanoguanidine Derivatives HIV-1 Protease Inhibitors by Artificial Neural Network. *Advances in Chemical Engineering and Science*, vol. 2, 82-100.
- [7] Dong Z. et. al. 2011. Injection Material Selection Method based on Optimizing Neural Network. *Advances in Intelligent and Soft Computing*. Vol. 104, 339-344.
- [8] Goodarzi M. and Freitas M.P. 2010. MIA-QSAR coupled to principal component analysis-adaptive neuro-fuzzy inference systems (PCA-ANFIS) for the modeling of the anti-HIV reverse transcriptase activities of TIBO derivatives. *European Journal of Medicinal Chemistry* vol. 45, 1352-1358
- [9] Graben P.B. and Wright J. 2011. From McCulloch-Pitts Neurons Toward Biology. *Society for Mathematical Biology*.
- [10] Hatzakis G.E. and Tsoukas C.M. 2002. Neural networks morbidity and mortality modeling during loss of HIV T-cell homeostasis. *Proc AMIA Symp*. 320-324.
- [11] Hatzakis G.E. et al. 2001. Neural networks in the assessment of HIV immunopathology. *Proc AMIA Symp*. 249-253.
- [12] Hatzakis G.E. et. al. 2005. Neural Network-longitudinal assessment of the Electronic Anti-Retroviral Therapy (EARTH) Cohort to follow response to HIV-treatment. *AMIA Annu Symp Proc*. 301-305.
- [13] Herman R.A. et. al. 1999. Comparison of a Neural Network Approach with Five Traditional Methods for Predicting Creatinine Clearance in Patients with Human Immunodeficiency Virus Infection. *PHARMACOTHERAPY* vol. 19, No. 6.
- [14] Kwak N. K. and Lee C. 1997. A neural network application to classification of health status of HIV/AIDS patients. *J. Med. Syst*. 21(2):87-97.
- [15] Lamers S.L. et. al. 2008. Prediction of R5, X4, and R5X4 HIV-1 Coreceptor Usage with Evolved Neural Networks. *IEEE/ACM transactions on computational biology and bioinformatics*. vol. 5, No. 2.
- [16] Larder et. al. 2009. Application of Artificial Neural Networks for Decision Support in Medicine. *Methods in Molecular Biology*. Vol. 458, 119-132.
- [17] Lee C.W. and Park J.A. 2000. Assessment of HIV/AIDS-related health performance using an artificial neural network. *Information & Management*, vol. 38, 231-238
- [18] Loannidis J.P.A. et. al. 1997. Use of Neural Networks to Model Complex Immunogenetic Associations of disease: Human Leukocyte Antigen Impact on the Progression of Human Immunodeficiency Virus Infection. *American Journal of Epidemiology*. vol. 147, Issue5. 464-471.

- [19] Nelwamondo et. al. 2007. Missing data: A comparison of neural network and expectation maximization techniques. *Current Science*, vol. 93, No. 11.
- [20] Pajares G. et. al. 2010. A Hopfield Neural Network for Combining Classifiers applied to Textured Images. *Neural Networks*. Vol. 23. Issue 1, 144-153.
- [21] Pandey B. et. al. 2012. Evolutionary Modular Neural Network Approach for Breast Cancer Diagnosis. *International Journal of Computer Science*. Vol. 9, Issue 1, No. 2, 1694-0814.
- [22] Pasomsub et. al. 2010. The Application of Artificial Neural Networks for Phenotypic Drug Resistance Prediction: Evaluation and Comparison with Other Interpretation Systems. *Japanese journal of infectious diseases*. Vol. 63, No. 2, 87-94.
- [23] Pradhan M. and Sahu R.K. Multilayer perceptron network in HIV/AIDS application. *International Journal of Computer Applications in Engineering Sciences (IJCAES)*, vol. I, Issue I.
- [24] Purwanto et. al. 2011. Improved Adaptive Neuro-Fuzzy Inference System for HIV/AIDS Time Series Prediction. *Communications in Computer and Information Science*, vol. 253, Part 1, 1-13.
- [25] Rebizant W. et. al. 2011. Fundamentals of System Analysis and Synthesis. *Digital Signal Processing in Power System Protection and Control Signals and Communication Technology*. 29-52.
- [26] Resino S. et al. 2011. An artificial neural network improves the non-invasive diagnosis of significant fibrosis in HIV/HCV coinfecting patients. *Journal of Infection*. 62, 77e86.
- [27] Reuter U. and Moller B. 2010. Artificial Neural Networks for Forecasting of Fuzzy Time Series. *Computer-Aided Civil and Infrastructure Engineering*. Vol. 25. Issue 5, 363-374.
- [28] Singh Y. and Mars M. 2010. Support vector machines to forecast changes in CD4 count of HIV-1 positive patients. *Scientific Research and Essays*. vol. 5(17), 2384-2390.
- [29] Sivaram G.S.V.S et. al. 2011. Mixture of Auto-associative Neural Networks for Speaker Verification. *Centre for Language and Speech Processing*. The John Hopkins University.
- [30] Su H. et. al. 2011. Stability Analysis for Stochastic Neural Network with Infinite Delay. *Neurocomputing*. Vol. 74. Issue 10, 1535-1540.
- [31] Wang D. et. al. 2009. A comparison of three computational modelling methods for the prediction of virological response to combination HIV therapy. *Artificial Intelligence in Medicine* vol. 47, 63-74.
- [32] Yoo P.D. et. al. 2010. Hierarchical kernel mixture models for the prediction of AIDS disease progression using HIV structural gp120 profiles. *BMC Genomics* vol. 11, (Suppl 4): S22
- [33] Yu-Dong Cai et. al. 1998. Artificial Neural Network Method for Predicting HIV Protease Cleavage Sites in Protein. *Journal of Protein Chemistry*. vol 17, No. 7.
- [34] azzi M. Et. al. 2011. Prediction of response to antiretroviral therapy by human experts and by the EuResist data-driven expert system (the EVE study). *HIV Medicine*, vol. 12, Issue 4, 211–218.