Forecasting Exchange Rates of Foreign Currencies Using Artificial Neural Networks

Submitted in partial fulfillment of
the requirements for the award of the degree of

Master of Science
in
Software Engineering

Submitted by

Shravan Venkataraman

Under the guidance of
N. Manikandan

School of Information Technology and Engineering
VELLORE INSTITUTE OF TECHNOLOGY, VELLORE
Vellore, Tamilnadu, India – 632 014

Winter Semester 2013

School of Information Technology and Engineering

Vellore Institute of Technology

Certificate

This is to certify that the project report entitled *Forecasting Exchange Rates of Foreign Currencies using Artificial Neural Networks* submitted by V. Shravan, 10MSE0236 to Vellore Institute of Technology University, Vellore in partial fulfillment of the requirement for the award of the degree of MS in Software Engineering is a record of bonafide work carried out by him under my guidance. The project fulfills the requirement as per the regulations of this institute and in my opinion meets the necessary standards of submission. The contents of this report have not been submitted and will not be submitted either in part or in full for the award of any other degree or diploma in this institute or any other institute or university and the same is certified.

N. Manikandan
Project Guide
School of Information Technology and Engineering
VIT University, Vellore, Tamilnadu, India.

Acknowledgments

I would like to thank Dr. G. Viswanathan, the chancellor of Vellore Institute of Technology University, for giving me the opportunity to pursue M.S Software Engineering.

I would also like to thank Prof.A.Jayaram Reddy, the program manager of M.S Software Engineering, for giving me adequate guidance to carry out my project.

I would like to extend my sincere thanks to Prof. N. Manikandan, my guide for his timely help and support and for guiding me to complete the project successfully.

Abstract

Forecasting currency exchange rates, stock prices, etc., based on time series data is one area where extensive research is going on for the last three decades. In the initial days, statistical models and methods solved the problems with financial analysis and prediction. In the last two decades, numerousartificial neural networks based learning models have been proposed to solve the problems of financial data and get accurate results in predicting the future trends and prices. This project uses the strengths of multivariate econometric time series models and artificial neural networks to provide an adaptive approach for predicting exchange rates. We propose a hybrid method for predicting exchange rates. In the first step, a time series model based on artificial neural networks generates the estimates of the currency exchange rates and other technical parameters that are used to forecast the exchange price of currency of our choice. In the second step, an error correction back-propagation neural network is used to correct the errors of the estimates. The proposed two-step model produces better accuracy in results than the single step models.

Contents

List of Figures

Chapter 1

Problem Statement

The historical data of exchange rates incorporates information used to capture the intricate dependency between the future exchange rates and the rates of the past. For over two decades, the Box-Jenkins ARIMA model was used in time-series forecasting and was also used as a benchmark to compare other models that were being created. However, ARIMA is a general uni-variate model which was developed on the assumption that the forecasted time series are linear and stationaryin nature. So, there was a need to create a non-linear model to use in the prediction of currency exchange rates.

Recent research has led to the usage of Artificial Neural Networks for the prediction of currency exchange rates because of the function approximating nature of ANNs in prediction and modelling, the reason being that the ANNs are of great assistance in multivariate analysis. The multivariate analysis involves not just the lagged time series as an indicator to predict future behaviour but also other indicators (such as technical, economic, and social indicators) combined along with the time series in the forecasting process. ANN is more appropriate for the time series forecasting problem because of its non-parametric and adaptive properties. Research has shown that ANNs can map any non-linear function without a prior assumption about the data and are hence used as universal function approximators.

Chapter 2

Introduction

2.1 Background and Recent Research

2.1.1 Background

Since the abandonment of the fixed rates in foreign exchange and implementing the floating exchange rate system, the explanation of movement of exchange rates was most sought after by researchers. Many kinds of forecasting methods were developed by the researchers and experts. Technical analysis and fundamental analysis are the two major forecasting methods that are used in finance.

The balance of payments method determined the foreign exchange rates in the beginning. It was a method to list payments and receipts in international transactions for a givencountry. The imports and exports of goods determined the balance. Therefore, foreign exchange rates prediction was not difficultat that time. Later on, when the foreign exchange wereliquidated, a lot of factors became correlated to each other in the currency market and therefore the prediction of foreign exchange rates became complex.

The forecasting of exchange rates involve two steps; first, the time series should be analysed and then, the best forecasting techniques are chosen for the given time series based on the analysis and out of those, the best method is chosen for the forecasting purpose. In modern time series forecasting involving exchange rates prediction, many factorsare correlated to each other namely, economic factors, political factors and even psychological factors that affect the foreign exchange rates by interacting in a complex fashion. Hence, the exchange rates are noisy, chaotic, and non-stationary. But research has shown that non-random and predictable behaviour can be emphasized in liquid market areas such foreign exchange market.

2.1.2 Literature Survey

Much research has been going on in forecasting the currency exchange prices in Forex Market for the last two decades using Artificial Neural Network as the frontrunner to perform the forecasting due to its non-linear, predictive and adaptive capabilities. A literature review on the previous works gave rise to following derivations.

The Currency exchange rate data in the foreign exchange market is chaotic, random, and noisyin nature. In earlier days, the Random Walk Model and the Efficient Market Hypothesis were the two most widely used models based on fundamental analysis for forecasting the currency rates. These models were based on the belief that the foreign exchange data was noisy and random. But then, research through statistical tests showed with a significance of 95%, that the forex time series isnot randomly distributed [7]. After much research using linear and then non-linear models, Artificial Neural Networks are said to perform better than ARIMA model in forecasting the foreign exchange rates due to the non linear nature of the time series data [5]. Useful prediction and significant paper profit can be made with simple technical indicators without the use of extensive market data or knowledge [1][5][6].

Models:

In [1], the author has used delayed time series data and technical indicators as input after performing R/S analysis on the time series data. He has also said the market is volatile anda significantly large amount of data is required for the neural network to identify the regularities in the data and make use of those regularities in the prediction process. In [3], the author created a model that is used to predict long-term trends, and also an alarm generator for short term forecasting to predict turning points in currency movements. In [5], the author discusses prediction modelling of foreign currency exchange rate using three algorithmsStandard Back-propagation, Scaled Conjugate Gradient, and Back-propagation with Bayesian regularization using Artificial Neural Networks as the method of implementation. In [9], the author uses a Cascaded Functional Link ANN to create a predictive model. The model uses an adaptive adjustment of connecting weights using a novel weight update learning rule after using trigonometric expansion to expand the input data in a non-linear manner.

Input:

Some researchers have used the delayed time series data as input while few other researchers have used Moving Average (MA) of the time series data as the input [5] to the neural networks since the Moving Average data tends to be a smoothed version of the delayed time series and with much less noise. Also, the MA technique is said to perform well only when the market follows a trend. However, it performs poorly when the index changes direction [7]. In [4], the author used daily time series data as input and used it to predict the Euro/USD exchange rate using genetic algorithm with Artificial Neural Networks up to three days ahead of the last data available. He used both macro-economic variables and market data for inputs from which it was learnt that the exchange rate of Euro-USD was conditional. Some researchers have shown that a few technical indicators influence the exchange rates strongly. The said indicators are Nasdaq Index, Gold Spot Price, Average returns of the government bonds, Crude oil price [4], Trade relation, Cost of imports and exports [5]. To tackle the market evolution, the input data should be kept consistent. Otherwise, after training, the network is said to degrade in performance. One way to achieve consistency is to periodically replace the past data with recent data [7] [8]. Increasing the number of inputs is said to have not much effect in improving the performance [5].

Layers:

As [2] specified in his work, the best activation function that can be used in the neural network design for prediction of time series data is a bipolar function [-1, 1] or a binary function [0, 1]. Reports [2][8][10] suggest that performance of the network does not improve when more than 2 hidden layers are used in the network. It has been reported that the presence ofmore than 2 hidden layers only make the network more complex. It also makes the training process difficult and a danger of over fitting is present in such networks having over two hidden layers. All the works in the neural network area has suggested the use of a maximum of two to three hidden layers as the optimum method to extract the best performance from the network.

Training Process:

To capture the regularities, [3] proposed a moving window model that uses a two layer back-propagation network with a fixed number of inputs modelling a window along the time series in fixed steps to capture the regularities in the underlying data. [5] For a large-scale problem back-propagation learns veryslowly and convergence islargelydependent on choosing suitable values of learning rate, momentum factor, step size. Literature review revealed that the testing and validation set should be exactly one fourth to one eighth of the training set. [7] Kaufman suggests that a balanced split 70-15-15 for training, validation, and testing sets. Sigmoid function commonly used transfer function since the time series data is non-linear in nature[8]. However, some researchers have suggested the use of hypertangent function and tansig function too, as transfer functions [4][10].

Training Process:

Usually the Normalised Mean Square Error is the most widely used metric [2][5][6] to measure the efficiency and the correctness of the trained neural network during the testing and validation process. But some researchers [2][5]have used a few other metrics too, to compare the performances and get the network to perform the best at any givensituation. The error metrics that are used alongside NMSE to measure the correctness of the trained network are MAE (Mean Absolute Error), DS(Directional Symmetry), CU(Correct Uptrend), CD(Correct Downtrend), PMAD(Percentage Mean Absolute Deviation), RMSE (Root Mean Squared Error), MAPE(Mean Absolute Percentage Error), and MAE(Mean Absolute Error). Some have also reported using hit rate as a measure of correctness of a network [6].Based on performance, it was reported that having small NMSE in validation and testing is more important than having a small NMSE for training [6].

Output:

Using models based on Artificial Neural Networks, reports show that a correctness of up to76% hasbeen achieved in the earlier works with the variants of Back-Propagation algorithm as the learning method [1] [6].

Challenges

Improving the accuracy and quality in time-series forecasting is an important and a difficult task for decision makers in many areas. There are many time series based models available. But still research for improving the effectiveness of those forecasting models has not stopped.

If non-linearities are present in the dataset and the sample sizes are restricted, then classical statistical techniques do not work well for such kinds of applications. Time series forecasting is one such application. So, we need a non-linear and an adaptive model which will not only analyse and forecast the trend but also the time series with a considerable amount of accuracy.

Essence of Approach

This project uses a sequential approach to problem solving. First, we construct the neural networks for the prediction of individual input variables that predict the future values using the historical values. Then we construct the final neural network which utilises all the input variables to give the prediction of the currency exchange rate value thatwe need.

2.2 Motivation

Earlier research in this area has proven that simple technical indicators are enough to obtain useful predictions and significant paper profit, using no extensive knowledge or data related to the market. Research has brought out the fact that Neural Networks have great potential for financial forecasting

Chapter 3

System Overview

3.1 System Description

The system is a forecasting model built using neural networks, where the input layer takes the input all the input variables involved in the forecasting, both technical and economic variables. The hidden layers process the input variables and add value to the system. The hidden layers contain 20 neurons. The output layer yields a univariate output.

3.2 Components Description

The network basically consists of artificial neurons which are interconnected to form the network. The different components that are used in building the forecasting model are as follows.

3.2.1 NAR Network for the prediction of input variables:

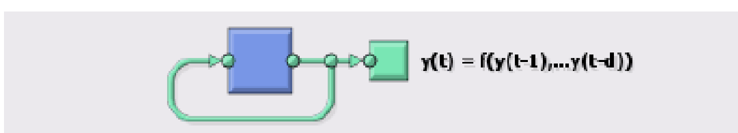

Figure 3.1: NAR Network

The NAR network takes the input variable y(t) where the network is trained with the past values of that time series to predict its future values. This NAR network is used to predict the variables close prices of the five currencies, gold and crude oil price, cpi and the inflation rate. Once trained with past values, the values are predicted from the point of prediction upto the point where a threshold level accuracy is present. To facilitate this, from the point of prediction, the predicted value is given as feedback input instead of using the original input values, using the closed loop backpropagation design. All the NAR networks are trained with a maximum of 35 hidden layer neurons and a delay value of 20. The maximum performances were reached at around 19 to 25 epochs.

- *Algorithms Used:* The network uses trainlm (a predefined algorithm function name for Levenberg Marquardt Backpropagation algorithm) to train the network.

- *Performance Measure:* Mean Squared Error (MSE) is used as the performance measure, calculating the mean squared difference between the expected target output and the actual output.

3.2.2 NARX Network for the prediction of output variable:

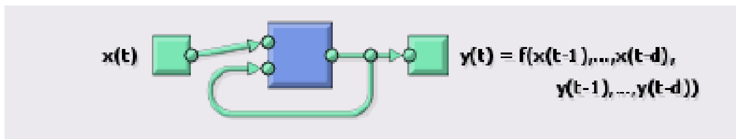

Figure 3.2: NARX Network

The following is the NARX network that is used for the prediction of the bid price of the desired currency to be forecasted, with the NAR network predictions as input to the trained network. The NARX network takes the input variables as exogenous inputs all the 10 input variables are taken for input in this network, and the network is trained by assigning the time series output variable that we want to predict as the target variable.

Once the network is trained using those exogenous input and required output, then, all the trained NAR network predictions of input variables can be supplied to this network as inputs to predict the target output with the presence of delay number of values from the point of prediction, to predict the next required number of steps that one wants to predict.

Usually the performance degrades with every step of prediction since we use the predicted values as feedback to targets instead of original values. Therefore a threshold is set to predict until a particular number of steps until which there is a good amount of accuracy.

The NARX network is trained with 25 number of hidden layer neurons and 20 as the delay value. The number of neurons is taken considerably high because of the high processing that a data such as a financial time series needs when in amounts of thousands. The delay is kept as 20 because of the high correlation of the target values with the values that belong to one month prior the current value in both input and target variables.

- *Algorithms Used:* The network uses trainlm (a predefined algorithm function name for Levenberg Marquardt Backpropagation algorithm) to train the network.

- *Performance Measure:* Mean Squared Error (MSE) is used as the performance measure, calculating the mean squared difference between the expected target output and the actual output.

3.2.3 Time-Series Data

Since the entire process of training and prediction is done in MATLAB, we use Excel input. We import the Excel file using the data import GUI and select the data which we need to train the network and the data used in prediction. Once the required data is selected, it is then converted into either of the following, to facilitate usage in the network construction. - Matrix - Cell Array - Column Vectors Usually a Matrix or Cell Array representation is used to store the required data in the workspace. The extracted data that is used for training/testing/validating/predicting are all stored in the workspace and the workspace is stored in a matlab file of extension .mat. This matlab file is accessed whenever the access to the required data is necessary. This is rather a simple way of storing data, since all of the data involved is numerical in nature, containing time steps of time series data. The data is retrieved by opening the mat file that has the required variables, in the workspace.

- *Dependencies* The NARX network depends on the other NAR network predictions for its input as it uses those predicted inputs to support the input layer and to predict the target output.

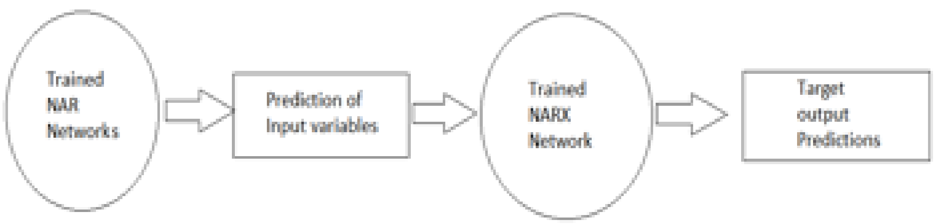

Figure 3.3: Dependencies

3.2.4 Learning Algorithm

Levenberg Marquardt Algorithm:

The Levenberg-Marquardt (LM) algorithm is the most widely used optimization algorithm. It outperforms simple gradient descent and other conjugate gradient methods in a wide variety of problems. The problem for which the LM algorithm provides a solution is called Nonlinear Least Squares Minimization. The LevenbergMarquardt algorithm blends the steepest descent method and the GaussNewton algorithm. Fortunately, it inherits the speed advantage of the GaussNewton algorithm and the stability of the steepest descent method. Its more robust than the Gauss Newton algorithm, because in many cases it can converge well even if the error surface is much more complex than the quadratic situation.

Chapter 4

Results & Discussion

4.1 Data Collection

The data used in this study are the daily foreign exchange rates of five different currencies against US Dollar, Prices of Crude Oil, Gold, US Inflation Rate and CPI from the period June 1993 to March 2013 made available by Oanda.com. We took into consideration the exchange rates of Australian Dollar (AUD), British Pound Sterling (GBP), Canadian Dollar (CAD), Swiss Franc (CHF) and Japanese Yen (JPY). The study took 4810 daily data into consideration of which the first 3367 rowswere used in training and the remaining 1443 rows were used for testing and evaluating the model.

4.2 Performance Metrics

$$\mathrm{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{Y}_i - Y_i)^2.$$

$$RSS = \sum_{i=1}^{n}(y_i - f(x_i))^2,$$

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|f_i - y_i| = \frac{1}{n}\sum_{i=1}^{n}|e_i|.$$

Figure 4.1: Performance Metrics used to evaluate the forecasting accuracy of the model

The performance of the above forecasting model is evaluated with the help of three statistical metric namely Mean Squared Error (MSE), Mean Absolute Error (MAE) and Sum Squared Error (SSE) also otherwise called as Residual Sum of Squares (RSS). Here, MSE and MAE measure the deviation between the actual value and the predicted value. SSE is a measure of discrepancy between the data and the forecasting model. Higher accuracy of prediction is indicated by the presence ofsmaller MSE and MAE values.

4.3 Simulation Results

CURRENCY	PERFORMANCE METRICS		
	MSE	MAE	SSE
Australian Dollar	9.74E-05	7.64E-03	3.45E-03
British Pound	2.35E-05	3.96E-03	8.94E-04
Canadian Dollar	3.20E-05	4.29E-03	1.23E-03
Swiss Franc	1.10E-04	8.62E-03	4.31E-03
Japanese Yen	2.21E+03	2.83E+01	9.65E+04

Figure 4.2: Measurement of Prediction Performance over 60 Day Prediction

A NARX neural network model was trained with 7 technical indicators and 2 economic indicators, a hidden layer and an output neuron unit to predict the exchange rate. The network uses Levenberg-Marquardt training algorithm which adaptively changes weights during each back propagation and stops training when the best performance for the giveninputsis obtained for both training and validation. The number of hidden neuron units was modified between 15 to 20 and the training was terminated at epochs between 60 to 100.

Based on the performance metrics measurements performed on the predicted data, we found out that the trained networks gave the best performance predictions with high rates of accuracy for GBP, CHF, AUD and CAD for 60 days from the point of prediction. But for the JPY currency exchange rate, the prediction accuracy lasted for only 15 days from the point of prediction which is shown in the Fig 2.5. This model is created for short term trend forecasting, hence 60 days period of prediction. The model can be extended to 120-150 days with a minimum loss of accuracy. The Levenberg-Marquardt algorithm of back-propagation works well for this application when compared to the performance of the previous works in foreign exchange rates prediction. The performance of such high accuracy is obtained due to the improved technique used for learning, in the Levenberg-Marquardt algorithm,

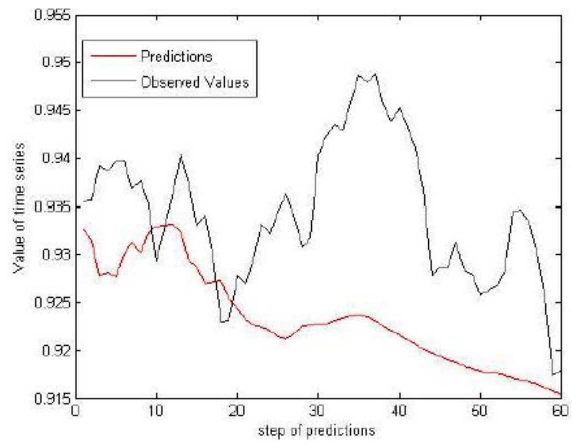

Figure 4.3: USD/CHF

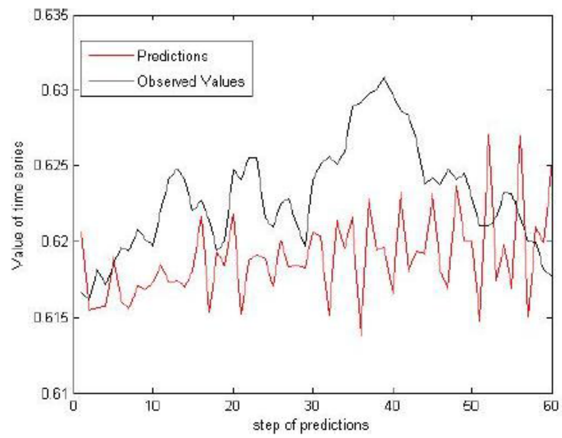

Figure 4.4: USD/GBP

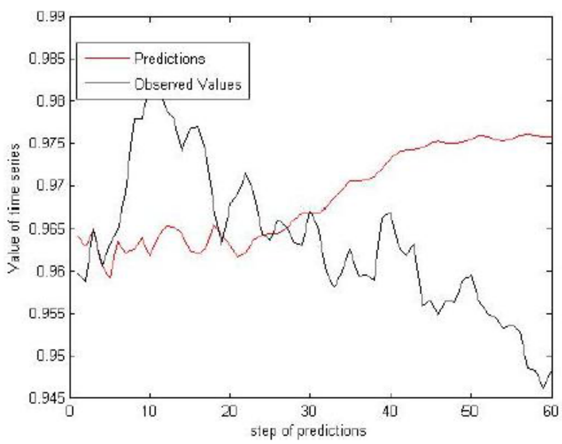

Figure 4.5: USD/AUD

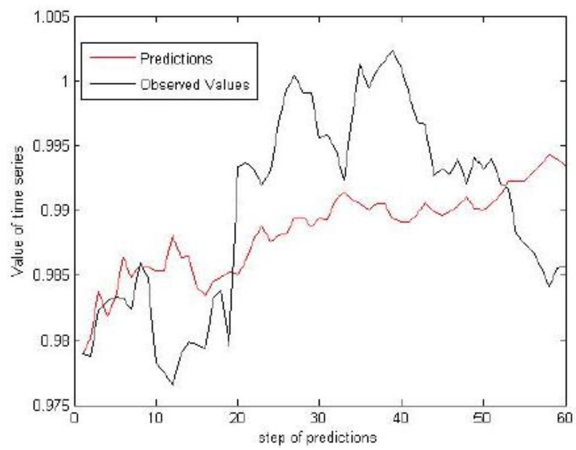

Figure 4.6: USD/CAD

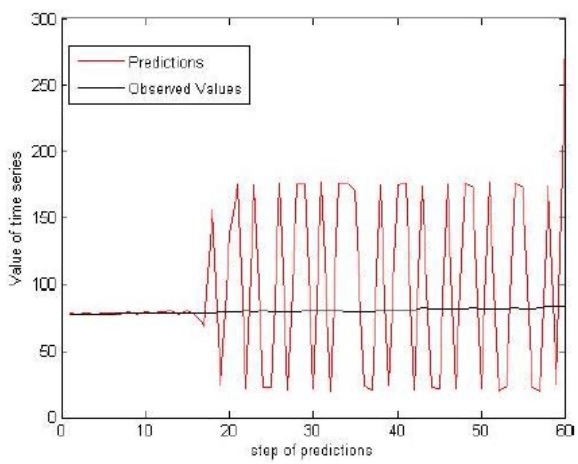

Figure 4.7: USD/JPY

combining the advantages of gradient descent and Gaussian-Newton methods. The MSE, MAE achieved by the trained Narx network is visibly higher than those obtained using other methods in the researches done in the past. [5]

The diagrams comparing the actual and the predicted exchange rate of the five currencies are shown in the Fig 4.3 to 4.7. The plots show that the forecasting follows the actual rates more closely in the cases of AUD, GBP and CAD. For CHF and JPY the prediction is relatively closer to the actual rates.

As you can see from the plots, the performance for USD/JPY degrades after 15 days. The other 4 currencies prediction shows significant accuracy around 90-95% for the first 120 days from the prediction start date and over 97% accuracy for the first 60 days of the prediction which is significantly higher when compared the other models used in the previous researches [1][6]. This means that for improved accuracy, the network has to be retrained every 120 days.

Chapter 5

Conclusion & Future Work

5.1 Conclusion

The prediction results are significantly promising for the four currencies GBP, AUD, CAD and CHF. The prediction performance for Japanese Yen is very poor. Instead of using MSE alone, we have used two other metrics along with that to measure the performance of the network. But, other additional metrics can be used to significantly measure the performance which can be used for comparisons. The Levenberg-Marquardt back-propagation algorithm that is used in this study to build and train the network has proved to be worthy in combining technical and economic indicators to perform the prediction. The above observations have confirmed the better performance of Artificial Neural Networks in the forecasting of currency exchange rates.

5.2 Future Work

Further research emphasis will be on using just the technical indicators in the NARX network and achieving a performance better than the models that were previously used for the purpose of forecasting the currency exchange rates, using the Levenberg Marquardt algorithm that was used for this study.

References

[1] Foreign Exchange Rates Forecasting with Neural Networks Jingtao Yao, Hean-lee Poh, Teo Jasic - 1996 - *International Conference on Neural Information Processing (Hong Kong).*

[2] Kumar, D.A.; Murugan, S., "Performance analysis of Indian stock market index using neural network time series model," *Pattern Recognition, Informatics and Medical Engineering (PRIME), 2013 International Conference on , vol., no., pp.72,78, 21-22 Feb. 2013*

[3] Refenes, A. N.; Azema-Barac, M.; Karoussos, S. A., "Currency exchange rate forecasting by error backpropagation," *System Sciences, 1992. Proceedings of the Twenty-Fifth Hawaii International Conference on , vol.iv, no., pp.504,515 vol.4, 7-10 Jan 1992*

[4] V. Pacelli, V. Bevilacqua and M. Azzollini, "An Artificial Neural Network Model to Forecast Exchange Rates," *Journal of Intelligent Learning Systems and Applications, Vol. 3 No. 2A, 2011, pp. 57-69. doi: 10.4236/jilsa.2011.32008.*

[5] Kamruzzaman, J.; Sarker, R.A., "Forecasting of currency exchange rates using ANN: a case study," *Neural Networks and Signal Processing, 2003. Proceedings of the 2003 International Conference on , vol.1, no., pp.793,797 Vol.1, 14-17 Dec. 2003 doi: 10.1109/ICNNSP.2003.1279395*

[6] Jingtao Yao, Chew Lim Tan, A case study on using neural networks to perform technical forecasting of forex, *Neurocomputing, Volume 34, Issues 14, September 2000, Pages 79-98, ISSN 0925-2312*

[7] Cain Evans, Konstantinos Pappas, Fatos Xhafa, Utilizing artificial neural networks and genetic algorithms to build an algo-trading model for intra-day foreign exchange speculation, *Mathematical and Computer Modelling, Volume 58, Issues 56, September 2013, Pages 1249-1266, ISSN 0895-7177*

[8] Iebeling Kaastra, Milton Boyd, Designing a neural network for forecasting financial and economic time series, *Neurocomputing, Volume 10, Issue 3, April 1996, Pages 215-236, ISSN 0925-2312*

[9] Ritanjali Majhi, G. Panda, G. Sahoo, Efficient prediction of exchange rates with low complexity artificial neural network models, *Expert Systems with Applications, Volume 36, Issue 1, January 2009, Pages 181-189, ISSN 0957-4174,*

[10] Woon-Seng Gan; Kah-Hwa Ng, "Multivariate FOREX forecasting using artificial neural networks," Neural Networks, 1995. Proceedings., IEEE International Conference on , vol.2, no., pp.1018,1022 vol.2, Nov/Dec 1995 doi: 10.1109/ICNN.1995.487560