Novelty Journals

ISSN 2394-7322

International Journal of Novel Research in Marketing Management and Economics
Vol. 11, Issue 2, pp: (29-37), Month: May - August 2024, Available at: www.noveltyjournals.com

Multivariate Foreign Exchange Rate Prediction with Long-Short Term Memory Deep Learning Networks

Saminu Umar[1], Aliyu M. Lamido[2], Muhammad K. Aminu[3]

[1, 2, 3]Department of Mathematics and Statistics, Umaru Ali Shinkafi Polytechnic, Sokoto, Nigeria

DOI: https://doi.org/10.5281/zenodo.11204022

Published Date: 16-May-2024

Abstract: **This study investigates and compared the performance of two distinct variants of Long Short-Term Memory (LSTM) networks—vanilla LSTM and bidirectional LSTM in forecasting multivariate exchange rate data. The dataset utilized in this study comprises the daily historical data of the official exchange rates of four foreign currencies: the United States Dollar ($), Swiss Franc (₣), Euro (€) and Pound Sterling (£) against the Nigerian Naira (N) spanning from January 2006 to May 2023. The data set was split into two subsets: the in-sample data, which were used for model estimation and the out-of-sample data, which were used for out-of-sample prediction evaluation. The data underwent preprocessing and formatting for LSTM input before training and testing varying architectures of the two LSTM variants. Furthermore, forecasting precision of the models were evaluated and compared through the use of Mean Absolute Error (MAE), Mean Squared Error (MSE) and Root Mean Squared Error (RMSE) performance metrics. The results revealed that the bidirectional LSTM outperformed vanilla LSTM in the forecast of multivariate exchange rate.**

Keywords: **Multivariate time series, Exchange rate forecasting, Long Short-Term Memory, Forecast accuracy.**

1. INTRODUCTION

Global financial landscape is in a perpetual state of flux, marked by the continuous oscillation of exchange rates. These rates serve as pivotal indicators that influence economic policies, investment strategies and trade relationships among nations (Frankel, 2019). In Nigeria, as in numerous other countries, the dynamics of exchange rates wield profound implications across various sectors of the economy. From trade facilitation to inflation control, interest rate management and overall economic stability, exchange rate fluctuations exert far-reaching effects (Obstfeld & Rogoff, 2009). Exchange rate movements in Nigeria significantly impact trade activities, determining the competitiveness of exports and imports, thereby shaping the balance of payments and trade deficits (Omotunde & Adepoju, 2018). Additionally, the oscillations in exchange rates exert a direct impact on domestic inflation rates, as evidenced by their influence on the pricing of imported commodities and services (Rudebusch & Svensson, 1999).

Financial time series data, such as exchange rates, are known to exhibit volatility, which can significantly erode investor confidence levels, thereby precipitating instances of capital flight and impeding the inflow of foreign direct investment (Rana & Pritchett, 2012). Volatility in financial time series is amplified by the interconnectedness of markets, leading to what is termed "volatility spillover" (Hull et al., 2015). This inherent volatility, coupled with the intricate nature of financial series, presents challenges for accurate forecasting (Boudt et al., 2013).

Sarno and Taylor (2003) stressed the importance of accurate forecasting for informed decision-making, risk mitigation and strategic optimization, all vital for navigating the complexities of the economic landscape. Exchange rate dynamics have diverse effects, highlighting the necessity for precise forecasting among stakeholders. Policymakers depend on forecasts to devise policies promoting stability and economic growth, while investors utilize predictions for asset allocation and risk hedging (Hansen & Lunde, 2017). Additionally, businesses use forecasts to assess trade feasibility and manage currency-related risks and financial institutions employ them for effective risk management and trading strategies (Dominguez & Tesar, 2006).

Despite the paramount importance of exchange rate forecasting, conventional econometric models frequently encounter challenges in adequately capturing the intricate patterns and nonlinear associations present within multivariate financial time series such as exchange rate data. Furthermore, the inherent unpredictability of global financial markets, coupled with the advent of novel factors influencing exchange rates, exacerbates the limitations of traditional forecasting techniques (Engle & Kelly, 2012). Deep Learning (DL) techniques, including Artificial Neural Networks (ANNs) and Support Vector Machine (SVM), serve as vital alternatives for addressing time series forecasting challenges. These methods consist of algorithms capable of autonomously learning patterns from extensive datasets, demonstrating potential in analyzing both univariate and multivariate financial time series (Gamboa, 2017).

2. LITERATURE REVIEW

Various DL techniques have been proposed and utilized to predict time series data including exchange rate such as with many empirical studies suggesting neural networks outperform statistical models, especially the LSTM (Manaswi et at., 2018). Chandar et al. (2016) utilized feedforward neural networks to forecast FOREX exchange rates in India against Euro, Japanese Yen, Pound Sterling and US Dollar from 2010 to 2015. Employing the Levenberg-Marquardt (LM) learning algorithm, they assessed performance using Root Mean Square Error (RMSE), Mean Absolute Error (MAE) and Forecasting Error (FE) indicators. Their findings indicate neural networks' effectiveness for FOREX prediction. Swagat et al. (2018) explored forecasting Nepalese currency against major currencies using four ANN architectures: Multilayer Perceptron (MLP), Simple Recurrent Neural Network (SRNN), Gated Recurrent Unit (GRU) and Long Short Term Memory (LSTM). Their study revealed LSTM's superiority over MLP, SRNN and GRU networks. Experimentation led to an LSTM configuration of 4-5-1 achieving the highest accuracy in MAE.

Umar & Aminu (2021) analyzed and forecasted Nigeria's monthly food inflation rate using the autoregressive integrated moving average (ARIMA) model and Extreme Learning Machine (ELM) algorithm. ARIMA (2, 1, 2) was identified as the best fit based on AIC and BIC criteria. Computational results demonstrated that ELM outperformed ARIMA in terms of MAE, MSE and RMSE when forecasting Nigerian food inflation rates. Zahrah et al. (2020) studied predicting foreign exchange rates amidst the COVID-19 pandemic's market complexity using LSTM networks. They forecasted exchange rates at hourly and daily intervals in 2020, optimizing hyperparameters with RMSE evaluations. Comparing predictions across 2018, 2019 and 2020, the 1-hour timeframe yielded the best RMSE result, with 2019 having the lowest error. Despite market volatility, LSTM effectively predicted 2020 exchange rates, shown by favorable RMSE outcomes.

Abdoli et al. (2020) applied LSTM to forecast Tehran stock prices and compared results with an ARIMA model. Empirical result showed that LSTM achieved lower prediction errors compared to ARIMA over a 20-day period and prediction accuracy decreases with longer time frames. Ju & Liu (2021) propose an ATT-LSTM model combining attention mechanism with LSTM. This model enhances LSTM's ability to analyze and predict multiple time data by screening sequences, removing irrelevant information and capturing sequence interactions. Experimental results on Nasdaq 100 and Beijing PM2.5 datasets demonstrate superior performance of ATT-LSTM over six other models based on MAE and RMSE evaluation. Further validation with actual data confirms the model's effectiveness.

Sirisha et al. (2023) propose a Deep Stacked Bidirectional LSTM model for forecasting future time window feature vectors of multiple sensors. Their model achieves robust results in test accuracy score, loss and mean absolute error without fine-tuning parameters. Compared to vanilla LSTM, Bi-LSTM and Deep LSTM, their model shows improvements in test accuracy, with a decrease in loss and mean absolute error. This research specifically employ different variants of (LSTM) Networks to forecast the intricate and uncertain patterns present in multivariate exchange rates of Naira vis-à-vis four different foreign currencies in order to assess their performance with the aim of identifying the most suitable for prediction exchange rate in Nigeria.

3. METHODOLOGY

3.1 Long Short Term Memory (LSTM)

Long Short Term Memory (LSTM) is a refined variation of the Recurrent Neural Network (RNN) developed by Hochreiter & Schmidhuber (1997) designed to improve the performance of RNN that has difficulties in terms of long-term dependencies because there is a vanishing gradient in which, the weights of previous output will decrease over time-steps by retaining both short-term and long-term memory states (Bengio et al., 1994). LSTM utilizes interconnected memory blocks to replace traditional hidden units in recurrent networks (Gers et al., 2000).

In LSTM, memory cells consist of self-loop units and multiplication units controlling the flow of information through three gate layers: 1) The Input Gate Layer (equation 1) decides what new information to store in the cell state (equation 2), with sigmoid (equation 3) and tanh (equation 4) layers determining the update values. 2) The Forget Gate Layer (equation 5) controls the retention of important information and deletion of irrelevant information from the cell state using a sigmoid layer, outputting a range from 0 to 1. 3) The Output Gate Layer (equation 6) determines the output based on the cell state, generating a hidden state (equation 7) by scaling the cell state values with a tanh function and multiplying them by the filtered output. Equation 8 represents the formula for the hidden state 'candidate,' calculated based on the current input and the previous hidden state (Dey & Salem, 2017). These mechanisms enable LSTM to effectively manage and process sequential data.

$$i_t = \sigma(W_i \cdot [h_{t-1}, x_t] + b_i) \quad (1)$$

$$C_t = f_t * C_{t-1} + i_t * g_t \quad (2)$$

$$\sigma_x = \frac{1}{1 + e^{-x}} \quad (3)$$

$$\tanh(x) = 2\sigma(2x) - 1 \quad (4)$$

$$f_t = \sigma(W_f \cdot [h_{t-1}, x_t] + b_f) \quad (5)$$

$$O_t = \sigma(W_o \cdot [h_{t-1}, x_t] + b_o) \quad (6)$$

$$h_t = O_t * \tanh(C_t) \quad (7)$$

$$g_t = \tanh\,(W_c \cdot [h_{t-1}, x_t] + b_c) \quad (8)$$

Where the symbol $*$ represents the Hadamard product (element-wise) product, i_t, f_t, o_t and g_t are the input, forget, output and cell activation vectors, respectively, x_t is the input matrix at time t. $\sigma_g$is the sigmoid activation function, which squashes the input values between 0 and 1, σ_c is the hyperbolic tangent (tanh) activation function, which squashes the input values between -1 and 1 and σ_h is the hyperbolic tangent (tanh) activation function applied to the cell state, used to compute the hidden state h_t. W_f, W_i, W_o and W_g are weight matrices Weight matrices corresponding to the input gate, forget gate, output gate and candidate cell state, respectively. b_i, b_f, b_o and b_g are bias vectors corresponding to the input gate, forget gate, output gate and candidate cell state, respectively. C_t is the cell state at time t, h_t is the hidden state or output at time t and h_{t-1} is the hidden state at time t−1. A structure of an LSTM cell is shows in figure 1.

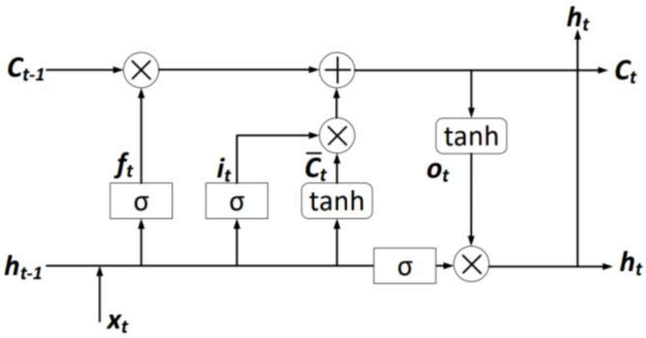

Fig. 1: LSTM cell diagram

International Journal of Novel Research in Marketing Management and Economics
Vol. 11, Issue 2, pp: (29-37), Month: May - August 2024, Available at: www.noveltyjournals.com

3.1.1 Vanilla LSTM

Vanilla LSTM also known as single LSTM was pioneered by Hochreiter & Schmidhuber (1997), is a fundamental component of recurrent neural networks (RNNs), designed to process sequential data effectively. It consists of memory cells and gating mechanisms including forget, input and output gates, regulated by sigmoid activation functions. The forget gate controls past information retention, the input gate manages new information integration and the output gate regulates information flow to the next hidden state.

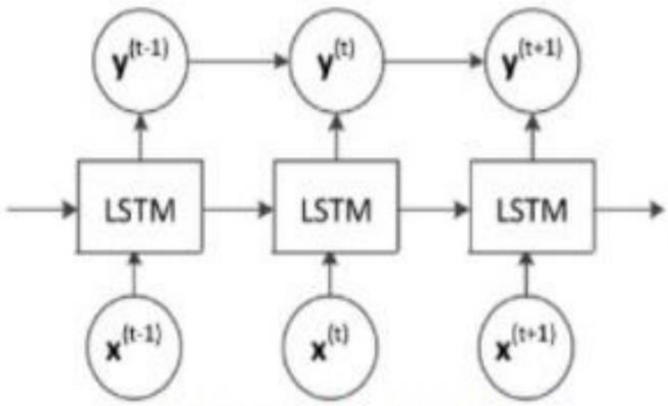

Fig. 2: Vanilla LSTM Architecture

3.1.2 Bidirectional LSTM

This variant of LSTM was introduced by Schuster & Paliwal (1997), enhances input information by storing sequence data in both forward and backward directions, allowing the network to learn from past and future states simultaneously.

A bidirectional LSTM for single time series forecasts estimation involves the implementation of RNNs cells that capture time-series data from left to right (standard RNNs cells) and RNNs cells that capture data in a reverse manner.

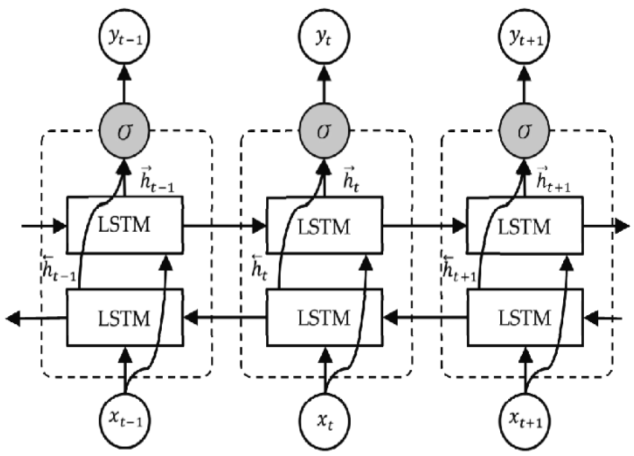

Fig 3: Bidirectional LSTM Architecture

3.2 Data Description

The dataset utilized in this study comprises the daily historical data of the official exchange rates of four foreign currencies: the United States Dollar ($), Swiss Franc (₣), Euro (€) and Pound Sterling (£) against the Nigerian Naira (N) spanning from January 2006 to May 2023. In this research, Long Short-Term Memory (LSTM) networks are employed as the predictive method for forecasting the foreign exchange rates of the Naira against the aforementioned currencies. Two distinct variants of LSTM are utilized in the analysis to assess their effectiveness in predicting exchange rate fluctuations over the specified time period.

Fig. 4: Time-series plot of the original dataset values

Fig. 4 displays a plot depicting the four time series prior to undergoing any processing.

Table 1 presents the summary statistics of the data set, providing insights into the central tendency, variability and distributional characteristics for each currency exchange rate.

Table I: Summary statistics for the exchange rate data

	Naira/USD	Naira/Swiss	Naira/Euro	Naira/Pound
Mean	230.2611	236.4745	275.5933	329.3351
Minimum	116.55	91.0600	145.03	173.03
Maximum	461.50	520.64	510.55	583.37
Std. Deviation	105.7218	117.6766	105.7382	112.1250
Skewness	0.6720	0.6698	0.7212	0.8117
Kurtosis	1.9921	2.1678	2.0703	2.3058

3.3 Data Preprocessing

Several steps were taken in preparing the data for constructing the LSTM. Initially the data was checked for missing values, and then Min-Max scaling technique described in equation (9) was utilized to normalize the data, ensuring uniformity across features.

$$\hat{x} = \frac{x - \min(x)}{\max(x) - \min(x)} \qquad (9)$$

Where $\hat{x}$ is normalized data, x is our data to be normalized, $\min(x)$ and $\max(x)$ are minimum and maximum values of all data.

After normalization, the dataset was divided into in to training (90%) and testing (10%) datasets, subsequently, the data was formatted for LSTM input, creating input-output pairs where each input sequence covered a specified number of past time steps, with the corresponding output representing the value at the subsequent time step.

3.4 Model Training and Testing

Model parameters were defined to guide the training process, the input data was reshaped to conform to the LSTM input format. All the LSTM architecture starts with a Sequential model incorporating different LSTM layers with varying units utilizing the Rectified Linear Unit (ReLU) activation function, a Dropout layer with a dropout rate of 0.2 was included in all the models to prevent overfitting. The models were compiled for training using the Adaptive Moment Estimation (Adam) optimizer with a learning rate of 0.0001 and the mean squared error (MSE) loss function to evaluate performance. Training was conducted over 50 epochs with a batch size of 64, employing a validation split of 0.1 to assess the model's performance

International Journal of Novel Research in Marketing Management and Economics
Vol. 11, Issue 2, pp: (29-37), Month: May - August 2024, Available at: www.noveltyjournals.com

during training. Optimal architectures for each model variant were determined through experimentation, involving adjustments to the model's parameters using the Keras deep learning library within Python environment.

Mean Absolute Error (MAE) mean squared error (MSE) and root mean square error (RMSE) are the performance metrics used for to evaluate the accuracy all the two variants of LSTM considered. The metrics are defined as;

$$\mathrm{MAE} = \frac{1}{n}\sum_{t=1}^{n} |(x_{it} - y_{it})| \tag{10}$$

$$\mathrm{MSE} = \frac{1}{n}\sum_{t=1}^{n} (x_{it} - y_{it})^2 \tag{11}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{t=1}^{n} (x_{it} - y_{it})^2} \tag{12}$$

Where x_{it} is the actual observed value of i^{th} variable in the period t and y_{it} is the predicted value of i^{th} variable in the same period and n is the number of data

4. RESULTS AND DISCUSSION

The summary of the architecture of the estimation LSTM variants presents in Table II. For the vanilla LSTM, the architecture consists of one LSTM layer followed by a Dense layer and a Dropout layer with their respective output shape. The architecture of the Bidirectional is similar to the vanilla LSTM. However, instead of a single LSTM layer, a Bidirectional LSTM layer was introduced followed by a dropout layer and a Dense layer matching the shape of the prediction target.

Table II: LSTM architecture summary and training metrics

	Vanilla		**Bidirectional**	
Layer (type)	Output Shape	Param	Output Shape	Param
LSTM	(None, 28, 512)	1058816		
Bidirectional			(None, 28, 1024)	2117632
Dropout	(None, 28, 512)	0	(None, 28, 1024)	0
Dense	(None, 4)	1026	(None, 4)	2050
Total params:	1,124,996		2,121,732	
Trainable params:	1,124,996		2,121,732	
Non-trainable params:	0		0	

Evaluation and comparison of the training and testing accuracy of the LSTM variants were conducted to assess their effectiveness in capturing temporal dependencies and forecasting accuracy using MAE, MSE and RMSE (Table III). Performance assessment helps identify the optimal LSTM architecture for forecasting exchange rates. The results indicate that the bidirectional LSTM shows better performance compared to the vanilla LSTM across all metrics. Fig 5, 6, 7 and 8 shows the graph of actual vs predicted value of different currencies according to bidirectional LSTM.

Table III: Training and Testing Accuracy Performance

Model	**Training**			**Testing**		
	MSE	MAE	RMSE	MSE	MAE	RMSE
Vanilla	76.6903	5.1301	8.7573	76.6903	6.0583	8.7573
Bidirectional	60.7565	4.5042	7.7946	58.9382	5.1301	7.6771

Novelty Journals
ISSN 2394-7322
International Journal of Novel Research in Marketing Management and Economics
Vol. 11, Issue 2, pp: (29-37), Month: May - August 2024, Available at: www.noveltyjournals.com

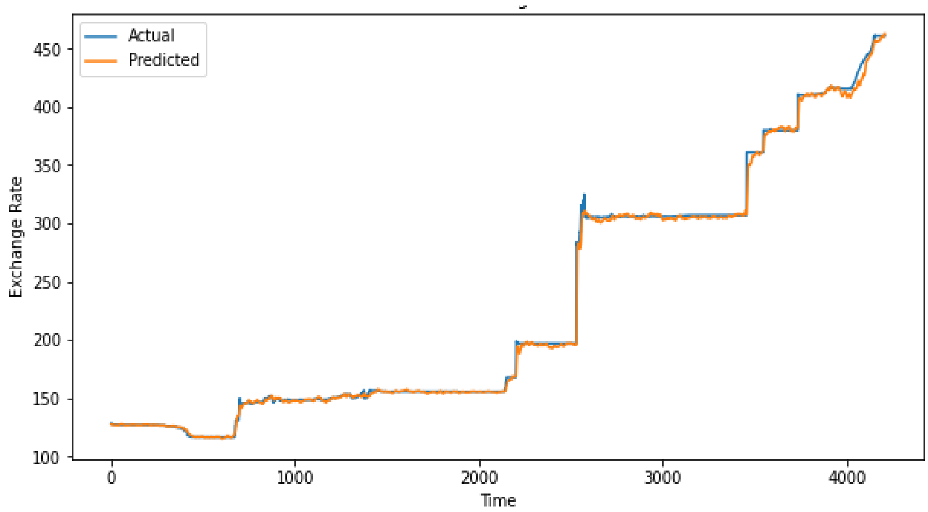

Fig 5: Actual vs predicted Value Graph of Naira/USD

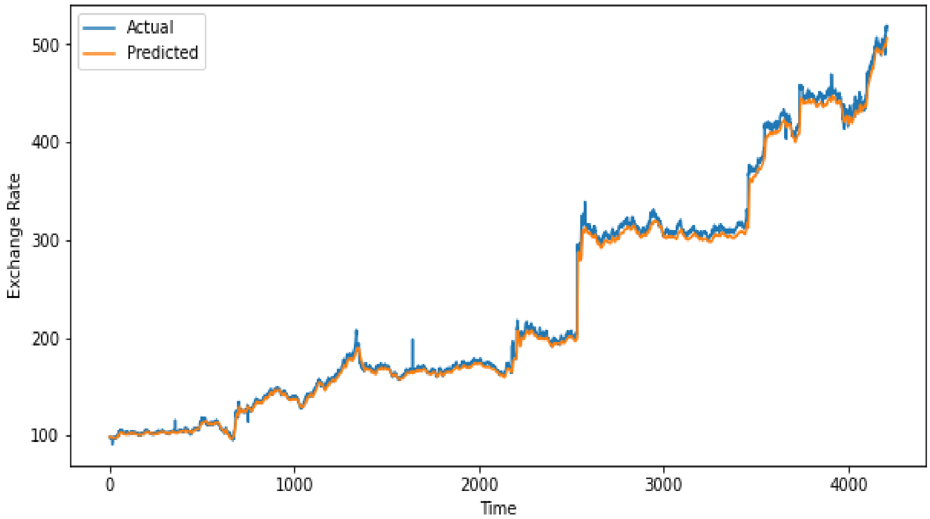

Fig. 6: Actual vs predicted Value Graph of Naira/Swiss

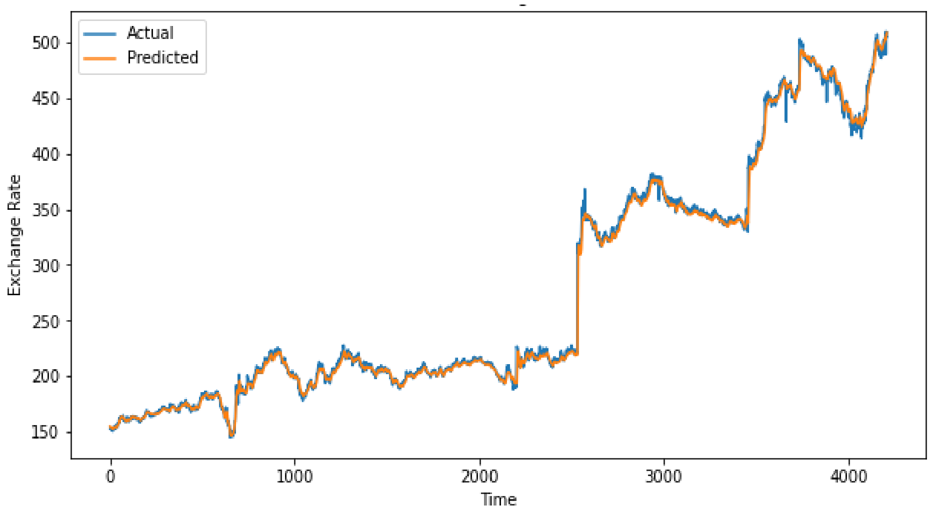

Fig. 7: Actual vs Predicted Value Graph of Naira/Euro

International Journal of Novel Research in Marketing Management and Economics
Vol. 11, Issue 2, pp: (29-37), Month: May - August 2024, Available at: www.noveltyjournals.com

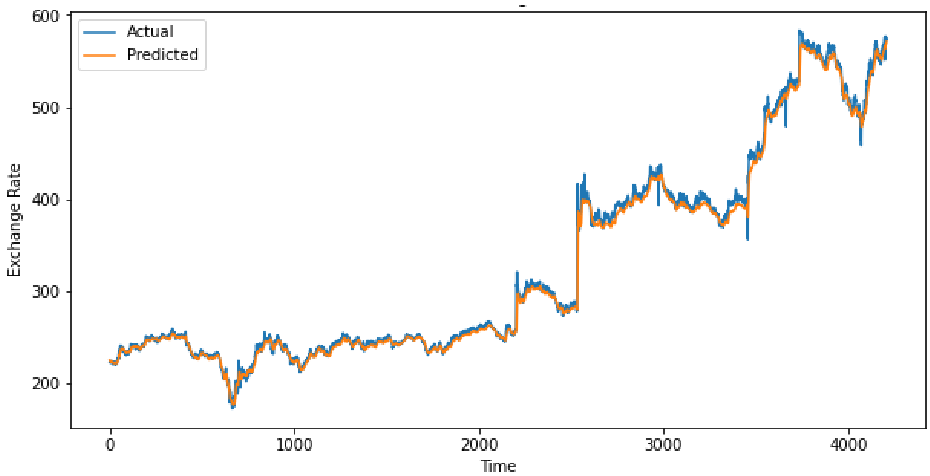

Fig. 8: Actual vs Predicted Value Graph of Naira/Pound

5. CONCLUSION

This paper investigates and compared the effectiveness of two distinct LSTM variants in ability to capture temporal dependencies and forecast accuracy of exchange rates of Naira against four different foreign currencies. A number of steps were taken to preprocess and prepare the data for constructing the LSTM. After several experiments with different varying parameters architectures, the results indicated that bidirectional LSTM outperformed the vanilla LSTM.

REFERENCES

[1] Abdoli, G. (2020). Comparing the prediction accuracy of LSTM and ARIMA models for time-series with permanent fluctuation. Periódico do Núcleo de Estudos e Pesquisas sobre Gênero e DireitovCentro de Ciências Jurídicas-Universidade Federal da Paraíba, 9.

[2] Bengio, Y., Simard, P., & Frasconi, P. (1994). Learning long-term dependencies with gradient descent is difficult. *IEEE transactions on neural networks*, *5*(2), 157-166.

[3] Boudt, K., Danielsson, J., & Laurent, S. (2013). Robust forecasting of dynamic conditional correlation GARCH models. *International Journal of Forecasting*, 29(2), 244-257.

[4] Dey, R., & Salem, F. M. (2017, August). Gate-variants of gated recurrent unit (GRU) neural networks. In *2017 IEEE 60th international midwest symposium on circuits and systems (MWSCAS)* (pp. 1597-1600). IEEE.

[5] Dominguez, K. M., & Tesar, L. L. (2006). Exchange rate exposure. *Journal of international Economics*, *68*(1), 188-218.

[6] Engle, R., & Kelly, B. (2012). Dynamic Equicorrelation. *Journal of Business & Economic Statistics*, 30(2), 212-228.

[7] Frankel, J. A. (2019). The exchange rate and inflation targeting. In The Exchange Rate in a Behavioral Finance Framework (pp. 147-162). Palgrave Macmillan, Cham.

[8] Gamboa, J. C. B. (2017). Deep learning for time-series analysis. *arXiv preprint arXiv:1701.01887*.

[9] Gers, F. A., Schmidhuber, J., & Cummins, F. (2000). Learning to forget: Continual prediction with LSTM. *Neural computation*, 12(10), 2451-2471

[10] Graves, A., Jaitly, N., & Mohamed, A. R. (2013, December). Hybrid speech recognition with deep bidirectional LSTM. In *2013 IEEE workshop on automatic speech recognition and understanding* (pp. 273-278). IEEE.

[11] Hansen, P. R., & Lunde, A. (2017). Forecasting exchange rates out of sample: Can economic fundamentals beat the random walk?.*Journal of International Economics*, 105, 81-113.

International Journal of Novel Research in Marketing Management and Economics
Vol. 11, Issue 2, pp: (29-37), Month: May - August 2024, Available at: www.noveltyjournals.com

[12] Ju, J., & Liu, F. A. (2021). Multivariate time series data prediction based on att-lstm network. *Applied sciences*, *11*(20), 9373.

[13] Obstfeld, M., & Rogoff, K. (2009). Global imbalances and the financial crisis: Products of common causes. In NBER International Seminar on Macroeconomics (Vol. 5, pp. 9-65).

[14] Omotunde, O. O., & Adepoju, O. A. (2018). Exchange Rate Volatility and Trade Flows in Nigeria. In Economics & Business Conference Proceedings (Vol. 2018, No. 1, p. 204). International Institute of Social and Economic Sciences.

[15] Rana, P. B., & Pritchett, L. (2012). Elephants in the bedroom: Financial flows and capital controls in the era of globalization. Center for Global Development Working Paper, (300).

[16] Rudebusch, G., & Svensson, L. E. (1999). Policy rules for inflation targeting. In *Monetary policy rules* (pp. 203-262). University of Chicago Press.

[17] Sarno, L., & Taylor, M. P. (2003). *The economics of exchange rates*. Cambridge university press.

[18] Schmidhuber, J., & Hochreiter, S. (1997). Long short-term memory. *Neural Computing*, 9(8), 1735-1780.

[19] Schuster, M., & Paliwal, K. K. (1997). Bidirectional recurrent neural networks. *IEEE transactions on Signal Processing*, *45*(11), 2673-2681.

[20] Sirisha, B., Naveena, S., Palanki, G., & Snehaa, P. (2023). Multivariate time series sensor feature forecasting using deep bidirectional LSTM. *Procedia Computer Science*, *218*, 1374-1383.

[21] Umar S, & M. K. Aminu (2021). Forecasting Food Inflation in Nigeria using ARIMA and ELM Techniques. International Journal of Progressive Sciences and Technologies (IJPSAT). Vol. 29 No. 2, pp.424-429

[22] Zahrah, H. H., Saâ, S., & Rismala, R. (2020). The foreign exchange rate prediction using long-short term memory. *International Journal on Information and Communication Technology (IJoICT)*, *6*(2), 94-105.