Asian Journal of Information Technology 10 (6): 209-217, 2011
ISSN: 1682-3915

An Optimally Configured Hybrid Model for Healthcare Time Series Prediction

[1,3]Purwanto, [1]Chikkannan Eswaran and [2]Rajasvaran Logeswaran
[1]Faculty of Information Technology, [2]Faculty of Engineering,
Multimedia University, 63100 Cyberjaya, Malaysia
[3]Faculty of Computer Science, Dian Nuswantoro University, 50131 Semarang, Indonesia

Abstract: The challenge in improving accuracy in time series prediction has motivated researchers to develop more efficient prediction models. Prediction of healthcare data such as mortality and morbidity assumes importance in healthcare management as these data serve as health indicators of a society. The accuracy rates obtained using linear models such as autoregressive integrated moving average and linear regression are not high as they have limitations in handling the non-linear relationships among the data. Neural network models are considered to be better in handling such non-linear relationships. Healthcare time series data consist of complex linear and nonlinear patterns and it may be difficult to obtain high prediction accuracy rates using only linear or neural network models. The researchers propose a hybrid method which combines the best linear model with an optimally configured neural network. Unlike other hybrid models which use a predetermined configuration for the linear and neural network components, the proposed method selects the best linear model and the optimum neural network configuration based on the type of input data. The proposed method is tested based on two types of healthcare data, namely infant mortality rate and morbidity of malaria data. Experiment results show that the proposed hybrid model yields more accurate prediction results compared to the other known models.

Key words: Hybrid method, neural network, moving average, weighted moving average, linear regression, Autoregressive Integrated Moving Average (ARIMA)

INTRODUCTION

The data related to infant mortality rate and morbidity of malaria serve as indicators of general health status of a community. Infant Mortality Rate (IMR) refers to the number of infant deaths in relation to the number of live births in a given year (Shiel and Stoppler, 2008). The Morbidity Of Malaria (MOM) refers to the number of people who suffer from malaria. The IMR reflects a broad range of social, economic and medical conditions. Communities with problems such as unemployment, poverty and low literacy tend to have higher infant mortality rates (Condran and Murphy, 2005). In Indonesia, mortality and morbidity statistics are important instruments to monitor and evaluate the national health development programs.

Time series forecasting is employed in various applications to predict the future value using the past time series data. Many models have been applied in time series forecasting such as autoregressive integrated moving average (Katimon and Demun, 2004), linear regression (Moghaddas-Tafreshi and Farhadi, 2008; Bing-Jun and Chun-Hua, 2007) and neural networks (Niska *et al*., 2004).

Time series prediction models have been used to predict mortality and morbidity (McNown and Rogers, 1992; Abeku *et al*., 2002; Nkurunziza *et al*., 2010). The ARIMA model has been applied to forecast the complete age profile of mortality (McNown and Rogers, 1992) and also malaria incidence (Abeku *et al*., 2002). Multiplicative exponential smoothing method has also been used for forecasting malaria cases (Nkurunziza *et al*., 2010). Unfortunately, the accuracy rates obtained by applying these models are not sufficiently high as they have restrictions in handling non-linear patterns. Neural networks are believed to be more suitable for handling non-linear patterns but they yield mixed results in handling linear patterns (Zhang, 2003). In the real world scenario, time series data comprise complex linear and nonlinear patterns. Thus, it is difficult to obtain high prediction accuracy rates by applying only linear or neural network models. It is expected that hybrid models combining both linear and neural network models can be employed to obtain high prediction accuracy rates.

Hybrid models combining ARIMA and neural networks have been used for some prediction applications (Aladag *et al*., 2009; Faruk, 2010; Diaz-Robles *et al*., 2008).

Corresponding Author: Purwanto, Faculty of Information Technology, Multimedia University, 63100 Cyberjaya, Malaysia

Hybrid models have been found to yield more accurate results for time series prediction compared to individual ARIMA and neural network models (Zhang, 2003; Faruk, 2010). Hybrid models using Exponential Smoothing (ES) and neural network models have been employed to predict financial time series and found to perform better than the individual models (Lai *et al*., 2006). However, from the results of use of various hybrids for different applications, it is clear that the performances of these models vary depending on the type of data used.

As the prediction model depends on the pattern of the data, it is necessary to determine the appropriate models for the linear and non-linear parts based on the pattern of data when developing a hybrid model.

In all the previous reports, predetermined linear and neural network models have been used without taking into account the type or complexity of the data (Zhang, 2003; Aladag *et al*., 2009; Diaz-Robles *et al*., 2008; Lai *et al*., 2006).

In this study, the researchers propose a hybrid method in which the best linear model as well as the best neural network configuration is identified based on the complexity of the input data. The best or optimum neural network configuration is selected with respect to the number of input and hidden layer neurons and the activation functions used for the hidden and output layers. The researchers also present a performance comparison of different models for a specific application, namely, prediction of Infant Mortality Rate (IMR) and Morbidity of Malaria (MOM). The prediction or forecasting of IMR and MOM assumes importance for health departments since, a good and accurate prediction is very helpful in devising appropriate action plans. Healthcare data in general, consist of complex linear and non-linear patterns and it may be difficult to obtain high prediction accuracy rates using only linear or neural network models. Hence, the researchers propose procedure hybrid model which combines the best linear model with the optimum neural network to obtain more accurate prediction.

MATERIALS AND METHODS

The basic concepts of Moving Average (MA), Linear Regression (LR), ARIMA, Neural Network and Hybrid models implemented are described.

A number of researchers still use traditional statistical methods such as moving average, exponential smoothing and ARIMA for time series forecasting (Diaz-Robles *et al*., 2008). Box and Jenkins developed an efficient fitting procedure based on the ARIMA methodology which is used nowadays as a standard for linear time series modeling (Box and Jenkins, 1976). In time series forecasting, it is known that the data arising out of many phenomena are nonlinear in nature. In other words, there exists a nonlinear relationship between the past and the present data. In such situations, the linear time series forecasting models are proved to be inadequate and as a consequence, nonlinear time series models such as neural networks have received a great deal of attention in the last few years. Examples of neural networks used for time series forecasting can be found in Niska *et al*. (2004).

Moving Average (MA) model: Average methods make use of past time series data to develop a system for forecasting. There are many moving average methods that are used for forecasting such as simple Moving Average (MA) and Weighted Moving Average (WMA). The MA method is one of the most widely used methods for prediction.

In simple MA (m) which uses past m consecutive observations, the future value $\hat{y}_{t+1}$ is computed as:

$$\hat{y}_{t+1} = \frac{1}{m}\sum_{i=t-m+1}^{t} y_i \tag{1}$$

WMA uses different weights for the past observations as shown in Eq. 2:

$$\hat{y}_t = w_1 y_{t-1} + w_2 y_{t-2} + w_3 y_{t-3} + \ldots + w_m y_{t-m} \tag{2}$$

where, $w_1, w_2, \ldots, w_m$ are the weights associated with the past observed values and:

$$\sum_{i=1}^{m} w_i = 1$$

Linear Regression (LR) model: Linear Regression (LR) is one of the most common forecasting models. The regression model describes the mean of the normally distributed dependent variable y as a function of the predictor or independent variable x (Le, 2003):

$$y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \tag{3}$$

Where:

y_i	=	The value of the response or dependent variable from the ith pair
β_0 and β_1	=	The two unknown parameters
x_i	=	The value of the independent variable from the ith pair
ε_i	=	A random error term

The predicted, estimated or fitted values of the regression model are calculated as (DeCoster, 2003):

$$\hat{y}_i = b_0 + b_1 x_i \tag{4}$$

The parameters b_0 and b_1 in the Eq. 5 are computed as:

$$\left.\begin{aligned} b_0 &= \frac{\sum_{i=1}^{n} y_i}{n} - b_1 \frac{\sum_{i=1}^{n} x_i}{n} \\ b_1 &= \frac{n\sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n\sum_{i=1}^{n} x_i^2 - (\sum_{i=1}^{n} x_i)^2} \end{aligned}\right\} \tag{5}$$

This model is referred to as the LR model. It is linear because the independent variable appears only in the first power and if we plot the mean of y versus x, the graph will be a straight line with intercept b_0 and slope b_1.

Autoregressive Integrated Moving Average (ARIMA) model: Time series analysis has several objectives such as modeling, forecasting and controlling. Forecasting deals with the issue of constructing models and methods that can be used to produce accurate short-term predictions (Chatfield, 2000).

The forecast in Autoregressive (AR) model is a function of its past observations and the forecast in a Moving Average (MA) model is a function of its past residuals. An Autoregressive (AR) model of order p, AR (p) is given by:

$$\hat{x}_t = \phi_1 x_{t-1} + \phi_2 x_{t-2} + \ldots + \phi_p x_{t-p} \tag{6}$$

A moving Average (MA) model of order q, MA (q) is given by:

$$\hat{x}_t = \theta_1 e_{t-1} + \theta_2 e_{t-2} + \ldots + \theta_q e_{t-q} \tag{7}$$

The order of the model is defined by the highest term present in the describing equation. The order of an autoregressive polynomial is denoted by p and that of a moving average polynomial is denoted by q.

Autoregressive Moving Average (ARMA) models are created from a finite, linear combination of past values of the series and a finite linear combination of past residuals. An ARMA (p, q) model is defined by (Brockwell and Davis, 2002):

$$x_t = \sum_{i=1}^{p} \phi_i x_{t-i} + e_t + \sum_{j=1}^{q} \theta_j e_{t-j} \tag{8}$$

Where:

e_t = The random error at time t, ϕ_i (i = 1, 2, ..., p)

θ_j (j = 1, 2, ..., q) = The model parameters are to be estimated. The orders of autoregressive and moving average polynomials are p and q, respectively

The autoregressive moving average ARMA (p, q) model could be expressed as:

$$\phi(B)x_t = \theta(B)e_t \tag{9}$$

where, ϕ (.) and θ (.) are the pth and qth degree polynomials given by (Brockwell and Davis, 2002):

$$\left.\begin{aligned} \phi(z) &= 1 - \phi_1 z - \phi_2 z^2 - \ldots - \phi_p z^p \\ \theta(z) &= 1 + \theta_1 z + \theta_2 z^2 + \ldots + \theta_q z^q \end{aligned}\right\} \tag{10}$$

ARMA model employed for time series data using ordinary differencing is called an Autoregressive Integrated Moving Average (ARIMA) models. An ARIMA model with differencing order d, ARIMA (p, d, q) can be expressed as (Brockwell and Davis, 2002):

$$\phi(B)(1-B)^d x_t = \theta(B)e_t \tag{11}$$

where, B represents the backshift operator.

Neural network model: In recent years, Neural Networks (NN) for forecasting have been investigated by many researchers (Aladag *et al*., 2009; Diaz-Robles *et al*., 2008). The motivation for using neural networks is based on the fact that these models are capable of handling non-linear relationships with considerable success in non-linear domains.

An NN model can consist of different layers which are connected to each other by connection weights or parameters. Between the extremities of the input layer and the output layer are the hidden layers. The nodes in each layer are connected by flexible weights which are adjusted based on the error or bias.

The function of the input layer is for data entry, data processing takes place in the hidden layer and the output layer functions as the data output result. Figure 1 shows the architecture of the very popular multilayer backpropagation neural network (MLP). The response value Y (x) is computed as:

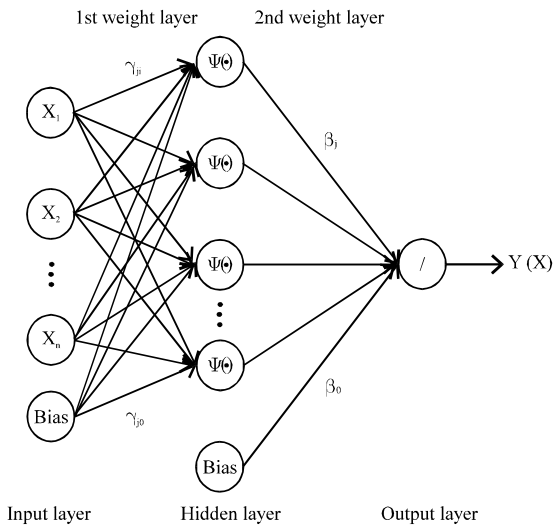


Fig. 1: The Multilayer Backpropagation Neural Network architecture

$$Y(x) = \beta_0 + \sum_{j=1}^{H} \beta_j \psi(\gamma_{j0} + \sum_{i=1}^{n} \gamma_{ji} x_i) \quad (12)$$

where, $(\beta_0, \beta_1, \ldots, \beta_H, \gamma_{10}, \ldots, \gamma_{Hn})$ are the weights or parameters of the NN. The non-linearity enters into the function Y (x) through the activation function ψ.

Information processing in every neuron is done by summing the multiplication result of connection weights with input data. The result is transferred to the next neuron through the activation function. There are several kinds of activation functions ψ such as sigmoid, bipolar sigmoid and hyperbolic tangent as given in Eq. 13-15, respectively (Kusumadewi and Hartati, 2006):

Sigmoid:

$$\psi(x) = \frac{1}{1+e^{-\alpha x}} \quad (13)$$

Bipolar sigmoid:

$$\psi(x) = \frac{1-e^{-x}}{1+e^{-x}} \quad (14)$$

Hyperbolic tangent:

$$\psi(x) = \tanh(x) = \frac{e^{x}-e^{-x}}{e^{x}+e^{-x}} \quad (15)$$

Hybrid models: Linear and NN models have achieved successes in their linear and non-linear domains, respectively. However, the use of linear models for complex non-linear problems may not yield accurate results. On the other hand, using NN models for linear problems has yielded mixed results (Zhang, 2003). Since, it is difficult to know the characteristics of the actual data in a real problem, hybrid methodology that has both linear and non-linear modeling capabilities can be a good strategy for obtaining accurate prediction.

In general, a time series data is composed of a linear autocorrelation structure and a non-linear component as shown in Eq. 16 (Zhang, 2003):

$$y_t = L_t + N_t \quad (16)$$

where, L_t and N_t represent the linear and non-linear components, respectively. The hybrid methodology consists of two steps (Zhang, 2003). In the first step, ARIMA is used to predict the linear part ($\hat{L}_t$) of the problem and in the second step, a NN model is used to predict the residuals ($\hat{N}_t$) from the linear model. The residual or error series e_t is obtained with Eq. 17:

$$e_t = y_t - \hat{L}_t \quad (17)$$

The overall prediction of the hybrid model is given in Eq. 18:

$$\hat{y}_t = \hat{L}_t + \hat{N}_t \quad (18)$$

Where $\hat{y}_t$ represents the combined forecast value of the hybrid model at time t.

Performance measures: The prediction models are evaluated in terms of their ability to predict the future values. Several measures are used to compare the forecasting performance of different models. The main measures are the Root Mean Square Error (RMSE) and the Mean Absolute Percentage Error (MAPE).

MAPE produces a measure of relative overall fit (Aladag *et al*., 2009) is computed as:

$$MAPE = \frac{\sum_{t=1}^{n} \frac{\left|Y_t - \hat{Y}_t\right|}{Y_t} * 100}{n} \quad (19)$$

RMSE measures the average magnitude of the error (Aladag *et al*., 2009; Rojas *et al*., 2008) and is computed as:

$$RMSE = \sqrt{\frac{\sum_{t=1}^{n} \left(Y_i - \hat{Y}_i\right)^2}{n}} \quad (20)$$

Both measures are used for comparison of model accuracy where lower values of RMSE and MAPE are desired for better prediction.

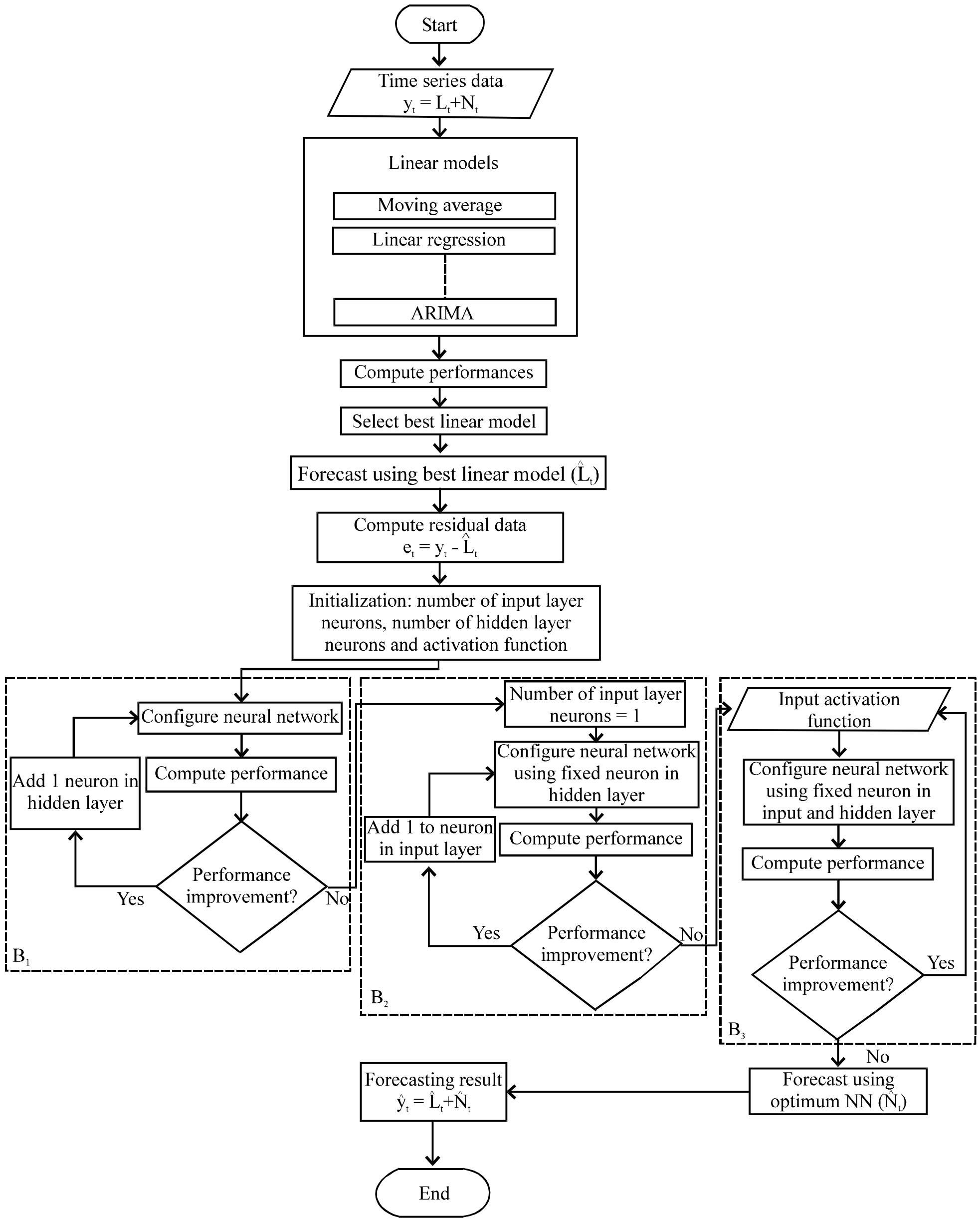


Fig. 2: The proposed Hybrid method using the best Linear and Neural Network models

Proposed hybrid method

Data: The healthcare data used in this study were collected from the Indonesian health profile and other established sources (MHRI, 2008; HMI, 2008).

Morbidity refers to the total number of people who suffer from a certain disease such as malaria, tuberculosis.

In this research, the reseachers use morbidity of Malaria (MOM-clinical Malaria case) in Indonesia (outer Java-Bali) during the period 1989-2007 and infant mortality rate during the period 1995-2008.

Proposed hybrid method: The proposed hybrid method for time series prediction is shown in Fig. 2. The method is explained as follows: Based on the time series data, linear models are first used to predict the linear component $\hat{L}_t$. Based on the performance measure RMSE, the best linear model among LR, ARIMA, MA and WMA is selected. The residuals or error series data e_t are obtained using Eq. 17 which are then applied to a neural network model (MLP) to determine $\hat{N}_t$ as shown in Eq. 12. To obtain the optimum architecture configuration for the neural network, different values of input Neurons (N_I) and hidden layer Neurons (N_H) as well as different Activation

Functions (AF) such as sigmoid, bipolar sigmoid and hyperbolic tangent functions are tested. Initially, the NN is configured with an ad hoc number of input layer neuron. From Suhartono, it is known that a larger number of input neurons are required in selecting the optimum configuration. For selecting the optimum configuration, the algorithm uses the following iterative steps sequentially as shown in Fig. 2. The iterative blocks B_1, B_2, B_3 in Fig. 2 are used to obtain the optimum values of N_H, N_I and AF, respectively. The overall prediction is obtained by combining the best linear model result ($\hat{L}_t$) with the optimum NN result ($\hat{N}_t$) as given in Eq. 18.

RESULTS AND DISCUSSION

The models tested were MA, WMA, Linear Regression, ARIMA, NN and hybrid and applied to IMR and MOM data. To select the best linear model and the optimum NN model, the performance measures achieved by each are discussed.

Moving average models: Moving Average (MA) and Weighted Moving Average (WMA) were used as the moving average model for forecasting. The performance measures using with different weight (m) are shown in Table 1. From Table 1, it is shown that MA (2) performed better than other MA models while WMA (2) yielded the minimum values of RMSE and MAPE for IMR and MOM data.

Linear regression model: The LR model used for time series forecasting were obtained experimentally though the solving of simultaneous equations using Eq. 4 during the training phase found to satisfy Eq. 22 for the IMR data and Eq. 23 for the MOM data as follows:

$$\text{IMR:} \quad y_t = 57.198 - 1.969 * t \tag{22}$$

Table 1: Performance measures using MA and WMA for IMR and MOM data

		Performance measures			
		IMR		MOM	
Models	m	RMSE	MAPE	RMSE	MAPE
MA	2	5.044	9.334	3.782	14.163
	3	5.407	11.626	4.184	15.731
	4	6.139	14.087	4.264	16.345
	5	6.528	15.664	4.103	14.679
	6	7.676	19.166	4.251	15.029
W MA	2	4.910	8.715	3.665	14.265
	3	5.137	10.170	3.902	14.952
	4	5.558	11.967	3.992	15.106
	5	5.601	12.756	3.943	14.524
	6	6.353	15.248	4.076	14.894

$$\text{MOM:} \quad y_t = 24 - 0.087895 * t \tag{23}$$

where, y_t is the predicted value for tth year.

The performance result obtained using LR model for IMR and MOM data are shown in Table 2. From the table, it is seen that the performance measures for both the data are different with smaller values for IMR as compared to MOM data.

ARIMA models: Prediction using ARIMA models with different parameter (p, d, q) values are computed in this section. The result of the time series forecasting with different ARIMA models using different parameter (p, d, q) values are shown in Table 3. Testing with other parameter values was also undertaken but only the significant results obtained are shown in the table. From Table 3, it is noted that the smallest values of RMSE and MAPE were obtained by the ARIMA (0, 1, 3) model for IMR data and ARIMA (0, 0, 3) for MOM data.

Subsequently, the researchers select the best linear model from the MA, WMA, LR and ARIMA models. From the performance shows in Table 1-3, the researchers can conclude that the best linear model for IMR data is LR and for MOM data is ARIMA (0, 0, 3). The model yields prediction of the linear component of the data $\hat{L}_t$.

Neural network models: NN is used to obtain the predicted value for the non-linear component of the data $\hat{N}_t$. The residuals series e_t is applied to the NN model as input. To determine the optimum configuration for the NN, the optimum number of hidden nodes, input units and activation function need to be determined. The NN configuration and performance measures for the IMR data residual are shown in Table 4. From the experimental results, it is found that the NN (5.9.1) model with 5 input neurons, 9 hidden layer neurons and using the hyperbolic tangent activation function for the hidden and output layers, yielded the minimum values for RMSE and MAPE.

Table 2: Performance measures for Linear Regression model

	Values	
Performance measures	IMR	MOM
MAPE	4.533	11.640
RMSE	2.767	3.523

Table 3: Result of the time series forecasting using ARIMA for IMR and MOM data

	Performance measures			
Models	RMSE	MAPE	RMSE	MAPE
ARIMA (0, 0, 1)	7.133	13.591	3.382	10.620
ARIMA (0, 0, 3)	6.861	10.789	3.204	10.237
ARIMA (1, 0, 1)	6.176	9.135	3.455	10.772
ARIMA (0, 1, 1)	3.375	5.412	3.833	13.914
ARIMA (0, 1, 2)	3.370	5.367	3.781	12.159
ARIMA (0, 1, 3)	3.270	4.983	3.416	12.528
ARIMA (1, 1, 0)	4.337	6.334	3.854	14.082
ARIMA (1, 1, 1)	3.450	5.413	3.688	12.244

Table 4: Performance measures using Neural Network models for the IMR data residual

Models (input, hidden, output)	Activation function	Performance measures RMSE	MAPE
NN (5, 2 ,1)	Hyperbolic tangent	0.3144	13.3244
NN (5, 3, 1)	Hyperbolic tangent	0.3048	10.1542
NN (5, 8, 1)	Hyperbolic tangent	0.3043	11.8960
NN (5, 9, 1)	Hyperbolic tangent	0.3013	9.0939
NN (5, 10, 1)	Hyperbolic tangent	0.3138	9.7359
NN (2, 9, 1)	Hyperbolic tangent	1.4361	135.4510
NN (3, 9, 1)	Hyperbolic tangent	1.1383	77.4425
NN (4, 9, 1)	Hyperbolic tangent	0.3093	28.6335
NN (5, 9, 1)	Bipolar sigmoid	2.3864	116.7463
NN (5, 9, 1)	Sigmoid	3.2515	111.7265

Table 5: Performance measures using Neural Network models for MOM data residual

Models	Activation function	Performance measures RMSE	MAPE
NN (6, 3, 1)	Hyperbolic tangent	0.3150	13.6759
NN (6, 4, 1)	Hyperbolic tangent	0.3139	10.7391
NN (6, 5, 1)	Hyperbolic tangent	0.3135	8.4341
NN (6, 6, 1)	Hyperbolic tangent	0.2956	7.2729
NN (6, 7, 1)	Hyperbolic tangent	0.3104	10.5811
NN (2, 6, 1)	Hyperbolic tangent	1.7664	46.1608
NN (3, 6, 1)	Hyperbolic tangent	1.1289	35.4807
NN (4, 6, 1)	Hyperbolic tangent	0.3361	16.5614
NN (5, 6, 1)	Hyperbolic tangent	0.3081	10.7627
NN (7, 6, 1)	Hyperbolic tangent	0.3031	8.6039
NN (6, 6, 1)	Bipolar sigmoid	2.4550	41.6670
NN (6, 6, 1)	Sigmoid	2.8567	72.3075

Table 6: Result of the time series forecasting using hybrid LRNN model for IMR data

Years	Actual data	LR[a]	Residual[b]	NN (5, 9, 1)[c]	LRNN[d]
2000	47	45.382	1.618	1.859	47.242
2001	50	43.413	6.587	5.848	49.261
2002	35	41.444	-6.444	-6.168	35.276
2003	38	39.475	-1.475	-1.590	37.885
2004	36	37.506	-1.506	-1.343	36.163
2005	36	35.536	0.464	0.453	35.990
2006	34	33.567	0.433	0.503	34.070
2007	32	31.598	0.402	0.381	31.979
2008	31	29.629	1.371	1.400	31.028
2009	-	27.659	-	2.413	30.073

a: $(\hat{L}_t)$; b: $(e_t = y_t - \hat{N}_t)$; c: $(\hat{N}_t)$; d: $(\hat{y}_t)$

The optimum NN configuration for MOM data residual was obtained as NN (6, 6, 1). This model is used to get the predicted values $\hat{N}_t$ that used to construct hybrid model as shown in Table 5. This NN also uses the hyperbolic tangent activation function.

Hybrid model: The hybrid models for IMR and MOM data were constructed based on the proposed method in Fig. 2. From the experimental results obtained, the best linear model was LR and the optimum NN was NN (5, 9, 1) using hyperbolic tangent activation functions for IMR data. As such both models were combined to obtain the hybrid model, namely hybrid LRNN (Linear Regression Neural Network) model. The time series forecast using hybrid LRNN model is shown in Table 6. The actual data for 2009 were not available as yet but prediction is undertaken using past data. Based on Table 6, the performance results for IMR data are obtained as shown in Table 7. For MOM data, the best linear model was ARIMA (0, 0, 3) and the optimum NN was NN (6, 6, 1) using hyperbolic tangent activation function. The time series prediction using hybrid ARNN (ARIMA Neural Network) model consisting of combination of ARIMA (0, 0, 3) and NN (6, 6, 1) is shown in Table 8 with the performance results shown in Table 9.

Table 7: Performance measures using hybrid LRNN model for IMR data

Performance measures	Values
MAPE	0.4366
RMSE	0.0808

Table 8: Time series forecasting using hybrid ARNN model for MOM data

Years	Actual data	ARIMA (0, 0.3)[a]	Residual[b]	NN (6, 6, 1)[c]	ARNN[d]
1995	19.40	21.485	-2.085	-2.209	19.276
1996	21.70	23.705	-2.005	-2.042	21.663
1997	16.10	21.130	-5.030	-4.405	16.725
1998	22.00	21.010	0.990	0.874	21.883
1999	24.90	21.294	3.606	3.511	24.805
2000	31.10	26.396	4.704	4.487	30.883
2001	26.20	26.207	-0.007	-0.068	26.139
2002	22.30	24.412	-2.112	-2.152	22.261
2003	21.80	20.665	1.135	1.053	21.718
2004	20.60	22.201	-1.601	-1.477	20.724
2005	24.80	23.849	0.951	0.836	24.684
2006	24.00	22.016	1.984	1.755	23.771
2007	19.70	24.800	-5.100	-4.348	20.452
2008		22.062		-5.489	16.574

a: $(\hat{L}_t)$; b: $(e_t = y_t - \hat{N}_t)$; c: $(\hat{N}_t)$; d: $(\hat{y}_t)$

Table 9: Performance measures using hybrid ARNN model for MOM data

Performance measures	Values
MAPE	0.9234
RMSE	0.0811

Table 10: Comparison of performance measures using individual and hybrid LRNN model for IMR data

Models	Performance measures RMSE	MAPE
Moving Average (MA)	5.044	9.334
Weighted Moving Average (WMA)	4.910	8.715
Linear Regression (LR)	2.767	4.533
ARIMA (0, 1, 3)	3.270	4.983
NN (6, 10, 1) [NN for IMR data input]	0.300	0.642
Proposed model: Hybrid LRNN	0.081	0.437

Comparison of models: The proposed method is compared with well known models including the MA, WMA, LR, ARIMA and NN models, based on the prediction performances for healthcare data. Table 10 shows a comparison of RMSE and MAPE values obtained using the linear, NN and the proposed hybrid models for IMR data.

From Table 10, it is seen that the hybrid LRNN model combining LR and NN (5, 9, 1) yields the best results compared to the other models for IMR data. It should be noted that NN (5, 9, 1) is the optimum NN using the IMR data residual and NN (6, 10, 1) is the optimum NN using IMR data. The improvements achieved with respect to RMSE and MAPE values by the hybrid LRNN model over

Table 11: Improvement achieved by proposed hybrid LRNN model over the other models for IMR data

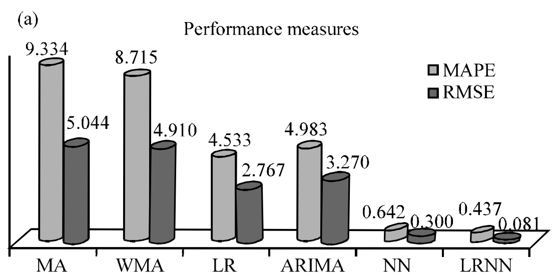
	Performance measures (%)	
Models	RMSE	MAPE
Moving Average (MA)	98.39	95.32
Weighted Moving Average (WMA)	98.35	94.99
Linear Regression (LR)	97.07	90.36
ARIMA (0, 1, 3)	97.52	91.23
NN (6, 10, 1)	73.00	31.93

Table 12: Comparison of performance measures using individual and hybrid ARNN model for MOM data

	Performance measures	
Models	RMSE	MAPE
Moving Average (MA)	3.782	14.163
Weighted Moving Average (WMA)	3.665	14.265
Linear Regression (LR)	3.523	11.640
ARIMA (0, 0, 3)	3.204	10.237
NN (8, 19, 1)	0.288	0.964
Proposed model: Hybrid ARNN	0.081	0.923

Table 13: Improvement achieved by proposed hybrid ARNN model over the other models for MOM data

	Performance measures (%)	
Models	RMSE	MAPE
Moving Average (MA)	97.86	93.48
Weighted Moving Average (WMA)	97.79	93.53
Linear Regression (LR)	97.70	92.07
ARIMA (0, 0, 3)	97.47	90.98
NN (8, 19, 1)	71.88	4.25



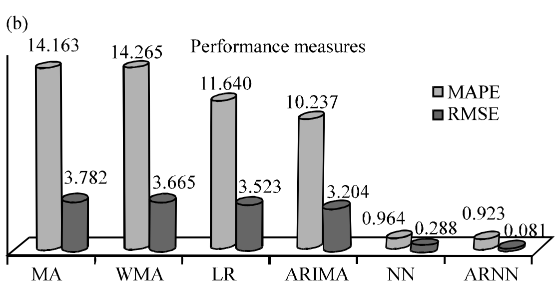


Fig. 3: Comparison of performance; (a) Infant Mortality Rate (IMR) data; (b) Morbidity of Malaria (MOM) data

other models are shown in Table 11. It is shown that the proposed hybrid LRNN model is able to achieve significant performance improvement over the other models for IMR data. The performance measures of the different models for MOM data are shown in Table 12. From Table 12, it is seen that the best results are achieved by the proposed hybrid ARNN model for MOM data. Table 13 shows the improvement achieved by the hybrid ARNN over other modes. It is clear from the table that the proposed hybrid ARNN model achieves significant performance improvement over other models for MOM data.

To illustrate the results further, a graphical representation of the performance results is shown in Fig. 3 for both IMR and MOM data. The horizontal axis represents the models and the vertical axis is the values of performance measure. It is clear that proposed approach is a significant improvement over the linear only models and also in all cases is an improvement over the NN only approach as well.

CONCLUSION

In this study, a new hybrid method is proposed for time series prediction. Unlike other hybrid models which use a predetermined configuration for the linear and neural network components, the proposed method selects the best linear model and the optimum neural network configuration based on the type of input data.

A performance comparison of the different models was carried out using measures such as RMSE and MAPE. From the experimental results, it was found that the proposed hybrid model performs better than the linear and neural network models. Hybrid LRNN model which combines Linear Regression and Neural Network models was found to be superior for IMR data. Hybrid ARNN was found to be superior for MOM data. It can be concluded that the proposed hybrid method is best suited for time series prediction of complex data such as IMR and MOM data which comprise both linear and non-linear patterns. Accurate forecasting of healthcare data assumes importance for decision making since a good and accurate forecasting is very helpful in devising appropriate action plans for healthcare professionals, authorities and governments.

REFERENCES

Abeku, T.A., S.J. de Vlas, G. Borsboom, A. Teklehaimanot and A. Kebede *et al*., 2002. Forecasting malaria incidence from historical morbidity patterns in epidemic-prone areas of Ethiopia: A simple seasonal adjustment method performs best. Trop. Med. Int. Health, 7: 851-857.

Aladag, C.H., E. Egrioglu and C. Kadilar, 2009. Forecasting nonlinear time series with a hybrid methodology. Applied Mathe. Lett., 22: 1467-1470.

Bing-Jun, L. and H. Chun-Hua, 2007. The combined forecasting method of GM(1,1) with linear regression and its application. Proceedings of IEEE International Conference on Grey Systems and Intelligent Services, Nov. 18-20, Nanjing, pp: 394-398.

Box, G.E.P. and G.M. Jenkins, 1976. Time Series Analysis, Forecasting and Control. Holden Day, Inc., Oakland, California.

Brockwell, P.J. and R.A. Davis, 2002. Introduction to Time Series and Forecasting. 8th Edn., Springer, Berlin, ISBN: 978-0-387-95351-9, Pages: 469.

Chatfield, C., 2000. Time-Series Forecasting. Chapman and Hall, London.

Condran, G.A. and J. Murphy, 2005. The infant mortality rate in late 19th and early twentieth century Philadelphia. The Annual Meeting of the American Sociological Association, Philadelphia.

DeCoster, J., 2003. Notes on applied linear regression. Department of Social Psychology, Free University Amsterdam, Netherlands. http://www.stat-help.com/linreg.pdf.

Diaz-Robles, L.A., J.C. Ortega, J.S. Fu, G.D. Reed, J.C. Chow, J.G. Watson and J.A. Moncada-Herrera, 2008. A hybrid ARIMA and artificial neural networks model to forecast particulate matter in urban areas: The case of Temuco, Chile. Atmosph. Environ., 42: 8331-8340.

Faruk, D.O, 2010. A hybrid neural network and ARIMA model for water quality time series prediction. Eng. Appl. Artificial Intel., 23: 586-594.

HMI, 2008. Current malaria situation in Indonesia and ACTmalaria activities. Directorate of Vector Borne Disease Control Ministry of Health Ministry, Indonesia.

Katimon, A. and A.S. Demun, 2004. Water use trend at Universiti Teknologi Malaysia: Application of Arima model. J. Teknol., 41: 47-56.

Kusumadewi, S. and S. Hartati, 2006. Neuro-fuzzy integrasi sistem fuzzy dan jaringan syaraf. Graha Ilmu, Yogyakarta, Indonesia. http://www.grahailmu.co.id/index/buku/detil/20/3/18/47/buku178.html.

Lai, K.K., L. Yu, S. Wang and W. Huang, 2006. Hybridizing Exponential Smoothing and Neural Network for Financial Time Series Predication. In: Computational Science-ICCS 2006, Alexandrov, V.N. *et al*. (Eds.). LNCS 3994, Springer Verlag, Berlin, Heidelberg, ISBN: 978-3-540-34385-1, pp: 493-500.

Le, C.T., 2003. IntroductoryBiostatistics. John Wiley and Sons, New Jersey .

MHRI, 2008. Indonesia Health Profile. Ministry of Health, RI, Jakarta.

McNown, R. and A. Rogers, 1992. Forecasting cause-specific mortality using time series methods. Int. J. Forecasting, pp: 413-432. http://www.forecasters.org/ijf/journal-issue/356/article/5512.

Moghaddas-Tafreshi, S.M. and M. Farhadi, 2008. Linear regression-based study for temperature sensitivity analysis of Iran electrical load. Proceeding of the IEEE International Conference on Industrial Technology, April 21-24, Chengdu, pp: 1-7.

Niska, H., T. Hiltunen, A. Karppinen, J. Ruuskanen and M. Kolehmainena, 2004. Evolving the neural network model for forecasting air pollution time series. Eng. Applied Artif. Intel., 17: 159-167.

Nkurunziza, H., A. Gebhardt and J. Pilz, 2010. Forecasting malaria cases in Bujumbura, World Academy of Science. Eng. Technol., 61: 253-258.

Rojas, I., O. Valenzuela, F. Rojas, A. Guillen and L.J. Herrera *et al*., 2008. Soft-computing techniques and ARMA model for time series prediction. Neurocomputing, 71: 519-537.

Shiel, W.C. and M.C. Stoppler, 2008. Webster's New World, Medical Dictionary. 3rd Edn., Wiley Publishing Inc., Hoboken, New Jersey.

Zhang, G.P., 2003. Time series forecasting using a hybrid arima and neural network model. Neurocomputing, 50: 159-175.