Dual Neural Paradigm: GRU-LSTM Hybrid for Precision Exchange Rate Predictions

Shamaila Butt
Faculty of Business, Sohar University, Sohar, Oman

***Abstract*—The USD/RMB exchange rate is significant when examining the structure of the Chinese financial system. Predicting the accurate USD/RMB exchange rate enables individuals to analyze the condition of the economy and prevent losses. We propose a novel hybrid approach of GRU-LSTM to improve the forecast of the future USD/RMB exchange rate. Deep learning techniques have become the cornerstone of numerous computer vision and natural language processing fields. This paper discusses various aspects and aims to show that they can help predict the exchange rate. We investigate how the newly developed hybrid GRU-LSTM model performs in terms of success rate and profitability compared with the LSTM and GRU models. The evaluation of the model is done on the USD/RMB currency pair and the forecasts made from September 13, 2023, to December 11, 2023. To increase the accuracy of the model, metrics like mean absolute error (MAE), mean square error (MSE), root mean square error (RMSE), and mean absolute relative error (MAPE) were introduced. The study found that the novel hybrid GRU-LSTM model was performing relatively well compared to the models of LSTM and GRU deployed in the survey for exchange rate prediction. This improvement can significantly benefit the analyst or trader in making the right decisions on the management of risks. The study further opens new possibilities for using the hybrid GRU-LSTM model by demonstrating the enhanced potential of this method, which can be more effective in the financial environment. Subsequent studies might improve the forecast by increasing the set of hybrid models and including more economic variables.**

***Keywords*—Prediction; LSTM; GRU; USD/RMB exchange rate; deep learning**

I. Introduction

The FOREX market is the largest market where the exchange of currencies takes place worldwide [1]. Trillions of dollars are exchanged by traders daily [2]. Since the currency prices are highly volatile, the FOREX market is referred to as a black box due to its intricacies and instabilities [3]. The authors in study [4] state that the exchange market is open 24/7. Nonetheless, the four primary time zones of Australia, Asia, Europe, and North America are used to verify the transactions. Each zone has its own hours of operation, and because it takes a lot of money to influence the exchange rate, scammers cannot easily manipulate the market [5]. Among other fields, foreign exchange market forecasting has attracted significant interest among researchers for several decades.

The international financial market is increasingly influenced by the exchange rate, a significant element affecting the global economy. Research has shown that whenever the volatility of the USD/RMB exchange rate exceeds a certain threshold, it has a negative impact on both national economic growth and the global economy. This underscores the importance of accurate USD/RMB exchange rate forecasting. The government has stated that rapid exchange rate changes intensify economic pressure, which high-precision exchange rate forecasting can help relieve. Such forecasting is not only essential for maintaining financial market stability but also aids in modifying the distribution of government resources, serving as a strong foundation for relevant administrative departments.

Moreover, deep learning has revitalized artificial neural network research. Deep neural networks (DNNs) have shown remarkable efficacy in various domains, such as computer vision and NLP. Enormous neural network optimization and control, the accessibility of extensive datasets, the computational capacity to train large networks, and approachable software libraries are associated with significant methodological advances [6]. Deep learning differs from traditional machine learning because it can independently extract discriminative features from raw data [7]. This capability reduces the need for human feature engineering while expanding the scope of deep learning applications. It also lowers the cost of deploying learning algorithms in the industry and facilitates model maintenance operations. Recurrent neural networks (RNNs) and convolutional neural networks (CNNs) are powerful deep learning techniques. RNNs are built to handle almost any type of time series, audio, and natural language data sequences.

Deep learning-based prediction models and their financial sector applications have been the subject of recent research [8], [9]. Nonetheless, not much research has been conducted on the forex market—research is scarce in this area for a few reasons. Determining the degree of accuracy with which market developments can be anticipated is a valuable academic and practical task. Moreover, exchange rates represent a highly stochastic, nonlinear, and non-stationary financial time series [10]. As such, they are challenging to anticipate and an exciting subject for forecasting studies [11]. Finally, since the foreign exchange market differs significantly from other financial markets, research studies on different financial markets, including the stock market, may not apply to the foreign exchange market.

Numerous studies have addressed the various features of the foreign exchange market. For instance, a lot of players in the foreign exchange market are trained by professional traders [12]. The foreign currency market has a higher percentage of short-term interdealer transactions than the stock market [13], [14]. Furthermore, the considerable fluctuations in exchange rates leave traders wanting to decide what to buy or sell. As a result, the fair value model needs to be more convincing to foreign exchange traders than to traders in the stock market [15].

Although there is a need for more research on advanced deep-learning algorithms for exchange rate prediction, this uniqueness makes it challenging to apply empirical results from other financial markets. Therefore, this study addresses critical gaps in exchange rate prediction by focusing on integrating microstructure variables such as bid-ask spread and order flow, which are often overlooked in traditional models. This approach enhances the accuracy of predictions and provides actionable insights for traders and policymakers. This study's primary objective is to add to the knowledge of exchange rate prediction by revising the deep learning algorithm for forex market dealers and reviewing the accuracy of the relevant models while predicting foreign currency movements. Given that this specificity is questioned by research on recent methods of in-depth learning to indicate the general state of other financial markets and the prediction of exchange rates, a concentrated study is encouraged. This thesis aims to retrofit the automatic learning procedure according to foreign exchange forecasts and review the accuracy of the relevant models while predicting foreign currency movements.

This study offers a novel method for predicting the USD/RMB exchange rate using microstructure variables, particularly high-frequency data at 1-minute intervals. Unlike previous studies that primarily rely on macroeconomic variables, this research pioneers the utilization of microstructure indicators—bid price, ask price, bid-ask spread, and order flow—which capture market inside information or private information inaccessible to the public. By incorporating these indicators, the model gains access to hidden or private information crucial for predicting exchange rate movements accurately. This novel approach forms the basis of research novelty and represents a significant departure from the conventional approach. To the best of the author's knowledge, no prior study had employed these microstructure and high-frequency indicators with a comparable level of granularity to predict exchange rates.

The novel approach and method employed align with the paper's goals. The primary objective of this study is to develop and validate a hybrid GRU-LSTM model that incorporates high-frequency microstructure indicators. To identify the exchange rate fluctuations, it is necessary to strengthen the internal structure of LSTM and integrate it with GRU. Then, the search conducts an empirical analysis of the given data concerning the USD/RMB exchange rate to extract helpful information like the bid price, ask price, bid-ask spread, and the flow of orders. These parameters act as dependent variables, extracting private or concealed information on the Forex market to determine exchange rates. The last objective of the study is to undertake exploratory research to evaluate the effectiveness of the proposed hybrid GRU-LSTM model against other distinct models, such as GRU and LSTM, drawn from the prediction models. All these tests will contain an analysis of the prediction errors and error margin for each of the given tests. Thus, based on the aims mentioned above, the study aims to propose a sophisticated and accurate forecasting model of the USD/RMB exchange rate. Additionally, we would like to extend our knowledge base and invest in the positive evolution of such sub-disciplines as financial modeling or other algorithmic trading.

The study's objectives are: (1) to apply microstructure indicators and optimize the internal LSTM architecture to create and calibrate the new GRU-LSTM model. Although this approach has limitations, its primary use is (2) to present selected USD/RMB exchange rate data characteristics, including bid price, ask price, bid-ask spread, and order flow. (3) to perform a research study to compare the ability of the hybrid GRU-LSTM model to the rest of the prediction models. As a result, it will be feasible to broadly assess the prediction model's accuracy and error margin. The research offers a vigorous and novel method for predicting the USD/RMB exchange rate. Additionally, the study contributes to advancements in algorithmic trading strategies and financial modeling.

The organization of the study is as follows: In Section II, relevant literature is compiled, and the paper's contribution to bridging research gaps is demonstrated. Section III discusses the empirical findings and outlines the experimental setup for the LSTM vs. GRU forecasting comparison. A summary and discussion of the results, policy implications, and opportunities for future research summarizes the paper.

II. Related Works

Many techniques have been used to forecast the Forex market in the past few years. Many such algorithms have been tried and experimented with, but the machine learning algorithm has a higher proportion. Depending on the specific study, some will use two or more methods, while others may only include one processing method. Thus, over the past few years, several ML models have been developed and used to predict the foreign currency market. Some of the approaches used in these divisions include regression analysis, decision trees, trading rule methodologies, fuzzy logic, and support vector machines. It was necessary to develop a hybrid model consisting of the cuckoo search algorithm and regression techniques for predictive modeling [16].

The authors in study [17] and [18] applied the USD/EUR currency pair, where they built the framework of a dataset derived from the autoregressive moving average (ARMA) model. The dataset has been instructed to use the following regression methods: SVR, PLS, CRT regression tree, and multiple linear regression. The four algorithms have developed weights, and the Cuckoo search algorithm uses it in its input data. The test used the pre-test data from two years to analyze. SVR, PLS, and CRT enhance the results obtained from MLR. Regression analysis has test results that are higher than those of other models, according to them. As in study [19] noted, a statistical and predictive analysis model has been developed to simplify autoregression in compressed vectors. They reduced a considerable amount of FOREX data using a random compression technique. Afterward, each random compressed data set was loaded into the Bayesian model averaging (BMA) method to find the intersecting parameters. The currency pairs show a significant mean squared error due to a condition including four lag-dependent variables and random compression of other Forex currency pairings. Their suggested approach has worked well for the following six currency pairs: AUD/JPY, CAD/CHF, EUR/DKK, CAD/JPY, EUR/MXN, and EUR/TRY. It has also outperformed the widely used Bayesian Autoregression benchmark. Previous studies such as study [20], [21] predicted foreign currencies used a similar methodology.

Many researchers have been working on creating prediction models based on trade regulations during the last few years.

The requirements for entry, exit, and currency management are outlined in the trading regulations. These rules are essential for judging if a deal will succeed or fail. The authors in [22] presented a model for exchange rate online prediction based on rules using the weighted majority (WM) approach for expert selection. Because the technology provides continuous estimates, they needed help maintaining a decent percentage of projections. As a solution, they have considered the recommendations made by websites and made plans based on them. The mean square error and realized profit figures were used to select the website. Data analysis showed that the intersection method offers a 30% higher accuracy in the 20-day prediction than the baseline. Some researchers have applied trade rules [23], [24]. Moreover, decision tree models have not been used as frequently as other methods. Authors in [25] have created a model that produces real-time FOREX market data. After that, these data would be converted into decision tables that must be bought, sold, or retained for specific attributes. Furthermore, the CART and C4.5 algorithms assessed the categorization's quality. Three device pairs and three files from each pair were used to create the system. Most data (86–98%) belong to the queue class, which makes data processing more difficult. Examining the decision tree's dimensions and the classification accuracy, it was found that the CART method produced the best results. Additionally, the researchers in study [26] use this method.

Furthermore, support vector machines, or SVMs, were another popular statistical prediction method. SVM with prediction capability was used for both individual and hybrid systems. An SVM-based model for foreign exchange prediction was presented in study [27]. They used the EUR/USD exchange rate to put their model into practice. They divided the results into positive and negative output categories using the cross-validation approach on their data set. To compare the outcomes, they used macro and micro averaging and both positive and negative accuracy rates. With the help of the Gaussian RBF approach, they got a difference in the training and test sets as high as 29.5%. Nevertheless, the difference was rather small with a polynomial model. Therefore, the kernel function derived using polynomial has given high performance. In addition, SVM raised the profit rate threefold when analyzing SVM transactions with a conventional strategy called the transaction model. Several scholars have also applied SVM in their research [28], [29].

In recent studies, scholars expressed their concern in natural language processing (NLP). Algorithms that are based on NLP have been used in foreign exchange, as in any other industry, for predicting the exchange rates. Besides, NLP has a high level of efficiency in both the prediction function and automatization of text-based functions [30], [31].

Numerous algorithms have attracted researchers' interest, Optimization Techniques stands as one of the most recognized algorithms. According to study [32], there has been a proposal of a model incorporating the following: the extreme learning machines along with the Jaya optimization to estimate device change rates. They employed USD/INR and USD/EUR as their two currencies in those analyses. To compare, they took their model against three models supported with NN, ELM, as well as FLANN. This led them to conclude that when it comes to optimization, ELM is superior to the other mentioned algorithms such as NN, FLANN. According to assessment data of ELM DE, the number of deficient errors affected rate of MAPE assessment. As observed above, the minimum value of MAE and maximum value of ARV and Theils U were obtained by employing ELM TLBO, ELM PSO, ELM Jaya. The researchers in study [33] employed a genetic algorithm to maximize the FOREX trading strategy and a variation-based ensembling method to produce a set of helpful trading rules. Their genetic algorithm creates rules by generating notably improved outcomes compared to the extensive search. The authors in study [34] suggested an additional hybrid learning approach. Their approach, linked to ELM-Jaya and ELM, uses a mix of technical indicators, statistical data, and both to forecast the prices of the currencies on a single exchange. Other researchers have also employed optimization strategies [35], [36].

It has been noted that chaos theory has received the interest of many researchers [37]. Several factors for exchange rate forecasting of the selected countries were incorporated, of which a novel method, Multivariate Adaptive Regression Splines (MARS), based on chaos theory, was employed apart from the above [35]. To assess their strategy, three primary FX pairs; JPY/USD, GBP/USD, and EUR/USD/KYM, utilized several types of the chaos-based forecasting model. When applying the Chaos + MARS method, they provided the most accurate forecasts of these currencies. Forecasting the global financial markets, [38] proposed a type-2 fuzzy neuro-oscillatory network with chaotic intervals CIT2-FNON. The chaotic discrete-time neural oscillator, or Lee-Oscillator, is the source of the CIT2-FNON. It is made up of short-lived fuzzy input neurons that are taken out of recurrent networks. To solve the complexity problem and create a highly Type-2 Fuzzy Logic System (T2FLS) with chaotic transient fuzzy qualities, the Chaotic Type-2 Transient Fuzzy Logic (CT2TFL) was added to their model. The FOREX price has been predicted by numerous models using chaos theory [39], [40].

Pattern-based models were also widely used and popular. The authors in study [41] created a multidimensional string model for statistical forecasting. By utilizing the D2-brane to create a 2-endpoint open string model, they enhanced the 1-endpoint open string model. They demonstrated how adding object attributes often changes the predictors' statistics by modeling several time series systems with them. They used demo simulations and four distinct currency pairs to evaluate the technology. They found that string model efficiency often increases with more extraordinary string lengths. The researchers in study [19] proposed a model that uses transformed models of general DMA and dynamic model averaging (DMA) to predict the FX rate. This method examined three forex pairs: USD/JPY, USD/EUR, and USD/GBP. Thirty percent of the entire data are used for data evaluation, and seventy percent are used to train the model. They found that the four-lag autoregression model (also known as AR (4)) and the time-varying autoregression model (also known as TVP-AR (4)) with four lags produce the best prediction result for USD/JYP based on the findings of their model proposal. The parsimonious model yields positive results for the EUR/USD data set. They predicted that the models that would perform best for this prediction would be those that use a stochastic process that evolves coefficients. Consistent DMA and DMS were found. Some researchers have used similar techniques

[42], [43].

An overview of foreign exchange rate prediction was provided by study [44]. They observed eighty-two hybrid systems that were used to predict the forex rate between 1998 and 2017. They found that hybrids with artificial neural network (ANN) foundations provide higher prediction rates with greater precision and stability. The analysis shows that hybrid models have overtaken single models. The hybrid models reduce uncertainty and offer more accuracy. After examining hybrid models based on artificial intelligence, they discovered that deep learning architecture was the least effective in predicting exchanges. Even though numerous neural network-based experiments have been conducted in the last few years [45], [46], [47]. In study [48] predicted the temporal sequence of the currencies using a hybrid C-RNN approach. A convolutional neural network (CNN) and a recurrent neural network (RNN) are combined to form a C-RNN. A data-driven strategy was used to alter the coin market's characteristics. The model was compared with CNN and LSTM methods. They discovered that the C-RNN model revealed fewer errors than the LSTM and CNN-based models using the RMSE performance evaluation approach.

The authors in study [49] use a neural network and genetic algorithm to anticipate the traded equities. This approach was used to assess the EUR/USD closing values. They contrasted their suggested model with various models, including NGD, NGW, MACD, and MA. They offered two distinct methods, one with a direction and the other with a weight. They discovered that their version Weighted, which showed a profit of 111%, outperformed version Direction, which showed a profit of 56%, after carrying out all 20 full experiences. To construct another hybrid model, [50] built another hybrid model by combining a computationally efficient functional link artificial neural network (CEFLANN) for prediction with an improved shuffling frog leaping (ISFL) model. The ISFL reduced the amount of error in the system. According to the study, three different currency pairs were employed in the method: In terms of pairs, such as USD/CAD, USD/CHF, and USD/JPY. The outcomes of the system performance were evaluated using the particle optimization method and the ISFL. The outcomes indicate that both the specialized and overall figures evince the effectiveness of the suggested model over the compared two algorithms. Analyzing the statistics of the USD/CHF currency pair, it was possible to establish that its error rate fluctuated within the range of 0. 03 to 0. 04, the USD/CAD and USD/JPY currency pairs error rates differing between 0. 04 and 0. 05.

Other researchers also took a similar approach [51], [52], [53]. The authors in study [54] examined three currency pairs to find the most accurate model for predicting exchange rates: USD/EUR, JPN/USD, and USD/GBP. An investigation of the performance of several ANN algorithms was also conducted in the study. The study used backpropagation to train the model to use neural network models after optimizing and preprocessing the raw input file. To project three distinct periods: quarterly, monthly, and daily. A multilayer perceptron with a 5-10-1 structure and a single one-step prediction mode was used for each currency pair in the study.

Similarly, other authors adopted the same strategy of analysis [51], [52], [53]. In [54] the authors examined three currency pairs to find the most accurate model for predicting exchange rates: These are USD EUR, JPN/USD, and USD/GBP. Performance analysis of several ANN algorithms was also carried out in the study as well. After optimizing and pre-processing the raw input file, the peculiarity of the study involved the use of backpropagation to train the model to use neural network models. To project three distinct periods: There are quarterly, monthly, and daily record keeping methods of releasing financial information to the public. A multilayer perceptron with 5-10-1 layers and single one-step prediction mode was used for all the currency pairs in the study.

A system that can forecast financial data and be used as an agent within the A-Trader system was proposed [55]. They examined the performance of deep learning methods and neural networks. They investigated the effectiveness of neural networks and deep learning methods. Four hidden layers, each containing 78, 64, 87, and 63 neurons, were used in the study. The B&H benchmark and the MLP agent were utilized for the performance evaluation. Their findings were split up over three distinct timeframes. Their suggested system did better for the combined primary and second periods. The authors in study [56] contrasted machine learning with statistical techniques. They investigated open, closed, high, and low variables. ASTAR produced superior results for close and high variables for one and one to five days of prediction, while GA-NN produced better results for open and low variables. The outcomes differed for predictions made over a more extended period (a month). In specific, for the remaining high and open conditions, GA-NN obtained better results than ASTAR; on the other hand, SVM yielded equal average values for both models. More recently, many more NN based systems have been realized [57], [58].

Thus, the literature analysis of the predictive models of the FOREX market indicates that future research should focus on more complex deep learning structures, mainly on the interaction between the GRU and LSTM. GRU and LSTM networks are still required to be actively used for the FOREX market prediction although prior studies were carried out with conventional machine learning algorithms such as regression, decision trees, and SVM. These new architectures of deep learning offer more specific advantages in analyzing difficult patterns and temporal correlations within FOREX data, leading to better accuracy of predictions. However, these models are relatively neglected despite their abilities to enhance the prediction precision due to integration of many predictors. Basically, interpretable models for the financial market are currently a dire necessity to come up with a better understanding of the existing principle behind predictions. Another way applicable for introducing the microstructure parameters, which include ask and bid prices, bid-ask spread, and order flow as independent variables, is an additional one, which also requires further investigation of the further increase of the models' accuracy and robustness.

This FOREX market research aims to close these gaps by focusing on analyzing and comparing the deep learning architectures like GRU and LSTM models and their combinations. These architectures possess different abilities that permit researchers to note such long-term relationships and sequential patterns in FOREX that can help them gain new perceptions and possess increased accuracy in forecasting. However, despite some limitations noted previously, adapting microstructure parameters for use in forecasting models is a plausible way of enhancing a model's realism while at the same time increasing its effectiveness in the face of real-time market data. Thus, our work can be seen to help narrow down gaps to foster better and more suitable prediction models for a FOREX market, thus aiding in the research on financial forecasting. While prior studies have explored LSTM and GRU models independently, few have systematically compared their performance against hybrid architectures. Our study bridges this gap by conducting a rigorous evaluation of the hybrid GRU-LSTM model against standalone models using multiple error metrics (MAE, MSE, RMSE, MAPE).

III. Methods

A. Long Short-Term Memory

In [59] researchers introduced the long-short-term memory (LSTM) model as a potential treatment to solve the gradient problem in models. LSTM has evolved into a novel neural network system that can manage sequential inputs during the last 20 years. Given that the widely used Python library Keras has the LSTM cell [60] and seems one of the most widely utilized LSTM designs in current research.

1) The Cell state: The cell state is a stream of data transmitted over time. The authors in [61] claim that the LSTM can memorize dependencies across time and bridge long-term delays by the cell state. The single LSTM cell contains all; however, the cell state pathway is grayed out, as depicted in Fig. 1.

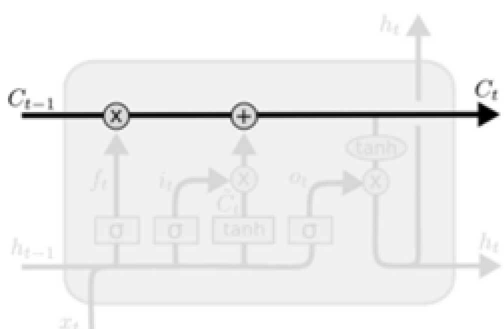

Fig. 1. The single LSTM cell contains all¸; however, the cell state pathway is grayed out.

2) Gate units: The LSTM cell is accessible because of its several gate structures. Two inputs are typically supported by an LSTM cell: the current input x_t and the recurrent input ($ht-1$, the hidden state of the previously executed time step). To read from and utilize the data contained in the cell state or to generate an updated cell state, C_t gate units regulate how these inputs change the cell state. The gating processes of the LSTM cell rely heavily on the logistic sigmoid function, which can be written as expressed as $\sigma(x) = \frac{1}{1+e^{-x}}$.

It maps the recurrent and weighted current inputs to the interval $[0, 1]$. This also clarifies the meaning of the term "gate," enabling the network to control the flow of information through the gates. Values of 0 and 1 can be interpreted as allowing all information to pass through a specific gate and preventing any information from doing so. In addition to the "gatekeeper" sigmoid, two LSTM gates use the hyperbolic tangent function, that is $\tanh(x) = \frac{e^x - e^{-x}}{e^x + e^{-x}}$. A sigmoid gatekeeper function is used to do this, as seen in Fig. 2.

The input and output gates in LSTMs usually use the tanh activation function, which pushes inputs into the interval $[-1, 1]$. The following are the derivatives of the logistic sigmoid and the hyperbolic tangent: $d(x)\sigma(x) = \sigma(x)(1 - \sigma(x))$ and d/d(x) tanh(x) = 1 - tanh2 (x). As a result, they may be used in network training, which comes after backpropagation.

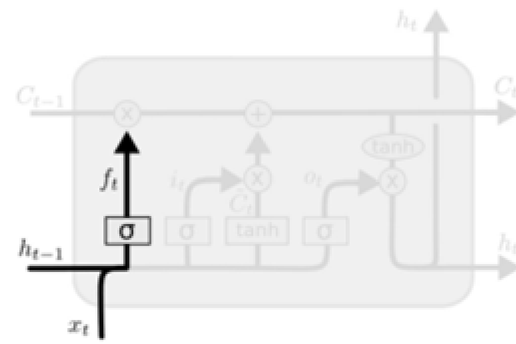

Fig. 2. Depicts the forget gate multiplied by the previous cell state $C_t - 1$ to ignore information [62].

3) The Forget gate: The portions of $C_t - 1$ that the cell state carried over from the previous time step are recognized and saved using the forget gate ft.

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

To further enable selective retention of information in memory, it is multiplied by $C_t - 1$. When $ft = 1$ or $ft = 0$, all the data from $C_t - 1$ is retained or deleted accordingly.

4) The Input gate: The input gate, indicated in Fig. 3, uses a sigmoid to regulate the flow of information:

$$i_t = \sigma(W_i[h_{t-1}, x_t] + b_i) \tag{2}$$

This gate's function prevents unnecessary updates from affecting the cell state data accumulated throughout earlier time steps. As a result, new information is selectively updated into the cell state by the input gate [61]. An activation function, typically a hyperbolic tangent, generates a new set of candidate values to accomplish this. $\tilde{C}_t$:

$$\tilde{C}_t = \tanh(W_C[h_{t-1}, x_t] + b_C) \tag{3}$$

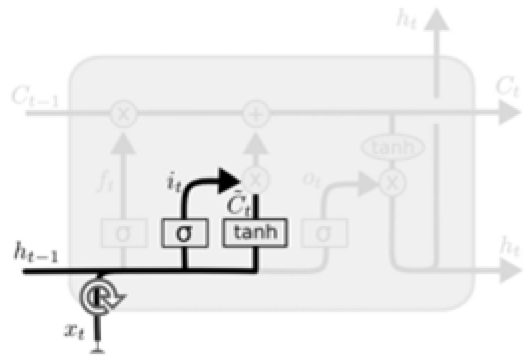

Fig. 3. The input gate 'it' directs where to update the cell state with new candidate values C_t [62]

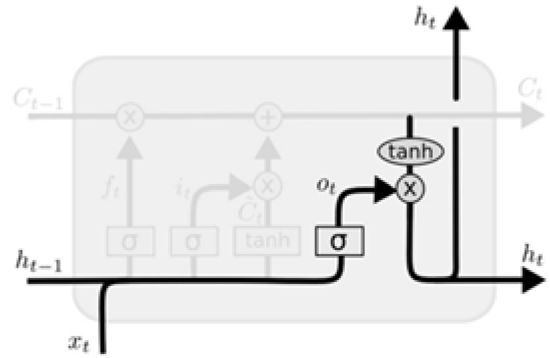

Fig. 4. The output gate 'ot' controls the network predictions [62].

5) The Updated cell state: The input produces the new cell state C_t and forgets gate mechanisms in two stages: first, it remembers (via f_t) a subset of the previous cell state $C_t - 1$ and updates (via i_t) with the new candidate values from $\tilde{C}_t$ when necessary. At the following time step, $t+1$, this changed cell state will be received:

$$C_t = f_t C_{t-1} + i_t \tilde{C}_t \tag{4}$$

Note that neither $i_t = f_t$ nor $i_t = 1 - f_t$ always holds. Not precisely the sections that were forgotten, nor the ones that were remembered, are updated. The forget gate and the input gate usually have separate weights and biases, even if they use the same arguments $(h_{t-1} and t_x)$ and an activation function of the sigmoid [62].

6) The Output gate: The actual prediction of LSTM depends on the current input (x_t) and cell state (C_t), controlled by the output gate. The current cell state data is subjected to a hyperbolic tangent, resulting in a scaled representation of the cell state within the interval $[-1, 1]$:

$$C_t^* = \tanh(C_t) \tag{5}$$

As seen in Fig. 4, the output gate (o_t) uses a sigmoid with the inputs h_{t-1} and x_t to choose which data to send to the output layer. The time steps in the newly hidden state (h_t) are then calculated.

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

The current time step's hidden output, ht, is then created by multiplying o_t and C_t^*.

$$h_t = o_t \circ C_t^* \tag{7}$$

This output is based on the prediction at time t and the recurrent input at time $t+1$. Predictions are calculated from the hidden state using an output activation in the last layer, just like in FNN.

7) The LSTM cell: Fig. 5 shows a typical LSTM cell with an input, output, and forget gate. Several gates and activations cooperate to save, hold, and produce information for the current task. When the gates and cell state are viewed as $h_t = o_t tanh(f_t C_(t-1) + i_t \tilde{C} t)$, ht can be thought of as a more complex activation function:

$$\begin{aligned} h_t &= \sigma(W_o[h_{t-1}, x_t] + b_o) \tanh\left(\sigma(W_f[h_{t-1}, x_t] + b_f) \cdot C_{t-1}\right. \\ &\quad \left. + \sigma(W_i[h_{t-1}, x_t] + b_i) \tanh(W_C[h_{t-1}, x_t] + b_C)\right) \\ &= g_h(W_h, h_{t-1}, x_t) \end{aligned} \tag{8}$$

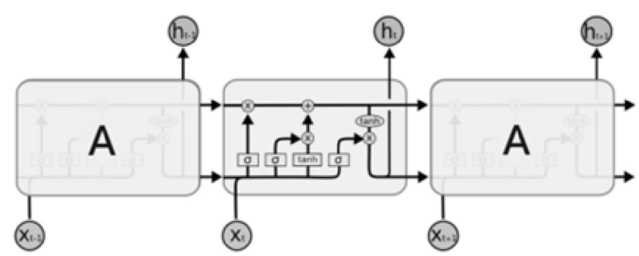

Fig. 5. A sequence of LSTM units through time [62].

This architecture, the most widely used configuration in the literature, is an improved version of the initial LSTM design. Fig. 6 shows how the cell state can transmit information over time. A series of LSTM cells are displayed throughout time. The hidden layer unit ("-" for closed and "O" for open) and the three gates to its left and above regulate which aspects of the cell state are updated, output, and forgotten at each time step.

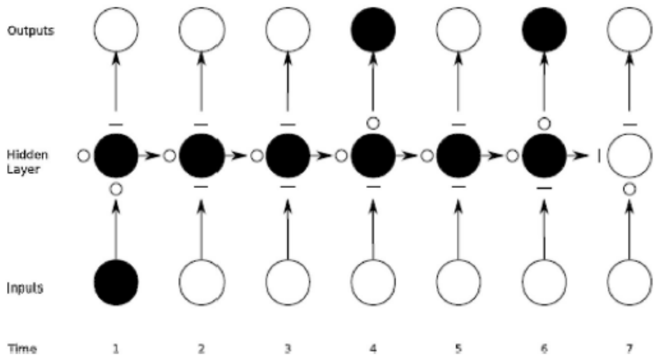

Fig. 6. Preservation of gradient information by LSTM.

B. Gated Recurrent Units

Gated recurrent units (GRUs) are an additional method for tackling the declining gradient problem in RNNs [63]. Even though they manage the cell state more straightforwardly, they still require gates. An update gate combined an LSTM forget and input functions, but two sigmoid gates managed a GRU's hidden state.

It determines the amount of recurring data that is retained:

$$z_t = \sigma(W_z[h_{t-1}, x_t] + b_z) \tag{9}$$

A reset gate regulates the degree of hidden recurring states that can be included in the present activation.

$$r_t = \sigma(W_r[h_{t-1}, x_t] + b_r) \tag{10}$$

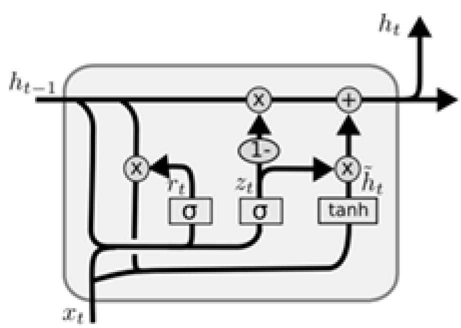

Fig. 7. Visualization of the recurrent hidden state, update gate, and reset gate over time.

Fig. 7 displays the recurrent hidden state (h_{t-1}), update gate (z_t), reset gate $r_t)$, new hidden state (h_t), and hidden state candidate vector (h_t) in addition to a detailed view of the GRU cell [62]. A memory cell with a closed reset gate $(r_t = 0)$ can act as if it were reading the first observation of a sequence, overlooking the recurrent state [64]. One possible way to compute the new activation is as follows.

$$\tilde{h}_t = \tanh(W_h[r_t \cdot h_{t-1}, x_t] + b_h) \tag{11}$$

and the new hidden state is

$$h_t = (1 - z_t) \cdot h_{t-1} + z_t \cdot \tilde{h}_t \tag{12}$$

Again, the GRU cell can be thought of as an advanced activation function:

$$\begin{aligned} h_t &= (1 - \sigma(W_z[h_{t-1}, x_t] + b_z)) \cdot h_{t-1} + \sigma(W_z[h_{t-1}, x_t] + b_z) \\ &\quad \tanh\left(W_h\left[\sigma(W_r[h_{t-1}, x_t] + b_r) \cdot h_{t-1}, x_t\right] + b_h\right. \\ &\qquad \left. \cdot C_{t-1} + \sigma(W_i[h_{t-1}, x_t] + b_i) \tanh(W_C[h_{t-1}, x_t] + b_C)\right) \\ &= g_h(W_h, h_{t-1}, x_t) \end{aligned} \tag{13}$$

Contrary to the LSTM, the GRU has no output activation function. Instead, the hidden cell state is constrained by the update gate, which links the input and forgets the gate. As shown in Fig. 7, GRUs contain fewer parameters than LSTMs, which should increase their computational efficiency. Regarding forecasting performance, prior research on GRUs against LSTMs is inconclusive [64], [65]. Therefore, this study focuses on both types of RNNs.

A hybrid GRU-LSTM model is shown in Fig. 8, where the input layer receives and processes the first set of data. The first hidden layer includes a pooling layer with a pool size of 1 and the same padding, a convolutional layer with 128 filters, a kernel size of 1, the ReLU activation function, and "same" padding to extract the essential features from the input. The 2nd Hidden Layer includes an LSTM layer with 128 units, depicting long-term dependencies in the data. The third hidden layer then uses the advantages of both architectures to learn complicated temporal patterns by switching between LSTM and GRU layers, each with 128 units. Finally, the output layer produces the model's predictions after processing the input data through the network's tiered structure.

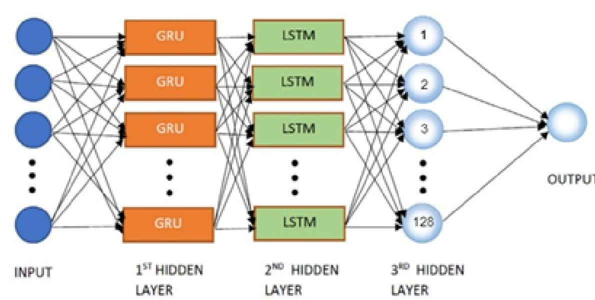

Fig. 8. Internal structure of hidden layers of hybrid model GRU-LSTM.

The flowchart shown in Fig. 9 presents the flow of the experiment, and it consists of generally arranged steps of establishing and validating a forecast model. Data collection is the first step followed by data preparation where pre-processing is completed to make the data more appropriate for training. During the pre-processing, the model is produced first, and the training data set is stored to it. This is the phase where the model for the forecast is being created; also, assessment of the model entails calculation of the loss function. If the ending criterion is not achieved, the parameters of the model are changed, and training continues in a cycle. There is creation of the model, and when the end condition is met, then there is saving of the model. When training of the model is complete, the model is checked, and an independent testing data set is used for prediction. This is succeeded by the assessment of the performance of the model when such forecast results are prepared and analyzed. This process is to establish an appropriate forecasting mode through the loops of train and test data.

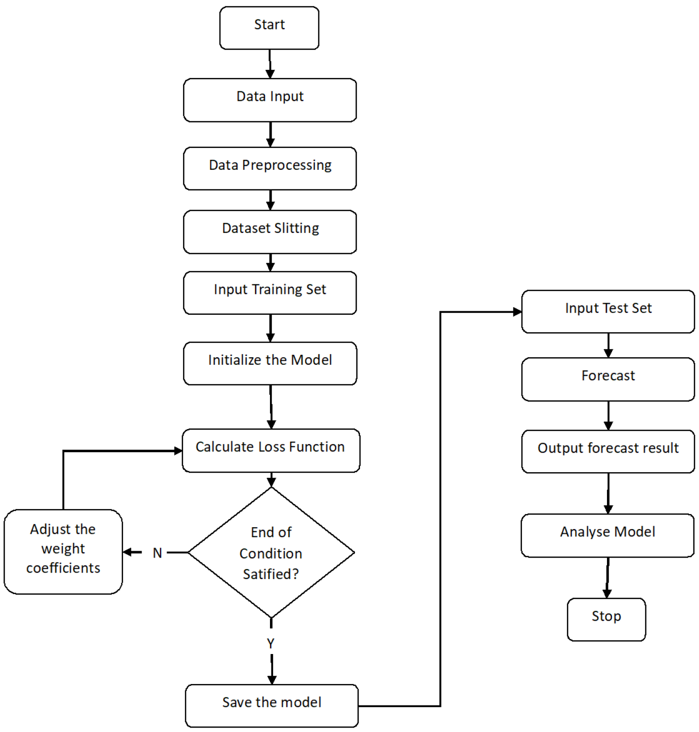

Fig. 9. Experimental process.

IV. EXPERIMENTS

This paper compares the hybrid model of the GRU-LSTM model with two individuals' models, the GRU and LSTM, to establish the efficiency of the hybrid approach. Training and the test data are identical in this case.

A. Experimental Environment

For the trials, Google Colab version with the option of the fastest GPU: NVIDIA Tesla P100, with 16 GB of GPU memory was used. The hardware architecture of the view MPS software involved a CPU that was an Intel Xeon with 25 GB of RAM. Environment software was Python 3. 8. 10, basic machine learning libraries for instance TensorFlow 2. 8. 0, PyTorch 1. 11. 0, and Scikit-learn 1. 0. 2, and data processing and visualization applications such as Pandas 1. 4. 2, NumPy 1. 21. 5, and Matplotlib 3. 5. 1.

B. Data Source

This paper presents a novel hybrid model for forecasting the USD/RMB exchange rate using the proposed GRU-LSTM model. The study uses the 1-minute USD/RMB exchange rate data set, which has 19807 observations. The parameters for this work are the bid price (Bid), ask price (Ask), order flow (OF), and bid-ask spread (BAS). The data was collected between September 13, 2023, and December 11, 2023.

In the forex market, the hidden daily information regarding macroeconomic fundamentals is communicated through the order flow [66]. The order flow values' sign might be both positive and negative. The sign denotes the buying and selling activity when a counterparty buys (+) at the dealer's offer or sells (-) at the dealer's bid. The Tick-test approach developed in study [67] and [68] methodology are the two main techniques used to speculate on the direction of currency transactions. Authors in study [68] compared the exchange rate and dealer quotes. Using this strategy, exchange rates higher or lower than the midpoint are classified as buy or sell. By comparing the current currency rate with the historical exchange rate, the tick-test examines fluctuations in exchange rates to determine the trade direction. A buy (uptick) is a rise in exchange rates or a transaction at a price more significant than the prior one. If not, a sell (down-tick) would be considered. At the same time, the transaction rates remain unchanged (zero-tick). According to [67], the transaction is classified according to the most recent difference between the current and exchange rates. Table I states the rules to differentiate between the buyers-initiated and the sellers-initiated trade.

TABLE I. IDENTIFICATION ALGORITHMS: TICK TEST

Specification	Conjecture for trade
$S_t > S_{t-1}$	Buyer-initiated
$S_t < S_{t-1}$	Seller-initiated
$S_t = S_{t-1}$	The conjecture for trade at t

Source: Adopted from [67]

The tick-test approach is used in this study because it is more accurate and adaptable [13], [12]. The order flow is calculated. Every trade is valued at $+1$ for purchases and -1 for sales. The daily trade is then the sum of all trade activities, whether buy or sell. Thus, researchers in [66] measure the daily order flow, defined as the buyer- and seller-initiated orders at the start of the working day. The related empirical studies used the same proxy for measuring the order flow [69], [70]. This research tracks tick-by-tick order flow one-minute data. The data is gathered from Bloomberg sources. The tick-test approach is used in this study because it is more accurate and adaptable [13], [12]. The order flow is calculated. Every trade is valued at $+1$ for purchases and -1 for sales. The daily trade is then the sum of all trade activities, whether buy or sell. Thus, [66] measures the daily order flow, defined as the buyer- and seller-initiated orders at the start of the working day. The related empirical studies used the same proxy for measuring the order flow [70], [69]. This research tracks tick-by-tick order flow one-minute data. The data is gathered from Bloomberg sources. The tick-test approach is used in this study because it is more accurate and adaptable [13], [12]. The order flow is calculated. Every trade is valued at $+1$ for purchases and -1 for sales. The daily trade is then the sum of all trade activities, whether buy or sell. Thus, [66] measures the daily order flow, defined as the buyer- and seller-initiated orders at the start of the working day. The related empirical studies used the same proxy for measuring the order flow [69], [70]. This research tracks tick-by-tick order flow one-minute data. The data is gathered from Bloomberg sources.

The ask price, bid price, and bid-ask spread are essential factors in predicting exchange rates. The bid price (buyer price) and ask price (Sell price) data are taken from Bank of China [71]. The realized bid-ask spread and the quoted bid-ask spread are the two ways to define the bid-ask spread [72]. The average difference between the purchase and sell prices the trader quotes when the buy and sale transactions occur at different times is known as the realized bid-ask spread. Alternatively, the difference between the purchase and sell prices that the dealers quote at trade time is known as the quoted bid-ask spread, and it is determined by the quantity transacted, stock price, and number of market makers (Chen, 2012). There is a strong correlation between exchange rate risk and the bid-ask spread, according to study [70] and [12]. The bid-ask spread data is computed as the difference between the ask (sell) price and the currency rate's bid (buy) price. Table II displays a subset of the data extracted from the Table I.

TABLE II. PARTIAL SAMPLE DATA

Date	ER	Bid	Ask	OF	BAS
2023-09-13 01:35:00	7.2835	7.2835	7.2835	0	0
2023-09-13 01:36:00	7.2820	7.2820	7.2820	-4	0
2023-09-13 01:38:00	7.2823	7.2823	7.2823	4	0
2023-09-13 01:39:00	7.2835	7.2835	7.2835	3	0
2023-09-13 01:40:00	7.2835	7.2835	7.2835	3	0

The complex nonlinear patterns of exchange rate fluctuations are too intricate to capture by traditional linear methods. Thus, researchers are adopting nonlinear methodologies that are intended to be more accurate [73], [74]. In times of economic turmoil and high volatility in the foreign currency market, when structural disruptions cause linear assumptions to be distorted, nonlinear models play a crucial role in predicting exchange rates [12].

The nonlinearities are efficiently managed by deep learning techniques such as GRU and LSTM. From September 13, 2023, to December 11, 2023, Fig. 10 illustrates the USD/RMB exchange rate for 1 minute dataset. It first displays an early appreciation of the RMB, followed by a sharp depreciation, stabilization, recovery, and a second depreciation phase. These oscillations underscore the necessity for nonlinear models to precisely forecast exchange rate movements and seize trading opportunities despite the extreme volatility and quick changes in the market. Economic data, policy changes, market sentiment, and geopolitical events impact these oscillations.

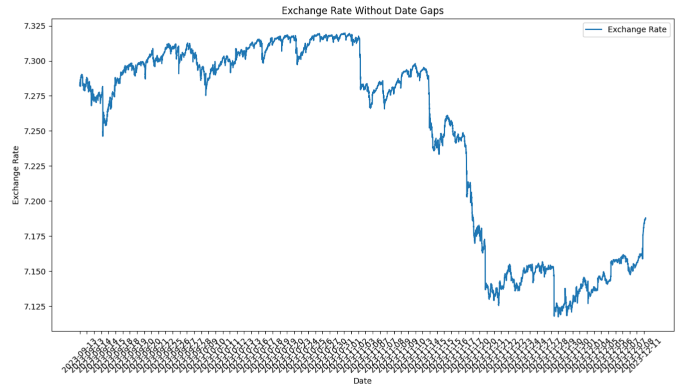

Fig. 10. Forex market dataset USD/RMB.

C. Descriptive Statistics

Table III shows a tabular representation of the descriptive statistics provided for the parameters such as RMB/USD exchange rate (ER), bid price (BID), ask price (ASK), order flow (OF), and bid-ask spread (BAS). This table summarizes the statistical measures for parameters based on the data from 19,808 observations. The mean values for bid and ask prices are very close, indicating a tight spread, confirmed by the mean BAS value being nearly zero. The standard deviation (std) across the bid, ask, and ER is similar, suggesting that the RMB/USD pair has been trading with relatively stable volatility. The order flow's large standard deviation points to significant buying and selling activity. The minimum and maximum values indicate the range of trade, while the distribution of the values is suggested by the 25th, 50th (median), and 75th percentiles.

TABLE III. DESCRIPTIVE STATISTICS

Statistic	Bid	Ask	ER	OF	BAS
count	19808	19808	19808	19808	19808
mean	7.24796	7.24783	7.24787	0.42392	-0.0001
std	0.0703	0.07035	0.07034	5.03654	0.00029
min	7.1175	7.1174	7.1175	-65	-0.0028
25%	7.1574	7.1574	7.1574	-3	0
50%	7.2839	7.2837	7.2838	2	0
75%	7.3045	7.3044	7.3044	3	0
max	7.3198	7.3198	7.3198	67	0

D. Data Preprocessing

The data preprocessing stages for exchange rate data involved several fundamental data transformations in preparing the data for a model. First, we converted the date column into a date-time format to assign it as the index and facilitate time series analysis. After that, the data was reset so that the date would appear as a column for further processing. To maintain chronological order, the dataset was sorted using dates to determine the ascending order.

Descriptive statistics were then produced to provide an overview of the dataset. "ER" was the target variable, and "Bid," "Ask," "OF," and "BAS" were the features chosen for the model. The features and the target variable were normalized using the MinMaxScaler to scale the values between 0 and 1 to 1 to improve the performance of models. Then, we established data sequences with a defined length of 4, meaning each sequence had four data points in a row. The target value for each sequence was the data point that came just after the sequence. The scaled data had to be iterated over to create sequences and their corresponding targets. The resulting sequences were then transformed into numpy arrays. Ultimately, the data was separated into training and testing sets. After training 80% of the data, the model was evaluated on 20% using an 80:20 split ratio. The model was trained on most of the data for this split, and its performance was assessed using an unknown component.

E. Experimental Parameter Settings

The critical parameter choices for a neural network model with convolutional and LSTM/GRU layers intended for sequential data analysis are shown in Table IV. Convolutional layers are set up with 128 filters, each using a ReLU activation function and a 1 x 1 kernel size to extract features while introducing nonlinearity effectively. Both convolutional and pooling layers utilize padding to maintain input/output dimensions. The LSTM/GRU layers are set to contain 128 hidden units, employing a Tanh activation function to capture temporal dependencies effectively. With a learning rate of 0.0001, the Adam optimizer improves the model, which guarantees adaptive changes to the weight parameters during training. Multiple loss functions, including MAE, MSE, RMSE, MSLE, Median Absolute Error, and MAPE, are employed to evaluate model performance from different perspectives. In any training, epochs are several complete cycles through the entire data set, and batch size is the number of training instances processed at each pass, therefore, training is performed over 50 epochs with a batch size of 54, which helps in optimizing the model for working on sequential data tasks. The purpose of such an elaborate set of parameters is to examine the sequential data, along with good performance and stability.

TABLE IV. MODEL PARAMETER SETTINGS

Parameters	Value
Convolution layer Filters	128
Convolution layer Kernel Size	1
Convolution layer Activation Function	ReLU
Convolution layer padding	Same
Pooling layer pool size	1
Pooling layer padding	Same
LSTM-GRU Hidden Units	128
LSTM-GRU Activation Function	Tanh
LSTM-GRU Optimizer	Adam
LSTM-GRU Learning Rate	0.0001
LSTM-GRU Loss Function	MAE, MAP, RMSE, MedAE, MAPE
LSTM-GRU Epochs	50
LSTM-GRU Batch Size	54

Note: The table shows the ideal parameters for each GRU-LSTM model layer. The acronyms MAE, MAP, RMSE, MedAE, and MAPE stand for Mean Absolute Error, Mean Average Precision, Root Mean Squared Error, Median Absolute Error, and Mean Absolute Percentage Error, respectively.

F. Experimental Results and Analysis

The primary objective of this study is to forecast the USD/RMB exchange rate with three different models' levels of accuracy: a hybrid GRU-LSTM model, a Gated Recurrent Unit (GRU), and a Long Short-Term Memory (LSTM) model. Various assessment criteria are used to assess these models, including training duration, MAE, MAP, RMSE, MAE, and MAPE. Before training, the dataset is standardized. Every model project the closing price for the subsequent trading day; the expected and actual values are then contrasted. A standardized Forex dataset is used in the studies for both evaluation and training. The LSTM, GRU, and hybrid GRU-LSTM models are trained on this dataset to forecast the closing values of the USD/RMB exchange rate. The model's parameters are adjusted to lower error metrics while the data is processed throughout several epochs during training.

Fig. 11 displays the performance of the LSTM model over ten epochs. The x-axis shows the epochs, while the y-axis displays the loss values. Included in the metrics are validation loss (Val-loss), training loss (Train-loss), mean absolute error (MAE), mean squared error (MSE), and root mean square error (RMSE). All metrics demonstrate a decreasing trend, indicating effective learning and improved predictions. The rapid initial decline in error values reflects the model's quick adaptation to data patterns. Despite minor fluctuations, the steady decrease in training loss and the general downward trend in validation loss suggest effective learning with some overfitting. The consistent decline in MSE, RMSE, and MAE confirms enhanced prediction accuracy.

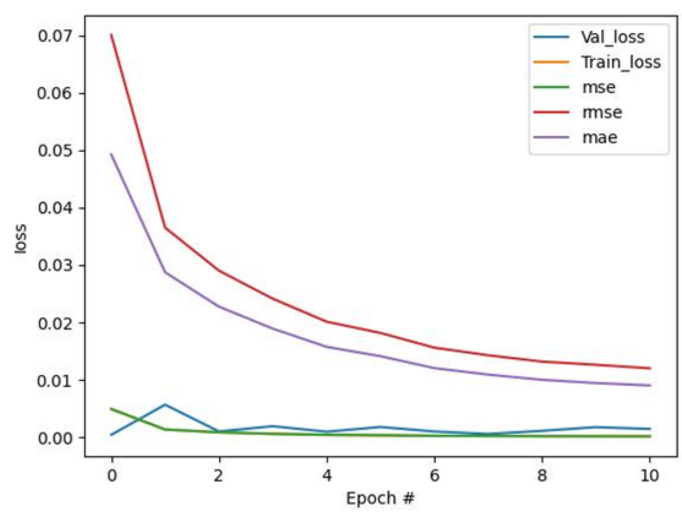

Fig. 11. LSTM Training and Validation loss for selected evaluation metrics.

Fig. 12 illustrates the performance of the GRU model over the same period. Like LSTM, the GRU model exhibits a rapid initial decline in error metrics, indicating efficient learning. The GRU model converges faster and displays stable validation performance with fewer fluctuations, indicating more consistent generalization. Both models effectively reduce errors and enhance predictive accuracy, with the GRU model demonstrating slightly faster and more stable convergence.

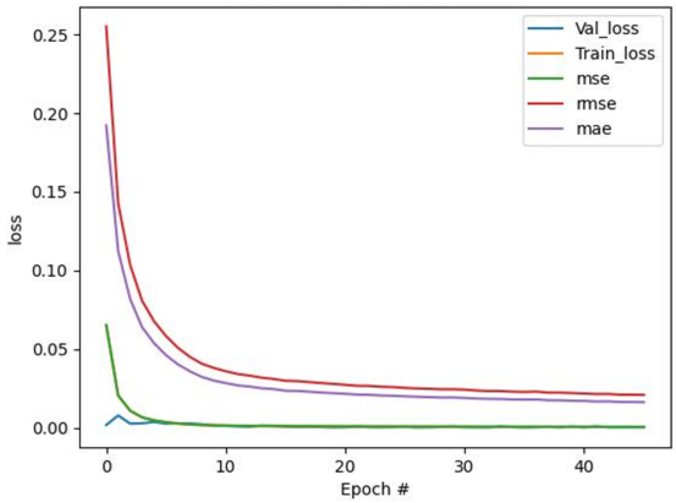

Fig. 12. Loss function for the training and validation for selected evaluation metrics.

The loss function for the combined LSTM and GRU model is shown in Fig. 13. To improve forecasting accuracy, the hybrid GRU-LSTM model combines the best features of the LSTM and GRU architectures. LSTM networks excel at learning long-term dependencies with their robust cell state and gate mechanisms, while GRUs offer computational efficiency and effective handling of short-term dependencies. In the hybrid model, data first passes through an LSTM layer to capture long-term patterns, then through a GRU layer to efficiently capture additional short-term patterns. This combination improves overall predictive performance, as demonstrated by the close alignment of the hybrid model's predictions with actual exchange rates in the provided plots. Since the hybrid model can account for both short-term volatility and long-term trends, it is a reliable method for forecasting the USD/RMB exchange rate.

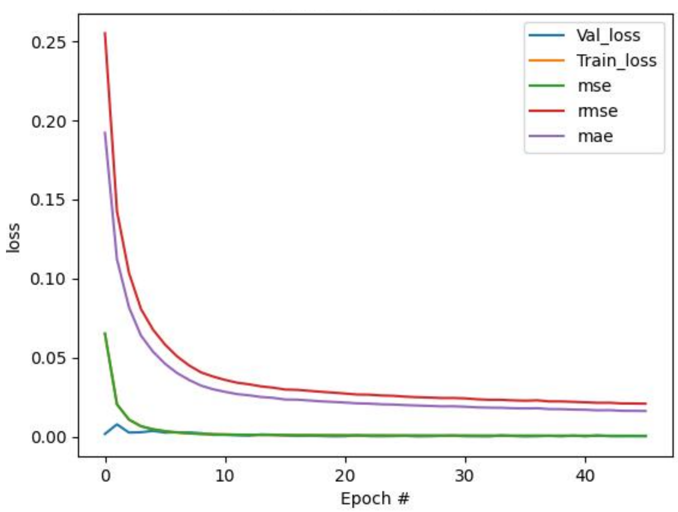

Fig. 13. The Loss function of the hybrid model for the training and validation of selected evaluation metrics.

The error metrics for the LSTM and GRU models trained on a Forex dataset. Both models exhibit a rapid initial decline in metrics such as training loss, validation loss, MSE, RMSE, and MAE, indicating efficient learning. The LSTM model shows steady improvement with minor fluctuations in validation loss, suggesting some variability in generalization. In contrast, the GRU model exhibits more consistent generalization due to its faster convergence and stable validation performance with fewer fluctuations. While both models can reduce errors and improve prediction accuracy, the GRU model shows quicker and more consistent convergence.

The prediction of the LSTM model is shown in Fig. 14, where it demonstrates an excellent predictive capability but is less consistent than the hybrid GRU-LSTM model, particularly during times of high volatility. Although some overfitting is seen because of small changes in validation loss, the model learns efficiently, as evidenced by the consistent decrease in training and validation loss. The model's predictions get more accurate with time, as evidenced by the steady drop in the MSE, RMSE, and MAE metrics.

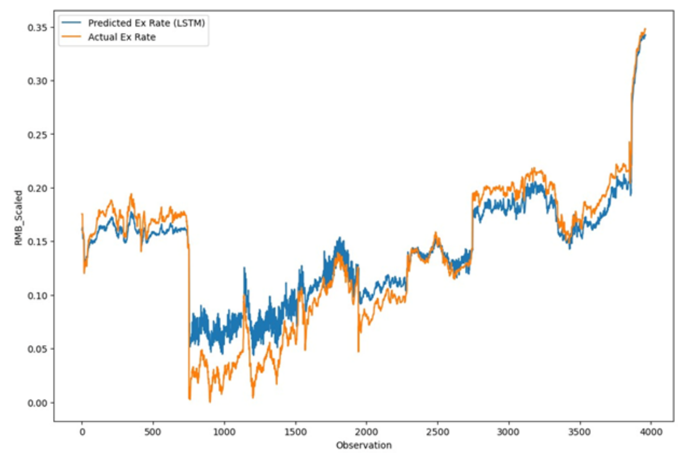

Fig. 14. LSTM Model's prediction for the test dataset.

The GRU model performs better than the LSTM model in terms of validation performance stability and convergence rate, as shown in Fig. 15. The validation loss fluctuating less suggests better generalization to unobserved data. Because of its strength, the GRU model is suitable for time series forecasting tasks where dependable performance is crucial.

Fig. 15. GRU Model prediction for the test dataset.

Fig. 16 demonstrates how a hybrid GRU-LSTM model better captures short—and long-term trends and fluctuations compared to the LSTM or GRU individual model. The hybrid model is a more accurate and reliable forecasting tool since it can use both architectures. This model closely aligns predicted values with actual exchange rates, demonstrating its robustness in handling financial time series data complexities. A high degree of accuracy is indicated by slight differences between expected and actual values; only occasional anomalies impact predictions.

The evaluation metrics of the GRU, LSTM, and hybrid GRU-LSTM models are displayed in Table V. The Mean Absolute Error (MAE) of 0.003368 indicates the average deviation from actual values for the Long Short-Term Memory (LSTM) model, created to represent long-term dependencies in sequence data. Its Root Mean Squared Error (RMSE) of 0.004109 indicates the impact of sporadic, more significant inconsistencies, even though its Mean Squared Error (MSE) of 0.000017 is relatively minor. The Mean Squared Logarithmic Error (MSLE) is zero, indicating robustness in handling logarithmic differences. The percentage accuracy is indicated by the Mean Absolute Percentage Error (MAPE) of 0.0472%, and resilience to outliers is indicated by the median absolute error of 0.002902.

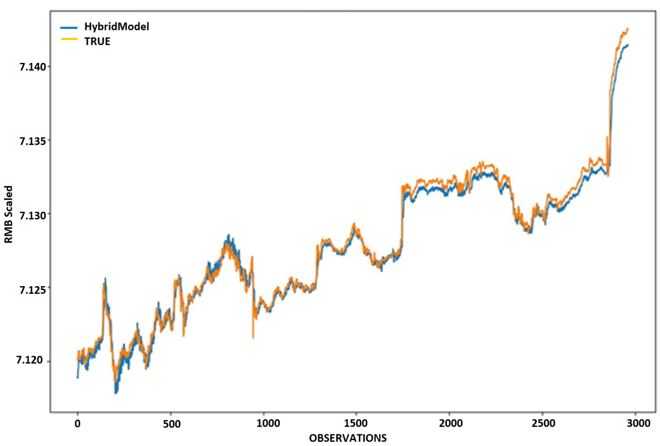

Fig. 16. Hybrid Model's prediction for the test dataset.

TABLE V. EVALUATION MATRICES

Model	MAE	RMSE	MedAE	MAPE
GRU-LSTM [75]	0.0012	0.0015	N/A	0.0010
LSTM [76]	0.0025	0.0032	N/A	0.0021
GRU [77]	0.0018	0.0023	N/A	0.0015
LSTM	0.0033	0.0041	0.0029	0.0004
GRU	0.0025	0.0037	0.0013	0.0004
Hybrid GRU-LSTM	0.0004	0.0005	0.0003	0.00006

The GRU model outperforms the LSTM in all criteria while maintaining similar functionality and simplifying the LSTM architecture. The MAE of 0.002505 indicates increased accuracy, with average predictions closer to actual values. A reduced error size is marked by an RMSE of 0.003746, highlighting improved overall prediction performance. A more consistent error distribution is indicated by the median absolute error of 0.001261, indicating that outliers impact the GRU less. The GRU model exhibits better accuracy than the LSTM model, with an average percentage error of 0.0351%, indicating a small average percentage error.

V. DISCUSSION

The findings of this study have significant implications for both practitioners and researchers. The hybrid GRU-LSTM model offers traders a reliable tool for predicting exchange rate movements, enabling them to manage risks more effectively. For researchers, the integration of microstructure variables highlights the importance of leveraging high-frequency data to uncover hidden patterns in financial markets. Moving forward, future studies could explore the application of hybrid models in other financial domains, such as stock price prediction or commodity trading.

The hybrid GRU-LSTM model outperforms the LSTM model on all evaluation measures by combining the best aspects of the GRU and LSTM architectures. The predictions of the hybrid model are surprisingly close to the actual values, with an MAE of only 0.000433. The model's outstanding performance is further supported by its RMSE of 0.000565, which shows the lowest error magnitude among the models. The lowest median absolute error, 0.000380, indicates that the model's predictions are less affected by anomalies and are, hence, stable. Finally, the hybrid model yields the most accurate and trustworthy forecasts, as evidenced by the MAPE of 0.0061%, the lowest average percentage deviation.

The comparison demonstrates that the GRU-LSTM hybrid model is the most accurate and consistent of the three models, surpassing the other two in each evaluated criterion. The GRU model outperforms the LSTM in all areas, demonstrating its accuracy and efficiency. Although the LSTM model performs well, the hybrid and GRU models outperform it. By utilizing the complementary advantages of both the LSTM and GRU architectures, the GRU-LSTM hybrid is the optimal option for attaining the highest prediction accuracy and consistency levels. This improved accuracy and dependability is essential when it comes to financial analysts and traders making well-informed decisions in the currency market.

The hybrid GRU-LSTM model not only improves predictive accuracy but also offers practical benefits for traders and financial analysts. By capturing both short-term volatility and long-term trends, the model enables more informed decision-making, particularly in high-stakes environments like the USD/RMB exchange market.

VI. CONCLUSION

This research aims to establish the reliability of the novel hybrid GRU-LSTM model to predict the USD/RMB exchange rate, which is extremely valuable in the Chinese financial sector. It will also help avoid possible financial risks associated with the exchange rate while providing critical economic information. Before developing a new hybrid GRU-LSTM network architecture, the goal is to combine the characteristics of the two GRU and LSTM models to increase the model's predictive capacity. Therefore, the study outlined the research's high efficiency based on the proposed novel hybrid model and the results of using separate LSTM and GRU models for the USD/RMB exchange rate prediction.

The findings demonstrated how the hybrid model, which combines LSTM and GRU components, can be distinguished from the precise prediction accuracy of the two models with a greater variety of parameter configurations. This is basically due to its cell state and gate mechanisms; LSTM has a mighty reliable cell state; GRUs, on the other hand, are efficient and effectively assert short-term dependency connections. When one works under the hybrid model, the actual data is well managed as it has short-term characteristics by going through a GRU layer once it has passed through an LSTM layer, which handles the features of long-term patterns.

The predictive plots presented demonstrate how this sort of combination enhances overall predictive accuracy by closely replicating the actual exchange rates among other techniques in the hybrid model. The hybrid model has the significant benefit of providing analysis over both the short and long terms; as a result, it may be used to forecast the USD/RMB exchange rate. This led to improved accuracy analysis, which is vital for traders and financial analysts before making any decisions in the foreign exchange market. Thus, although the presented LSTM-based model exhibits a high level of predictability, its capability becomes relatively volatile at extreme levels of volatility. In contrast, the GRU model has a similar training performance for different epochs with less fluctuation in the validation performance and takes fewer epochs to converge. Lastly, the hybrid model, which solely captures the short-term and extended-term changes, achieves the highest predicted accuracy given by the equation of the superiority of LSTM over GRU and vice versa.

A. Policy Implications

The study's findings have several policy implications.

- The newly proposed GRU-LSTM model provides numerous benefits to financial institutions and policy-making bodies in the projection of exchange rates by improving the model's expandability while maintaining the needed degrees of instability to account for long-term dependencies. This concept indicates that institutions can reduce currency risk by effectively employing hedging mechanisms and risk management structures that are less vulnerable. This helps avoid unfavorable currency movements, which result in declining performance on the economic portfolios, stabilizing the overall performance.
- In keeping the market stable, prompt and accurate forecasting models such as the hybrid model GRU-LSTM are vital as they highlight the exchange rate volatility. The market volatility is seen before it reaches extreme levels, which means that the financial institutions and the policymakers can respond as soon as the signs of volatility become apparent by using available instruments like the regulation of the monetary supply or using forex reserves. It acts as a shield that minimizes the disruption that comes about due to a sharp fluctuation in exchange rates, thereby minimizing shocks to financial markets, investors' confidence, and the sustainability of economic growth.
- The hybrid forecasting models' performances are highly relevant in determining policy decisions, especially in controlling foreign exchanges and monetary policies. The hybrid predictive model can be valuable for strategic planning of the economy and advantageous because it effectively facilitates anticipation of future market trends as influenced by exchange rate predictions. The projections by the predictive model facilitate the management of the monetary policies and foreign exchange reserves, thus helping counter-economic volatility and adopting stability as well as growth. Therefore, policymakers rely on predictive models when making decisions, reducing uncertainty and creating a favorable environment for the economy's long-run growth.
- The results also indicate that using deep learning improves the required models and introducing them into financial markets is the next step to achieving innovation. Financial organizations are in a diverse position to embrace innovation in their operations to enhance the viability and effectiveness of the financial services delivery systems. The models particularly excel in handling large volumes of data and more extensive and intricate patterns, thus enabling accurate prediction and decision-making. Hence, the use of advanced technology leads financial institutions to achieve a competitive market position and, at the same time, makes financial institutions more responsive to the prevailing conditions. Such flexibility enhances the financial system's effectiveness with a stressed and improved capacity to adapt to adversity and change and boosts the sector's resilience, efficiency, and innovation
- A microstructure model that integrates order flow, bid price, ask price, and the bid-ask spread is quite complex in Forex trading. However, when used, it can offer deep insights into the market that are not provided with traditional models. Essentially, through processing such data, the model predicts future variations of the RMB/USD exchange rate so that trading players can devise informed trading policies. Real-time data are easily interpreted to make it a source of trading signals that are key in responding quickly to market changes. For example, when the model indicates that the RMB is likely to appreciate against the USD, the trader may find it valuable to open a long position.
- The model also provides predictions and relative probabilities for preparing to manage risks. Trading professionals can employ these projections to place stop-loss orders and hence manage the capital that one is willing to risk per trade, not forgetting the profit targets. Probability forecasts allow traders to accurately reposition according to short-term fluctuations and long-term trends to effectively get the right timing for entering or exiting the market to earn the best profits.
- Several parameters affect placing an order, such as bid ask spread and order flow, which creates the opportunity to form positive order limits, making profit expectations high. The model outcome can also be carried out in the algorithms for actual trading since its

applicability is relevant, especially in high-frequency trading. The model's value is attributed mainly to the realization that you can obtain the probability outcomes from the model, which can then best be used in conjunction with sentiment analysis for the clients and the broker or dealers to determine the best time for trading. The traders that received this information have continued to be relevant through functional communication strategies such as swift notifications in the messaging and trading platform.

B. Future Recommendations

Future research could consider other hybrid models or the use of ensembles for better precision in forecasting exchange rates. Also, further development of the model with more elaborate deep learning techniques, such as attention techniques, and including external economic parameters, can be helpful.

Declaration of Competing Interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

REFERENCES

[1] Farid Irani, Abobaker Al. Al. Hadood, Korhan K Gokmenoulu, and Seyed Alireza Athari. Impact of financial market uncertainty and financial crises on dynamic stock—foreign exchange market correlations: A new perspective. *SAGE Open*, 15(1):21582440251314719, 2025.

[2] Kenneth Rogoff. *Our Dollar, Your Problem: An Insider's View of Seven Turbulent Decades of Global Finance, and the Road Ahead*. Yale University Press, 2025.

[3] Shamsul Arefeen, Md Shah Ali Dolon, Nigar Sultana, Mahmud Hasan, Mohammad Hasan Sarwer, Tui Rani Saha, Shariar Islam Saimon, and Intiser Islam. Comparative analysis of currency exchange and stock markets in brics using machine learning to forecast optimal trends for data-driven decision making. *Journal of Economics, Finance and Accounting Studies*, 7(1):26–48, 2025.

[4] Asif Khan, Muhammad Abid Hussain Shah Jillani, Maseeh Ullah, and Muneeb Khan. Regulatory strategies for combatting money laundering in the era of digital trade. *Journal of Money Laundering Control*, 28(2):408–423, 2025.

[5] Svyatoslav Yu Biryukov, Vadim N Perekrestov, Marina Yu Fadeeva, and Vladimir M Shinkaruk. Procedural and tactical aspects of investigating embezzlements committed on the forex currency market. In *Remote Investment Transactions in the Digital Age: Perception, Techniques, Law Regulation*, pages 355–365. Springer, 2024.

[6] Eleonora Bernasconi and Stefano Ferilli. Enhancing symbol recognition in library science via advanced technological solutions. *Information*, 16(2):119, 2025.

[7] Mohammad Zoynul Abedin, Mahmudul Hasan Moon, M Kabir Hassan, and Petr Hajek. Deep learning-based exchange rate prediction during the covid-19 pandemic. *Annals of Operations Research*, 345(2):1335–1386, 2025.

[8] Mohamed Elhoseny, Noura Metawa, Gabor Sztano, and Ibrahim M El-Hasnony. Deep learning-based model for financial distress prediction. *Annals of operations research*, 345(2):885–907, 2025.

[9] Alisa Kim, Yaodong Yang, Stefan Lessmann, Tiejun Ma, M-C Sung, and Johnnie EV Johnson. Can deep learning predict risky retail investors? a case study in financial risk behavior forecasting. *European Journal of Operational Research*, 283(1):217–234, 2020.

[10] Emre Urkmez. A comparative analysis of artificial neural networks and time series models in exchange rate forecasting. In *Machine Learning in Finance: Trends, Developments and Business Practices in the Financial Sector*, pages 71–85. Springer, 2025.

[11] Txus Blasco, J Salvador Sanchez, and Vicente Garcia. A survey on uncertainty quantification in deep learning for financial time series prediction. *Neurocomputing*, 576:127339, 2024.

[12] Shamaila Butt, Muhammad Ramzan, Wing-Keung Wong, Muhammad Ali Chohan, and Suresh Ramakrishnan. Unlocking the secrets of exchange rate determination in malaysia: A game-changing hybrid model. *Heliyon*, 9(8), 2023.

[13] Shamaila Butt. *Hybrid Model of Exchange rate Determinants in Malaysia*. PhD thesis, Universiti Teknologi Malaysia, 2020.

[14] Kimihiko Sasaki and Daisuke Yokouchi. An artificial market model for the forex market. *Humanities and Social Sciences Communications*, 12(1):1–16, 2025.

[15] Ahmet Umur Ozsoy and Omur Ugur. The qlbs model within the presence of feedback loops through the impacts of a large trader. *Computational Economics*, pages 1–28, 2025.

[16] Francesco Zito, Vincenzo Cutello, and Mario Pavone. Deep learning and metaheuristic for multivariate time-series forecasting. In *International Conference on Soft Computing Models in Industrial and Environmental Applications*, pages 249–258. Springer, 2023.

[17] Peng Yaohao and Pedro Henrique Melo Albuquerque. Nonlinear interactions and exchange rate prediction: Empirical evidence using support vector regression. *Applied Mathematical Finance*, 26(1):69–100, 2019.

[18] Bruno Miranda Henrique, Vinicius Amorim Sobreiro, and Herbert Kimura. Stock price prediction using support vector regression on daily and up to the minute prices. *The Journal of finance and data science*, 4(3):183–201, 2018.

[19] Paponpat Taveeapiradeecharoen and Nattapol Aunsri. Dynamic model averaging for daily forex prediction: A comparative study. In *2018 International conference on digital arts, media and technology (ICDAMT)*, pages 321–325. IEEE, 2018.

[20] Chanakya Serjam and Akito Sakurai. Analyzing predictive performance of linear models on high-frequency currency exchange rates. *Vietnam Journal of Computer Science*, 5:123–132, 2018.

[21] Shamaila Butt, Suresh Ramakrishnan, Muhammad Ali Chohan, and Suresh Kumar Punshi. Prediction of malaysian exchange rate using microstructure fundamental and commodities prices: A machine learning method. *International Journal of Recent Technology and Engineering (IJRTE)*, 8(2), 2019.

[22] Jia Zhu, Jing Yang, Jing Xiao, Changqin Huang, Gansen Zhao, and Yong Tang. Online prediction for forex with an optimized experts selection model. In *Web Technologies and Applications: 18th Asia-Pacific Web Conference, APWeb 2016, Suzhou, China, September 23-25, 2016. Proceedings, Part I*, pages 371–382. Springer, 2016.

[23] Sasika Roledene, Lakna Ariyathilaka, Nadun Liyanage, Prasad Lakmal, and Jeewanee Bamunusinghe. Genibux-event based intelligent forex trading strategy enhancer. In *2016 IEEE International Conference on Information and Automation for Sustainability (ICIAfS)*, pages 1–6. IEEE, 2016.

[24] Tuchsanai Ploysuwan and Roungsan Chaisricharoen. Gaussian process kernel crossover for automated forex trading system. In *2017 14th International Conference on Electrical Engineering/Electronics, Computer, Telecommunications and Information Technology (ECTI-CON)*, pages 802–805. IEEE, 2017.

[25] Juszczuk Przemyslaw, Kozak Jan, and Trynda 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*, pages 127–138. Springer, 2016.

[26] Dadabada Pradeepkumar and Vadlamani Ravi. Forex rate prediction using chaos and quantile regression random forest. In *2016 3rd International Conference on Recent Advances in Information Technology (RAIT)*, pages 517–522. IEEE, 2016.

[27] Thuy Nguyen Thi Thu and Vuong Dang Xuan. Using support vector machine in forex predicting. In *2018 IEEE International Conference on Innovative Research and Development (ICIRD)*, pages 1–5. IEEE, 2018.

[28] Tran Thai Hoa, Thanh Manh Le, and Cuong H Nguyen-Dinh. Hybrid model of 1d-cnn and lstm for forecasting ethereum closing prices: a case study of temporal analysis. *International Journal of Information Technology*, pages 1–13, 2025.

[29] Mustafa Onur Özorhan, İsmail Hakkı Toroslu, and Onur Tolga Şehitoğlu. A strength-biased prediction model for forecasting exchange rates using support vector machines and genetic algorithms. *Soft Computing*, 21:6653–6671, 2017.

[30] Adedoyin Tolulope Oyewole, Omotayo Bukola Adeoye, Wilhelmina Afua Addy, Chinwe Chinazo Okoye, Onyeka Chrisanctus Ofodile, and Chinonye Esther Ugochukwu. Automating financial reporting with natural language processing: A review and case analysis. *World Journal of Advanced Research and Reviews*, 21(3):575–589, 2024.

[31] Ruizhuo Gao, Zeqi Zhang, Zhenning Shi, Dan Xu, Weijuan Zhang, and Dewei Zhu. A review of natural language processing for financial technology. In *International Symposium on Artificial Intelligence and Robotics 2021*, volume 11884, pages 262–277. SPIE, 2021.

[32] Ankit Thakkar and Kinjal Chaudhari. Applicability of genetic algorithms for stock market prediction: A systematic survey of the last decade. *Computer Science Review*, 53:100652, 2024.

[33] Svitlana Galeshchuk and Sumitra Mukherjee. Forex trading strategy optimization. In *Decision Economics: In the Tradition of Herbert A. Simon's Heritage: Distributed Computing and Artificial Intelligence, 14th International Conference*, pages 69–76. Springer, 2018.

[34] Smruti Rekha Das, Debahuti Mishra, and Minakhi Rout. A hybridized elm using self-adaptive multi-population-based jaya algorithm for currency exchange prediction: an empirical assessment. *Neural Computing and Applications*, 31(11):7071–7094, 2019.

[35] Dadabada Pradeepkumar and Vadlamani Ravi. Forecasting financial time series volatility using particle swarm optimization trained quantile regression neural network. *Applied Soft Computing*, 58:35–52, 2017.

[36] Spyros K. Chandrinos and Nikos D. Lagaros. Construction of currency portfolios using an optimized investment strategy. *Operations Research Perspectives*, 5:32–44, 2018.

[37] Md Saiful Islam, Emam Hossain, Abdur Rahman, Mohammad Shahadat Hossain, and Karl Andersson. A review on recent advancements in forex currency prediction. *Algorithms*, 13(8):186, 2020.

[38] Raymond ST Lee. Chaotic interval type-2 fuzzy neuro-oscillatory network (cit2-fnon) for worldwide 129 financial products prediction. *International Journal of Fuzzy Systems*, 21(7):2223–2244, 2019.

[39] Vadlamani Ravi, Dadabada Pradeepkumar, and Kalyanmoy Deb. Financial time series prediction using hybrids of chaos theory, multi-layer perceptron and multi-objective evolutionary algorithms. *Swarm and Evolutionary Computation*, 36:136–149, 2017.

[40] Raymond ST Lee. Cosmos trader-chaotic neuro-oscillatory multiagent financial prediction and trading system. *The Journal of Finance and Data Science*, 5(2):61–82, 2019.

[41] Erik Bartos and Richard Pincak. Identification of market trends with string and d2-brane maps. *Physica A: Statistical Mechanics and its Applications*, 479:57–70, 2017.

[42] Antonio V Contreras, Antonio Llanes, Alberto Pérez-Bernabeu, Sergio Navarro, Horacio Pérez-Sánchez, Jose J López-Espín, and José M Cecilia. Enmx: An elastic network model to predict the forex market evolution. *Simulation Modelling Practice and Theory*, 86:1–10, 2018.

[43] João Carapuço, Rui Neves, and Nuno Horta. Reinforcement learning applied to forex trading. *Applied Soft Computing*, 73:783–794, 2018.

[44] Dadabada Pradeepkumar and Vadlamani Ravi. Soft computing hybrids for forex rate prediction: A comprehensive review. *Computers & Operations Research*, 99:262–284, 2018.

[45] Jan Rindell. *Exploring the utilization of neural networks in Forex market forecasting: a comparative analysis of nonlinear autoregressive neural network and autoregressive integrated moving average*. Bachelor thesis, LUT University, 2024.

[46] Mohammad Zoynul Abedin, Mahmudul Hasan Moon, M Kabir Hassan, and Petr Hajek. Deep learning-based exchange rate prediction during the covid-19 pandemic. *Annals of Operations Research*, 345(2):1335–1386, 2025.

[47] Rajashree Dash. Performance analysis of an evolutionary recurrent legendre polynomial neural network in application to forex prediction. *Journal of King Saud University-Computer and Information Sciences*, 32(9):1000–1011, 2020.

[48] Lina Ni, Yujie Li, Xiao Wang, Jinquan Zhang, Jiguo Yu, and Chengming Qi. Forecasting of forex time series data based on deep learning. *Procedia computer science*, 147:647–652, 2019.

[49] Jacek Mańdziuk and Piotr Rajkiewicz. Neuro-evolutionary system for forex trading. In *2016 IEEE Congress on Evolutionary Computation (CEC)*, pages 4654–4661. IEEE, 2016.

[50] Rajashree Dash. An improved shuffled frog leaping algorithm based evolutionary framework for currency exchange rate prediction. *Physica A: Statistical Mechanics and its Applications*, 486:782–796, 2017.

[51] Yoke Leng Yong, Yunli Lee, and David Ngo. An investigation into the recurring patterns of forex time series data. In *2015 IEEE International Symposium on Robotics and Intelligent Sensors (IRIS)*, pages 313–317. IEEE, 2015.

[52] Smruti Rekha Das, Debahuti Mishra, and Minakhi Rout. A hybridized elm-jaya forecasting model for currency exchange prediction. *Journal of King Saud University-Computer and Information Sciences*, 32(3):345–366, 2020.

[53] Piotr Czekalski, Michal Niezabitowski, and Rafal Styblinski. Ann for forex forecasting and trading. In *2015 20th International Conference on Control Systems and Computer Science*, pages 322–328. IEEE, 2015.

[54] Svitlana Galeshchuk. Neural networks performance in exchange rate prediction. *Neurocomputing*, 172:446–452, 2016.

[55] Jerzy Korczak and Marcin Hemes. Deep learning for financial time series forecasting in a trader system. In *2017 Federated Conference on Computer Science and Information Systems (FedCSIS)*, pages 905–912. IEEE, 2017.

[56] Sitti Wetenriajeng Sidehabi, Sofyan Tandungan, et al. Statistical and machine learning approach in forex prediction based on empirical data. In *2016 International Conference on Computational Intelligence and Cybernetics*, pages 63–68. IEEE, 2016.

[57] Anastasios Petropoulos, Sotirios P Chatzis, Vasilis Siakoulis, and Nikos Vlachogiannakis. A stacked generalization system for automated forex portfolio trading. *Expert systems with applications*, 90:290–302, 2017.

[58] Yoke Leng Yong, Yunli Lee, Xiaowei Gu, Plamen P Angelov, David Chek Ling Ngo, and Elnaz Shafipour. Foreign currency exchange rate prediction using neuro-fuzzy systems. *Procedia computer science*, 144:232–238, 2018.

[59] Klaus Greff, Rupesh Srivastava, Jan Koutník, Bas Steunebrink, and Jürgen Schmidhuber. Lstm: A search space odyssey. *IEEE transactions on neural networks and learning systems*, 28, 03 2015.

[60] François Chollet et al. Keras: The python deep learning library. *Astrophysics source code library*, pages ascl–1806, 2018.

[61] Sepp Hochreiter and Jürgen Schmidhuber. Long short-term memory. *Neural computation*, 9(8):1735–1780, 1997.

[62] Christopher Olah. Understanding lstm networks, 2015. Online; accessed 2024-07-02.

[63] Kyunghyun Cho, Bart Van Merriënboer, Caglar Gulcehre, Dzmitry Bahdanau, Fethi Bougares, Holger Schwenk, and Yoshua Bengio. Learning phrase representations using rnn encoder-decoder for statistical machine translation. *arXiv preprint arXiv:1406.1078*, 2014.

[64] Junyoung Chung, Caglar Gulcehre, KyungHyun Cho, and Yoshua Bengio. Empirical evaluation of gated recurrent neural networks on sequence modeling. *arXiv preprint arXiv:1412.3555*, 2014.

[65] Rafal Jozefowicz, Wojciech Zaremba, and Ilya Sutskever. An empirical exploration of recurrent network architectures. In *International conference on machine learning*, pages 2342–2350. PMLR, 2015.

[66] Martin DD Evans and Richard K Lyons. Order flow and exchange rate dynamics. *Journal of political economy*, 110(1):170–180, 2002.

[67] Richard W Sias and Laura T Starks. Institutions and individuals at the turn-of-the-year. *The Journal of Finance*, 52(4):1543–1562, 1997.

[68] Charles MC Lee and Mark J Ready. Inferring trade direction from intraday data. *The Journal of Finance*, 46(2):733–746, 1991.

[69] Abolaji Daniel Anifowose, Izlin Ismail, and Mohd Edil Abd Sukor. Currency order flow and exchange rate determination: Empirical evidence from the malaysian foreign exchange market. *Global Business Review*, 19(4):902–920, 2018.

[70] Lorna Katusiime, Abul Shamsuddin, and Frank W Agbola. Macroeconomic and market microstructure modelling of ugandan exchange rate. *Economic Modelling*, 45:175–186, 2015.

[71] Bank of China. The people's bank of china. http://pbc.gov.cn, 2023.

[72] Hans R Stoll. Inferring the components of the bid-ask spread: Theory and empirical tests. *the Journal of Finance*, 44(1):115–134, 1989.

[73] Sanyam Jain, Aswani Kumar Cherukuri, and Amit Kumar Tyagi. 13 analysis of forex forecasting using machine learning and deep learning techniques. *AI and Blockchain in Smart Grids: Fundamentals, Methods, and Applications*, page 223, 2025.

[74] G Pandeeswari. Progressive time series analysis: Cnn-lstm models for stock market trend prediction. In *AIP Conference Proceedings*, volume 3204. AIP Publishing, 2025.

[75] Md Saiful Islam and Emam Hossain. Foreign exchange currency rate prediction using a gru-lstm hybrid network. *Soft Computing Letters*, 3:100009, 2021.

[76] Yu Chen, Ruixin Fang, Ting Liang, Zongyu Sha, Shicheng Li, Yugen Yi, Wei Zhou, and Huilin Song. Stock price forecast based on cnn-bilstm-eca model. *Scientific Programming*, 2021(1):2446543, 2021.

[77] Abba Suganda Girsang and Stanley Hudson. Cryptocurrency price prediction based social network sentiment analysis using lstm-gru and finbert. *IEEE Access*, 2023.