BRICS JOURNAL OF ECONOMICS
DOI 10.3897/brics-econ.6.e141556
2025 Volume 6 Number 2

Predicting Currency Crises in Emerging Markets: A Case Study on South Africa Using Artificial Neural Networks

Gladys Fernandes-Gondoza[1], Ronney Ncwadi[2], John Manuel Fernandes[3], Farai Nyika[4]

1 *Department of Economics, Nelson Mandela University, University Way, Summerstrand, Gqeberha, 6019 (South Africa)* **2** *Department of Economics, Nelson Mandela University, University Way, Summerstrand, Gqeberha, 6019 (South Africa)* **3** *Department of Marine Engineering, Nelson Mandela University, University Way, Summerstrand, Gqeberha, 6019 (South Africa)* **4** *School of Public Administration, MANCOSA, Cape Town (South Africa)*

Corresponding author: Farai Nyika (farai.nyika@mancosa.co.za)

Academic editor: Sheresheva M. | Received 11 November 2024 | Accepted 17 December 2024 | Published 13 June 2025

Citation: Fernandes-Gondoza, G., Ncwadi, R., Fernandes, J. M., & Nyika, F. (2025). Predicting Currency Crises in Emerging Markets: A Case Study on South Africa Using Artificial Neural Networks. *BRICS Journal of Economics, 6*(2), 91–116. https://doi.org/10.3897/brics-econ.6.e141556

Abstract

Purpose: In this paper, we study the potential of using Artificial Neural Network (ANN) models to predict currency crises in emerging markets, with a specific focus on the South African economy. South Africa's rand is one of the most volatile currencies in the world and is prone to crises.

Methodology: We built two ANN models, where Model 1 uses ten economic indicators and Model 2 uses four. These models were assessed for statistical significance (using probit analysis), and their performance in predicting major South African currency crises (e.g., 1998, 2001, and 2008) was tested with both in-sample and out-of-sample data.

Results: The first model was much more accurate than Model 2 in predicting early warning signs nearly two years before the currency crises occurred. Model 1's higher accuracy is attributed to its inclusion of a greater number of economic variables. Both models occasionally produced false positives, though overall, they were very accurate in predicting crises.

Originality: Our paper highlights the importance of ANNs in capturing nonlinear patterns in economic data, demonstrating their strength as early warning tools for financial crises.

We recommend that ANN methods continue to be researched and advanced to further reduce false positives and improve predictive performance.

Keywords
Currency crises, Early Warning System, Artificial Neural Networks, South Africa, Emerging markets.

JEL: C45, C53, F37, G01.

1. Introduction

Currency crises are not only a problem of the past, but also of the present and future. Since the fall of the Bretton Woods system in 1973, emerging and developing markets have experienced an increased frequency of crises, which continue to pose a threat today (Bernoth & Herwartz, 2021). Financial crises have affected numerous countries in the past and are likely to happen again in the future, as predicted in a prescient study by Bordo, Eichengreen, Klingebiel, Martinez-Peria and Rose (2001). Notable currency crises include the Latin American crises of the 1970s/80s, crises in Europe from 1992 to 1993 counteracted by the European monetary exchange rate mechanism (ERM), the Mexican peso crisis of 1994 to 1995 (also known as the tequila crisis), the East Asian crisis of 1997 to 1998 which affected Indonesia, Korea, Malaysia, Philippines, and Thailand, Brazil in 1999, Argentina in 2002, the Zimbabwean currency crisis from 1990 to 2009, the USA in 2007, the GFC, and the 2015 Greek currency crisis (International Monetary Fund Staff Report, 2009). Given the great impact of sovereign debt and currency crises on economic activity, the study of international financial crisis events, particularly their forecasting, has garnered substantial attention from economists in the last twenty years (Alaminos Aguilera et al., 2021).

South Africa, as the economic powerhouse of Africa, is an interesting case to examine. The country operates a floating exchange rate system, which makes it, in theory, less vulnerable to currency crises (Juhro et al., 2022; Guei & Choga, 2022). This is because market expectations for adjustments prevent the buildup of excessive pressure in its foreign exchange markets (Ghosh et al., 2015; Georgiadis & Zhu, 2021). South Africa's foreign currency market is known for its volatility and frequent bouts of turbulence (Laubscher, 2020). Instances of currency crises were documented in 1996, 1998, 2001, and 2008. The 2007/2008 GFC was the most severe economic downturn the globe had encountered since the Great Depression of the 1930s (Mirzaei, 2023). The GFC had a significant negative impact on South Africa's economy, revealing the country's fragility, as mineral exports declined leading to a growing trade imbalances (Bhundia & Ricci, 2005; Raudzingana, 2020). Furthermore, the economy experienced a recession for the first time in 17 years. Despite the South African Reserve Bank's confidence in the economy's strong foundations, its macroeconomic predictions had to be significantly adjusted downwards (South African Reserve Bank, 2012; Beyers et al., 2024).

South Africa as an emerging market economy (EME) is gaining global importance due to its strategic partnership with the BRICS+ association. Since EMEs have historically been more susceptible to crises (Alonso-Alvarez & Molina, 2023), it is essential that fiscal and monetary policymakers understand the extent, to which their financial systems are vulnerable to disturbances that can easily spill over into the key partner economies, and be ready to respond appropriately. South Africa suffers from frequent strikes and violent protests (Kali, 2023); severe currency crises and economic hardships could trigger even more populist events. All this highlights the importance of Early Warning Systems (EWSs). We are living in an era of increased risks of future pandemics, climate change induced global boiling, and global violent conflict (Goniewicz et al., 2023). These challenges further raise the risk of global financial crises and contagion.

The South African economy needs an alternative early warning system (EWS) model, which is to be developed and evaluated. This model should accurately detect the early signs of crisis events before they occur, allowing authorities to implement optimal policies to prevent or mitigate their impact. One tool that has garnered significant interest from the International Monetary Fund (IMF), policymakers, and academics is the Artificial Neural Networks (ANN) approach (Gupta & Kumar, 2022a; Ataman & Kahraman, 2022). ANNs probability prediction is a type of dynamic filtering in which past values of one or more time series are used to predict future values (Pokou et al., 2024). ANN models are commonly applied in management fields such as sales and demand forecasting, logistics and supply chain (Khan et al., 2020; Huang et al., 2021; Elalem et al., 2023). However, there are increasing studies of ANN's use to forecast economic growth, downturns, and financial crises (Tölö, 2020; Longo et al., 2022; Liu et al., 2022; Bluwstein et al., 2023; Barthélémy et al., 2024; Raj et al., 2024). Most of the recent literature in this field does not concern the developing world, let alone sub-Saharan Africa. We assert that using ANN as described in this paper is superior to other methodologies such as the conventional machine learning models, shallow neural networks or traditional econometric models, which cannot properly analyze the challenging multidimensional and highly volatile nonlinear financial market data (Heo et al., 2020; HongXing et al., 2022). To the best of our knowledge, this study is one of the first in South Africa to use the ANN methodology to evaluate its ability to forecast business cycles.

2. Materials and Methods

2.1. Description of Artificial Neural Networks

The architecture and functionality of ANNs bear resemblance to those of the human brain, albeit with reduced processing speed, computational capacity, and storage capabilities (Dastres & Soori, 2021). The fundamental premise of ANNs lies in their emulation of biological neural network operations (Montesinos López et al., 2022). A biological neural system comprises a rudimentary structure that executes three primary functions: signal reception (input) from other neurons, signal processing, and subsequent signal transmission to additional neurons. In the ANN paradigm, an artificial neuron receives environmental stimuli and, upon activation, propagates a signal to all interconnected artificial neurons (Liu et al., 2023).

A neural network is a realization (R) of a nonlinear mapping from R^I to R^K i.e

$$f_{NN} : R^I \rightarrow R^K \tag{1}$$

where I and K are respectively the dimension of input and the desired output space. f_{NN} is usually a complex function of a set of nonlinear functions, one for each neuron in the network. The artificial neuron receives vector input signals,

$$z_i = (z, z_2, \ldots, z_I) \tag{2}$$

either from the environment or from other artificial neurons. To each input signal, z_i is associated a weight, v_i, to strengthen or deplete the input signal. The artificial neuron computes the net signal and uses an activation function (f_{AN}) to compute the output signal, y, given the net input. This is shown in Figure 1 below.

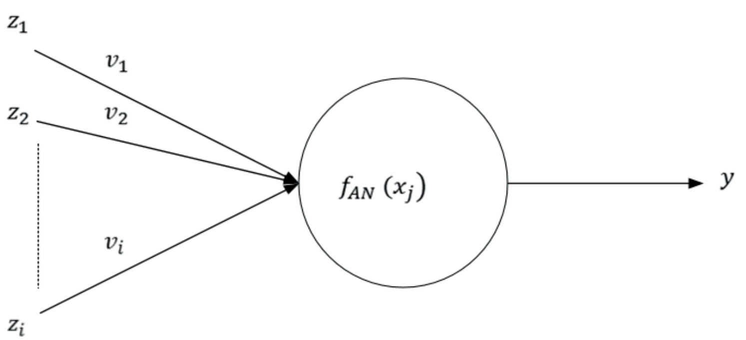

Figure 1. An Artificial Neuron. *Source*: Adapted from Engelbrecht (2007:6)

ANN models can be divided into three or more layers - an input layer, an output layer and an intermediate layer/s (i.e. a hidden layer/s between input and output layers (See Figure 2 below) (Madhiarasan & Louzazni, 2022). All neurons in these layers are fully connected to other neurons in the next layers but no connections exist between neurons within a layer and each connection is associated with a weight (Kabir et al., 2022).

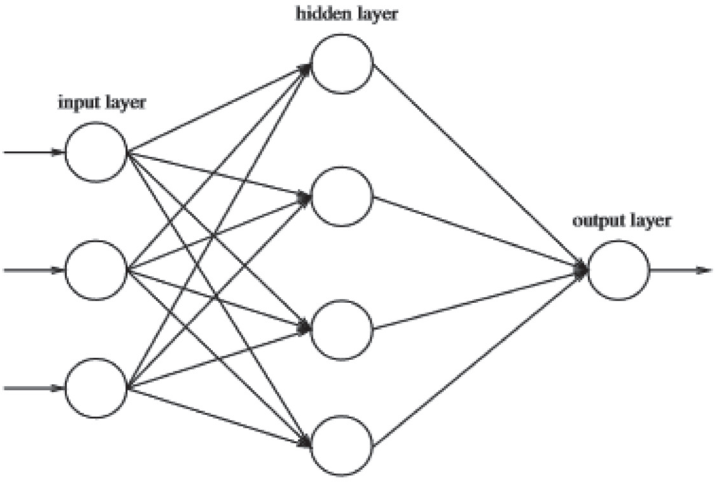

Figure 2. An Artificial Neural Network. *Source*: Adapted from Engelbrecht (2007:6)

The initial phase in the development of an artificial neural network involves the judicious selection of appropriate indicators for input variables. This critical step is essential to ensure the accuracy and precision of the probability prediction of the output (Chen et al., 2020). The more relevant are the explanatory variables regarding the output variable, the less the learning time for the model (Biecek & Burzykowski, 2021).

A critical consideration in artificial neural network design is the determination of the optimal number of hidden layers and the appropriate quantity of neurons within each of these layers. While the utilization of multiple hidden layers is permissible, such architectural decisions necessitate computational resources that may exceed the capabilities of standard computing systems (Capra et al., 2020). Furthermore, the implementation of additional hidden layers introduces an associated increase in the temporal requirements for the learning process to achieve convergence. This trade-off between model complexity and computational efficiency presents a significant challenge in the optimization of neural network architectures, requiring careful balancing of performance gains against practical constraints.

Our study therefore uses only one hidden layer. Regarding the number of neurons in the hidden layer, note that too many neurons will affect the learning period of the model, making it longer; this can result in over-fitted data and poor model performance (Boloukian & Safi-Esfahani, 2020). On the other hand, too few neurons in the hidden layer may result in the neural networks' incapacity to deal with a complex dataset (Karsoliya, 2012; Uzair et al., 2020). The output neuron in the output layer corresponds with the predicted variable (Belciug, 2020). In this study, the expected output is only one, the probability of a crisis in South Africa having a value in the range of 0 to 1. To train this model in the supervised learning neural network, the output neuron is compared to the target (Zhang et al., 2020).

Several neural networks have been developed. For example, the most successful applications in terms of business analyses were the multi-layered feed-forward neural network used for prediction and classification (Kurani et al., 2023) and the self-organizing map with unsupervised learning for clustering. As suggested earlier, the artificial neural network models using the feedforward artificial neural network with the backpropagation learning algorithm is the best type of network when working with macroeconomic and financial data for the prediction of currency crises (Šestanović & Arnerić, 2021).

Net Input Signal

Once all indicators have been selected, all signals from all neurons in the input layer, their corresponding weights that are set to small random values and a bias neuron are sent to the neurons in the hidden layer. Two processes occur in the hidden layer, first the summation of the input signals followed by using the activation function to fire a signal to the output layer. Thus, all incoming signals (net input signal into the hidden neuron), are summed as,

$$x_i = v_{oj} + \sum_{i=1}^{I} z_i v_{ij} \tag{3}$$

where x_i is the signal from a hidden neuron and v_{oj} is the bias (weight) attached to each hidden neuron (Engelbrecht, 2007).

Activation Functions

An activation function receives the net signal and bias and thus determines the output (or firing strength). In general, activation functions are monotonically increasing mappings, where,

$$f_{AN}(-\infty) = 0 \text{ or } f_{AN}(-\infty) = -1 \tag{4}$$

and

$$f_{AN}(\infty) = 1 \tag{5}$$

This excludes the linear function (Engelbrecht, 2007). There are various types of activation functions that can be used. This study employs the sigmoid function (for $\theta = 0$):

$$f_{AN}(net - \text{q}) = \begin{cases} \dfrac{1}{1 + e^{-1(net - \text{q})}} \end{cases} \tag{6}$$

The sigmoid function is a continuous version of the ramp function with, $f_{AN}(net) \in (0,1)$. The parameter λ controls the steepness of the function and is usually equated to one (Engelbrecht, 2007). In this study, the preferred output of the artificial neural network model is the probability of a crisis occurring, which is valued in the range of 0 for "no crisis" to 1 for the "presence of crisis", and thus applied the logistic sigmoid activation function to keep its value within this range.

The Learning Algorithm

An artificial neural network model, very much like the human brain, can be trained to make its prediction more accurate and powerful. The artificial neuron learns the best values for v_{ij} and the threshold from the given data by adjusting weights and threshold values until a certain criterion/s is/are satisfied (Goodfellow et al., 2016). Three available types of learning are supervised, unsupervised and reinforcement learning. In this study, the focus is on supervised learning as this learning implies that the artificial neural network is provided with a data set consisting of input vectors and a target (desired output) associated with each vector input. The aim of this type of learning is to adjust the weight values in such a way that the error between the real output, Y, of the neuron and the target output, T, is minimized (Farizawani et al., 2020). Under supervised learning there are numerous learning algorithms. The most popular and commonly applied within the field of business and financial analyses is the back propagation (Kurani et al., 2023). This algorithm can be broken into three distinctive steps, the feed-forward mechanism, back propagation of error and adjustment of associated weights.

Feed-forward

In the feed-forward step, all signals travel from the input layer with one bias neuron to the hidden layer with their respective initial weights and in the hidden layer all input signals are summed as in Equation 3 and transformed to binary values within the range of 0 and 1 via the logistic sigmoid activation function and then sent to the output layer. The logistic sigmoid activation function in Equation 6 can be rewritten simply as,

$$x_j = f\left(x_j\right) = \frac{1}{1+\exp\left(-x_j\right)} \tag{7}$$

where x_j is the activation signal from the hidden neuron (Engelbrecht, 2007). The activation signals x_j, are sent to the next layer, the output layer. In the output neuron a similar process is applied as with the hidden neuron, that is, all incoming signals from hidden neurons are summed using the following equation,

$$y_k = w_{ok} + \sum_{j}^{I} x_i w_{jk} \tag{8}$$

where w_{jk} is the weight of input signals from the hidden layer and w_{ok} is the bias on output unit (Widodo et al., 2021). Then, as the expected value of the output neuron is the probability of a crisis in the range of 0 to 1, Equation 8 is transformed to any value in the given range by the logistic sigmoid activation function:

$$Y_k = f(y_k) = \frac{1}{1 + \exp(-y_k)} \tag{9}$$

(Engelbrecht, 2007).

Back-propagation of Error

The next step is to compare the network's output (Y_k) with the target (T_k) and all errors from this comparison are backpropagated from the output to the input layer via a hidden layer as follows:

$$E = \frac{1}{2}\sum_k [T_k - Y_k]^2 \tag{10}$$

As the error (E) is a function of weights, the network will minimize this error by adjusting the connection weight of each neuron in the entire layer.

Adjusting the Associated Weights

The last step in the process of back propagation is the gradual adjustment of the weights associated with each connection among the neurons. The idea is to gradually adjust the error in a backward direction. To do this, the model updates the weights on connections between output neurons and hidden neurons and between hidden neurons and input neurons. The new connection weight between output and hidden neurons can be given by

$$w_{jk}^{*} = w_{jk} + \mathrm{D}w_{jk} \tag{11}$$

where Δw_{jk} is the weight correction term.

The back-propagation algorithm is the optimizing technique using the steepest descent that requires a step size or learning rate (a) to be specified (Abdolrasol, et al. 2021), so the weight correction term can be obtained as follows:

$$\mathrm{D}w_{jk} = -\mathrm{a}\frac{dE}{dw_{jk}} \tag{12}$$

Equation 12 indicates that the error will be minimized by adjusting the connection weight (Li et al., 2021) and the minus sign in a direction in which the function decreases more rapidly. This learning process continues until the network meets one condition, that the net output of the model must converge to its target or until the minimum threshold of the error is achieved (Engelbrecht, 2007).

Limitations of ANNs

ANNs are commonly referred to as 'black-box models' that do not have interpretability, which is a result of their complex internal structures and large number of parameters. It is not easy to make sense of how an ANN justifies a decision and prediction (He et al., 2020; Kuok et al., 2024). Furthermore, ANNs require precise tuning of numerous hyperparameters, including the learning rate, activation functions, and number of hidden layers (Rosa et al., 2020; Abdolrasol et al. 2021). Incorrect tuning can lead to unsatisfactory performance and requires extensive experimentation (Abdolrasol et al. 2021). Other commonly used time series methods for currency prediction such as the Autoregressive Integrated Moving Average (ARIMA), Generalized Autoregressive Conditional Heteroskedasticity (GARCH) and Vector Autoregression (VAR) may be interpretable, but their performance and accuracy deteriorate when they are applied to highly complex, volatile, nonlinear datasets (Singh et al., 2020; Rast et al., 2022; Kuok et al., 2023). The authors of this paper believe that the advantages of ANNs for currency prediction outweigh the disadvantages and the next section explains how the ANN EWS used in their research was developed.

2.1.1. Application And Construction of The Artificial Neural Network EWS Model

In constructing the artificial neural network model, firstly, an output neuron had to be determined. The output neuron was the currency crises in South Africa. This study defines a crisis as the actual depreciation or devaluation of a currency that takes place coupled with interventions by the central bank to counteract this depreciation (Dominguez, 2020). Therefore, if exchange rates, interest rates and foreign exchange reserves are combined in one index, namely the EMP index, then a crisis can be identified by observing the behavior of the index. A crisis is then indicated if the following condition holds:

$$CRISIS = \begin{cases} 1, if\ EMP_{i,t} > \mathrm{m}_{EMP} + \mathrm{d} \cdot \mathrm{s}_{EMP} \\ \quad 0, otherwise \end{cases} \tag{13}$$

That is to say, a crisis is indicated if the value of the EMP index exceeds the mean of the EMP time series (μ_{EMP}) plus the standard deviation of the EMP time series (σ_{EMP}) multiplied by a weight (δ) (Jones et al., 2019). The EMP index was constructed and a crisis was defined when the EMP index exceeded a threshold set at 1.645 times the standard deviation; this was true for the crises of July 1998, December 2001 and October 2008 in South Africa.

The next step was to select the set of variables for the input neurons to ensure the precision of the prediction of output. No specific methods have been recommended in the literature on how to select the input variables for this model. This study employed the NSR for selecting a set of input variables. Ten variables were chosen as leading indicators of a currency crisis in South Africa. Using fewer input variables is superior to using too many as the inclusion of noise in the data set degrades the performance of the artificial neural networks model (Thike et al., 2020).

On the other hand, Hashemi, (2019) found that using many input neurons creates a model superior to that with fewer input neurons. For this reason, the present study created a second artificial neural networks model using the statistically significant variables attained by the application of a simple probit model employing the ten chosen leading indicators. Tables 1 and 2 present the input variables that were used for Model 1 and Model 2, respectively.

Table 1. Artificial Neural Networks Model 1 Input Variables

No	Variables
1	ratio of budget deficits to GDP
2	change in the international liquidity position
3	growth rate of merchandise exports
4	growth rate of merchandise imports
5	growth rate of the ratio of domestic credit to GDP
6	growth rate of bank deposits
7	growth rate of foreign debt
8	changes in the gold price
9	change in the domestic interest rate
10	inflation differential to the USA

Table 2. Artificial Neural Networks Model 2 Input Variables

No	Variables
1	change in the international liquidity position
2	growth rate of foreign debt
3	change in the domestic interest rate
4	inflation differential to the USA

All variables were stationary, and no multicollinearity existed when models were constructed. To provide statistical distribution and equal proportional contributions, and also remove any biases in the forecasting model, Hamid (2004) suggested normalization of the data and addition of one bias neuron. The normalization of data in the range -1 to 1 was obtained by applying the following formula:

$$x_{it} = 2\left[\frac{x_{it} - \min(x_{it})}{max\left(x_{it)} - \min(x_{it}\right)}\right] - 1 \tag{14}$$

Source: Hamid (2004).

Concerning the number of hidden neurons, it has been noted that if the number of hidden neurons increases, there is a tendency to overfitting, which means that the model mimics a "parrot" and will thus fail in its predictions. At the same time, too few hidden neurons decrease the ability of the model to manage a more complex data set. As one hidden layer is sufficient for the model to solve all problems (Xu, et al. 2021) this study implemented only one hidden layer. The specific-to-general method, as described by Mekkaoui et al. (2023), was employed to determine the optimal number of hidden neurons. This approach involves selecting the number of hidden neurons at the point where the model exhibits the smallest training error across various specified maximum iterations. The models utilized the logistic sigmoid activation function to generate the target output, which represents the probability of a crisis and is valued between 0 and 1. Regarding the optimal learning rate and momentum rate, the study adopted the default settings for neural networks, setting them to 0.01 and 0.8, respectively.

3. Results

The ANN models used a backpropagation learning algorithm to determine model performance (Heng et al., 2022). For training[1] the network, this study considered two periods, an in-sample period from February 1993 to December 2004 and out of sample period from January 2005 to March 2017. Simulations were conducted on the models with various numbers of iterations to achieve the smallest training error. To this effect one hundred networks were simulated using one to 20 neurons in maximum iterations of 10000 (0.130), 20000 (0.131), 30000 (0.126), 40000 (0.131) and 50000 (0.137). With thirty thousand iterations the smallest error 0.126 was found using 19 hidden neurons. Following this, both models were trained using parameters set in Table 3 below.

1 The process in which the weights of interconnected neurons are updated to emulate any nonlinear relationship from input to the target output.

Table 3. Parameters of the Artificial Neural Network Models

No.	TRAINING INFORMATION	MODEL 1	MODEL 2
1	Type of network	Multi-layer perception	Multi-layer perception
2	Number of layers	3	3
3	Number of hidden layers	1	1
4	Number of input neurons	10	4
5	Number of hidden neurons	19	19
6	Number of output neurons	1	1
7	Activation function	Logistic Sigmoid	Logistic Sigmoid
8	Performance function	Mean square error	Mean square error
9	Training algorithm	Back-propagation	Back-propagation
10	Starting weights and biases	Random	Random
11	Number of iterations	30 000	30 000
12	Training error	0.126	0.148
13	Learning rate	0.010	0.010
14	Momentum factor	0.8	0.8

Source: Author's own computation.

The process began when both models stochastically assigned their initial connection weights between neurons and bias neurons in the input layer and the hidden neurons in the hidden layers. Feed-forward was employed to ensure that signals constantly flow in a forward direction from the input layer to the output layer through the hidden layer. Within the concealed neuron, the models aggregate all these signals and convert them into a numerical range of 0 to 1 by means of the logistic sigmoid activation function. The initial weights connecting the hidden neuron and bias neuron to the output neuron in the hidden layer are randomly assigned. Subsequently, the models transmit modified signals from the concealed neuron to the output neuron, employing the logistic sigmoid activation function once more. The learning stage commences by contrasting the output neuron with its designated target. The mean square error is propagated backwards from the output layer to the input layer through the hidden layer using the backpropagation learning technique. This mechanism enables the establishment of suitable connection weights for all neurons. Consequently, the models stop the learning process whenever they reach the minimum error or the maximum number of repetitions.

To establish the most statistically significant input variables, the study employed a popular method using connection weights and the feed-forward back-propagation algorithm (Yadav et al., 2022). This method determines the relative importance of the predictor variables of the model as a function of the neural network synaptic weights, according to the mathematical expression:

$$R_{ij} = \sum_{H=1}^{k} W_{ik}.W_{kj} \tag{15}$$

Where R_{ij} is the relative importance of the variable x_i with respect to the output neuron *j*, *H* is the number of neurons in the hidden layer, W_{ik} is the synaptic connection weight between the input neuron *i* and the hidden neuron *k* and W_{kj} is the synaptic weight between the hidden neuron *k* and the output neuron *j* (Nguyen et al., 2021).

Table 4 below presents the contribution of input neurons to the output neurons for Model 1. It shows that the most significant contributor is the ratio of budget deficits to GDP, followed by growth rate of the ratio of domestic credit to GDP and then a change in the international liquidity position. For more details on the other variables, see Table 4 below.

Table 4. Average Contributions of Input Variables to Output Variable for Artificial Neural Networks Model 1

NO.	VARIABLES	CONTRIBUTION
1	ratio of budget deficits to GDP	22.68%
2	change in the international liquidity position	13.38%
3	growth rate of merchandise exports	4.33%
4	growth rate of merchandise imports	12.76%
5	growth rate of the ratio of domestic credit to GDP	16.23%
6	growth rate of bank deposits	3.08%
7	growth rate of foreign debt	12.37%
8	changes in the gold price	4.58%
9	change in the domestic interest rate	3.31%
10	inflation differential to the USA	7.28%

Source: Author's own computation.

The contributions of the input variables in the reduced model (Model 2) were calculated. As shown in Table 5, the primary contributor to the model's predictions was the change in the international liquidity position, followed by the inflation differential to the USA and then the growth rate of foreign debt. The change in the domestic interest rate was found to be the least significant input variable in Model 2.

Table 5. Average Contributions of Input Variables to Output Variable for Artificial Neural Networks Model 2

NO.	VARIABLES	CONTRIBUTION
1	change in the international liquidity position	53.13%
2	growth rate of foreign debt	18.47%
3	change in the domestic interest rate	2.64%
4	inflation differential to the USA	25.76%

Source: Author's own computation.

3.1. Predicting the South African Currency Crises

Both models were employed to simulate and evaluate their ability to predict the probability of the South African currency crises of July 1998, December 2001, and October 2008. To analyze the probability predictions, this section was divided into in-sample probability predictions (Figure 3) and out-of-sample probability predictions (Figure 4). The study utilized a 24-month crisis window. This means that any signals detected within the 24 months preceding a crisis were considered as valid indicators of an impending event.

3.1.1. In-Sample Probability Prediction of Currency Crisis, 1993/02 – 2004/12

Figure 3 presents the in-sample probability predictions for both Model 1 and Model 2 in three separate graphs. While the two models' predictions were initially plotted together in Figure 3(a), it became evident that their plots were not overlapping, unlike the Probit models. To facilitate analysis, the probability predictions were separated into individual graphs: Figure 3(b) for Model 1 and Figure 3(c) for Model 2. In Figure 3(b), Model 1 accurately predicts the July 1998 crisis. As early as July 1996 (24 months prior), warning signals were emitted, reaching a magnitude of approximately 42%. Although these signals subsided to around 5% in early 1997, they intensified again in the latter part of 1997, reaching approximately 36%. Following a brief decline in early 1998, the signals gradually increased until the occurrence of the crisis in July 1998, with a predicted probability of approximately 59%. Model 2 (Figure 3(c) also correctly identified the July 1998 crisis, with a predicted probability of around 66%. Similar to Model 1, Model 2 emitted warning signals as early as January 1996, although with a higher probability of almost 60%. While these early signals were outside the 24-month crisis window, the model continued to send signals throughout the period leading up to the crisis, albeit with lower magnitude compared to Model 1.

Regarding the second in-sample crisis of December 2001, Model 1 emitted warning signals from as early as the latter part of 1999. The signal strength reached approximately 58% in December 1999, which falls within the 24-month crisis window. After a decline to around 9% in the early months of 2000, the signals intensified again, reaching heights of approximately 70% in late 2000 and early 2001. By December 2001, the predicted probability had risen to 82%. Model 2 also sent signals indicating the December 2001 crisis, but the signal strengths were not as strong as those from Model 1. As December 2001 approached, Model 2's predicted probability was approximately 62%, which is 20% lower than that of Model 1.

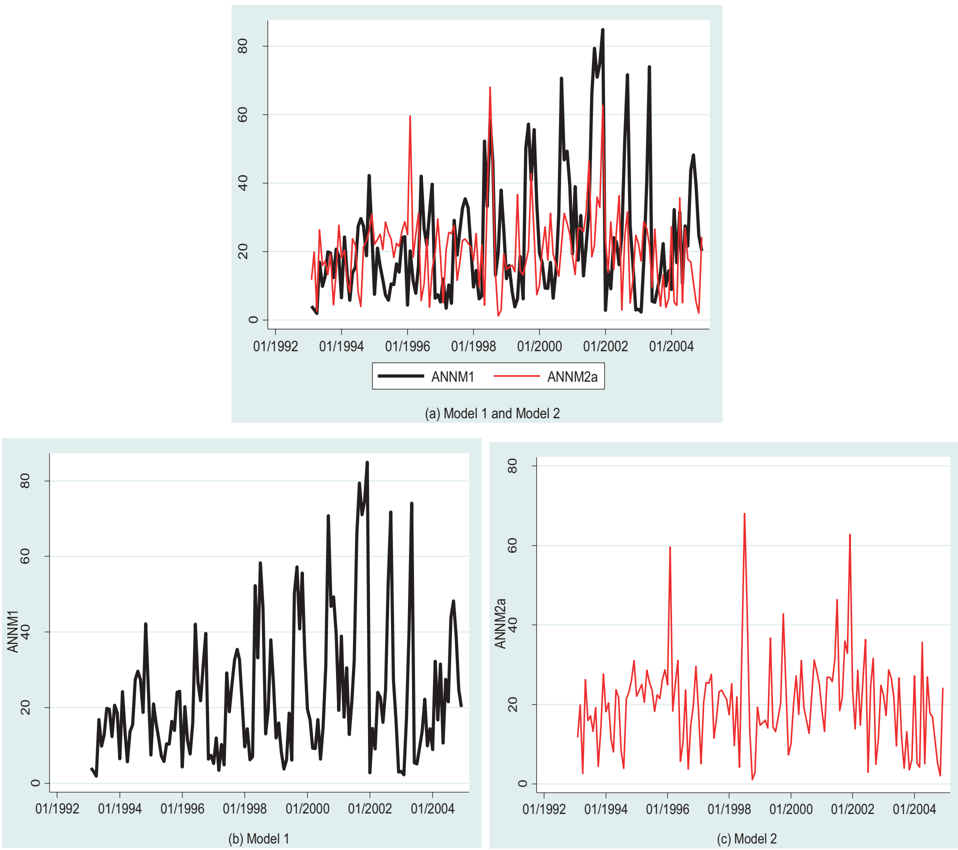

Figure 3. The Artificial Neural Networks models: In-Sample Predictions, 1993/02 – 2004/12. *Source*: Authors own calculations from data set.

Despite these models' abilities to correctly point to the two in-sample crises, many signals are also emitted outside the 24-month crisis window. This indicates that both models sent many false alarms.

3.1.2. Out-Of-Sample Prediction Currency Crisis, 2005/01 – 2017/03

Figure 4 below presents the out-of-sample probability predictions for currency crises in the period 2005/01 to 2017/03. In Figure 4 (a) both Model 1 and Model 2 predictions are plotted. Once again it is clear that the model plots do not overlap. Although it is observed that both models called the October 2008 GFC, Model 1 was sending more intense signals than model 2. For the analysis, the probability predictions of the models and model plots were separated and presented in Figure 4 (b), Model 1 and Figure 4 (c), Model 2. As depicted in Figure 4(b), Model 1 accurately predicts the October 2008 GFC. Early warning signals were emitted as early as the end of 2005 and the beginning of 2006. However, these signals were classified as false alarms because they occurred outside of the 24-month crisis window. Model 1 continued to send signals throughout the latter part of 2006 and early 2007 (within the 24-month window), reaching a peak probability of approximately 90%. Although the probability subsided to around 30%, it rose again to 90% in early 2008. By October 2008, the predicted probability of a crisis remained above 40%.

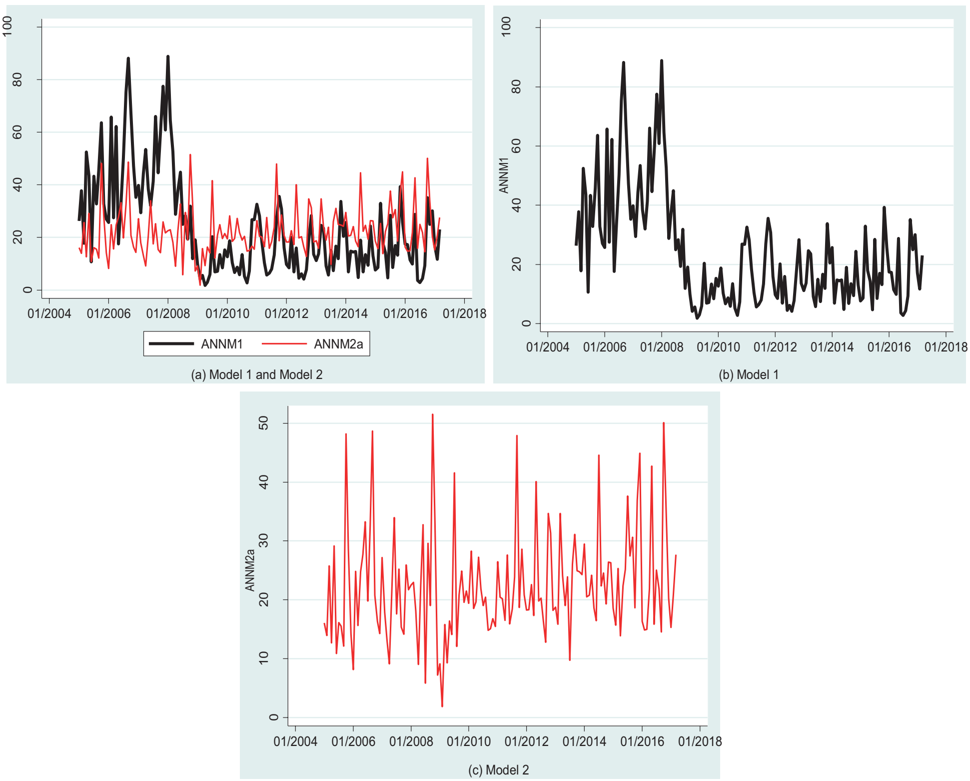

Figure 4. The Artificial Neural Networks models: Out-of-Sample Predictions, 2005/01 – 2017/03. *Source*: Authors own calculations from data set.

In contrast, Model 2 (Figure 4(c)) also correctly predicted the 2008 crisis, but with an approximate probability of 52%. However, the signals emitted 24 months prior to the crisis only reached magnitudes of 30%. For the out-of-sample period of 2005/01 to 2017/03, specifically after the October 2008 crises, Model 1 sends signals of an impending crisis of between 20% to 40%. Model 2 emits signals during the same time period to just below 50%. During that period South Africa never formally reported any occurrence of a currency crisis: fluctuating signals could have simply been attributed to South Africa's extremely volatile foreign market fluctuations or they could have pointed to some volatility of an impending crisis or simply regarded as false signals.

4. The Artificial Neural Network Model's Performance Evaluation

The primary objective of this study was to develop an early warning system (EWS) model capable of effectively predicting financial crises, rather than merely identifying tranquil periods. To achieve this, the study utilized various cut-off probabilities (20%, 30%, 40%, and 50%) as thresholds for crisis prediction. To evaluate the performance of the EWS model in predicting South African financial crises, both in-sample and out-of-sample, six assessment methods were employed (Berg and Pattillo, 1999):

(i) Percentage of Observations Called Correctly: Measures the overall accuracy of the model.net-

(ii) Percentage of Pre-Crisis Periods Called Correctly: Assesses the model's ability to accurately identify periods leading up to crises. (iii) Percentage of Tranquil Periods Called Correctly: Evaluates the model's effectiveness in distinguishing tranquil periods from those with potential crisis indicators.

(iv) False Alarms as a Percentage of Total Alarms: Measures the rate of incorrect crisis predictions.

(v) Diebold and Rudebusch's (1989) Quadratic Probability Score (QPS): Quantifies the accuracy of probabilistic forecasts.

(vi) Global Score Bias (GSB): Measures the overall bias in the model's predictions.

For a more detailed explanation of these methods, please refer to the appendix. The results of the performance evaluation are presented in Table 6. Following Berg and Pattillo (1999), the performance of the models can be seen from its ability to capture the whole observation when there is a crisis and when there is no crisis. Thus, from Table 6, it is seen that as the threshold increases to from 20% to 50%, the percentage of observations correctly called also increases in both models. With regard to the first out-of-sample probabilities, as the threshold increases from 20% to 50%, the percentage of observations correctly called decreases. The same is true for the second out-of-sample period. To conduct a more comprehensive evaluation, it is useful to consider the percentage of pre-crisis periods correctly identified as well as the percentage of tranquil periods correctly classified. In the in-sample periods, as the cut-off probability increases, a decrease in the percentage of pre-crisis periods correctly called is observed, while the percentage of tranquil periods correctly called follows a similar trend. For the first out-of-sample period, both Model 1 and Model 2 demonstrated the ability to correctly identify crises within the 24-month window prior to their occurrence. Model 1 exhibited a higher success rate, correctly calling crises in 96% of cases compared to Model 2's 52%. However, both models also generated a high percentage of false alarms, indicating the need for further refinement. Comparable results were observed in the second out-of-sample period, although with slightly reduced performance. In terms of false alarms, both in-sample and out-of-sample periods showed a decreasing trend in false alarm rates as the cut-off probability threshold increased, aligning with the observations depicted in Figure 4. In summary, the non-linear relationship shows the delicate balance between correctly detecting crises (true positives) and not misclassifying calm times (false alarms). The percentage of false alarms goes down as the threshold goes up, but some crises are missed. This is why the percentage of false alarms goes down in a way that is not linear.

Table 6. Artificial Neural Networks Models' Performance Evaluation

Thresholds (Pr*)	Assessment methods	In-sample		Out-of-sample			
		1993/02-2004/12		2005/01-2008/10		2008/11-2017/03	
		M1	M2	M1	M2	M1	M2
20%	% of observations correctly called	59%	55%	61%	52%	34%	56%
	% of pre-crisis periods correctly called	56%	60%	96%	52%	23%	57%
	% of tranquil periods correctly called	40%	48%	81%	48%	32%	48%
	% of false alarms of total alarms	60%	60%	41%	43%	32%	22%
	QPS	0.643	0.839	1.000	0.957	0.520	1.140
	GSB	0.113	0.189	0.417	0.004	0.080	0.520
30%	% of observations correctly called	69%	65%	59%	43%	27%	32%
	% of pre-crisis periods correctly called	44%	18%	80%	12%	7%	15%
	% of tranquil periods correctly called	18%	9%	67%	19%	12%	16%
	% of false alarms of total alarms	44%	54%	41%	57%	38%	27%
	QPS	0.364	0.434	0.739	0.783	0.120	0.340
	GSB	0.006	0.017	0.160	0.185	0.002	0.024
40%	% of observations correctly called	63%	65%	54%	43%	25%	26%
	% of pre-crisis periods correctly called	20%	6%	56%	4%	0%	5%
	% of tranquil periods correctly called	14%	3%	48%	10%	0%	12%
	% of false alarms of total alarms	57%	50%	42%	67%	0%	43%
	QPS	0.336	0.378	0.365	0.870	0.080	0.200
	GSB	0.006	0.061	0.009	0.306	0.005	0.010
50%	% of observations correctly called	66%	66%	50%	48%	25%	26%
	% of pre-crisis periods correctly called	16%	4%	36%	4%	0%	1%
	% of tranquil periods correctly called	6%	1%	33%	0%	0%	0%
	% of false alarms of total alarms	43%	33%	44%	0%	0%	0%
	QPS	0.350	0.392	0.200	0.870	0.080	0.120
	GSB	0.028	0.077	0.024	0.378	0.005	0.006

Source: Authors own calculations from data set.

Both models have performed well on the whole, as indicated by the small values of the QPS and GSB in which a value of zero indicates perfection and a value of 2 indicates model failure (Handoyo et al., 2020; Gupta & Kumar, 2022b). Furthermore, based on these two measures it is obvious that the Model 1 outperforms Model 2. This confirms the results of Zhang et al. (2021), showing that an artificial neural network model with more input neurons is better than that with fewer input neurons.

5. Conclusion

The study has developed and estimated an artificial neural network (ANN) model as an alternative EWS for predicting currency crises in South Africa. Two models were constructed: a comprehensive model (Model 1) incorporating ten input variables and a reduced model (Model 2) using only four statistically significant variables. The purpose of estimating the reduced model was to investigate the impact of the number of input variables on model performance, addressing a common point of contention in the literature. Performance evaluations confirmed that models with more input variables outperform those with fewer. The study also identified the key contributors to South Africa's currency crises: the ratio of budget deficits to GDP, followed by the growth rate of the ratio of domestic credit to GDP, and then changes in the international liquidity position.

The ANN models demonstrated their effectiveness by accurately predicting currency crises up to 24 months in advance, both in-sample and out-of-sample. Both models successfully identified all three notable currency crises in South Africa during the study period. However, Model 1, with its larger number of input variables, was able to provide earlier warning signals, up to 24 months prior to crisis occurrences, compared to Model 2. Additionally, both models correctly predicted the absence of currency crises after October 2008, a period during which South Africa did not report any such events. A particularly noteworthy feature of the ANN model was its ability to predict the probability of the out-of-sample GFC without requiring prior training with a target output. This demonstrates the model's capacity to identify emerging trends and patterns that may indicate potential crises. While the ANN models demonstrated promising results, they also generated a considerable number of false alarms during the 1998 and 2001 periods. This suggests a potential limitation of the models, as they were unable to effectively differentiate between currency crises and other economic vulnerabilities. Further research is needed to explore potential relationships between various economic vulnerabilities and their impact on currency crises. Overall, the study concludes that ANNs offer a promising alternative to traditional EWS tools for predicting currency crises in South Africa. The models' ability to provide early warnings and accurately identify the key contributing factors highlights their potential value in risk management and policymaking.

Author contributions

Fernandes-Gondoza, G: Conceptualization, Formal analysis, Data curation, Investigation, Methodology, Writing – original draft, review & editing. Ncwadi, R: Project administration, Supervision, Validation, Writing – original draft. Fernandes, J: Data curation, Software, Resources, Methodology, Data Visualization, Writing – original draft. Nyika, F: Writing – Review & editing, Updating of manuscript literature.

Use of AI tools declaration

Artificial Intelligence (AI) tools were not used to write this manuscript.

Acknowledgments

No funding was used in the writing of this article.

Conflict of interest

All authors declare no conflicts of interest in this article.

References

Abdolrasol, M. G., Hussain, S. S., Ustun, T. S., Sarker, M. R., Hannan, M. A., Mohamed, R., (...) & Milad, A. (2021). Artificial neural networks-based optimization techniques: A review. *Electronics*, *10*(21), 2689. https://doi.org/10.3390/electronics10212689

Alaminos Aguilera, D., Peláez Sánchez, J. I., Salas Compás, M. B., & Fernández Gámez, M. Á. (2021). Sovereign Debt and Currency Crises Prediction Models Using Machine Learning Techniques. *Symmetry*, *13*(4), 652. https://doi.org/10.3390/sym13040652

Alonso-Alvarez, I., & Molina, L. (2023). How to foresee crises? A new synthetic index of vulnerabilities for emerging economies. *Economic Modelling*, *(125)*, 106304. https://doi.org/10.1016/j.econmod.2023.106304

Ataman, G., & Kahraman, S. (2022). Stock market prediction in BRICS countries using linear regression and artificial neural network hybrid models. *The Singapore Economic Review*, *67*(02), 635-653. https://doi.org/10.1142/S0217590821500521

Barthélémy, S., Gautier, V., & Rondeau, F. (2024). Early warning system for currency crises using long short-term memory and gated recurrent unit neural networks. *Journal of Forecasting*, 43(5), 1235-1262. https://doi.org/10.1002/for.3069

Belciug, S. (2020). Logistic regression paradigm for training a single-hidden layer feedforward neural network. Application to gene expression datasets for cancer research. *Journal of Biomedical Informatics*, (*102)*, 103373. https://doi.org/10.1016/j.jbi.2019.103373

Berg, A., & Pattillo, C. (1999). Predicting currency crises:: The indicators approach and an alternative. *Journal of international Money and Finance*, *18*(4), 561-586. https://doi.org/10.1016/S0261-5606(99)00024-8

Bernoth, K., & Herwartz, H. (2021). Exchange rates, foreign currency exposure and sovereign risk. *Journal of International Money and Finance*, *(117),* 102454. https://doi.org/10.1016/j.jimonfin.2021.102454

Beyers, C., Essel-Mensah, K. A., & Tsomocos, D. P. (2024). A computable general equilibrium model of the monetary policy implications for financial stability in South Africa. *South African Journal of Economics*, 92(4), 415-433. https://doi.org/10.1111/saje.12383

Bhundia, A. J., & Ricci, L. A. (2005). The Rand Crises of 1998 and 2001: What have we learned. In *Post-apartheid South Africa: The first ten years (*pp. 156-173). International Monetary Fund.

Biecek, P., & Burzykowski, T. (2021). *Explanatory model analysis: explore, explain, and examine predictive models*. Chapman and Hall.

Bluwstein, K., Buckmann, M., Joseph, A., Kapadia, S., & Şimşek, Ö. (2023). Credit growth, the yield curve and financial crisis prediction: Evidence from a machine learning approach. *Journal of International Economics*, *(145),* 103773. https://doi.org/10.1016/j.jinteco.2023.103773

Boloukian, B., & Safi-Esfahani, F. (2020). Recognition of words from brain-generated signals of speech-impaired people: Application of autoencoders as a neural Turing machine controller in deep neural networks. *Neural Networks, (121),* 186-207. https://doi.org/10.1016/j.neunet.2019.07.012

Bordo, M., Eichengreen, B., Klingebiel, D., & Martinez-Peria, M. S. (2001). Is the crisis problem growing more severe? *Economic policy*, *16*(32), 52-82.

Capra, M., Bussolino, B., Marchisio, A., Masera, G., Martina, M., & Shafique, M. (2020). Hardware and software optimizations for accelerating deep neural networks: Survey of current trends, challenges, and the road ahead. *IEEE Access*, *(8),* 225134-225180. https://doi.org/10.1109/ACCESS.2020.3039858

Chen, Y., Song, L., Liu, Y., Yang, L., & Li, D. (2020). A review of the artificial neural network models for water quality prediction. *Applied Sciences*, *10*(17), 5776. https://doi.org/10.3390/app10175776

Dastres, R., & Soori, M. (2021). Artificial neural network systems. *International Journal of Imaging and Robotics*, *21*(2), 13-25.

Diebold, F. X., & Rudebusch, G. D. (1989). Scoring the Leading Indicators. *Journal of Business,* 62(3), 369-391. https://doi.org/10.1515/9780691219585-017

Dominguez, K. M. (2020). Revisiting exchange rate rules. *IMF Economic Review*, *68*(3), 693-719. https://doi.org/10.1057/s41308-020-00120-6

Elalem, Y. K., Maier, S., & Seifert, R. W. (2023). A Machine Learning-Based Framework For Forecasting Sales Of New Products With Short Life Cycles Using Deep Neural Networks. *International Journal Of Forecasting*, *39*(4), 1874-1894. https://doi.org/10.1016/j.ijforecast.2022.09.005

Engelbrecht, A. P. (2007). *Computational Intelligence: An Introduction*. John Wiley & Sons.

Farizawani, A. G., Puteh, M., Marina, Y., & Rivaie, A. (2020). A review of artificial neural network learning rule based on multiple variants of conjugate gradient approaches. *Journal of Physics: Conference Series*, 1529(2), 022040. https://doi.org/10.1088/1742-6596/1529/2/022040

Georgiadis, G., & Zhu, F. (2021). Foreign-currency exposures and the financial channel of exchange rates: Eroding monetary policy autonomy in small open economies? *Journal of International Money and Finance*, *(110)*, 102265. https://doi.org/10.1016/j.jimonfin.2020.102265

Ghosh, A. R., Ostry, J. D., & Qureshi, M. S. (2015). Exchange rate management and crisis susceptibility: A reassessment. *IMF Economic Review*, *63*(1), 238-276. https://doi.org/10.1057/imfer.2014.29

Goniewicz, K., Khorram-Manesh, A., & Burkle, F. M. (2023). Beyond boundaries: Addressing climate change, violence, and public health. *Prehospital and Disaster Medicine*, *38*(5), 551-554. https://doi.org/10.1017/S1049023X23006271

Goodfellow, I., Bengio, Y., & Courville, A. (2016). *Deep Learning*. MIT Press.

Guei, K. M., & Choga, I. (2022). Exchange rate volatility and regional trade agreements in Southern Africa. *Economic Change and Restructuring*, *55*(2), 635-652. https://doi.org/10.1007/s10644-021-09323-x

Gupta, N., & Kumar, A. (2022a). Artificial neural networks for developing early warning system for banking system: Indian context. *International Journal of Economics and Business Research*, *23*(2), 229-254. https://doi.org/10.1504/IJEBR.2022.120652

Gupta, N., & Kumar, A. (2022b). Comparing parametric, semi parametric and non-parametric early warning systems for banking crisis: Indian context. *Global Business and Economics Review*, *26*(2), 111-134.

Hamid, S. A. (2004). Primer On Using Neural Networks For Forecasting Market Variables. *Working Paper Of The Centre For Financial Studies, Southern New Hampshire University*, 2004-3, 1-48.

Handoyo, R. D., Erlando, A., & Astutik, N. T. (2020). Analysis of twin deficits hypothesis in Indonesia and its impact on financial crisis. *Heliyon*, *6*(1), e03248. https://doi.org/10.1016/j.heliyon.2020.e03248

Hashemi, M. (2019). Enlarging smaller images before inputting into convolutional neural network: zero-padding vs. interpolation. *Journal of Big Data*, *6*(1), 1-13. https://doi.org/10.1186/s40537-019-0263-7

He, C., Ma, M., & Wang, P. (2020). Extract interpretability-accuracy balanced rules from artificial neural networks: A review. *Neurocomputing*, *(387)*, 346-358. https://doi.org/10.1016/j.neucom.2020.01.036

Heng, S. Y., Ridwan, W. M., Kumar, P., Ahmed, A. N., Fai, C. M., Birima, A. H., & El-Shafie, A. (2022). Artificial neural network model with different backpropagation algorithms and meteorological data for solar radiation prediction. *Scientific reports*, *12*(1), 10457. https://doi.org/10.1038/s41598-022-13532-3

Heo, W., Lee, J. M., Park, N., & Grable, J. E. (2020). Using artificial neural network techniques to improve the description and prediction of household financial ratios. *Journal of Behavioral and Experimental Finance*, *(25)*, 100273. https://doi.org/10.1016/j.jbef.2020.100273

HongXing, Y., Hafiz, M. N., Answer, M. U., Memon, B. A., & Akhtar, M. (2022). Evaluation optimal prediction performance of MLMs on high-volatile financial market data. *International Journal of Advanced Computer Science and Applications*, *13*(1), 239-246.

Huang, L., Xie, G., Zhao, W., Gu, Y., & Huang, Y. (2021). Regional logistics demand forecasting: a BP neural network approach. *Complex & Intelligent Systems*, 1-16. https://doi.org/10.1007/s40747-021-00297-x

International Monetary Fund Staff Report. (2009). *Zimbabwe: Staff Report for the 2009 Articla IV Consultation*. International Monetary Fund. https://www.imf.org/external/pubs/ft/scr/2009/cr09139.pdf

Jones, J., Deininger, M., & Pandey, S. (2019). Response and recovery: Does the delay between a crisis and an IMF loan affect the length of recovery? *International Advances in Economic Research, (25),* 481-482. https://doi.org/10.1007/s11294-019-09756-1

Juhro, S.M., Siregar, R.Y., & Trisnanto, B. (2022). Exchange rate policy and capital flow management. In *Central Bank Policy Mix: Issues, Challenges, and Policy Responses: Handbook of Central Banking Studies* (pp. 51-72). Springer Nature Singapore.

Kabir, H.D., Abdar, M., Khosravi, A., Jalali, S.M.J., Atiya, A.F., Nahavandi, S., & Srinivasan, D. (2022). Spinalnet: Deep neural network with gradual input. *IEEE Transactions on Artificial Intelligence*, *4*(5), 1165-1177. https://doi.org/10.1109/TAI.2022.3185179

Kali, M. (2023). A comparative analysis of the causes of the protests in Southern Africa. *SN Social Sciences*, *3*(2), 28. https://doi.org/10.1007/s43545-023-00613-x

Kaminsky, G.L., Lizondo, S., & Reinhart, C.M. (1998). Leading Indicators of Currency Crises. *IMF Staff Papers,* 45(1), 1–48. https://doi.org/10.2307/3867328

Karsoliya, S. (2012). Approximating number of hidden layer neurons in multiple hidden layer BPNN architecture. *International journal of engineering trends and technology*, *3*(6), 714-717.

Khan, M.A., Saqib, S., Alyas, T., Rehman, A.U., Saeed, Y., Zeb, A., (...) & Mohamed, E.M. (2020). Effective demand forecasting model using business intelligence empowered with machine learning. *IEEE access*, *(8),* 116013-116023. https://doi.org/10.1109/ACCESS.2020.3003790

Kuok, R.U.K., Koo, T.T., & Lim, C. (2023). Economic policy uncertainty and international tourism demand: A global vector autoregressive approach. *Journal of Travel Research*, *62*(3), 540-562. https://doi.org/10.1177/00472875211072551

Kuok, K.K., Said, K.A.M., & Yun, C.M. (2024). Neural Network–A Black Box Model. *Metaheuristic Algorithms and Neural Networks in Hydrology*, *(1),* 1-34.

Kurani, A., Doshi, P., Vakharia, A., & Shah, M. (2023). A comprehensive comparative study of artificial neural network (ANN) and support vector machines (SVM) on stock forecasting. *Annals of Data Science*, *10*(1), 183-208. https://doi.org/10.1007/s40745-021-00344-x

Laubscher, P. (2020). A history of South Africa's modern business cycle. *Business Cycles and Structural Change in South Africa: An Integrated View*, 111-166.

Li, Y., Deng, S., Dong, X., Gong, R., & Gu, S. (2021). A free lunch from ANN: Towards efficient, accurate spiking neural networks calibration. In *International conference on machine learning* (pp. 6316-6325). PMLR.

Liu, H., Qin, Y., Chen, H.Y., Wu, J., Ma, J., Du, Z., (...) & Wang, H. (2023). Artificial neuronal devices based on emerging materials: neuronal dynamics and applications. *Advanced Materials*, *35*(37), 2205047. https://doi.org/10.1002/adma.202205047

Liu, L., Chen, C., & Wang, B. (2022). Predicting financial crises with machine learning methods. *Journal of Forecasting*, *41*(5), 871-910. https://doi.org/10.1002/for.2840

Longo, L., Riccaboni, M., & Rungi, A. (2022). A neural network ensemble approach for GDP forecasting. *Journal of Economic Dynamics and Control*, *(134),* 104278. https://doi.org/10.1016/j.jedc.2021.104278

Madhiarasan, M., & Louzazni, M. (2022). Combined Long Short-Term Memory Network-Based Short-Term Prediction of Solar Irradiance. *International Journal of Photoenergy*, *2022*(1), 1004051. https://doi.org/10.1155/2022/1004051

Mekkaoui, S. E., Benabbou, L., & Berrado, A. (2023). Rule-Extraction Methods from Feedforward Neural Networks: A Systematic Literature Review. *arXiv preprint arXiv:2312.12878*.

Mirzaei, N. (2023). 2007–2008 Financial Crisis. In *The Palgrave Encyclopedia of Global Security Studies* (pp. 1-6). Springer International Publishing.

Montesinos López, O. A., Montesinos López, A., & Crossa, J. (2022). Fundamentals of artificial neural networks and deep learning. In *Multivariate statistical machine learning methods for genomic prediction* (pp. 379-425). Springer International Publishing.

Nguyen, T., Nguyen-Phuoc, D. Q., & Wong, Y. D. (2021). Developing artificial neural networks to estimate real-time onboard bus ride comfort. *Neural Computing and Applications*, *33*(10), 5287-5299. https://doi.org/10.1007/s00521-020-05318-3

Pokou, F., Sadefo Kamdem, J., & Benhmad, F. (2024). Hybridization of ARIMA with learning models for forecasting of stock market time series. *Computational Economics*, *63*(4), 1349-1399. https://doi.org/10.1007/s10614-023-10499-9

Raj, A. S. A., Kumarasankaralingam, L., Balamurugan, M., Maheswari, B., Gowri, J., & Dutta, A. (2024). Forecasting the Economic Crisis of Sri Lanka: Application of Machine Learning Algorithms for Time Series Data. *Procedia Computer Science, (235),* 1087-1096. https://doi.org/10.1016/j.procs.2024.04.103

Rast, P., Martin, S. R., Liu, S., & Williams, D. R. (2022). A new frontier for studying within-person variability: Bayesian multivariate generalized autoregressive conditional heteroskedasticity models. *Psychological methods*, *27*(5), 856. https://psycnet.apa.org/doi/10.1037/met0000357

Raudzingana, D. (2020). *Determinants of corporate default: The case of South Africa*. Wits University.

Rosa, J. P., Guerra, D. J., Horta, N. C., Martins, R. M., Lourenço, N. C., Rosa, J. P., (...) & Lourenço, N. C. (2020). Overview of artificial neural networks. *Using artificial neural networks for analog integrated circuit design automation*, 21-44.

Šestanović, T., & Arnerić, J. (2021). Can recurrent neural networks predict inflation in euro zone professional forecasters? *Mathematics*, *9*(19), 2486. https://doi.org/10.3390/math9192486

Singh, R. K., Rani, M., Bhagavathula, A. S., Sah, R., Rodriguez-Morales, A. J., Kalita, H., (...) & Kumar, P. (2020). Prediction of the COVID-19 pandemic for the top 15 affected countries: Advanced autoregressive integrated moving average (ARIMA) model. *JMIR public health and surveillance*, *6*(2), e19115. https://doi.org/10.2196/19115

South African Reserve Bank. (2012). *Introduction to Central Banking.* South African Reserve Bank College.

Thike, P. H., Zhao, Z., Shi, P., & Jin, Y. (2020). Significance of artificial neural network analytical models in materials' performance prediction. *Bulletin of Materials Science*, *43*(1), 211. https://doi.org/10.1007/s12034-020-02154-y

Tölö, E. (2020). Predicting systemic financial crises with recurrent neural networks. *Journal of Financial Stability*, *(49),* 100746. https://doi.org/10.1016/j.jfs.2020.100746

Uzair, M., & Jamil, N. (2020). Effects of hidden layers on the efficiency of neural networks. In *2020 IEEE 23rd international multitopic conference (INMIC)* (pp. 1-6). IEEE.

Widodo, I., Pramono, H. S, Aswin, M. (2021). Analysis of the Compatibility of Student Satisfaction Index with Course Value Using Significance Correlation and Backpropagation Approach. *International Journal of Computer Applications Technology and Research*, *10*(7), 192-197.

Xu, A., Chang, H., Xu, Y., Li, R., Li, X., & Zhao, Y. (2021). Applying artificial neural networks (ANNs) to solve solid waste-related issues: A critical review. *Waste Management, (124),* 385-402. https://doi.org/10.1016/j.wasman.2021.02.029

Yadav, A., Chithaluru, P., Singh, A., Joshi, D., Elkamchouchi, D. H., Pérez-Oleaga, C. M., & Anand, D. (2022). An enhanced feed-forward back propagation Levenberg–Marquardt algorithm for suspended sediment yield modeling. *Water*, *14*(22), 3714. https://doi.org/10.3390/w14223714

Zhang, C., Bengio, S., Hardt, M., Recht, B., & Vinyals, O. (2021). Understanding deep learning (still) requires rethinking generalization. *Communications of the ACM*, *64*(3), 107-115. https://doi.org/10.1145/3446776

Zhang, M., Wu, J., Belatreche, A., Pan, Z., Xie, X., Chua, Y., (...) & Li, H. (2020). Supervised learning in spiking neural networks with synaptic delay-weight plasticity. *Neurocomputing, (409),* 103-118. https://doi.org/10.1016/j.neucom.2020.03.079

Appendix

In evaluating the performance of the four EWS models, suggest the following assessment criteria: the percentage of observations correctly called; the percentage of pre-crisis periods correctly called; the percentage of tranquil periods called and the percentage of false alarms to total alarms. In order to carry out these assessment methods, use is made of the two-by-two matrix in Table A1, whereby the crisis signal was classified into four categories, namely, *A*, *B*, *C* and *D* (See Table A1 below for explanation of categories). The assessment methods are given by the following equations:

Table A1. Classifying crisis signals into categories

• The percentage of observations called correctly = (A + D)/(A + B + C + D)	(16)
• The percentage of pre-crisis periods called correctly = A/(A + C)	(17)
• The percentage of tranquil periods called correctly = B/(B + D)	(18)
• The false alarms as a percentage of the total alarms = B/(A + B)	(19)
• The Quadratic probability score or $QPS = \frac{1}{T}\sum_{t=1}^{T} 2\left(P_t - R_t\right)^2$	(20)
• The Global score bias or $GSB = 2\left(\bar{P} - \bar{R}\right)^2$	(21)

If the values in Equations 16, 17 and 18 are high, the model performs better. In contrast, in Equation 19, the lower the ratio the better the indictor. The QPS and GSB allow for the assessment of the average closeness of the realisation of a crisis (R_t) and the probability prediction of a crisis (P_t) during the signalling horizon. Recall that if the event is a crisis, it scores 1 and 0 otherwise. The QPS statistic lies in the range between zero and two. The performance quality of the composite index is better the closer the test statistic is to zero. The performance quality of the composite index can be affected by the sample size. The larger the sample size, the more robust the test statistics (Diebold & Rudebusch, 1989).

The ten leading indicators used in this study were selected using the signal approach. The performance of an indicator in predicting a crisis can be shown in the value of its noise-to-signal ratio (NSR). The NSR can be presented by taking the ratio of the percentage of bad signals over the percentage of good signals (Kaminsky et al., 1998),

$$NSR_t = \left[\frac{B_t}{B_t + D_t}\right] / \left[\frac{A_t}{A_t + C_t}\right] \tag{15}$$

In Equation 15, if the $NSR \geq 1$ for an indicator, this indicates excessive noise and thus contributes less to the probability prediction of a currency crisis. The desired outcome is to have an *NRS* that is less than 1. Kaminsky et al. (1998) specified that the lower the value of the *NRS*, the more powerful the indicator in predicting crises.