

## Comparative Analysis of LSTM, GRU, and ANN Models for USD/NGN Exchange Rate Forecasting

Aminu Bashir Suleiman<sup>1</sup>, Stephen Luka<sup>2</sup>, and Ahmed Ibrahim Mahmud<sup>3</sup>

<sup>1</sup> Department of Cyber Security, Federal University Dutsin-ma, Katsina State

<sup>2,3</sup> Department of Software Engineering, Federal University Dutsin-ma, Katsina State

\*Corresponding Author: [ameenu.basheer10@gmail.com](mailto:ameenu.basheer10@gmail.com)

**ABSTRACT** This study performs a comparative examination of three deep learning models: Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Artificial Neural Network (ANN) to predict the USD/NGN exchange rate utilizing a decade-long historical dataset obtained from Investing.com. The dataset underwent thorough pre-processing and normalization, subsequently followed by an 80/20 train-validation division. The performance of each model was evaluated using conventional regression metrics: Mean Absolute Error (MAE), Root Mean Squared Error (RMSE), and R-squared ( $R^2$ ). Of the three models, the GRU exhibited the most performance, attaining the lowest MAE (0.609), the lowest RMSE (0.103), and a commendable  $R^2$  score of 0.629. These findings indicate that GRU provides a superior approach for short-term exchange rate forecasting, especially in turbulent financial markets such as that of the Nigerian Naira.

**KEYWORDS:** Machine Learning, Model, Algorithm, Deep Learning, Forex.

### I. INTRODUCTION

Currency systems have shaped worldwide trade and finance. Barter was used in primitive societies until commodity money like barley and pearls evolved into metallic currencies like Lydian silver coins. Modern monetary systems were built on banking, fiat currency, and the gold standard [1]. The Forex market became the world's largest financial market in 2016 with \$5.1 trillion in daily trading volume (Bank for International Settlements). Forex changes may be unpredictable, say researchers [2]. To improve Forex forecasting and decision-making, traders utilize economic indicators and machine learning models to discover recurring trends, classification models to predict price direction, and regression models to estimate price values [3]. Classification models predict price direction, while regression models estimate exact values, enabling data-driven strategies to improve Forex forecasting and decision-making [4]. Conventional econometric techniques, although analytically sound, frequently fail to account for the complex temporal relationships present in currency markets. Conversely, deep learning models, specifically recurrent neural networks (RNNs) and their derivatives, provide enhanced flexibility and predictive accuracy in modeling these dependencies [5]. Despite the growing body of literature exploring deep learning techniques for Forex forecasting, most studies focus on major currency pairs such as EUR/USD or GBP/USD, with limited attention given to emerging market currencies like the Nigerian Naira (NGN). Moreover, comparative studies evaluating the performance of different deep learning architectures on the same dataset remain scarce. This study seeks to systematically evaluate the prediction accuracy of three leading deep learning algorithms LSTM, GRU, and ANN in portraying the exchange rate of the United States Dollar (USD) against the Nigerian Naira (NGN). This research aims to determine the most successful method for short-term Forex price prediction by utilizing a decade of historical data and applying stringent pre-processing, feature extraction, and model validation procedures.

## II. RELATED WORKS

The utilization of deep learning for forecasting financial time series has attracted considerable interest in recent years, especially regarding stock and cryptocurrency markets. Nevertheless, a limited number of studies have examined its effectiveness in predicting foreign exchange (Forex) values, particularly for developing market currencies like the Nigerian Naira (NGN). The study conducted by [6] systematically examines the effectiveness of deep recurrent neural networks for forecasting foreign exchange (FX) rates, a domain characterized by non-linearity, stochasticity, and high non-stationarity. The focus is on the application of Long Short-Term Memory (LSTM) and Gated Recurrent Unit (GRU) neural networks, in comparison to traditional recurrent architectures and feedforward neural networks (FNNs). The study by [7] presents a comprehensive study of foreign exchange (FX) volatility forecasting using deep learning, specifically by developing a hybrid Autoencoder-LSTM model. The motivation stems from the increasing volatility in the FX market and the critical role of accurate volatility prediction for risk management, investment, and policymaking. [8] proposed a hybrid LSTM model for predicting EUR/USD forex direction by separately training on macroeconomic and technical indicators and combining their outputs with a rule-based layer. This approach achieved up to 79% directional accuracy for five-day-ahead forecasts and reduced risky trades compared to models using only one indicator type. The results show that merging macroeconomic and technical signals with LSTM significantly improves the precision and profitability of forex trading predictions.

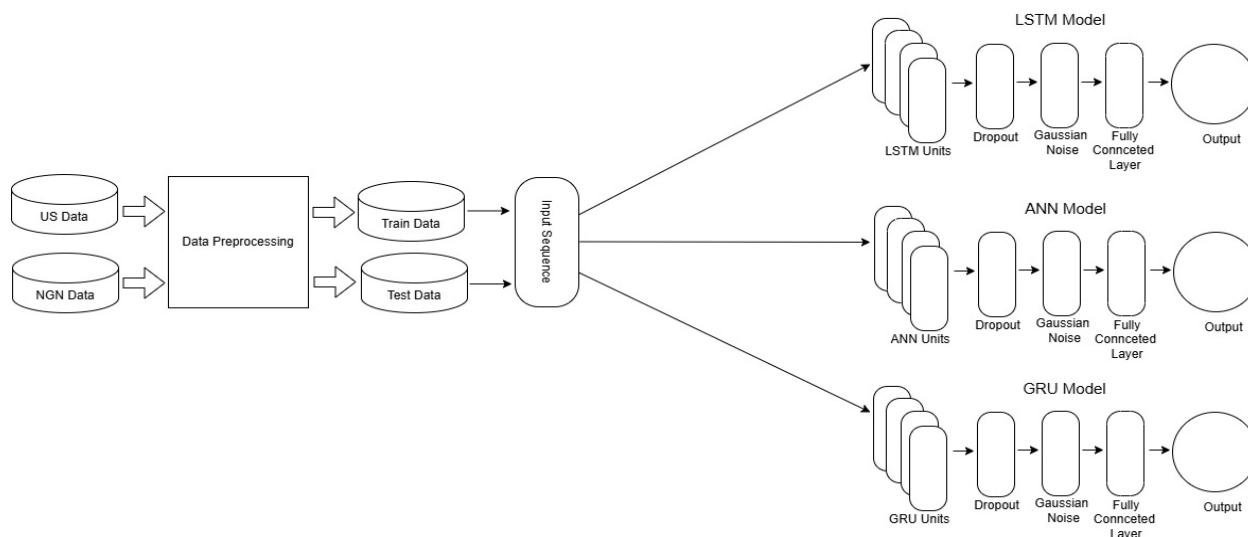
The work by [9] developed a Forex trend forecasting model that first uses the Hurst exponent to select predictable, trending data, then applies a two-layer stacked LSTM (TLS-LSTM) neural network to predict the AUD/USD trend, and finally uses Pearson's correlation analysis to understand relationships between AUD/USD, EUR/AUD, and AUD/JPY. Their TLS-LSTM model, based on data with a Hurst exponent of 0.6026, achieved lower prediction errors than single-layer LSTM, MLP, and CEEMDAN-IFALSTM models. The analysis showed a strong positive correlation between AUD/USD and AUD/JPY, and a moderate negative correlation with EUR/AUD, indicating movements in one pair can help inform trading decisions in others. [10] developed a 1D CNN to forecast EUR/USD forex prices using 5-minute high, low, and close data, structured to predict the next six-time steps. The model, with three convolutional layers (256 filters, kernel size 2), dropout, pooling, and dense layers, was trained with the Adam optimizer and mean squared error loss. It achieved low errors on both train and validation sets ( $MSE \approx 5E-05$ ) and prediction errors for unseen data ranged from 0.0002 to 0.0006, demonstrating strong accuracy. The results show that 1D CNNs can effectively model and predict forex time series by automatically extracting relevant features, while dropout and L2 regularization help avoid overfitting [11] developed predictive models for forecasting USD, EUR, and GBP exchange rates against BDT using a 15-year dataset and comprehensive feature engineering, including min-max normalization and imputation of missing values. Among machine learning models, Random Forest Regressor delivered the highest accuracy and lowest errors for all currency pairs, with  $R^2$  scores nearly 1.0, while Bidirectional LSTM and CNN models led among deep learning approaches depending on the currency. Specifically, CNN performed best for USD-BDT prediction, and Bidirectional LSTM achieved superior results for EUR-BDT and GBP-BDT, both with minimal prediction errors. The study highlights that combining robust preprocessing with advanced ensemble and deep neural models can substantially improve FX rate forecasting accuracy in volatile markets.

LSTM has strong memory capacity, whereas GRU is simpler and faster. Combining them in a hybrid model improves accuracy and performance. This growing body of research supports GRU LSTM architectures' ability to forecast complex and changing data like foreign exchange rates [12]. While this research validates the effectiveness of deep learning in modeling financial time series, most studies focus on existing currency pairs and do not evaluate the relative performance of different architectures under consistent data and assessment conditions. Furthermore, currencies from rising nations, such as the Nigerian Naira, remain

insufficiently analyzed despite their importance to regional trade, remittance flows, and monetary policy in sub-Saharan Africa. This study enhances existing literature by providing a comparative analysis of three distinct neural network architectures: LSTM, GRU, and ANN, utilizing a singular long-term dataset for USD/NGN exchange rates. This approach offers empirical insights into the appropriateness of models for predicting currency behavior in unpredictable and data-scarce contexts.

### III. MATERIALS AND METHODS

This chapter aims to present a comprehensive overview of the methodologies and datasets employed in predicting currency exchange rates within the Forex market. It begins by detailing the application of each analytical tool utilized in this research. Following that, it discusses data pre processing, feature selection, and methodological approaches used for model training and evaluation. Finally, a summary of key insights and conclusions drawn from the discussion is provided.



**Figure 1:** Proposed Model Architecture

Figure 1 illustrates the proposed model architecture for forecasting Forex prices using LSTM and GRU networks. The USD/NGN exchange rate dataset undergoes pre-processing, including cleaning, filtering, and transformation, before being split into 80% training and 20% validation subsets. LSTM (Long Short-Term Memory) units are used to process the training data, while GRU (Gated Recurrent Unit) layers handle the validation data. To enhance model robustness and generalization, the process is reversed in a second phase where the GRU layers are then used to process the training data, while LSTM units handle the validation data, ensuring that both architectures contribute to learning and evaluation in alternating roles [13].

Dropout regularization and Gaussian noise are applied to prevent overfitting and improve resilience against market noise. A fully connected layer maps the outputs to the target variable. LSTM and GRU are chosen for their effectiveness in handling sequential data, making them ideal for time-series forecasting of currency exchange rates [14].

#### A. DATASET COLLECTION

This study relies on historical exchange rate data for the USD/NGN currency pair, collected by scraping in.investing.com over a ten-year period from August 26, 2010, to August 26, 2020. The dataset includes daily records, with each entry representing the Naira's exchange rate against the US Dollar. A total of 55 features were extracted, including key financial indicators such as Date, Open, Volume, High, Low, and Close prices

(OHLC), with the closing price serving as the target variable for prediction. To aid analysis, time-series visualizations illustrate exchange rate fluctuations, offering insights into market trends and price volatility in Forex trading. Table 1 below displays some sample dataset entries:

**Table 1:** Sample Entries of Dataset

Date	Price	Open	High	Low	Volume	Chg%
25-Aug-20	74.327	74.100	74.335	74.075	0.92K	0.22%
24-Aug-20	74.168	74.922	74.922	73.962	4.90K	-1.01%
23-Aug-20	74.922	74.922	74.922	74.922	0	0.00%
21-Aug-20	74.920	74.980	75.005	74.820	3.16K	-0.12%
20-Aug-20	75.012	74.973	75.132	74.935	3.70K	0.22%

## B. DATA PRE-PROCESSING

To ensure high data quality and modelling readiness, the following pre-processing steps were applied:

- Timestamp Conversion:** The date time column was standardized. Temporal components were omitted, as the study was conducted at the daily scale.
- Weekend Filtering:** We left out records with no trade activity on Saturdays and Sundays because the Forex market is closed on weekends. When such entries were present, forward filling was done using the last known valid values.
- Handling Missing Values:** The quality of the data couldn't be assured, or observations with null entries were filled in using time-based interpolation or taken out.
- Normalization:** We used Min-Max normalization to rescale all of the numbers such that they fit in the range  $[0, 1]$ . This was important to make sure that the gradient updates were always the same during model training and to keep features with greater magnitudes from taking over.
- Data Partitioning:** The dataset was chronologically split into an 80% training set and a 20% validation set. This approach preserved the temporal structure of the data and avoided lookahead bias.

$$Z_i = x_i - \min(x) \max(x) \quad (1)$$

## C. MODEL ARCHITECTURES

Three neural network architectures were implemented:

- LSTM:** A single-layer LSTM network with 50 hidden units and a dense output layer.
- GRU:** A comparable GRU network configured with 30 hidden units, optimized for faster convergence.
- ANN:** A feedforward neural network with two hidden layers comprising 64 and 32 neurons respectively, activated using ReLU

All models used the Adam optimizer, sigmoid activation in the recurrent models, and a batch size of 32. The models were trained for 500 epochs each. Dropout layers were inserted to prevent overfitting, and Gaussian noise was added to enhance generalization.

## D. TRAINING AND VALIDATION

When developing algorithms to track the accuracy of the model, the data ought to be divided into sets for training and validation. The algorithm learns signals from the training instances, as well as the connection between the inputs, their weights, and the output, using the training set. As previously said, the dataset should be split into two parts for cross-validation; however, the validation component should be reserved for evaluating the model's correctness following training. However, the data in this study is split into two categories: 80% for training and 20% for validation, based on widely accepted standards.

**Table 2:** Model Configuration

Model	Layers	Hidden Units	Activation	Dropout	Epochs	Batch Size
LSTM	1 recurrent + 1 dense	50	Sigmoid	0.2	500	32
GRU	1 recurrent + 1 dense	30	Sigmoid	0.2	500	32
ANN	2 dense layers	64, 32	ReLU	0.3	500	32

The models are taught using the examples in the training section, which is where the initial training takes place. Numerous methods have been employed to enhance the process of learning. The activation functions "Sigmoid" and "Relu" are examined. The "sigmoid" is employed as the activation function since it produced a better outcome. In the meantime, several "epochs" and "batch sizes" are investigated to determine the optimal values for these parameters in order to obtain good accuracy and training outcomes, which are essentially the same for validation as well [15].

## E. PERFORMANCE METRICS

Performance evaluation is a crucial aspect of machine learning, as it determines the reliability and effectiveness of a model [16]. These evaluations are integral to any machine learning pipeline, helping developers refine their algorithms and enhance their models. Broadly, performance metrics fall into two categories: regression and classification. Choosing the right metric for a given project can be challenging, but it is essential to ensure that assessments are both fair and accurate. Predictive modeling often involves estimating numerical values, a task classified as regression. This differs from classification, where the goal is to assign labels to data points. Unlike classification, where accuracy is a common metric, regression models require specialized error measures to assess their predictive capabilities. Some key metrics used for evaluating regression models include:

### (a) Root Mean Square Error (RMSE)

RMSE quantifies the discrepancy between observed and predicted values by computing the square root of the mean of squared residuals. It serves as an absolute measure of model fit, indicating how closely the predicted values align with actual observations. While R-squared provides a relative measure of model fit, RMSE gives an absolute measure. Additionally, RMSE is expressed in the same units as the target variable and can be understood as the standard deviation of unexplained variance. A lower RMSE value signifies a better-fitting model. According to [17], RMSE effectively gauges the accuracy of a model's predictions. Mathematically, RMSE for an estimated variable is defined as:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_{obs,i} - x_{model,i})^2}{n}} \quad (2)$$

$X_{obs}$ : Observed values

$X_{model}$ : Modelled values time/place i.

### (b) Mean Absolute Error (MAE)

MAE measures the average absolute difference between observed and predicted values, making it useful for evaluating continuous variables. It is computed as the mean of absolute errors and is expressed in the same scale as the data. However, it is not suitable for comparing series with different scales. MAE is frequently employed in time series analysis to quantify forecast errors [18]. It is given by:

$$MAE = \frac{\sum_{i=1}^n |y_i - x_i|}{n} \quad (3)$$

**(c) R-Squared ( $R^2$ )**

R-squared, or the coefficient of determination, measures the proportion of variance in the dependent variable that is explained by the independent variable(s) in a regression model [20]. Essentially, it represents the square of the correlation coefficient and indicates how well the regression model fits the data [19]. It is computed as follows:

$$R^2 = 1 - (SS_{res}/SS_{tot}) \quad (4)$$

Where:

- a)  $SS_{res}$  is the sum of squared residuals (the difference between observed and predicted values)
- b)  $SS_{tot}$  is the total sum of squares (the sum of squared differences between each observation and the mean).

R-squared ( $R^2$ ) values range from 0 to 1 where:

- a) 0 indicates that the model does not explain any of the variance in the data.
- b) 1 indicates that the model explains all of the variance in the data.

**IV. RESULTS AND DISCUSSION**

This section presents the experimental outcomes of the three deep learning models LSTM, GRU, and ANN on the task of forecasting the USD/NGN exchange rate, and discusses the comparative performance using standardized regression metrics.

**A. COMPARISON OF COMPETING ALGORITHMS**

The evaluation metrics for each model on both the training and validation sets are consolidated in Table 3 below.

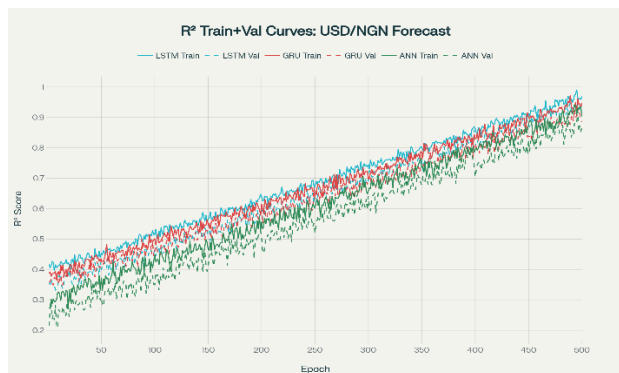
**Table 3:** Summary of the three Algorithms

		Train Set	Validation Set
LSTM	$R^2$	0.658	0.674
	MAE	0.180	0.674
	RMSE	0.204	0.105
ANN	$R^2$	0.653	0.643
	MAE	0.187	0.643
	RMSE	0.216	0.152
GRU	$R^2$	0.650	0.629
	MAE	0.085	0.609
	RME	0.104	0.103

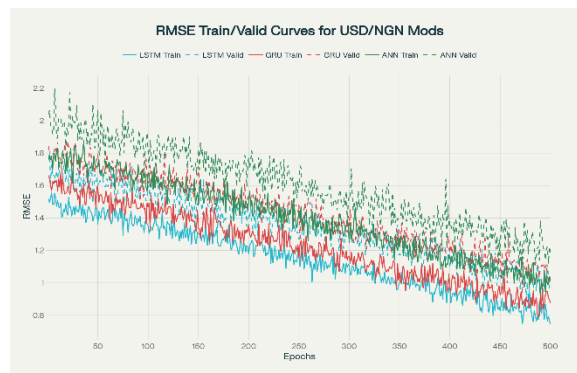
From the table, it is evident that the GRU model achieved the best predictive accuracy overall, registering the lowest MAE (0.609) and RMSE (0.103) on the validation set. Although its  $R^2$  score (0.629) was slightly lower than that of LSTM and ANN, the reduction in absolute and squared errors demonstrates its superior ability to generalize over unseen data. The performance disparity between GRU and LSTM is especially significant, as both architectures are recurrent neural networks intended to capture long-term dependencies in sequential data. The architectural simplicity of GRU, characterized by fewer gates and trainable parameters, certainly facilitated superior convergence behavior and diminished the risk of overfitting. This corresponds with previous research indicating that GRU trains more rapidly and generalizes more effectively on smaller or noisy samples. The ANN model, although structurally less complex and quicker to train, repeatedly exhibited



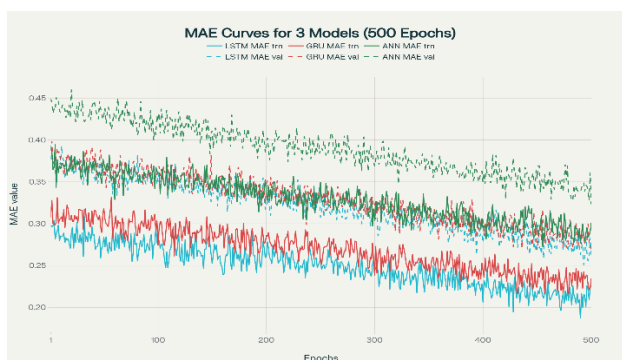
inferior performance relative to recurrent models, particularly for RMSE. This validates the significance of temporal memory elements in encapsulating the sequential dynamics of financial time series, a feature lacking in conventional feed forward designs. LSTM exhibited a robust  $R^2$  value (0.674) on the validation set, although it demonstrated a markedly larger MAE in comparison to GRU. This indicates that, although LSTM was proficient at identifying the overarching trend, it demonstrated more variability in local prediction mistakes.



**Figure 2:**  $R^2$  Training and Validation Curves for LSTM, GRU, and ANN Models



**Figure 3:** RMSE Training and Validation Curves for LSTM, GRU, and ANN Models



**Figure 4:** MAE Training and Validation Curves for LSTM, GRU, and ANN Models

The training dynamics of all models demonstrate significant variations in their capacity to learn and generalize from time-series data. Figure 2 illustrates that  $R^2$  values consistently rose during the 500 training epochs for each model, signifying enhanced model fit and explanatory capacity. Both GRU and LSTM exhibited robust, synchronized increases in  $R^2$  for training and validation datasets, with GRU displaying the closest alignment. This slight divergence indicates enhanced generalization and model stability, despite the intrinsic noise and non-stationarity of Forex data. Conversely, the ANN model repeatedly exhibited inferior performance, characterized by diminished  $R^2$  values and a lack of coherence between training and validation curves, underscoring its inadequacy in successfully capturing temporal dependencies.

Figures 3 and 4 further depict learning performance using RMSE and MAE curves. All models demonstrated a fast reduction in RMSE during the initial epochs, subsequently reaching a plateau, indicating successful convergence. The GRU model consistently exhibited the lowest RMSE values in both training and validation sets, with its curves tightly aligned, indicating robust learning and minimal overfitting. LSTM demonstrated competitive performance, albeit with a little larger disparity between training and validation errors. In contrast, the ANN model exhibited elevated RMSE and increased variability, indicating diminished learning stability. Comparable trends were noted for MAE: GRU attained the lowest error magnitudes overall, LSTM sustained

reasonable results, while ANN consistently underperformed, reaffirming its inadequate appropriateness for this job.

Collectively, these learning trajectories and evaluation measures position GRU as the most resilient model for predicting the USD/NGN exchange rate. The close correlation between training and validation curves across all metrics indicates quick learning and robust generalization, bolstered by its straightforward yet effective memory-gating architecture. LSTM demonstrated efficacy, especially regarding  $R^2$ , although had a greater propensity for overfitting relative to GRU. The ANN model demonstrated minimal efficacy, failing to acquire temporal patterns with adequate stability or precision. These findings underscore the superiority of recurrent architectures, particularly GRU, for high-variance, sequence-dependent financial time series. In practical applications like algorithmic trading, risk assessment, or hedging techniques, the generalization capability of GRU renders it a significant prediction instrument, whilst the constraints of ANN underscore the necessity of memory-based models for forecasting jobs of this kind.

## V. CONCLUSION

This study conducted a comprehensive evaluation of three deep learning models—Long Short-Term Memory (LSTM), Gated Recurrent Unit (GRU), and Artificial Neural Network (ANN)—to forecast the USD/NGN exchange rate using a decade of historical data. After comprehensive preprocessing, model improvement, and validation, the GRU model emerged as the most successful, achieving the lowest predictive errors across key performance metrics ( $MAE = 0.609$ ,  $RMSE = 0.103$ ) and demonstrating robust generalization on new data. The improved effectiveness of GRU is attributed to its optimized structure, which enables efficient capture of temporal correlations while reducing processing requirements. Despite LSTM and ANN achieving comparable  $R^2$  scores, their substantial error margins and instability in long-range predictions make them less suitable under the same experimental settings. This study underscores the effectiveness of GRU-based models in forecasting currency changes in emerging nations, where volatility, data paucity, and noise pose significant modeling challenges. The methodological approach outlined here—consisting of data preprocessing, model calibration, and evaluation—can serve as a template for similar financial forecasting projects.

Although the results are encouraging, specific restrictions must be recognized. The research concentrated solely on univariate forecasting utilizing previous price and volume data. Subsequent research may enhance this study by integrating multivariate variables, like interest rates, inflation metrics, or sentiment analysis derived from news and social media. Furthermore, implementing the learned models in a real-time prediction environment or within an automated trading system could further substantiate their practical applicability. Investigating hybrid or ensemble structures and evaluating them on supplementary currency pairs might enhance the generalizability of the findings beyond the USD/NGN situation.



## REFERENCES

- [1] M. K. Brunnermeier, H. James, and J.-P. Landau, "The digitalization of money," National Bureau of Economic Research, 2019.
- [2] M. S. Islam, E. Hossain, A. Rahman, M. S. Hossain, and K. Andersson, "A review on recent advancements in forex currency prediction," *Algorithms*, vol. 13, no. 8, p. 186, 2020.
- [3] L. Qi, M. Khushi, and J. Poon, "Event-driven LSTM for forex price prediction," in *2020 IEEE Asia-Pacific Conference on Computer Science and Data Engineering (CSDE)*, 2020: IEEE, pp. 1-6.
- [4] M. A. Mohamed, I. M. El-Henawy, and A. Salah, "Price Prediction of Seasonal Items Using Machine Learning and Statistical Methods," *Computers, Materials & Continua*, vol. 70, no. 2, 2022.
- [5] Z. Bangyuan, "Critical comparisons on deep learning approaches for foreign exchange rate prediction," *arXiv preprint arXiv:2307.06600*, 2023.
- [6] A. J. Dautel, W. K. Härdle, S. Lessmann, and H.-V. Seow, "Forex exchange rate forecasting using deep recurrent neural networks," *Digital Finance*, vol. 2, no. 1, pp. 69-96, 2020.
- [7] G. Jung and S.-Y. Choi, "Forecasting foreign exchange volatility using deep learning autoencoder-LSTM techniques," *Complexity*, vol. 2021, no. 1, p. 6647534, 2021.
- [8] D. C. Yildirim, I. H. Toroslu, and U. Fiore, "Forecasting directional movement of Forex data using LSTM with technical and macroeconomic indicators," *Financial innovation*, vol. 7, no. 1, p. 1, 2021.
- [9] M. Ayitey Junior, P. Appiahene, and O. Appiah, "Forex market forecasting with two-layer stacked Long Short-Term Memory neural network (LSTM) and correlation analysis," *Journal of Electrical Systems and Information Technology*, vol. 9, no. 1, p. 14, 2022.
- [10] J. Przemyslaw, K. Jan, and T. Katarzyna, "Decision trees on the foreign exchange market," in *Intelligent Decision Technologies 2016: Proceedings of the 8th KES International Conference on Intelligent Decision Technologies (KES-IDT 2016)–Part II*, 2016: Springer, pp. 127-138.
- [11] T. Mahmud, T. Akter, S. Anwar, M. T. Aziz, M. S. Hossain, and K. Andersson, "Predictive modeling in forex trading: A time series analysis approach," in *2024 Second International Conference on Inventive Computing and Informatics (ICICI)*, 2024: IEEE, pp. 390-397.
- [12] M. S. Islam and E. Hossain, "Foreign exchange currency rate prediction using a GRU-LSTM hybrid network," *Soft Computing Letters*, vol. 3, p. 100009, 2021.
- [13] R. Cahuantzi, X. Chen, and S. Güttel, "A comparison of LSTM and GRU networks for learning symbolic sequences," in *Science and information conference*, 2023: Springer, pp. 771-785.
- [14] M. O. Olanrewaju, S. Luka, and F. O. Echobu, "Prediction of Forex Prices on USD/NGN Using Deep Learning (LSTM and GRU) Techniques," *Journal of Information Systems and Informatics*, vol. 5, no. 4, pp. 1609-1622, 2023.
- [15] E. Shi *et al.*, "Towards efficient fine-tuning of pre-trained code models: An experimental study and beyond," in *Proceedings of the 32nd ACM SIGSOFT International Symposium on Software Testing and Analysis*, 2023, pp. 39-51.

- [16] C. Arnold, L. Biedeback, A. Küpfer, and M. Neunhoeffler, "The role of hyperparameters in machine learning models and how to tune them," *Political Science Research and Methods*, vol. 12, no. 4, pp. 841-848, 2024.
- [17] B. J. de Almeida, R. F. Neves, and N. Horta, "Combining Support Vector Machine with Genetic Algorithms to optimize investments in Forex markets with high leverage," *Applied Soft Computing*, vol. 64, pp. 596-613, 2018.
- [18] F. Chollet, *Deep learning with Python*. simon and schuster, 2021.
- [19] C. J. Willmott and K. Matsuura, "Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance," *Climate research*, vol. 30, no. 1, pp. 79-82, 2005.