A Review of Epidemic Forecasting Using Artificial Neural Networks

Philemon Manliura Datilo1,2*, Zuhaimy Ismail1*, Jayeola Dare3

1Department of Mathematical Sciences, Universiti Teknologi Malaysia, Johor, Malaysia
2Department of Information Technology, Modibbo Adama University of Technology, Yola School of Management and Information Technology, Adamawa State, Nigeria
3Adekunle Ajasin University, Department of Mathematical Sciences, Faculty of Science, Ondo State, Nigeria

Abstract

Background and aims: Since accurate forecasts help inform decisions for preventive health-care intervention and epidemic control, this goal can only be achieved by making use of appropriate techniques and methodologies. As much as forecast precision is important, methods and model selection procedures are critical to forecast precision. This study aimed at providing an overview of the selection of the right artificial neural network (ANN) methodology for the epidemic forecasts. It is necessary for forecasters to apply the right tools for the epidemic forecasts with high precision.

Methods: It involved sampling and survey of epidemic forecasts based on ANN. A comparison of performance using ANN forecast and other methods was reviewed. Hybrids of a neural network with other classical methods or meta-heuristics that improved performance of epidemic forecasts were analysed.

Results: Implementing hybrid ANN using data transformation techniques based on improved algorithms, combining forecast models, and using technological platforms enhance the learning and generalization of ANN in forecasting epidemics.

Conclusion: The selection of forecasting tool is critical to the precision of epidemic forecast; hence, a working guide for the choice of appropriate tools will help reduce inconsistency and imprecision in forecasting epidemic size in populations. ANN hybrids that combined other algorithms and models, data transformation and technology should be used for an epidemic forecast.

Keywords: Epidemic, Epidemic forecasting, Artificial neural network, Hybrid ANN

Introduction

Epidemic forecasting is the application of statistical and machine learning techniques to forecast the outbreaks of infectious diseases. Epidemic forecast uses past incidence data to forecast the future epidemic size, peak times and duration. It makes use of demographic, immunological or geographic data of outbreaks. The process does not have to understand the details of disease dynamics, but simply aims at forecasting accurately the future epidemics using suitable methods. Epidemic forecasting is crucial for deploying efficient prevention and control measures of infectious diseases.

Forecasting problems arise in all spheres of life. Predicting future events, on the basis of past history, is carried out using different methods for the purpose of planning and evaluating disease control. Forecasts allow users to have insight before making decisions and taking actions that influence the future of an epidemic.

Different types of linear or nonlinear models are used to forecast epidemics; these models use epidemiological time series data to perform short-term and/or long-term prediction of a plague. Every forecast technique aims at attaining a high precision in foretelling the future, so high confidence is reposed on generalization.

Although variability may occur in performance-measurement tools for forecasting models, preventive healthcare and control decisions for containing an outbreak are ensured by high-precision forecasts. Attaining high forecast precision depends on the use of the right tools and methods in analysis. Making the right choice among different methods of analysing epidemic growth is challenging. Methodologies that should fully incorporate nonlinearity (exhibit and inhibit) to represent the reality of disease dynamics with accurate forecast are often preferred. Therefore, providing insight into the characteristic impact of a particular method for epidemic
forecast will help resolve the problems of selecting epidemic forecast methods. It is the purpose of this study to provide an overview of artificial neural network (ANN) and its appropriateness for epidemic forecasting. This will help analysts make critical decisions in carrying out epidemic forecasts using the artificial neural network.

The study surveyed epidemic forecasts that applied standalone or hybrid ANN\textsuperscript{5,6} in predicting epidemics, with suggested resolution for issues that affect the forecast process. The application of ANN in forecasting deadly infectious disease outbreaks was reviewed. Although most studies considered the multi-layer perceptron feed forward neural network (MLPFFNN), that is not to say that the epidemic forecast cannot be modeled according to other network architectures.\textsuperscript{7} In order to enhance the accuracy of ANN performance, hybrids of ANN are usually used. This study is organized into four sections. Section one provides an introduction to the necessity of making the right choice for epidemic forecast methodology. In section two, an overview of the application of ANN in epidemiological studies is provided, including a summary of epidemic forecasts based on ANNs. Section three explains the critical issues for analysis that require the analyst to make a decision in order to achieve desirable forecast results. The performance of ANN as compared to traditional forecast methods is also summarized. Finally, section four attempts to review ANN hybrids involving a combination of meta-heuristics and classical statistical methods.

Overview of Artificial Neural Network

Neural networks are computational modeling tools that have recently emerged and are widely accepted in many disciplines for modeling complex real-world problems. ANN model is a mathematical evolutionary problem-solving approach that makes decisions (provides solutions) based on the organization concept of signals transmission in the human nervous system. It is a flexible processing model that accepts arbitrary input and generates the output based on nonlinear relationships among variables and parameters.\textsuperscript{8} ANN model learns what to do directly from the problem data, and then renders the output. They are adaptable to problem and data characteristics. The advantages of ANN include ease of optimization, cost-effectiveness, flexibility in non-linear modeling of large datasets, and enhanced accuracy of prediction generalization. Most of the literature on ANN explains the mechanisms and operations of the model.

ANN is developed based on the biological nervous system. Natural neurons receive signals through synapses located on the dendrites or membrane of the neuron. When signals received are strong enough (surpass a certain threshold), the neuron is activated and emits a signal through the axon. These signals might be sent to another synapse and activate other neurons until the final result\textsuperscript{9} (Figure 1). Since ANN is developed based on principles of the natural neuron system, it is an intelligent forecast model.

ANN consists of a set of processing units (nodes) interconnected with arcs indicating relationships between these units, thereby forming a network with input and output terminals. Nodes are simple computing elements using weights of incoming signals, while arcs are the relative weighted influence of each factor on the solution process. Each node has a real-valued bias associated with it. ANN has a parallel and distributed processing structure, as it is presented in Figure 1. ANN was defined and summarized\textsuperscript{10,11} to consist of:

i. A directed graph, which is known as the network topology
ii. A state variable associated with each node
iii. A real-valued weight associated with each link
iv. A real-valued bias associated with each node
v. A transfer function for each node, which determines the state of a node as a function of its bias (or threshold), weights of incoming links, and states of the nodes connected to it by these links.

The basic model of ANN consists of three interconnected layers including the input, hidden, and output layers. Unlike the input or output, the hidden layer could have more than one layer depending on the complexity of the problem. Neurons of input and hidden layer connect to all neurons in the subsequent layer. Neurons of input layers take input data-set from the real environment that is modeled.

The input vector (\(x\)) is transmitted using a connection that multiplies its strength by a weight (\(w_i\)) to produce the product. Each neuron has a bias \(\theta_i\) (see Figure 2). The output of the neuron is generated by transfer and activation functions.

An activation function consists of linear and nonlinear algebraic formulas. These two functions enable the neural network to find the relationships between

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure1.png}
\caption{A Simple Biological Neuron.}
\end{figure}

\begin{figure}
\centering
\includegraphics[width=\textwidth]{figure2.png}
\caption{A Simple Biological Neuron.}
\end{figure}
input and output. The outputs either are sent to other interconnected neurons or directly to the environment. The difference between the output and neural network output is considered as an error.9

Using ANN requires the analysts to decide on algorithms and parameters to facilitate effectiveness and efficiency of model solution strategy. ANN is applied in different fields of knowledge to solve different problems of optimization, classification, and prediction. In the medical domain, ANN has been efficient for clinical diagnosis, image/signal analysis and interpretation and drug development.12-13 In epidemiology, ANN has successfully been used to study dynamics, risk, growth, severity, and control of infectious diseases.6,14 In epidemic forecasting network architecture, all layers and nodes are fully connected, that is, each neuron in one layer is connected to all neurons in the following layer.

Overview of Epidemiology

Epidemiology is broadly termed the science of public health involving quantitative investigation of disease outbreak with the aim of controlling epidemics. Epidemiology provides essential quantitative and analytical methods, principles of logical inquiry, and rules for evidence for disease management. It helps in measuring and projecting the health needs of community and populations and determining how to allocate and manage health care resources. It assesses intervention strategies and evaluates the impact of required health services.15 Canzani and Lechner16 explored mathematical models in epidemiology, with emphasis on theories and methodologies. The study identified core building blocks of models and research patterns to model diseases dynamics in epidemic situations. Moreover, the mathematical modelling of epidemics dynamics was explored.

Different techniques and methods have been developed to investigate epidemics including classification, dynamics, forecast, and intervention or control strategy optimization. Epidemiological information is vital for planning and evaluating epidemic control; therefore, assessing the risk and prevalence of a disease in a population through forecast is necessary. Forecasters have to choose, based on efficiency and performance, among forecasting methods, in order to be able to perform epidemic forecasts. Therefore, a review that highlights potentials and solution approach to forecasting epidemics using the artificial neural network is essential.

The Applications of ANN in Epidemiology

In this section, the study highlights the applications of ANN in epidemiology. In a survey,12 wide application of neural networks in diagnostic approaches in medicine including image analysis, drug design, biochemical analysis, and diagnostic systems was considered. Different clinical studies of diseases using neural networks have also been studied by Dybowski and Gant.17 ANN was used to estimate parameters for a deterministic or stochastic epidemic model of disease control16. Antoniou and Mentzelopoulou18 used neural network to estimate the transmission rate for susceptible-infected-recovered (SIR) epidemic model. However, Samarasinghe and Waidyarathne19 used neural networks to classify and determine the severity of diseases in a population. A radial basis function-based neural network (RBFNN) was developed by Rajalakshmi and Mala20 to diagnose diseases.

Epidemic Classification: In a mathematical sense, this involves dividing an n-dimensional space into various regions, with a given point in the space one should tell which region it belongs to using ANN. Each pattern is transformed into a multi-dimensional point and is classified to a group, each of which represents a known pattern. Clinical observations of diseases tend to be similar at some early stages; further refined characteristics of diseases have to be identified in order to correctly classify a patient or victim of specific infectious diseases. ANN has had tremendous achievements in epidemic classification.20 ANN is an important tool for data mining of medical records for classification and prediction purposes.22 In a large number of previous studies, neural network was used for classifying such diseases as dengue fever,23-25 chest or heart diseases,26,27 West Nile virus diseases,28 tuberculosis,29,30 gestational diabetes mellitus,31 swine flu,32 and pancreatic cancer.33 These studies had helped in diagnosis and case management of epidemic victims.

Epidemic Forecasting: Different methods of epidemic forecasting include the use of ANN, time series analysis, hybrid, and data mining to predict outbreaks of epidemics.2 Forecasting prevalence of diseases using ANN algorithm is a widely accepted approach for epidemiologic investigation of the risk and outbreak size of diseases.34 The recurrent neural network is the most commonly applied model in epidemic forecasting against non-recurrent architecture. The incidence and prevalence of a disease can be predicted35 in order to assess the impact and requirement of control measures. A summary of epidemic forecasts using ANN is provided in Table 1.

ANN Issues for Epidemic Forecasting

Applying ANN in epidemic forecast requires analysts’ decisions that enhance forecast performance. These decisions required choosing from approaches or techniques of ANN algorithm. They include decisions on data pre-processing, network architecture or structure, the number of input, hidden layers or output nodes, training algorithms, specification of training parameter, the number of epoch runs, and accuracy measurement tools.

Data Pre-processing and Normalization

There are various challenges in collecting, recording, and reporting epidemiologic data. If data are limited,
imbalanced, incomplete, highly dimensional,\textsuperscript{36} they have to be pre-processed before performing a forecast with them. Data pre-processing removes noise or outliers, input dimensionality through data transformation, and treatment of non-normal distributed values. For some epidemics that are severe and have a short generation time, ongoing data collection is often marred by under- or over-reporting. Others may be characterized as high-volume data. On the other hand, short epidemic datasets are insufficient for being partitioned for training, testing, and validation. Therefore, random noise can be injected to generate new ones; this is to enhance the robustness of ANN against measurement errors. However, Priddy and Keller\textsuperscript{37} explained some techniques of dealing with limited data problem, and Pasini\textsuperscript{38} developed a counterpart ANN tool for the epidemic forecast with a small dataset.

Data pre-processing include transformation, scaling, and normalization. Normalization is the scaling of data within a uniform range, (0,1) or (-1,1) to reduce the effect of outliers. Different normalization methods are used based on different characteristics of data and problems being investigated. Common normalization methods in performing forecast are the z-score and minimax, sigmoid or soft normalization.\textsuperscript{37,39} Input variables are normalized using z-score normalization as

\[ x'_i = \frac{x_i - \mu_i}{\sigma_i} \]

In an interval \([\lambda_1, \lambda_2]\) input and output, variables \((z)\) can be normalized using minimax as:

\[ x'_i = \lambda_{\text{min target}} + (\lambda_{\text{max target}} - \lambda_{\text{min target}}) \left( \frac{z_i - z_{\text{min target}}}{z_{\text{max target}} - z_{\text{min target}}} \right) \]

The sigmoid normalizations of input variables within the intervals \((0,1)\) and \((-1,1)\) are

\[ x'_i = \frac{1}{1 + \exp \left( -\frac{z_i - \mu_i}{\sigma_i} \right)} \text{ for } (0,1) \]

\[ x'_i = \frac{1}{1 + \exp \left( \frac{\lambda_1 - z_i}{\lambda_2 - \lambda_1} \right)} \text{ for } (-1,1) \]

<table>
<thead>
<tr>
<th>Table 1. Summary of Some Epidemics’ Forecasts Using ANNs</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Epidemic</strong></td>
</tr>
<tr>
<td>Influenza</td>
</tr>
<tr>
<td>HIV/AIDS</td>
</tr>
<tr>
<td>Dengue</td>
</tr>
<tr>
<td>Tuberculosis</td>
</tr>
<tr>
<td>Malaria</td>
</tr>
<tr>
<td>SARS</td>
</tr>
<tr>
<td>Cholera</td>
</tr>
<tr>
<td>Infectious diarrhoea</td>
</tr>
<tr>
<td>Ebola</td>
</tr>
<tr>
<td>Swine Flu</td>
</tr>
<tr>
<td>Cardiovascular disease</td>
</tr>
<tr>
<td>Dementia</td>
</tr>
<tr>
<td>Typhoid fever</td>
</tr>
</tbody>
</table>

BP (back propagation); CF (criterion function); MLP (multi-layer perceptron); FC (fully connected); FF (forward feed); GA (genetic algorithms); GMDH-PNN (group method of data handling-polynomial neural network); LM (Levenberg Marquardt); MAPE (mean absolute percentage error); ME (mean Error); MRE (mean relative error); MSE (mean squared error); PNN (probabilistic neural network); R² (coefficient of determination); RBFS (radial basis forward network); REP (relative error of prediction); RMSE (root mean squared error); # number.
and \( x'_i = \frac{1 - \exp \left( \frac{-x_i}{\sigma_i} \right)}{1 - \exp \left( \frac{-x_i}{\mu_i} \right)} \) for \((-1,1)\).

However, binary and linear transformations are widely applied in the epidemic classification survey. Aburas et al used a statistical method, while Husin et al applied the minimax method for forecasting dengue growth.

**Network Architecture**

ANN architecture for epidemic forecast is problem-dependent, based on the decision variables and parameters of the problem investigated. The number of input nodes, hidden layers and nodes, and output nodes would determine the network architecture for forecasting. The mutual information between variables can help inform the choice of structure.

In most epidemic forecasts, neural network structures are in the form of multi-layered perceptron feed-forward fully-connected network architecture. The interrelationship of causative, disease transmission and control factors in the development of epidemics tends to inform the choice of network architecture. Since the objective of epidemic control is to stop disease growth at a minimal cost, output target should be a single layer/node.

The number of input nodes corresponds to a number of variables in the input vector used to forecast future values. Too few or too many input nodes can affect learning or prediction capability of the network. The sizes of these inputs depend on vector, disease transmission, and control. Environmental, climatic/weather, demographic, time and type intervention and control measures, disease dynamics, risk factors, vector/reservoir, and symptoms or signs affect the outbreak and spread of a disease. These factors are combined with cases data to perform outbreak forecast. A number of hidden layers and nodes influence training and forecast generalization. Epidemic incidence data do not indicate clinical or pathological information about diseases; therefore, hidden layers and nodes are expected to explain the nonlinearity in the epidemic forecast model. Output node for epidemic forecast generates incidence, recovery, and death. Forecast output will spell dataset required to train the entire network. Forecasting infective cases would require infection incidence data, while death or recovered cases will require mortality or control data.

**Activation Function**

Activation function, also called transfer function, can

<table>
<thead>
<tr>
<th>Name</th>
<th>Function</th>
<th>Characteristics</th>
</tr>
</thead>
<tbody>
<tr>
<td>Binary threshold</td>
<td>( \delta(x_j) = \begin{cases} 1 &amp; x_j \geq 0 \ 0 &amp; x_j &lt; 0 \end{cases} )</td>
<td>Non-differentiable step-like ( x_j \in {0,1} )</td>
</tr>
<tr>
<td>Bipolar threshold</td>
<td>( \delta(x_j) = \begin{cases} 1 &amp; x_j \geq 0 \ -1 &amp; x_j &lt; 0 \end{cases} )</td>
<td>Non-differentiable step-like ( x_j \in {-1,1} )</td>
</tr>
<tr>
<td>Linear</td>
<td>( \delta(x_j) = \alpha x_j )</td>
<td>Differentiable unbounded ( x_j \in {-\alpha,\alpha} )</td>
</tr>
<tr>
<td>Linear threshold</td>
<td>( \delta(x_j) = \begin{cases} 0 &amp; x_j &lt; 0 \ \alpha x_j &amp; 0 &lt; x_j &lt; x_m \ 1 &amp; x_j \geq x_m \end{cases} )</td>
<td>Differentiable, monotonic, smooth ( x_j \in {0,1} )</td>
</tr>
<tr>
<td>Sigmoid</td>
<td>( \delta(x_j) = \frac{1}{1 + e^{-\lambda x_j}} )</td>
<td>Differentiable, monotonic, smooth</td>
</tr>
<tr>
<td>Hyperbolic tangent</td>
<td>( \delta(x_j) = \tanh(\lambda x_j) ) or ( \delta(x_j) = \frac{e^{x_j} - e^{-x_j}}{e^{x_j} + e^{-x_j}} )</td>
<td>Differentiable, monotonic, smooth ( \delta(x_j) = \sin(x_j) )</td>
</tr>
<tr>
<td>Sine or cosine</td>
<td>( \delta(x_j) = \sin(x_j) ) or ( \delta(x_j) = \cos(x_j) )</td>
<td>Differentiable, monotonic, smooth ( x_j \in {0,1} )</td>
</tr>
<tr>
<td>Gaussian</td>
<td>( e^{(x_j-x_m)^2/2\sigma_j^2)} )</td>
<td>Differentiable, monotonic, smooth ( x_j \in {0,1} )</td>
</tr>
<tr>
<td>Stochastic</td>
<td>( \delta(x_j) = \frac{1}{-1} )</td>
<td>Non-differentiable step-like ( x_j \in {0,1} ) or {-1,1}</td>
</tr>
</tbody>
</table>
be linear or nonlinear. The function determines the relationship between inputs and outputs of a node or output of an entire network. In general, the function introduces a degree of nonlinearity, thereby enhancing the performance of neural networks. There are a number of activation functions that have been used to train a neural network to perform epidemic forecast, optimization, classification, pattern or speech recognition, image compression, and selection. Although Zhang et al argued that there are no straight rules for making choice, Ittiyavirah et al provided a guideline for using a particular type of activation function for specific types of forecast models in neural network problems. The use of one or more activation functions in the same or different layers depends on the characteristics of forecast problems. Table 2 summarizes different types of activation functions that can be used in neural networks modeling.

The nonlinearity of the problems of epidemic forecast associated with multiple factors and complex relationships in analyzing epidemic growth or dynamics warrants the use of nonlinear activation functions. Epidemic forecast deals with real and future occurrence; hence, nonlinear nonnegative interpretable activation functions should be used for growth parameters of the disease. The sigmoid (logistic) transfer function is commonly used. This function is preferred because it is bounded, monotonically increasing, and differentiable. In most of the previous studies, the sigmoid (logistic) activation function was generally used at all levels for forecasting epidemics, as summarized in Table 1.

Training Algorithms

The neural network training is an unconstrained nonlinear minimization problem in which weights of networks are iteratively modified to minimize overall mean or total squared error between desired and actual output values for all nodes over all inputs. One of the factors that affect the performance of neural networks is the learning algorithm. Several training algorithms can be applied to train ANNs for different time series data. However, most popular ones implemented in various neural network software tools include Back Propagation, Levenberg-Marquardt, Quick Propagation, Conjugate Gradient, Competitive Learning, and Quasi-Newton. Convergence rate, lack of robustness, and local optima of unconstrained searches are often used to compare performances of learning algorithms. Training can be carried out under a supervised, unsupervised, or reinforcement activity.

Learning or training includes specifying a number of hidden nodes of layers, learning rate, momentum, training cycles, and the size of training or test datasets. Over or under-specification of these parameter values has a direct effect on training performance. Correct specification of training parameter is crucial to timeliness or global convergence, accuracy, and generalization of the forecast using ANN model. Various training algorithms are commonly used in software platforms for neural network applications. The most widely used learning algorithms in the epidemic forecast are back propagation, Levenberg Marquardt, and genetic algorithm for hybrid ANN. Back propagation is a gradient descent algorithm that requires specifying step size or learning rate for the learning process and a slow convergence rate, inefficiency and lack of robustness. Levenberg Marquardt algorithm is an efficient training algorithm with minimal error and faster rate of convergence, but its computational complexity of the thrust region mars its wide adoption. The implementation of a genetic algorithm in training neural networks (optimizing weights) enhances the performance of ANN in forecasting compared to some traditional statistical forecast methods. Hybrid training algorithms have more advantages in improving the performance of ANN.

Underfitting and Overfitting Problem

The learning of ANN forecast model affects precision; in order to achieve higher precision, training has to be done using the appropriate techniques. Underfitting and overfitting affect neural network generalization.

Underfitting occurs when the model is not able to obtain a sufficiently low error value on the training set.

Overfitting occurs when the gap between the training error and test error is too large. Analysts can control whether a model is more likely to overfit or underfit by altering its capacity (data, algorithm, and architecture).

An under-trained model often has high sum of squared error (SSE) values for either or both training and validation datasets. Under-training often occurs when there are insufficient data for model fitting. Overfitting (over-trained) occurs when the network fits the in-sample data well but produces poor out-of-sample results. Forecasters can use the following equation to assess the absence of underfitting.

\[ \epsilon_1 = \frac{(SSE_k - SSE_{k-1})}{SSE_k} \leq \epsilon_1^* \]

To avoid overfitting, Ghiassi et al recommended that at each iteration \( k \), \( k > 1 \), MSE values should be computed for both training (MSET) and validation (MSEV) datasets and then use \( \epsilon_2 \) relation below, to guard against overfitting.

\[ \epsilon_2 = \frac{|MSE_T - MSE_V|}{MSE_T} \leq \epsilon_2^* \]

The accuracy levels of \( \epsilon_1 \) and \( \epsilon_2 \) are problem-dependent and should be determined experimentally. The analyst should consider fully trained network when the user-specified accuracy criteria and the over-fitting constraint are both satisfied.

A modern approach for solving over-fitting problem is the...
cross validation technique. There are different techniques for cross validation, among which are leave-one-out, leave-
p-out, k-fold, holdout, and Monte Carlo cross validation. In forecasting practices, the k-fold cross validation is commonly used. K-fold cross-validation involves random classification of the dataset into k-partitions, the kth subset is used as a test set and the model is trained on all the (k-
1) partitions. The test sum error on the selected partition is evaluated and then averaged to produce a single estimation. Though k is arbitrary, 10-fold cross validation is commonly used.

Forecast Performance Measurement
ANN performance is assessed with different error measurement tools in training, validation and testing data sets. It is notable that there can be many performance measures for an ANN forecast, including modeling and training time. For instance, prediction accuracy is considered as the most important measure of performance that the model can achieve beyond training data. Forecast accuracy measurement has been a growing concern among forecast theorists and practitioners. There is no singular accuracy measurement that is universally accepted; therefore, different accuracy measurements are used in epidemic forecasting, with their own strengths and weaknesses. The calculation speed, interpretability, asymmetry, impact of outliers, biasedness, scale dependence, and division by zero (in the case of equal

<table>
<thead>
<tr>
<th>Table 3. Error Measurement Types</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean root error (MRE)</td>
</tr>
<tr>
<td>Root mean square error (RMSE)</td>
</tr>
<tr>
<td>Mean square error (MSE)</td>
</tr>
<tr>
<td>Mean absolute error (MAE)</td>
</tr>
<tr>
<td>Mean absolute percentage error (MAPE)</td>
</tr>
<tr>
<td>Pearson Method (R^2)</td>
</tr>
<tr>
<td>Root error percentage (REP)</td>
</tr>
</tbody>
</table>

Where \( e_i' \) denotes the difference between forecast and realized values.

ANN Methodology for Epidemic Forecasting
There are many different ways to construct and implement a neural network for epidemic forecasting. Epidemic forecasts help monitor the growth of epidemic and work to reduce the spread of disease. In most of the previous forecasts as well as this survey, the networks are fully connected, and numerous input nodes against a single output node with hidden layers and nodes are considered. No definite number can suit the peculiarity of variables in every disease forecast. For finding the optimal architecture of an ANN, selection of these parameters is usually quite complex in nature and is difficult to implement. Up to now, there has been no clear-cut method for determining these parameters, except by trial and error. However, Zhang et al. provided guidelines based on heuristics and simulations.

Forecasting epidemics with neural networks involves training an ANN structure with a portion of epidemiological dataset to validate, and then, the remainder of the dataset was used to test and generalize outbreak size in the future. The process of the epidemic forecast is summarized in the following pseudo algorithm. The methods for architectures that implement hybrids for forecast differ and thus the hybrid technique should be incorporated at the right step according to the adopted hybrid technique (i.e., technology, data, model, and algorithm).

Undertaking an epidemic forecast using ANN is challenging because while epidemics could be ongoing, it is demanded to assess their threat; hence, findings will only provide an interim guide to the control of epidemics. Until the recent study of Pasini, larger datasets were preferred in applying ANN since epidemic forecast using small dataset suffered from generalization.

Hybrid ANN for Epidemic Forecast
The hybrid model involved a combination of two or more independent models or algorithms. In order for ANN to computationally be efficient in solving complex and computationally expensive problems, Basheer and Hajmeer and Nedjah et al. suggested that ANN be hybridized. In order to hybridize ANN for the epidemic
forecast, other forecast models or algorithms can be combined with ANN for high-precision performances. Hybrid ANN could include technology, models, algorithms, or data. Model hybrid is required when the mathematical models include different typology functions or even different statistical learners. Algorithms hybrid is required when the learning procedures take advantage of traditional methods and heuristics, while data hybrid of ANNs is obtained from heterogeneous data structures. Technology hybrid with ANN involves the development of ANN forecasting platform based on information technologies paradigm. This hybrid takes advantage of the technology and performs fast epidemic classification and forecast. Technology hybrid ANN solves the problem of machine (both software and hardware) factor in the model; it has been found to be efficient. Cloud-based technology implementing ANN model at the core was applied to analyze medical data to evaluate and classify epidemics. Similar technology is used in dengue tracking and prediction by Makridakis and Hwang et al. ANNs are recommended for every real-life problem for a high forecast performance index.

Hybrid ANN involves using other nonlinear forecast models such as Fuzzy, ARIMA, and logistics regression. High performance forecast models should be chosen to combine with a neural network to attain high precision. Hybrids support combination techniques of forecast models proposed by M3 competition. Hybrid ANN model overcomes weaknesses of incorporated models, minimizes forecast error, and fastens error optimization convergence. Combination forecasts are useful when analysts are uncertain about which method is the most accurate, while they want to avoid large errors.

Hybrid ANN algorithm for epidemic forecast will require incorporating other improved search or optimization algorithms in ANN structure, weight determination algorithms, learning algorithm or error minimization routines and time-dependent activation functions. Algorithm hybrids use heuristics or combinatorial algorithms in achieving high convergence for desirable performance, within optimal run time. Algorithms from a different field or model can be used to train a forecast process. The use of metaheuristic optimization algorithms including evolutionary algorithms, genetic algorithms, evolutionary programming, particle swarm optimization, ant colony optimization, bird mating optimizer (BMO), and bat algorithm improved ANN training procedure in epidemic forecast.

Hybrid ANN was implemented by Kanjamapornkul to transform non-stationary Ebola data to stationary data using Hilbert-Huang Transform (HTT) before using grey-model and neural network to forecast the prevalence of the virus in the most affected countries. Arifianto et al. used group method data handling (GMDH) with the polynomial neural network to predict malaria incidences in South East Asia.

Several infectious and non-infectious diseases have been forecasted using hybrid ANN in achieving high precision. Genetic algorithms were applied to train ANN in forecasting dengue fever in Malaysia, the model was better than the traditional approaches. Similarly, Belciug and Gorunescu implemented genetic algorithm weight training upon multi-layer perceptron neural network to assess the detection and recurrence of breast cancer. Gan et al. combined the grey model (GM) and back propagation (BP) neural network to predict the growth of hepatitis B, where the results proved that the hybrid has an advantage over GM (1,1) and GM (2,1) in performance evaluations.

In forecasting schistosomiasis, Zhou et al. used ARIMA-NARNN to predict prevalence and found that hybrid ANN outperformed both individual models. However, Yan et al. used seasonal ARIMA and generalized regression neural network to predict bacillary dysentery. Similar hybrid is applied for prediction of tuberculosis with comparable performance.

The weaknesses of independent traditional forecast method are often overcome in hybridization; however, modelers have to ensure that better models with relatively high performance are chosen to hybridize ANN. If more hybrid techniques (data, algorithm model, and technology) can be combined in a single forecast project, epidemic forecast precision will be highly boosted.

**Performance of ANN in Epidemic Forecasts**

ANNs have had considerable advantages over traditional statistical forecast methods. The comparison of the performance of ANN model with other methods in forecast generalization has been contentious because of differences in the nature of problem and data type, run time and implementation of payload, parameterization requirement of algorithms (learning and training activation functions), and ease of adoption. The M3 and NN3 competition organizers generalize that:

- the characteristics of the data series are important factors in determining the relative performances of different methods;
- the sampling variability of the performance measures renders comparisons which are unreliable based on single time series: comparisons based on multiple time origins are recommended;
- sophisticated or complex methods do not necessarily provide more accurate forecasts than simpler ones;
- the relative ranking of the performance of the various methods varies according to the accuracy measure being used;
- combinations of prediction methods tend to be quite accurate, and often outperform an individual method;
- the accuracy of various methods depends upon the length of forecasting horizon involved.
In line with these guidelines and for purpose of improving the precision of epidemic forecast, hybridizing ANN is necessary. Comparative studies of different forecasting techniques facilitate best model selection for forecasting future epidemic behavior in specific diseases. Several ANN techniques (e.g., Radial Basis Function NN, Probabilistic NN, and back propagation NN) have been compared to determine the efficiency and suitability of ANN models and their variants in epidemic forecast generalization. While Zhang et al compared ANN variants with SARIMA and established that RBFNN, ERNN, and BPNN outperform SARIMA forecast technique, they summed that ANN can outperform traditional methods when forecasting with increased horizon, using short memory series, and having more input nodes in its architecture. A similar conclusion was drawn when Paliwal and Kumar reviewed the performance of ANNs in comparison with several traditional statistical methods of forecasting in health and medical sciences. When ANN was compared with other nonlinear regression models in predicting dengue outbreak, it was established that the neural network performed better, and hybrid neural network tends to supersede. Similar conclusions were drawn in studies on malaria. ANN was found to perform better compared to support vector regression (SVR), random forests regression (RFR) and multivariate linear regression (MLR) in predicting infectious diarrhea in China. The application of hybrids of ANN based on different technologies, models, algorithms, and data has enhanced ANN performance in forecasting diseases more than classical or traditional forecast methods, as summarized in Table 4.

### Conclusion

The measure of accuracy in predicting epidemic cases or mortality is important in deciding on the intervention or control strategy. Therefore, clarifying model characteristics would facilitate the selection of a model for forecast application. In this review, we presented the use of ANN in the epidemic forecast. Several disease forecasts were summarized based on characteristic issues in performing a forecast using ANN. A considerable amount of research has been done in the epidemic forecast that supports the use of ANN in epidemic studies. The use of ANN is not limited to prediction, but it includes classification, optimization, parameter estimation, and optimal disease control in epidemic management.

Critical issues that require the analysts to decide among optimal options for enhancing ANN performance in the study of epidemic growth were elaborated. In generalizing the prediction of epidemic growth, ANN has outperformed traditional forecast methods; therefore, developments of hybrid neural network based on different techniques were applied to several diseases to improve the accuracy of the epidemic forecast. Latest re-emergent epidemics like Ebola, Zika, Middle East Respiratory Syndrome (MERS), and Lassa fever can be forecasted using ANN

#### Table 4. Summary of Hybrids in Predicting Some Epidemics

<table>
<thead>
<tr>
<th>Disease</th>
<th>Ref.</th>
<th>Hybrid ANN</th>
<th>Hybrid Technique</th>
<th>Component Models</th>
<th>Component Algorithms</th>
<th>Accuracy Measurement Tools</th>
<th>Hybrid Prediction Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Diabetes</td>
<td>83</td>
<td>ANFIS</td>
<td>Algorithm</td>
<td>Fuzzy logic</td>
<td>Gradient and combined algorithm</td>
<td>RMSE</td>
<td>10% and 5% increase in validity and testing</td>
</tr>
<tr>
<td>Tuberculosis</td>
<td>84</td>
<td>SARIMA-NNAR</td>
<td>Model</td>
<td>ARIMA</td>
<td>MSE, RMSE, MASE, MAE, MAPE</td>
<td>Lower AIC</td>
<td>94.88% classification accuracy</td>
</tr>
<tr>
<td></td>
<td>26</td>
<td>GA and MLNN</td>
<td>Algorithm</td>
<td>Genetic Algorithm</td>
<td>BP and gray wolf optimization</td>
<td>RMSE</td>
<td>High quality performance with shortest epochs</td>
</tr>
<tr>
<td>Heart disease</td>
<td>6</td>
<td>ANN-GWO</td>
<td>Algorithm</td>
<td></td>
<td>R, MAE, MAPE, SSE, RMSE, MSE</td>
<td>Advantageous over GM(1,1) and GM(2,1)</td>
<td></td>
</tr>
<tr>
<td>Hepatitis B</td>
<td>81</td>
<td>Grey Model-NN</td>
<td>Model</td>
<td>Grey Model</td>
<td>GM and Back propagation</td>
<td>RAE, MAE, RMSE, MAE/MAPE</td>
<td>Better than their compositional models</td>
</tr>
<tr>
<td>Hepatitis</td>
<td>85</td>
<td>ARIMA-GRNN</td>
<td>Model</td>
<td>ARIMA</td>
<td>MAE/MAPE, MSE/MAE</td>
<td>Best with lower MAE, MAE, MSE</td>
<td></td>
</tr>
<tr>
<td>Schistosomiasis</td>
<td>32</td>
<td>ARIMA-NARN</td>
<td>Model</td>
<td>ARIMA</td>
<td>Back-propagation and ARIMA (p, d, q)</td>
<td>MAE, MSE, MAPE</td>
<td></td>
</tr>
<tr>
<td>Pima Indian diabetes</td>
<td>59</td>
<td>GA-BPN</td>
<td>Algorithm</td>
<td></td>
<td>GA-correlation based feature selection (GA-CFS)</td>
<td>MAE, MSE, MAPE</td>
<td>84.73% precision</td>
</tr>
<tr>
<td>Cardiometabolic risk</td>
<td>86</td>
<td>EANN-EA</td>
<td>Model</td>
<td>EANN</td>
<td>Evolutionary Algorithm</td>
<td>MAE, MSE, MAPE</td>
<td>&gt;90% precision</td>
</tr>
<tr>
<td>Malaria</td>
<td>36</td>
<td>GMDH-PNN</td>
<td>Data</td>
<td>Polynomial NN</td>
<td>Group method data handling</td>
<td>MAE, MSE, MAPE</td>
<td>88.02% precision with 72% reduced training time</td>
</tr>
<tr>
<td>Ebola</td>
<td>42</td>
<td>HHT-ANN</td>
<td>Data</td>
<td></td>
<td>Hilbert Huang transform</td>
<td>Percentage relative error (PRE)</td>
<td></td>
</tr>
</tbody>
</table>
with improved hybrid techniques to provide accurate quantitative information for epidemic control decisions.

**Ethical Approval**
Not applicable.

**Conflict of Interest Disclosures**
None.

**References**


