Matrik: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer
Vol. 25, No. 3, July 2026, pp. 509∼522
ISSN: 2476-9843, accredited by Kemenristekdikti, Decree No: 10/C/C3/DT.05.00/2025
DOI: 10.30812/matrik.v25i3.6226

Optimized BiLSTM and GRU Models Using QHBM for Forex Price Prediction

Febrianto Alqodri, Triyanna Widiyaningtyas, Didik Dwi Prasetya
Universitas Negeri Malang, Malang, Indonesia

Article Info

Article history:

Received March 06, 2026
Revised March 26, 2026
Accepted May 02, 2026

Keywords:

Deep Learning;
Foreign Exchange;
Prediction;
Hyperparameter Optimization.

ABSTRACT

The foreign exchange market is highly volatile and complex, making accurate price prediction challenging. This study aims to develop an optimized deep learning framework for predicting daily closing prices of seven major currency pairs (AUDUSD, EURUSD, GBPUSD, USDCAD, USDCHF, USDCNY, and USDJPY) by integrating Bidirectional Long Short-Term Memory (BiLSTM) and Gated Recurrent Unit (GRU) models with optimization strategies. Historical data from the Federal Reserve Economic Data were evaluated using Fixed Date Split and Walk Forward Validation (WFV), where WFV consistently achieved better performance than the fixed date. To enhance model performance, hyperparameter optimization was conducted using the Queen Honey Bee Migration (QHBM) algorithm, a metaheuristic approach inspired by the migration behavior of queen bees, divided into two characteristics: high learning rate and low learning rate. The optimized models achieved performance improvements of approximately 10-70% in MAPE and RMSE compared to the baseline models, while maintaining high R^2 values. The results indicate that optimal configurations are pair-specific, where most currency pairs perform best with a high learning rate and high unit settings, while AUDUSD achieves superior performance with a low learning rate and low unit configuration. This study contributes a novel integration of WFV and QHBM-based optimization. Adaptive deep learning models with proper validation significantly improve forecasting accuracy, robustness, and generalization for financial decision-making and algorithmic trading applications.

Corresponding Author:

Triyanna Widiyaningtyas,
Department of Electrical Engineering and Informatics, Faculty of Engineering,
niversitas Negeri Malang, Malang, Indonesia,
Email: triyannaw.ft@um.ac.id

How to Cite:
F. Alqodri, T. Widiyaningtyas, and D. D. Prasetya, ”Optimized BiLSTM and GRU Models Using QHBM for Forex Price Prediction”, *MATRIK: Jurnal Manajemen, Teknik Informatika, dan Rekayasa Komputer*, Vol. 25, No. 3, pp. 509-522, July, 2026.

1. INTRODUCTION

The foreign exchange (forex) market is the largest and most liquid financial market in the world, with daily transaction volumes exceeding USD 7.5 trillion as reported by the Bank for International Settlements (BIS) in its 2025 Triennial Survey [1–3]. Throughout the 21st century, the forex market has experienced extreme volatility driven by major global economic events, including the 2007–2011 Subprime Mortgage Crisis, the 2011–2014 European Debt Crisis, monetary policy normalization during 2014–2020, and the severe market disruptions caused by the COVID-19 pandemic and the Russia-Ukraine war [4]. These crises have reinforced the fragility of global financial systems, as reflected in volatility spikes in the U.S. Dollar Index (DXY) of up to 300%. High liquidity, narrow spreads, and strong macroeconomic sensitivity make forex trading attractive for both individual traders and large institutions such as central banks, hedge funds, and multinational corporations. Among the most actively traded forex pairs are AUD/USD, EUR/USD, GBP/USD, USD/CAD, USD/CHF, USD/CNY, and USD/JPY, which together represent over 80% of global trading volume and exhibit strong correlation with the DXY [5–7]. The highly dynamic, nonlinear, and non-stationary nature of forex price movements poses significant challenges for accurate prediction [8]. Approaches to predicting forex prices have evolved from conventional methods [9]. Previous studies have also used various models such as ARIMA, machine learning methods, and LSTM-based architectures to predict forex movements. While these models show potential, their performance is often inconsistent across different market conditions and currency pairs. Most prior research also relies on default or manually selected hyperparameters, limiting model adaptability for highly volatile forex data. Therefore, further studies are needed to explore alternative deep learning models and apply systematic hyperparameter optimization to achieve more accurate and robust predictions [10, 11].

In recent years, deep learning has emerged as a superior alternative due to its ability to extract hierarchical temporal patterns from complex financial data autonomously [12]. Among these, particularly Long Short-Term Memory (LSTM), its bidirectional extension (BiLSTM), and Gated Recurrent Unit (GRU) have demonstrated strong performance in modeling nonlinear sequential dependencies [13, 14]. The complex temporal relationships in forex time series data are often difficult to capture. In contrast, deep learning approaches such as Long Short-Term Memory (LSTM) have demonstrated superiority in handling fluctuating and non-linear time series data. A bidirectional LSTM (BiLSTM) model was developed that is capable of reading sequential data in two directions, namely forward and backward, in order to understand the temporal context more comprehensively [15]. BiLSTM has been proven to be superior to basic deep learning [16] and excels in capturing bidirectional temporal context, making it suitable for high-complexity markets. LSTM was also developed under the name Gated Recurrent Unit (GRU), which has faster capabilities with a higher level of accuracy [17]. GRU offers computational efficiency with comparable accuracy in certain conditions.

Despite their strong predictive capability, deep learning models such as BiLSTM and GRU are highly sensitive to hyperparameters, including learning rate, batch size, number of neurons, dropout rate, and sequence length [18]. An improper configuration may cause the model to overfit, underfit, or require excessively long training time, leading to unstable or suboptimal forecasting results. The challenge becomes even more critical because each forex currency pair exhibits unique market behavior [19] for example, USD/JPY is highly reactive to Bank of Japan's monetary intervention, while USD/CAD is strongly influenced by oil price fluctuations, and EUR/USD tends to display longer consolidation phases [20]. These distinct market characteristics require tailored algorithmic treatment, rather than a one-size-fits-all model. Therefore, systematic hyperparameter optimization is essential not only to enhance predictive accuracy but also to ensure model adaptiveness, robustness, and computational efficiency in real-world financial environments.

Previous studies highlight that hyperparameter optimization is essential for improving time series prediction models. However, most research focuses on modifying LSTM architectures or applying specific optimization techniques rather than conducting systematic tuning [21]. For example, derivative-free optimization combined with parallel LSTM has been used to improve efficiency and accuracy, but lacks comprehensive comparison and scalability analysis [22]. Other approaches, such as Genetic Algorithm and Heap-Based Optimizer, demonstrate improvements in prediction accuracy and error reduction, yet are often limited in generalizability across datasets and model configurations [23]. This indicates a clear research gap in the development of adaptive, efficient, and scalable hyperparameter optimization frameworks for diverse time-series applications.

To address this gap, this study employs the Queen Honey Bee Migration (QHBM) algorithm, a nature-inspired metaheuristic, to optimize hyperparameters within a deep learning framework integrating BiLSTM and GRU models [24]. QHBM is a metaheuristic algorithm that mimics the natural behavior of honey bee colonies in identifying the most optimal queen candidate to lead the swarm. When applied to deep learning models, QHBM systematically explores the hyperparameter search space to identify the best configuration that yields the highest model accuracy and stability [25]. This approach is particularly effective for complex prediction tasks in which manual tuning or grid search becomes inefficient due to high computation cost and the risk of falling into a local optimum. In this research, QHBM is utilized to systematically explore the hyperparameter space and determine configurations that maximize accuracy, stability, and generalization. Compared to conventional approaches such as grid search or manual tuning, this method provides a more adaptive and computationally efficient optimization strategy for complex financial time-series prediction.

This study focuses on developing an optimized deep learning framework that integrates BiLSTM and GRU models with hyperparameter optimization to improve predicting accuracy and robustness. This is expected to improve the accuracy of daily price predictions (daily close price) for the seven most active forex currency pairs in the 21st century. The research dataset uses historical data from Federal Reserve Economic Data (FRED) for the period between January 1999 and June 2025. This study will compare the performance of BiLSTM and GRU, optimize their hyperparameters, and evaluate the results using metrics such as Mean Absolute Percentage Error (MAPE), Root Mean Square Error (RMSE), and R-squared (R^2) [26]. Accurate forex price prediction offers substantial benefits across multiple financial domains [27]. For traders and financial institutions, it enables more precise entry-exit timing, risk mitigation, and portfolio optimization, thereby minimizing potential losses caused by sudden market volatility.

2. RESEARCH METHOD

This study proposes a deep learning framework to predict the daily closing prices of seven major forex currency pairs, AUD/USD, EUR/USD, GBP/USD, USD/CAD, USD/CHF, USD/CNY, and USD/JPY, covering the period from January 1999 to June 2025. The proposed framework consists of five main stages: data collection, preprocessing, modeling, hyperparameter optimization, and performance evaluation. The research flow is shown in Figure 1.

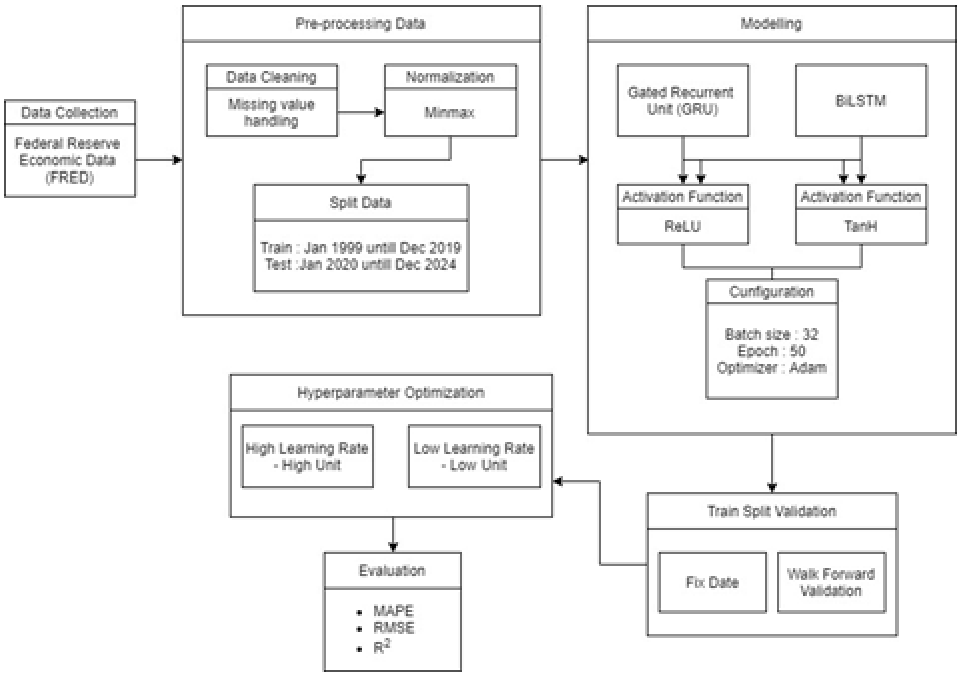

Figure 1. Research Flow

2.1. Data Collection

The data used in this study consist of historical daily foreign exchange prices from seven major currency pairs that represent the most actively traded instruments in the 21st century: AUD/USD (Australian Dollar–US Dollar), EUR/USD (Euro–US Dollar), GBP/USD (British Pound–US Dollar), USD/CAD (US Dollar–Canadian Dollar), USD/CHF (US Dollar–Swiss Franc), USD/CNY (US Dollar–Chinese Yuan), and USD/JPY (US Dollar–Japanese Yen). These seven pairs were selected because they account for the largest share of global forex trading volume and collectively reflect the economic interactions among the world's dominant financial regions: North America, Europe, and Asia-Pacific. The dataset was obtained from the Federal Reserve Economic Data (FRED) repository, managed by the Federal Reserve Bank of St. Louis. The dataset spans from January 1999 to June 2025, ensuring comprehensive coverage of multiple economic cycles, including periods of global stability, financial crises, and post-crisis recoveries. This study adopts a univariate approach, focusing solely on the closing price as the primary variable. This choice is well-suited for deep learning models and is widely used in prior research, as it effectively captures overall price dynamics. Technical indicators are not included because they require parameter tuning for each currency pair and may not generalize well, particularly in markets with frequent sideways movements, and because they increase computational complexity. Similarly, macroeconomic variables are excluded due to potential data revisions, delays, and external interventions. Therefore, using the closing price alone provides a consistent and reliable representation of market behavior, including trends, volatility, and extreme events [28, 29]. The raw data and the atrribute of this study can be seen in Table 1.

Table 1. The Sample of the Dataset

Pairs (String)	Date (Date)	Close (Float)
AUDUSD	04/01/1999	0.6182
EURUSD	04/01/1999	11.812
GBPUSD	04/01/1999	16.581
USDCAD	04/01/1999	15.268
USDCHF	04/01/1999	13.666
USDCNY	04/01/1999	82.793
USDJPY	04/01/1999	112.15

For model development, the closing price was selected as the main predictive feature. The daily closing price represents the final recorded price of a currency pair at the end of a trading day. In the forex market, which operates 24 hours a day across different global sessions (Asia, Europe, and North America), the closing price is typically defined at the end of the New York trading session, marking the completion of a 24 hour trading cycle. It reflects the last transaction price before market rollover and is considered the most stable and reliable indicator of market sentiment for that day. For time series modeling, using the closing price as the target variable enables algorithms to identify meaningful temporal dependencies and long-term patterns. Consequently, in this study, the daily closing price serves as the key predictive feature for each currency pair, forming the foundation for evaluating the performance of BiLSTM and GRU models in forecasting forex market trends. The dataset descriotion can be seen in Table 2.

Table 2. Dataset Description

Pairs	Amount of Data	Min	Max	Mean	Skew	Kurt	SD	CV (%)
AUDUSD	6520	0.48	110.260	0.7580	0.4603	-0.2984	0.1382	182.267
EURUSD	6520	0.83	160.100	11.847	0.0706	-0.3737	0.1559	131.565
GBPUSD	6520	10.703	211.040	15.308	0.4465	-0.5458	0.2200	143.705
USDCAD	6520	0.92	161.280	12.580	0.0027	-0.8994	0.1717	136.460
USDCHF	6520	0.73	182.500	11.195	11.394	0.1825	0.2498	223.111
USDCNY	6521	60.402	828.000	71.796	0.3505	-14.027	0.7762	108.119
USDJPY	6520	757.200	16.173.000	1.109.733	0.3069	0.6440	164.891	148.587

After data collection, to ensure data quality, normality testing is conducted using the Shapiro-Wilk and Jarque-Bera tests. These tests assess whether the data follow a normal distribution and provide information on skewness and kurtosis, which are important for understanding the dataset's statistical properties. Outlier handling in this study is performed using a missing value approach by excluding non-trading days from the dataset. Since the foreign exchange market operates only on active trading days, Monday to Friday, weekend data are not considered missing but are removed entirely. The dataset is then arranged sequentially based on trading days only, ensuring that the time series reflects actual market activity and preserves temporal consistency without artificial data imputation. For model evaluation, statistical methods are applied to compare forecasting performance. Confidence intervals are used to measure the uncertainty of predictions. The Diebold Mariano test is employed to assess the significance of differences in predictive accuracy between models, while the Model Confidence Set method is used to identify a group of models with statistically equivalent performance.

2.2. Pre-processing Data

The data preprocessing phase is a crucial step that ensures the forex time series data are properly formatted and optimized for deep learning model training [30]. After collecting daily closing price data for seven major currency pairs from FRED, the dataset undergoes a series of preprocessing steps.

1. First, the data are cleaned to handle missing or inconsistent records using interpolation and forward-fill methods, ensuring that each time series remains continuous across the entire observation period. The dataset contained a total of 1,916 missing values across all currency pairs. These missing entries were identified during preprocessing and subsequently removed.

2. Next, the cleaned dataset is normalized using the Min-Max scaling technique, which rescales all price values to a range between 0 and 1. This step is essential because BiLSTM and GRU models are sensitive to data magnitude; normalization helps stabilize the learning process and accelerates convergence. The data are then structured into supervised learning format, where a specific number of previous time steps (n) are used as input features to predict the next closing price, effectively converting sequential data into model readable input-output pairs. The formula for min max normalization can be seen in equation 1 [31].

$$X' = \frac{X - Xmin}{Xmax - Xmin} \tag{1}$$

3. Finally, the processed data are split for validation according to two static segments training data from January 1999 to December 2019 and testing data from January 2020 to December 2024. To prevent data leakage, this study applies a strict time based train test split to preserve temporal order. A sliding window approach is used to construct sequential inputs, while feature scaling is performed using only training data parameters and then applied to validation and testing sets. Additionally, a short warm-up phase in the initial iterations is introduced to stabilize training and avoid early prediction spikes, ensuring a robust and unbiased forecasting process.

2.3. Modelling

In this stage, deep learning models were constructed to predict forex price movements. BiLSTM and GRU architectures were employed due to their proven ability to capture complex non-linear patterns in sequential financial data. The architecture of BiLSTM and GRU used in this study shown in Table 3.

Table 3. BiLSTM adn GRU Architecture

Parameter	BiLSTM Architecture	GRU Architecture
Input Shape	(batch_size, 1, 1)	(1, 1)
Layer	Bidirectional LSTM (activation=None, return_sequences=False)	GRU (int(units), activation=None, return_sequences=False)
Learning Rate	0.001	0.001
Unit	50 ($\times$2 direction)	50
Dropout	0.2	0.2
Dense Layer	Dense(1)	units=1
Activation Function	tanh or ReLU	tanh or ReLU
Batch_size	32	32
Optimizer	Adam Optimizer	Adam Optimizer
Number of Epoch	50	50
Loss Function	MAE	MAE

The selection of activation functions in this study is based on the characteristics of sequential financial data and the architecture of recurrent networks. The tanh activation function is primarily used within the BiLSTM and GRU layers because it provides bounded outputs in the range of [-1, 1], which helps stabilize gradients and is well-suited for capturing both positive and negative temporal dependencies in normalized time-series data. This is particularly important for forex data, which exhibits fluctuating patterns around a central tendency. In contrast, ReLU is used in dense (fully connected) layers due to its computational efficiency and ability to accelerate convergence by mitigating vanishing-gradient issues. The combination of tanh in recurrent layers and ReLU in dense layers enables the model to learn temporal dynamics while maintaining efficient training effectively.

The model training strategy incorporates several regularization techniques to ensure optimal performance and prevent overfitting. Specifically, Dropout is applied to the network structure to reduce model complexity and improve generalization by randomly deactivating neurons during training. Early Stopping is used to control training duration by halting training when no further improvement is observed on the validation set, thereby avoiding overfitting. In addition, statistical evaluation methods, MCS, and DM test are employed as a final validation step.

2.4. Split Data Validation

In this research, two types of data-split validation techniques are applied to ensure comprehensive model evaluation and assess the robustness of predictive performance: Fix Date Validation and Walk Forward Validation (WFV). The Fix Date Validation method divides the dataset into two fixed periods: training and testing data. This approach provides a clear and consistent boundary between training and testing phases, allowing for straightforward comparison of model performance. The justification for using the Fixed Date split based on the COVID-19 period is supported by observable structural changes in market behavior. Prior to COVID-19, forex markets exhibited relatively stable movements with less significant fluctuations. However, during the pandemic, the shift toward remote work and increased financial activity led to a substantial rise in trading demand and volume. This change is reflected in several currency pairs. For example, the average value of USDJPY increased from approximately 106.93 before COVID-19 to 128.02 during the pandemic period. Similarly, USDCAD rose from around 1.24 to 1.32, while EURUSD increased from a minimum value of 0.48 to 0.57, indicating higher trading activity. Additionally, USDCHF experienced an increase of about 5% in its minimum trading value [32]. The split of training and testing data for the fixed date can be seen in Table 4.

Table 4. Fix Date Split

Pairs	Training Data Period	Testing Data Period
AUDUSD	01-01-1999 untill 31-12-2019	01-01-2020 untill 31-12-2024
EURUSD	01-01-1999 untill 31-12-2019	01-01-2020 untill 31-12-2024
GBPUSD	01-01-1999 untill 31-12-2019	01-01-2020 untill 31-12-2024
USDCAD	01-01-1999 untill 31-12-2019	01-01-2020 untill 31-12-2024
USDCHF	01-01-1999 untill 31-12-2019	01-01-2020 untill 31-12-2024
USDCNY	01-01-1999 untill 31-12-2019	01-01-2020 untill 31-12-2024
USDJPY	01-01-1999 untill 31-12-2019	01-01-2020 untill 31-12-2024

Meanwhile, the Walk Forward Validation (WFV) technique introduces a dynamic evaluation process through the expanding window approach. In this method, the training window starts with an initial subset of data, and as the process continues, the window expands by adding more recent observations before predicting the next period. After each iteration, the model is retrained using all available past data and tested on the subsequent unseen segment. The split data validation, using both Fix Date and Walk Forward Validation, is applied only to the best-performing model to ensure its predictive capability is thoroughly evaluated under different market conditions.

The Walk Forward Validation (WFV) process is implemented using an expanding-window approach with a window size of 120 days and 10 iterations. The initial training set begins with a fixed amount of training data, and in each subsequent iteration, the training window is expanded by an additional 120 days until all testing data is exhausted. At each step, the model is trained on the expanded dataset and evaluated on the next unseen segment while preserving temporal order.

2.5. Hyperparameter Optimization

The hyperparameter optimization stage is a crucial process in this research, aimed at improving the predictive performance and computational efficiency of the BiLSTM and GRU models. Hyperparameters such as the number of units, learning rate, dropout rate, batch size, and number of epochs significantly influence how well the model learns from data, converges during training, and generalizes to unseen market conditions. Instead of relying on manual or trial-and-error methods, this study employs the Queen Honey Bee Migration (QHBM) to search for the most effective hyperparameter combinations automatically. QHBM mimics the intelligent migration and selection behavior of queen bees, balancing exploration and exploitation to efficiently identify optimal configurations. The mathematical formulation and the QHBM optimization process can be shown in Table 5.

Optimization process focuses on identifying the most effective hyperparameter configuration by comparing combinations of learning rate, unit size, and dropout rate. The study evaluates two main configurations, low learning rate low unit and high learning rate high unit, to determine which setting better captures the temporal dynamics of each currency pair while maintaining model stability. Low learning rate low unit configuration represents parameter values that are intentionally reduced below the baseline to encourage more gradual weight updates and lower model complexity, which can enhance stability and generalization. In contrast, the high learning rate high unit configuration uses parameter values increased above the baseline, enabling faster adaptation and greater representational capacity to capture more complex temporal patterns. Thus, the optimization process evaluates whether decreasing or increasing parameter magnitudes relative to the baseline leads to improved forecasting accuracy and stability, allowing the model to adapt more effectively to the unique characteristics of each currency pair. The pseudo code of the hyperparameter optimization and hyperparameter search space can bee seen in and Table 6, 7, and 8.

Table 5. QHBM Optimization Process

QHBM Optimization Pseudocode
Initialization : Set all variables
Decision: Pole Selection
Calculate the excitement of scout bee in sector.
$c_j = \frac{1}{n} \sum_{j=1}^{n} c_{1(j)}$
Calculate probability (weigh of information)
$P_j = \frac{c_j}{\sum_{j=1}^{8} c_j}$
Move to the selected pole
Calculate Migration Length
$r_m^{(k+1)} = \left(1 - g_m^{(k)}\right) \cdot r_s$
Update Natural Factor
Update Queen Position
$g_m^{(k+1)} = g_m^{(k)} \cdot rand(1)$
$x_s^{(k+1)} = x_s^{(k)} + r_m^{(k+1)} \cos\theta^{(k+1)}$
$y_s^{(k+1)} = y_s^{(k)} + r_m^{(k+1)} \sin\theta^{(k+1)}$
$z_s^{(k+1)} = z_s^{(k)} + r_m^{(k+1)} \cos\theta^{(k+1)}$
If current fitness improves the best fitness then
Update best solution
End if
Return best solution

Table 6. Optimization Pseudocode

Optimization Pseudocode
Initialize position, radius, g, and k
Scan nodes into 8 sectors
Compute energy and probability for each sector
Select the target sector with the highest probability (p)
Update distance traveled and position using the formula:
$X_s^{(k+1)} = X_s^{(k)} + r_m^{(k+1)}.cos\theta^{(k+1)}$
$Y_s^{(k+1)} = y_s^{(k)} + r_m^{(k+1)}.sin\theta^{(k+1)}$
Update the number of iterations
Repeat until completion or until the stopping condition based on iterations is met

Table 7. Hyperparameter Search Space

Parameter	Low LR and Unit	High LR and Unit
Hyperparameter Search Space		
Learning Rate	0.000010 - 0.001000	0.000010 - 0.001000
Units	50 - 250	250 - 500
Dropout Rate	0.2 - 0.5	0.2 - 0.5
Epoch	Oct-50	Oct-50
Stability Test	10	10
QHBM Search Space		
Radius	5	
Epoch	10	
Scout Count	100	
Patience Level	250	

2.6. Evaluation

To ensure a rigorous and reliable evaluation, this study employs three statistical tests. The Model Confidence Set (MCS) is used to identify a subset of models that are statistically equivalent in predictive performance, based on their average loss. The procedure begins by computing the loss series for each model, followed by iteratively eliminating inferior models using hypothesis testing until only the best-performing set remains. A model ranked first in MCS indicates it has the lowest average loss and is among the statistically superior models. In addition, this study MAE is also used as an amplifier, because the model confidence set (MCS) will assess how small the average loss that occurs.

To further validate pairwise differences, the Diebold-Mariano (DM) test is applied to compare forecast errors between competing models. The DM statistic is calculated from the difference in performance between two models and evaluates whether the mean difference is significantly different from zero. A higher number of DM wins indicates that a model consistently outperforms others with statistical significance. Additionally, confidence intervals (CI) are used to assess prediction stability by measuring the variability of model errors across testing periods. The CI is computed from the distribution of prediction errors, with narrower intervals indicating more consistent and reliable performance.

3. RESULT AND ANALYSIS

The Results and Discussion section systematically presents the findings of this research, beginning with the performance evaluation of the BiLSTM and GRU models under both baseline and QHBM-optimized settings. Using daily closing prices for seven forex currency pairs. The results are analyzed using multiple accuracy metrics, including MAPE, RMSE, and R^2.

3.1. Modeling (BiLSTM and GRU)

In this section, the BiLSTM and GRU models are evaluated to determine which architecture provides the highest predictive capability. To identify the best configuration among the experiments, this study adopted a Multi-Criteria Decision Making (MCDM) method using Simple Additive Weighting (SAW). SAW was selected because it effectively merges several performance indicators that may conflict in financial prediction tasks, producing a more balanced and unbiased assessment [33]. The results for the best model with the activation function are shown in Table 8.

Table 8. BilSTM and GRU Evaluation

Pairs	Model Activation	MAPE (35%)	RMSE (35%)	R^2 (30%)	Final Score
AUD/USD	BiLSTM_tanh	0.007516	0.006691	0.975233	0.867339
EUR/USD	BiLSTM_relu	0.008353	0.011083	0.963488	0.875558
GBP/USD	GRU_tanh	0.00563	0.009067	0.980721	0.876332
USD/CAD	BiLSTM_relu	0.003728	0.006427	0.984056	0.881828
USD/CHF	BiLSTM_tanh	0.014409	0.014285	0.836299	0.840828
USD/CNY	BiLSTM_relu	0.004593	0.038307	0.985259	0.871531
USD/JPY	GRU_relu	0.012241	2.088.819	0.986513	0.652144

The modelling results reveal several notable patterns related to model performance, activation functions, and the intrinsic characteristics of each currency pair. Overall, BiLSTM consistently outperforms GRU across most major currency pairs, particularly in terms of MAPE, RMSE, and R^2. This aligns with the theoretical understanding that BiLSTM captures both forward and backward temporal dependencies, making it more suitable for markets with highly complex and nonlinear price movements [34]. Across the seven major pairs, AUD/USD, EUR/USD, USD/CAD, and USD/CNY show exceptionally strong performance, with BiLSTM achieving MAPE values near or below 0.005 and R^2 values consistently above 0.90. These pairs are known for high liquidity and relatively stable macroeconomic drivers, enabling the models to learn clearer temporal patterns. Conversely, USD/JPY and USD/CHF exhibit significantly lower accuracy, reflected in higher MAPE and RMSE values and noticeably reduced R^2 scores. An interesting finding is that the choice of activation function influences performance differently across pairs. ReLU generally performs better for highly liquid, trending pairs (EUR/USD, USD/CAD, USD/CNY, USD/JPY), while TanH provides more stability for pairs with mixed volatility patterns (AUD/USD, GBP/USD, USD/CHF). This distinction reflects the sensitivity of activation functions to the distribution and scale of financial time series data.

The distinct statistical characteristics of each market can explain the variation in model performance across currency pairs. Currency pairs such as GBPUSD, USDCAD, USDCHF, USDCNY, and USDJPY exhibit relatively dynamic price movements, higher variability, or larger-scale values, which require models with higher learning rates and greater representational capacity (more units) to capture complex and nonlinear patterns effectively. In contrast, AUDUSD exhibits a more moderate, smoother trend, where a lower learning rate and fewer units yield more stable learning and prevent over-adjustment to noise. This confirms that the model's effectiveness is strongly dependent on the underlying data distribution and the temporal dynamics of each currency pair.

In addition, the role of activation functions contributes to these performance differences. The use of tanh in recurrent layers helps stabilize learning by constraining outputs within a bounded range, which is particularly beneficial for capturing oscillatory patterns in normalized forex data. Meanwhile, ReLU in dense layers accelerates convergence and enhances the model's ability to learn complex nonlinear relationships, especially for pairs with higher volatility and larger fluctuations. The interaction between activation functions and hyperparameter configurations (learning rate and units) enables the model to adapt to different market behaviors, thereby influencing overall predictive performance across currency pairs.

3.2. Train Split Validation

After selecting the best model and confirming the most suitable activation function based on the evaluation, these findings are carried forward to the split-data validation stage to further examine the model's reliability across varying temporal data partitions. The results of the train-split validation are shown in Table 9. After selecting the best model and confirming the most suitable activation function based on the evaluation, these findings are carried forward to the split-data validation stage to further examine the model's reliability across varying temporal data partitions. The results of the train-split validation are shown in Table 9.

Table 9. Split Validation Result

Pairs	Model Activation	MAPE		RMSE	
		Fix Split	WFV	Fix Split	WFV
AUDUSD	BiLSTM tanh	0.00752	0.0090	0.00669	0.0074
EURUSD	BiLSTM relu	0.00835	0.0063	0.01108	0.0082
GBPUSD	GRU tanh	0.00563	0.0093	0.00907	0.0135
USDCAD	BiLSTM relu	0.00373	0.0045	0.00643	0.0074
USDCHF	BiLSTM tanh	0.01441	0.0061	0.01429	0.0067
USDCNY	BiLSTM relu	0.00459	0.0042	0.03831	0.0330
USDJPY	GRU relu	0.01224	0.0062	2.08882	0.9909

The comparison between Fix Date and WFV shows that WFV consistently outperforms Fix Date across almost all currency pairs and MAPE evaluation metrics. For pairs such as AUDUSD, GBPUSD, USDCAD, USDCHF, USDCNY, and USDJPY, WFV yields significantly lower error values, indicating better predictive accuracy and higher model stability. This improvement is especially clear in the MAPE and RMSE columns, where the WFV scores are notably lower than those for Fix Date. The expanding window mechanism of WFV allows the model to continuously retrain using cumulative historical information, making it more responsive to structural changes and long-term dependencies present in real exchange rate movements.

WFV is a more reliable method because it mimics real forecasting conditions: the model is updated sequentially using only past data before predicting the next step. This rolling-window mechanism prevents information leakage and provides a realistic representation of how models behave in live environments. As reflected in the results, WFV captures temporal dynamics more effectively than a fixed train–test split, leading to more robust generalization. This is in line with several studies that state WFV is the most appropriate approach for time-series modeling because it maintains chronological consistency and avoids data leakage, which is common in traditional K-Fold validation [35, 36]. The results confirm that WFV is the superior validation approach, offering better accuracy and more realistic performance for time-series prediction, making it the preferred method for subsequent modeling and optimization steps.

3.3. Hyperparameter Optimization

After obtaining the best split validation results, which are WFV, hyperparameter optimization is conducted further to enhance the model's predictive accuracy and computational efficiency. In this study, hyperparameter optimization used a sweat spot approach, where each parameter was evaluated not only at its baseline value but also across systematically higher and lower ranges around that baseline. To provide a more comprehensive evaluation, this study also presents a model comparison matrix incorporating several statistical criteria. The result of the hyperparameter optimization (WFV-HPO) can be seen in Table 7 and the best parameter from HPO WFV is shown in Table 10.

The results indicate that the optimal hyperparameter configuration is closely related to the statistical characteristics of each currency pair. Currency pairs with high volatility (CV $>$ 18%) and high skewness tend to perform better under the HPO-WFV-HighLRUnit configuration, as higher learning rates and larger model capacity enable faster adaptation to strong fluctuations and asymmetric patterns. In contrast, pairs with high volatility but moderate skewness are better modeled with HPO-WFV-LowLRUnit, where lower learning rates yield more stable convergence and reduce the risk of overshooting. Additionally, datasets with highly negative kurtosis (<-0.8), indicating flatter distributions, benefit from HPO-WFV-HighLRUnit, which allows more efficient learning despite the lack of extreme variations. The best parameter is shown in Table 11.

Table 10. HPO WFV Result

Pairs	Code Mode	MAPE	RMSE	R^2	CI	MCS	DM	MAPE Improvement
	Baseline	0.007516	0.006691	0.975233	0.062811	2	1	
AUDUSD	High LR and Unit	0.007895	0.007047	0.970844	0.065247	3	0	29.19%
	Low LR and Unit	0.005322	0.004779	0.987367	0.050434	1	2	
	Baseline	0.008353	0.011083	0.963488	0.0712	3	0	
EURUSD	High LR and Unit	0.003717	0.005388	0.991385	0.03683	1	2	55.50%
	Low LR and Unit	0.003987	0.005703	0.990339	0.037306	2	1	
	Baseline	0.00563	0.009067	0.980721	0.047575	3	0	
GBPUSD	High LR and Unit	0.004322	0.007346	0.987378	0.044724	1	2	23.23%
	Low LR and Unit	0.004436	0.007511	0.986803	0.045681	2	1	
	Baseline	0.003728	0.006427	0.984056	0.032182	2	0	
USDCAD	High LR and Unit	0.003347	0.005909	0.986644	0.032115	1	2	10.22%
	Low LR and Unit	0.00346	0.006096	0.985773	0.032912	3	0	
	Baseline	0.014409	0.014285	0.836299	0.075597	3	0	
USDCHF	High LR and Unit	0.003911	0.004758	0.981865	0.037561	1	2	72.86%
	Low LR and Unit	0.004477	0.005231	0.978078	0.039721	2	1	
	Baseline	0.004593	0.038307	0.985259	0.037961	3	0	
USDCNY	High LR and Unit	0.00203	0.020093	0.996079	0.023146	1	2	55.80%
	Low LR and Unit	0.003552	0.030428	0.990988	0.027553	2	1	
	Baseline	0.012241	2.088.819	0.986513	0.092053	3	0	
USDJPY	High LR and Unit	0.004661	0.892123	0.997642	0.050104	1	2	61.92%
	Low LR and Unit	0.007258	1.248.165	0.99515	0.057222	2	1	

Table 11. HPO WFV Best Parameter

Pairs	Code Mode	Learning Rate	Units	Dropout Rate	Epoch
AUDUSD	Low LR and Unit	0.000082	217	0.387535	41
EURUSD	High LR and Unit	0.000223	437	0.347868	43
GBPUSD	High LR and Unit	0.000405	439	0.308281	18
USDCAD	High LR and Unit	0.000899	425	0.293775	14
USDCHF	High LR and Unit	0.000224	372	0.301103	32
USDCNY	High LR and Unit	0.000454	438	0.324223	26
USDJPY	High LR and Unit	0.000493	434	0.343516	17

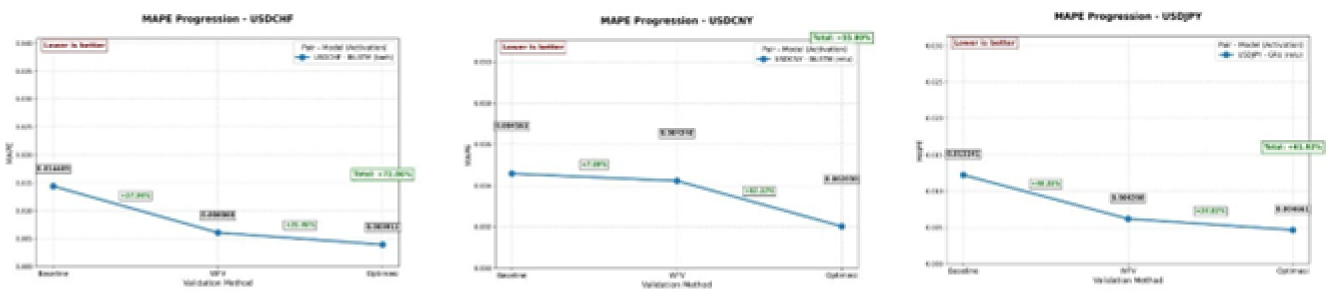

Figure 2. Sample of MAPE Improvement Chart

Figure 2 provides a visual representation of the MAPE improvement achieved by the optimized models across different currency pairs. As shown in Figure 2, the results demonstrate that optimal hyperparameter configurations vary across currency pairs, confirming that each market exhibits distinct temporal dynamics and volatility characteristics. This variation is consistently reflected in forecasting accuracy metrics (MAPE, RMSE, and R^2), statistical comparison through the Diebold–Mariano (DM) test, and the Model Confidence Set (MCS) selection procedure. For most currency pairs, namely EURUSD, GBPUSD, USDCAD, USDCHF, USDCNY, and USDJPY, the strongest performance is achieved using the High Learning Rate and High Units configuration. These pairs exhibit comparatively more dynamic price movements requiring faster adaptation and higher representational capacity. The increased learning speed and larger network structure allow the model to better capture nonlinear fluctuations that cannot be adequately modeled using conservative parameter settings. This is evidenced by higher MAPE and RMSE values and high R^2 scores, indicating stable and reliable predictive behavior. AUDUSD represents the only exception, where the Low Learning Rate and Low Units configuration yields the best overall performance. This suggests that AUDUSD generally presents smoother trend evolution where gradual weight updates help the model capture temporal dependencies without overreacting to short-term noise [37]. The reduced number of hidden units also limits model complexity, preventing overfitting and supporting better generalization.

The Confidence Interval (CI) values in Table 6 reflect prediction stability across testing windows. Models with narrower CI widths indicate more consistent forecasting behavior and lower uncertainty. In this study, most of the optimized configurations exhibit lower CI values than the baseline, indicating improved stability after hyperparameter adjustment. For instance, the optimized models for EUR/USD, USD/CAD, USD/CHF, and USD/CNY show notably tight CI ranges, confirming that their predictive performance remains stable across different temporal segments. Conversely, slightly wider CI values in some baseline models suggest greater variability and less reliable consistency in forecasting. The Model Confidence Set (MCS) column ranks models by average loss, with 1 indicating the best-ranked model for each currency pair. The results show that optimized configurations frequently achieve MCS rank 1, highlighting their superiority in overall predictive accuracy. Specifically, the High LR-High Unit configuration attains the top MCS rank for EURUSD, GBPUSD, USDCAD, USDCHF, USDCNY, and USDJPY. At the same time, AUDUSD reach MCS rank 1 under the Low LR-Low Unit. This pattern confirms that hyperparameter tuning successfully identifies the most reliable configuration for each pair by minimizing average forecasting loss. The Diebold Mariano (DM) test results further validate these findings by measuring pairwise predictive superiority among competing models. A DM value of 1 indicates one significant win, while 2 denotes two significant wins against alternative configurations. Across experiments, optimized models consistently achieve higher DM win counts than the baseline, indicating statistically significant improvements in forecast accuracy. All optimized configurations with High LR-High Unit or Low LR-Low Unit, record two DM wins, reinforcing their dominance over competing settings.

The impact of the learning rate is related to gradient descent behavior: smaller values enable more stable, precise convergence, while larger values allow faster adaptation to rapidly changing patterns. The number of units determines the representational capacity of the model, where larger architectures can capture complex nonlinear relationships but increase the risk of overfitting, whereas smaller architectures promote better generalization. This behavior is consistent with the bias variance trade off, where model performance depends on balancing underfitting and overfitting based on data characteristics. Therefore, the variation in optimal hyperparameters across currency pairs can be attributed to differences in volatility, noise level, and temporal complexity, which require different learning dynamics and model capacities. Overall, optimized configurations consistently outperform the baseline models in terms of accuracy, stability, and statistical significance. Models ranked first in the MCS, supported by stronger DM wins and narrower CI widths, demonstrate superior robustness across evaluation windows. These findings confirm that although the Low LR Low Unit configuration dominates across most pairs, optimal forecasting performance remains pair specific, highlighting the importance of adaptive hyperparameter tuning aligned with each currency pair's market behavior.

4. CONCLUSION

This study set out to optimize deep learning models, BiLSTM and GRU, for predicting the daily closing prices of seven major forex currency pairs throughout the 21st century. Rather than comparing model performance alone, this study demonstrates that predicting accuracy in forex is fundamentally dependent on the alignment among model configuration, validation strategy, and the underlying characteristics of each currency pair. The findings confirm that BiLSTM generally achieves higher predictive accuracy than GRU because it captures bidirectional temporal dependencies, while Walk-Forward Validation (WFV) consistently yields more reliable results than the Fix Date approach. More importantly, the study highlights that temporally consistent validation is critical in time-series forecasting. Building upon the WFV results, hyperparameter optimization using the QHBM algorithm further enhances model performance. The key insight is that optimal configurations are not universal but pair-specific: most currency pairs benefit from high learning rates and high unit settings, whereas more dynamic pairs, such as AUD/USD, require lower learning rates and smaller model capacities. This finding emphasizes that model adaptability is more important than model complexity alone in achieving robust forecasting performance.

Despite these contributions, several limitations should be acknowledged. This study relies solely on historical price data and does not incorporate macroeconomic variables such as interest rates, inflation, or geopolitical indicators, which may further improve predictive capability. Although regularization techniques were applied, the risk of overfitting remains inherent in deep learning models, especially when dealing with highly volatile financial data. The use of metaheuristic optimization such as QHBM introduces additional computational complexity, which may limit scalability in real-time or large-scale applications. Future research can address these limitations by integrating macroeconomic and sentiment-based features, exploring hybrid or ensemble models, and investigating more efficient optimization techniques to reduce computational cost. Additionally, extending the framework to high-frequency data and real-time forecasting environments would further enhance its practical applicability.

5. ACKNOWLEDGEMENTS

During the preparation of this work, the authors used Grok to assist with model No acknowledgments were provided for this study

6. DECLARATIONS

AI USAGE STATEMENT

During the preparation of this work, the authors used Grok to assist with model implementation and ChatGPT (OpenAI) and Claude AI to improve the manuscript's language and clarity. After using these tools, the authors reviewed and edited all content as necessary and took full responsibility for the final version of the publication.

AUTHOR CONTIBUTION

Febrianto Alqodri contributed to the research conceptualization, methodology design, data collection, model development, experimentation, and manuscript drafting. Triyanna Widiyaningtyas contributed to supervision, methodological validation, and critical review of the manuscript. Didik Dwi Prasetya contributed to data analysis, interpretation of results, and refinement of the manuscript's structure and clarity. All authors have read and approved the final version of the article.

FUNDING STATEMENT

This study was conducted without any financial support.

COMPETING INTEREST

The authors declare no conflict of interest regarding the publication of this article.

REFERENCES

[1] B. Eichengreen, R. Lafarguette, A. Mehl, and M. Ferrari Minesso, "Technology and the geography of the foreign exchange market," *Journal of International Money and Finance*, vol. 131, p. 102802, Mar. 2023, https://doi.org/10.1016/j.jimonfin.2023.102802.

[2] C. Kookanok, S. Lertsakulbunlue, M. Poochanasri, and N. Aiumtrakul, "WCN26-4181 Us Economic and Kidney Transplant Awareness: Fred and Google Time Series Database from 2004 to 2025," *Kidney International Reports*, vol. 11, no. 4, p. 104987, Apr. 2026, https://doi.org/10.1016/j.ekir.2026.104987.

[3] K. C. Guyard and M. Deriaz, "Predicting Foreign Exchange EUR/USD Direction Using Machine Learning," in *Proceedings of the 2024 7th International Conference on Machine Learning and Machine Intelligence (MLMI)*. Osaka Japan: ACM, Aug. 2024, pp. 1–9, https://doi.org/10.1145/3696271.3696272.

[4] Y. Qiu, X. Sun, X. Xiong, and S. Si, "Systemic financial risk analysis of the U.S. based on the complex network," *Journal of Management Science and Engineering*, vol. 10, no. 3, pp. 414–433, Sep. 2025, https://doi.org/10.1016/j.jmse.2025.05.002.

[5] Y. Kitamura, "The price discovery in the renminbi/USD market: Two spot, two swap, and three forward FX rates," *International Review of Financial Analysis*, vol. 95, p. 103362, Oct. 2024, https://doi.org/10.1016/j.irfa.2024.103362.

[6] Y.-L. Chen and J. J. Yang, "SDR adjustment and FX liquidity," *Pacific-Basin Finance Journal*, vol. 91, p. 102711, Jun. 2025, https://doi.org/10.1016/j.pacfin.2025.102711.

[7] C. G. Maté, "Forecasting in FOREX the spot price interval of tomorrow with the same information of today. An analysis of the seven majors using a linear regression model based on interval arithmetic," *Knowledge-Based Systems*, vol. 258, p. 109923, Dec. 2022, https://doi.org/10.1016/j.knosys.2022.109923.

[8] S. Perla, R. Bisoi, and P. Dash, "A hybrid neural network and optimization algorithm for forecasting and trend detection of Forex market indices," *Decision Analytics Journal*, vol. 6, p. 100193, Mar. 2023, https://doi.org/10.1016/j.dajour.2023.100193.

[9] P. K. Sarangi, M. Chawla, P. Ghosh, S. Singh, and P. Singh, "FOREX trend analysis using machine learning techniques: INR vs USD currency exchange rate using ANN-GA hybrid approach," *Materials Today: Proceedings*, vol. 49, pp. 3170–3176, 2022, https://doi.org/10.1016/j.matpr.2020.10.960.

[10] P. Albrecht and E. Kočenda, "Volatility connectedness on the central European forex markets," *International Review of Financial Analysis*, vol. 93, p. 103179, May 2024, https://doi.org/10.1016/j.irfa.2024.103179.

[11] F. Alqodri, T. Widiyaningtyas, and D. D. Prasetya, "Comparative Analysis of ARIMA, Random Forest, and BiLSTM for Dollar Index Prediction in the 21 st Century," pp. 1–6, 2025, https://doi.org/10.1109/ICEEIE66203.2025.11251582.

[12] N. Mirza, S. K. A. Rizvi, B. Naqvi, and M. Umar, "Inflation prediction in emerging economies: Machine learning and FX reserves integration for enhanced forecasting," *International Review of Financial Analysis*, vol. 94, p. 103238, Jul. 2024, https://doi.org/10.1016/j.irfa.2024.103238.

[13] T. Fischer, M. Sterling, and S. Lessmann, "Fx-spot predictions with state-of-the-art transformer and time embeddings," *Expert Systems with Applications*, vol. 249, p. 123538, Sep. 2024, https://doi.org/10.1016/j.eswa.2024.123538.

[14] I. H. Sarker, "Deep Learning: A Comprehensive Overview on Techniques, Taxonomy, Applications and Research Directions," *SN Computer Science*, vol. 2, no. 6, p. 420, Nov. 2021, https://doi.org/10.1007/s42979-021-00815-1.

[15] M. Yang and J. Wang, "Adaptability of Financial Time Series Prediction Based on BiLSTM," *Procedia Computer Science*, vol. 199, pp. 18–25, 2022, https://doi.org/10.1016/j.procs.2022.01.003.

[16] A. Chennupati, B. Prahas, B. A. Ghali, B. D. Jasvitha, and K. Murali, "Comparative Analysis of Bitcoin Price Prediction Models: LSTM, BiLSTM, ARIMA and Transformers," in *2024 15th International Conference on Computing Communication and Networking Technologies (ICCCNT)*, Jun. 2024, pp. 1–7, https://doi.org/10.1109/ICCCNT61001.2024.10724247.

[17] P. T. Yamak, L. Yujian, and P. K. Gadosey, "A Comparison between ARIMA, LSTM, and GRU for Time Series Forecasting," in *Proceedings of the 2019 2nd International Conference on Algorithms, Computing and Artificial Intelligence*. Sanya China: ACM, Dec. 2019, pp. 49–55, https://doi.org/10.1145/3377713.3377722.

[18] P. Lalwani and R. Ganeshan, "A Novel CNN-BiLSTM-GRU Hybrid Deep Learning Model for Human Activity Recognition," *International Journal of Computational Intelligence Systems*, vol. 17, no. 1, p. 278, Nov. 2024, https://doi.org/10.1007/s44196-024-00689-0.

[19] S. Hansun, F. P. Putri, A. Q. M. Khaliq, and H. Hugeng, "On searching the best mode for forex forecasting: Bidirectional long short-term memory default mode is not enough," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 4, p. 1596, Dec. 2022, https://doi.org/10.11591/ijai.v11.i4.pp1596-1606.

[20] L. Vancsura, T. Tatay, and T. Bareith, "Navigating AI-Driven Financial Forecasting: A Systematic Review of Current Status and Critical Research Gaps," *Forecasting*, vol. 7, no. 3, p. 36, Jul. 2025, https://doi.org/10.3390/forecast7030036.

[21] Y. Zhao, W. Zhang, and X. Liu, "Grid search with a weighted error function: Hyper-parameter optimization for financial time series forecasting," *Applied Soft Computing*, vol. 154, p. 111362, Mar. 2024, https://doi.org/10.1016/j.asoc.2024.111362.

[22] H. Dhake, Y. Kashyap, and P. Kosmopoulos, "Algorithms for Hyperparameter Tuning of LSTMs for Time Series Forecasting," *Remote Sensing*, vol. 15, no. 8, p. 2076, Apr. 2023, https://doi.org/10.3390/rs15082076.

[23] M. M. Öztürk, "Hyperparameter Optimization of a Parallelized LSTM for Time Series Prediction," *Vietnam Journal of Computer Science*, vol. 10, no. 03, pp. 303–328, Aug. 2023, https://doi.org/10.1142/S2196888823500033.

[24] A. Aripriharta, M. S. Al Rasyid, M. C. Bagaskoro, I. Fadlika, S. Sujito, A. N. Afandi, S. Omar, and N. Rosmin, "Queen honey bee migration (QHBM) optimization for droop control on DC microgrid under load variation," *Journal of Mechatronics, Electrical Power, and Vehicular Technology*, vol. 15, no. 1, pp. 12–22, Jul. 2024, https://doi.org/10.55981/j.mev.2024.742.

[25] D. Irmanto, A. Aripriharta, and S. Sujito, "Multi-Objective Optimization of Tactical Rescue Strategies Using QHBM and IoT in Urban Conflict Zones," *Journal of Wireless Mobile Networks, Ubiquitous Computing, and Dependable Applications*, vol. 16, no. 2, pp. 833–851, Jun. 2025, https://doi.org/10.58346/JOWUA.2025.I2.051.

[26] 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)*. Bangalore, India: IEEE, Jun. 2024, pp. 390–397, https://doi.org/10.1109/ICICI62254.2024.00070.

[27] 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, Jul. 2020, https://doi.org/10.3390/a13080186.

[28] 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, Dec. 2022, https://doi.org/10.1186/s43067-022-00054-1.

[29] B. M. Henrique and E. Santos, "Cryptocurrencies trading using Parrondo's Paradox," *International Review of Economics & Finance*, vol. 105, p. 104859, Jan. 2026, https://doi.org/10.1016/j.iref.2025.104859.

[30] F. A. Rahman, T. Widiyaningtyas, H. R. Adha, A. A. Smaragdina, L. A. Chisbiyah, and L. L. Pramita, "Comparative Performance Analysis of Decision Tree and K-Nearest Neighbors (KNN) Algorithms for Malformed Food Aroma Classification," in *2024 IEEE 2nd International Conference on Electrical Engineering, Computer and Information Technology (ICEECIT)*. Jember, Indonesia: IEEE, Nov. 2024, pp. 153–157, https://doi.org/10.1109/ICEECIT63698.2024.10859612.

[31] M. Shantal, Z. Othman, and A. A. Bakar, "A Novel Approach for Data Feature Weighting Using Correlation Coefficients and Min–Max Normalization," *Symmetry*, vol. 15, no. 12, p. 2185, Dec. 2023, https://doi.org/10.3390/sym15122185.

[32] J. Gao, C. Fan, T. Liu, X. Bai, W. Li, and H. Tan, "Embracing market dynamics in the post-COVID era: A data-driven analysis of investor sentiment and behavioral characteristics in stock index futures returns," *Omega*, vol. 131, p. 103193, Feb. 2025, https://doi.org/10.1016/j.omega.2024.103193.

[33] A. M. Hagag, L. S. Yousef, and T. F. Abdelmaguid, "Multi-Criteria Decision-Making for Machine Selection in Manufacturing and Construction: Recent Trends," *Mathematics*, vol. 11, no. 3, p. 631, Jan. 2023, https://doi.org/10.3390/math11030631.

[34] M. F. B. Hossain, L. Z. Lamia, M. M. Rahman, and M. M. Khan, "FinBERT-BiLSTM: A deep learning model for predicting volatile cryptocurrency market prices using market sentiment dynamics," *Applied Intelligence*, vol. 56, no. 3, p. 89, Feb. 2026, https://doi.org/10.1007/s10489-026-07086-4.

[35] S. Suradhaniwar, S. Kar, S. S. Durbha, and A. Jagarlapudi, "Time Series Forecasting of Univariate Agrometeorological Data: A Comparative Performance Evaluation via One-Step and Multi-Step Ahead Forecasting Strategies," *Sensors*, vol. 21, no. 7, p. 2430, Apr. 2021, https://doi.org/10.3390/s21072430.

[36] M. L. Hossain, S. M. N. Shams, and S. M. Ullah, "Time-series and deep learning approaches for renewable energy forecasting in Dhaka: A comparative study of ARIMA, SARIMA, and LSTM models," *Discover Sustainability*, vol. 6, no. 1, p. 775, Aug. 2025, https://doi.org/10.1007/s43621-025-01733-5.

[37] S. Hansun, F. P. Putri, A. Q. M. Khaliq, and H. Hugeng, "On searching the best mode for forex forecasting: Bidirectional long short-term memory default mode is not enough," *IAES International Journal of Artificial Intelligence (IJ-AI)*, vol. 11, no. 4, p. 1596, Dec. 2022, https://doi.org/10.11591/ijai.v11.i4.pp1596-1606.

[This page intentionally left blank.]