SLJoAS Sri Lankan Journal of Applied Sciences

Research Article

www.sljoas.uwu.ac.lk/index.php/sljoas

# Forecasting USD/ LKR Daily Exchange Rate with Long-Short Term Memory Neural Network

W.K.R. Madhubhashitha [a, *], G.M.L.M. Aponsu [a]

[a] Department of Mathematical Sciences, Wayamba University of Sri Lanka, Kuliyapitiya, 60202, Sri Lanka

*** Corresponding author email address**: rameshmadhubhashitha01@gmail.com

(Received 17th April 2022; accepted 24rth February 2023)

## Abstract

As Sri Lanka is an open economic country since 1977, a large share of the country's economy is depending on imports as of today. International trade is often based on the US dollar and the changes in US dollar rate relative to Sri Lankan Rupee (USD/LKR exchange rate) is significantly affected on the shape of the country's economy. Having a prior understanding about the fluctuations of the exchange rate is very important for both government and private sectors as their monetary policy-making process, investments decisions and import-export activities are highly depending on the exchange rate. The main objective of this study is to identify an accurate forecasting model to forecast volatility of the daily USD/LKR exchange rate using Recurrent Neural Network (RNN) based Long Short-Term Memory (LSTM) Neural Network. Daily data of USD/LKR exchange rate from 1st January 2015 to 30th April 2021 obtained from the Central Bank of Sri Lanka (CBSL) were used in the study and it was found the existence of the volatility clustering over the study period. The data set was divided in to two parts viz. Training set and Testing set, where each containing 70% and 30% of the whole data respectively. The "ADME" LSTM model was selected as the optimization algorithm and "TanH" was selected as the activation function. The hyperparameters that lead to a better generalization of the model were identified while validating the model. The accuracy of the identified model was evaluated by using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Mean Squared Error (MSE) which results in 0.002072, 0.001071 and 4.293544e-06 respectively.

Keywords: Exchange Rate, Forecasting, LKR, LSTM, USD, Volatility

## 1. Introduction

A comparative value between currencies is required in international trade to exchange one currency with other currencies and this comparative value is defined as the exchange rate. Exchange rate is usually determined by the supply and demand for a foreign currency in the foreign exchange (forex) markets. When the demand of some currencies changes with time in the forex market, then its relative value changes with other currencies. Specially it is very important to identify the factors influencing the variation of the exchange rate in developing countries. Raja Sher Ali Khan [1] identified that the interest rate, oil price, imports, exports as influencing factors for exchange rate in Pakistan. The strict increase in the exchange rate significantly impacts the developing open economic countries like Sri Lanka. According to statistics of CBSL, USD/LKR exchange rate has been increasing over the last forty years. All key sectors of Sri Lanka have significantly affected by the continuous increment of the exchange rate and finally it determines the economic growth of the country. Government and non-government monetary policymaking process, banks, financial institutions, investors, entrepreneurs, importers and exporters risk their business if they do not have an understanding of the future fluctuation of the exchange rate volatility. However, most of the researchers have identified traditional linear methodology like Autoregressive Integrated Moving Average (ARIMA) and Seasonal Autoregressive Integrated Moving Average (SARIMA) for predicting the exchange rate variation. Kamruzzaman and Sarker [2] have found that Artificial Neural Network (ANN) based models outperform the ARIMA model performance in exchange rate forecasting. It is difficult to capture the non-linear behavior of the daily exchange rate variation using traditional linear methodology. Therefore, the objective of this study is to identify a model to forecast daily USD/LKR exchange rate volatility using Long Short-Term Memory (LSTM) architecture which is non-linear in nature. Researchers have proposed different methodologies to forecast the exchange rate and some have compared prediction performances of the traditional linear methodologies and ANN models. Mogari and Diteboho [3] examined the forecasting performance of South African exchange rate with ARIMA and ANN models. The MSE and MAE were used measure the forecasting performance of the models. The results obtained from ARIMA model and ANN model showed that ANN model performance was superior than ARIMA model. Escudero et.al. [4] compared the accuracy of the forecasting models of EUR/USD exchange rate with the use of ARIMA, RNN of the Elman type, and LSTM techniques. Results showed that the performance of LSTM is much better compared to the other two models. Qu and Zhao [5] investigated the best exchange rate prediction model among the RNN and LSTM. By comparing the evaluation indexes of two deep learning models, the optimal neural network model was selected. The existing foreign exchange rate and technical analysis indexes were taken as input parameters. The experimental results showed that the LSTM neural network model has smaller RMSE and MAE than the RNN network model, and the predicted price was very much accurate. Fiore et.al. [6] analyzed EUR/USD foreign exchange rate and proposed a LSTM model to forecast the future values. In the proposed model, there were three classes viz. no_action, decrease, and increase. TI_LSTM and ME_LSTM models were built and compared for the EUR/USD exchange rate. Finally, all of the features were combined into a single LSTM model called ME_TI_LSTM and concluded that hybrid model does not significantly increase the accuracy over the single models. Kumar and Patil [7], tried to identify a model to forecast the volatility of stock index by using time series and ANN techniques. Finally, it was found that the performance of LSTM techniques outperformed ARIMA, ARFIMA and other neural network-based techniques.

As Sri Lanka is a developing country and the economy of the country depends on imports, it is vital to have a pre-judgement about the USD/LKR exchange rate. Therefore, the main objective of this study is to identify an accurate forecasting model to forecast volatility of the daily USD/LKR exchange rate using Recurrent Neural Network (RNN) based Long Short-Term Memory (LSTM) Neural Network.

## 2. Materials and Methods

### 2.1. Data Collection

Daily USD/LKR exchange rate data from 1st January 2015 to 30th April 2021 were obtained from CBSL. Return series was calculated from the following Eq. (01).

$$R_t = \ln\left(\frac{x_t}{x_{t-1}}\right) = \ln(x_t) - \ln(x_{t-1}) \qquad (01)$$

, where $R_t$ is the daily return of exchange rate of USD/LKR at time t. $X_t$ is the daily exchange rate of USD/LKR at time $t$. The daily return is divided in to three sets as shown in the following Table 1 to avoid the over fitting of the forecasting models that will be developed by ANN.

**Table 1:** Breakdown of Data

| | |
|---|---|
| Training set | 1st January 2015 to 31th May 2019 (1062 observations) |
| Validation set | 3th June 2019 to 30th July 2019 (40 observations) |
| Testing set | 31th July 2020 to 30th April 2021 (418 observations) |

### 2.2. Theoretical Background

#### 2.2.1. Neural Network Architecture

ANN is a collection of artificial neurons that behave like biological neural networks in human brain. In 1943, Warren McCullough and Walter Pitts developed a mathematical algorithm called threshold logic, the model of which was the beginning of neural networks. An ANN contains thousands of artificial neurons which transmit the signals among others as human neurons. ANN architecture is divided into two models according to data flow pattern viz. Feedforward Networks and Feedback (Recurrent) Networks [8]. Feedforward networks flow data from the input layer to the output layer only in one direction, and therefore it is not possible to retain memory of previous feed data. But feedback architecture enable the data to flow both directions during the model training and possible to retain memory about the previous feed data.

#### 2.2.2. Recurrent Neural Network (RNN)

The idea of a recurrent neural network (RNN) was proposed by David Rumelhart in 1986 [9] and this provides the best answers to some problems which were unable to reach by the traditional neural networks. Recurrent neural network is a feedback architecture, which contains a data recurrent or looping system.

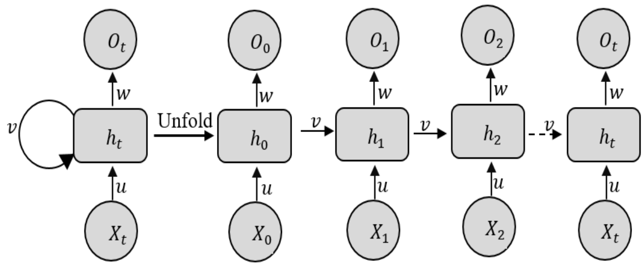


**Fig. 1.** An unrolled RNN

Fig. 1 illustrates the architecture of the RNN with self-looping or recursive pointer to itself. $X_t, O_t$ and $h_t$ are the input, output and hidden state respectively and $U, W$ and $V$ are the model parameters.

#### 2.2.3. The LSTM Architecture

Hochreiter and Schmidhuber [10] introduced LSTM architecture which is an enhancement of RNN and it contains a collection of special recurrent memory unit. Fig. 2 shows the structure of the typical LSTM memory unit that mainly contains the gates, the input signal $(X_t)$, the output $(H_t)$, a cell state $(C_t)$, the activation functions and peephole connections.

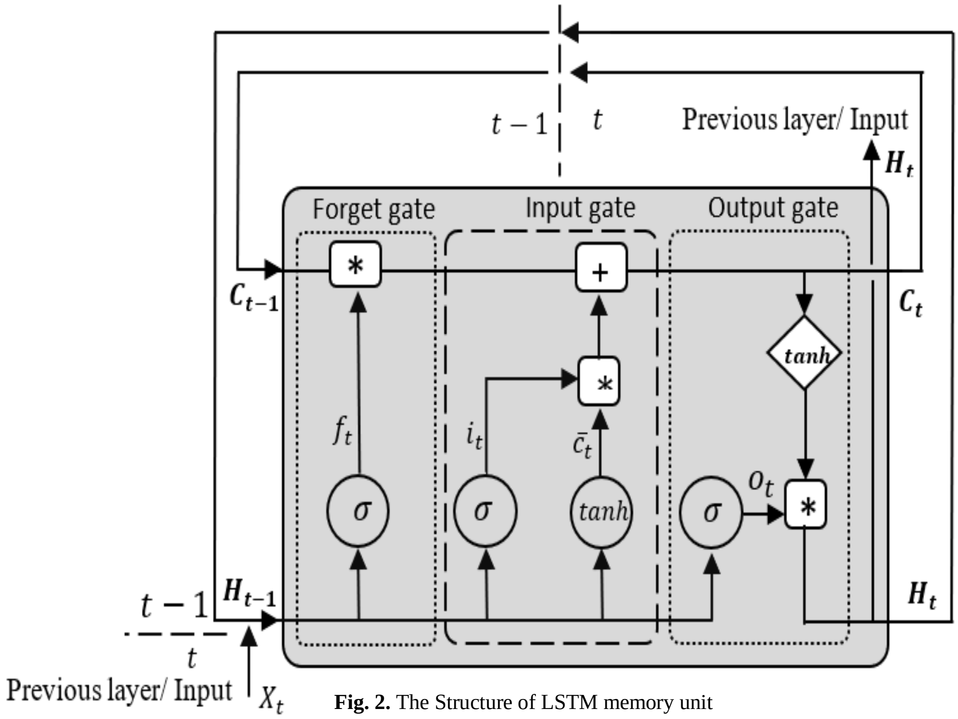


**Fig. 2.** The Structure of LSTM memory unit

#### 2.2.3.1. Forget gate

LSTM forget gate ($f_t$) takes input from the current input ($X_t$) at time $t$ and the output of previous memory unit($H_{t-1}$) at time $t-1$. The sigmoid function was selected as the activation function of the forget gate that identifying and getting rid of inputs data and determine which part should be get rid of from the previous output, where $\sigma$ is the sigmoid function, $w_f$ is the weight and $b_f$is the bias. Outputs of the forget gate lie between 0 and 1.

$$f_t = \sigma\left(w_t * [H_{t-1}, X_t] + b_f\right)$$

#### 2.2.3.2. Input gate

The input gate ($i_t$) contains two layers, one based on sigmoid activation function and other based on the Tanh activation function. Firstly, sigmoid layer takes input from the current input ($X_t$) at time $t$ and the previous output of memory unit ($H_{t-1}$) at time $t-1$. That decides which values will be updated by transforming the values to be between 0 and 1.

$$i_t = \sigma\left(w_i * [H_{t-1} * X_i] + b_i\right)$$

, where $\sigma$ is the sigmoid function, $w_i$ is the weight and $b_i$is the bias of the input gate.
Secondly the tanh layer takes same inputs and transforms the values to be between 1 and -1. , where $\underline{c}_t$ is the temporary state layer, $w_c$ is the weight and $b_c$is the bias of the temporary state layer.

$$\bar{c}_i = \tanh\left(w_c * [H_{t-1}, X_t] + b_c\right)$$

#### 2.2.3.3. Cell state

Previous cell state was updated to current state using the output of the current forget and input gates.

$$C_t = f_t * C_{t-1} + i_t * \bar{c}_t$$

#### 2.2.3.4. Output gate

Both Tanh and sigmoid functions are used as the activation function of the output gate ($o_t$), which determines the current output of the LSTM model. Sigmoid layer takes input from the current input ($X_t$) at time $t$ and the previous output of memory unit($H_{t-1}$) at time $t-1$.

$$o_t = \sigma\left(w_0 * [H_{t-1}, X_t] + b_0\right)$$

Tanh layer takes input from the current cell state ($C_t$) at time $t$.

$$H_t = o_t * \tanh\left(C_t\right)$$

, where $\sigma$ is the sigmoid function, $w_0$ is the weight and $b_0$ is the bias of the temporary state layer.

### 2.3. **Structural Framework**

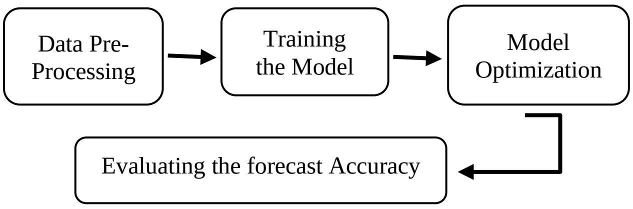


**Fig. 3.** The structural framework of the model identification

The main four steps of model identification is shown in Fig. 3 and the key features of each step is given below.

#### 2.3.1. **Data Pre-processing**

Daily USD/LKR exchange rate data were checked for outliers, but there were no outliers in the data set. Next, as shown in Eq (1), the return series was taken to make the daily exchange rate series stationary. The LSTM model is sensitive to the scale of the input data, especially when the Tanh activation function is used. Therefore, before inserting the data into the input layer of the network, data were normalized using 'sklearn.preprocessing.MinMaxScaler'. MinMaxScaler in scikit-learn used the Eq. (2) to transform return series.

$$X_{new} = \frac{X - X_{max}}{X_{max} - X_{min}} \qquad (02)$$

, where $X_{new}$ is daily return of exchange rate between -1 and 1. $X_{max}$ and $X_{min}$ are the maximum and minimum values of scale of the daily return of exchange rate respectively.

#### 2.3.2. **Training the Model**

Normalized daily USD/LKR exchange rate return series was divided into a training set (70%), validation set and testing set (30%) according to Table 1. Normalized series was inserted into input layer of LSTM network as a three dimensional numpy array; 1062 number of training sequences, 1 sequence length and 1 number of features of each sequence were defined as a dimension and was entered to LSTM model as input data. The hyperparameters such as batch size, epoch, dropout, activation function, loss function, number of hidden layers and number of neurons in each layer were defined next. These hyperparameters play a significant role in optimizing the LSTM network.

#### 2.3.3. **Model Optimization**

The loss function, optimization algorithm and learning rate were used to optimize the model. The hyperparameters were tuned in order to minimize the loss function. The MSE was considered as a loss between forecasted and actual returns of the daily exchange rate. It was evaluated for every batch of predicted and actual output. The "ADAM" was selected as the optimization algorithm of the model that is efficient for training LSTM architecture and it is an extension version of the Stochastic Gradient Descent (SGD) optimizer. Finally, the learning curve was constructed and used to diagnose problems such as an underfit or overfit in the model during training. The shape of the learning curve was considered by changing the hyperparameters such as batch size and epoch until the model has converged.

#### 2.3.4. **Evaluating the forecast Accuracy**

After forecasting the daily USD/LKR exchange rate using best fitted LSTM architecture, the forecast accuracy was evaluated by using Mean square error (MSE), Root mean square error (RMSE) and Mean Absolute Error (MAE).

## 3. **Results and Discussion**

When the model was converged, the optimal technical specification of the LSTM network of this study is shown in Table 2. The optimal LSTM architecture contains two hidden layers where the first layer and the second layer contain 40 and 30 neurons respectively. The learning rate and dropout of the 'ADAM' optimization algorithm were selected as 0.001 and 0.2 respectively.

**Table 2:** Technical specifications of optimal LSTM architecture

| Hyperparameter | Values |
|---|---|
| Number of hidden layers | Two hidden layers |
| Number of neurons in each layer | first layer – 40, second layer – 30 |
| Activation function | TanH |
| Number of epochs | 30 |
| Batch size | 50 |
| Loss function | MSE loss |
| Dropout | 0.2 for each layer |
| Learning rate | 0.001 |
| Optimization Algorithm | Adam |

After constructing the adequate LSTM model, the number of epochs and batch size were adjusted and learning curve was checked. Fig. 4 illustrates the learning curve of the LSTM model at the converged point.

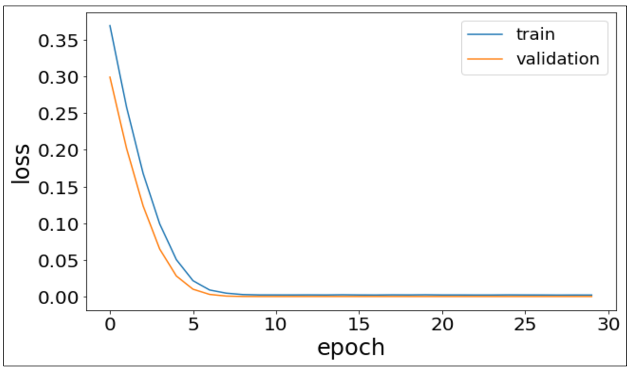


**Fig. 4.** Learning Curve

Learning curve shows that both the training loss and validation loss decreases to a point of stability. Therefore, model was converged and best fitted to the daily exchange rate returns of USD/LKR. When the model converged, the number of epochs were 30 and batch size was 50. The forecasting accuracy of the selected LSTM model was checked by using the data from 31st July 2020 to 30th April and Fig. 5 shows the actual vs forecasted exchange rate return value of LSTM architecture.

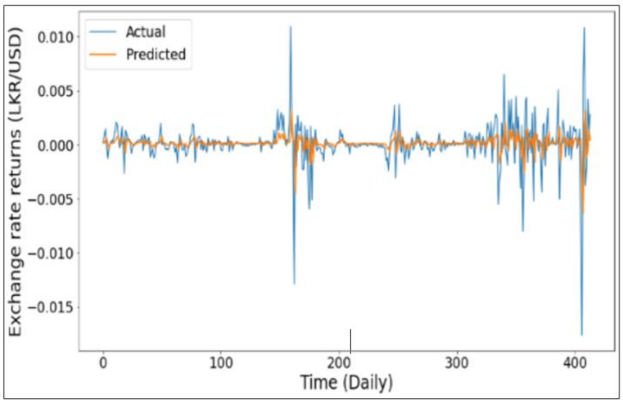


**Fig. 5.** The Actual value vs Forecasted value of LSTM architecture.

**Table 3:** Forecasting Performance of daily returns of LKR/USD exchange rate

| Performance Metrics | Result |
|---|---|
| Mean square error (MSE) | 4.293544e-06 |
| Root mean square error (RMSE) | 0.002072 |
| Mean Absolute Error (MAE) | 0.001071 |

Forecast performance of the fitted LSTM model of returns of USD/LKR exchange rate was measured by using MAE, RMSE and MAE and the results are shown in Table 3. According to the results, it is clear that the all-error values are very low indicating a higher accuracy of the identified model.

## 4. Conclusions

The main purpose of this paper was to identify an accurate forecasting model for the daily USD/LKR exchange rate returns. The LSTM architecture is more capable of capturing complex dependencies within time series data and therefore it is used to forecast the volatility of daily USD/LKR exchange rate. The model convergence point was attained when epochs and batch size were reached 30 and 50 respectively and after evaluating the forecast performance of LSTM models, the best model was identified as follows. There were two hidden layers where 40 neurons in the first hidden layer and 30 neurons in the second hidden layer. Dropout and Learning rate were equal to 0.2 and 0.001 respectively and TanH function and ADAM algorithm were selected as the activation function and Optimization Algorithm of the final best fitted model. Moreover, the forecast performance of the model was evaluated and MSE, RMSE and MAE were found to be 4.293544e-06, 0.002072, and 0.001071 respectively. Therefore, the identified model can be used to forecast the daily USD/LKR exchange rate with high accuracy.

Based on the findings of the study, it is recommended to use LSTM models to forecast other exchange rates such as YEN/LKR, GBP/LKR, etc., as they directly affect some economic activities of Sri Lanka. Furthermore, it is suggested to use hybrid LSTM models to forecast exchange rates in future studies to enhance the performance.

## References

[1] Khan, R. S. A. (2014, June). Analysis if the factors affecting exchange rate variability in Pakistan. *IOSR Journal of Business and Management (IOSR-JBM)*, 16 (6), 115-121.

[2] Kamruzzaman, J., & Sarker, R. (2003). Comparing ANN Based Models with ARIMA for Prediction of Forex Rates. ASOR Bulletin, 22(2).

[3] Mogari, I. R., & Diteboho, X. (2017). A Comparative Analysis of Artificial Neural Network and Autoregressive Integrated Moving Average Model on Modeling and Forecasting Exchange Rate. *World Academy of Science, Engineering and Technology International Journal of Economics and Management Engineering*, 11(11).

[4] Escudero, P., Alcocer, W., Paredes, J. (2021).Recurrent Neural Networks and ARIMA Models for Euro/Dollar Exchange Rate Forecasting. *Applied Sciences*.11 (12). https://doi.org/10.3390/app11125658

[5] Qu, Y., & Zhao, X. (2019). Application of LSTM Neural Network in Forecasting Foreign Exchange Price. *Journal of Physics: Conference Series*. 1237. https://doi.org/10.1088/1742-6596%2F1237%2F4%2F042036

[6] Yıldırım, D., Toroslu, I., & Fiore, U. (2021). Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators. *Financial Innovation*. http://dx.doi.org/10.1186/s40854-020-00220-2

[7] Kumar, H. P., & Patil, B, S., (2016). Volatility Forecasting - A Performance Measure of Garch Techniques with Different Distribution Models. *International Journal of Soft Computing, Mathematics and Control*. http://dx.doi.org/10.14810/ijscmc.2016.5301

[8] Gencay, R., & Liu, T. (1997, September, 15). Nonlinear modelling and prediction with feedforward and recurrent networks. Physica D, 1-2(108), 119-134. https://doi.org/10.1016/S0167-2789(97)82009-X

[9] Rumelhart, D. E., Smolensky, P., Mcclelland, J. L., & Hinton, G. E. (1986). Sequential thought processes in PDP models. *In: V 2*, 3 – 57.

[10] Hochreiter, S., & Schmidhuber, J. (1997). Long short-term memory. *In: Neural computation* 9(8), 1735 – 1780.