Proceedings of the 2023 International Conference on Machine Learning and Automation
DOI: 10.54254/2755-2721/32/20230177

Exchange rate prediction research based on LSTM-ELM hybrid model

Xintong Cao

Boda College, Jilin Normal University, 1999 East Feng Road, Tiexi District, Siping City

m18844160940@163.com

Abstract. The fluctuation of exchange rates holds paramount importance for a country's economic and trade activities. Due to the non-stationary and nonlinear structural characteristics of exchange rate time series, accurately predicting exchange rate movements is a challenging task. Single machine learning models often exhibit lower precision in exchange rate prediction compared to combined machine learning models. Hence, employing a combined model approach aims to enhance the predictive performance of exchange rate models. Both Long Short-Term Memory (LSTM) and Extreme Learning Machine (ELM) exhibit intricate structures, making their direct integration challenging. To address this issue, an innovative weighted approach is adopted in this study, combining LSTM and ELM models and further refining the combination weights using an improved Marine Predators Algorithm. This paper encompasses both univariate and multivariate prediction scenarios, employing two distinct allocation strategies for training and testing datasets. This is done to investigate the influence of different dataset allocations on exchange rate prediction. Finally, the proposed LSTM-ELM weighted combination exchange rate prediction model is compared with SVM, Random Forest, ELM, LSTM, and LSTM-ELM average combination models. Experimental results demonstrate that the LSTM-ELM weighted combination exchange rate prediction model outperforms the others in both univariate and multivariate prediction settings, yielding higher predictive accuracy and superior fitting performance. Consequently, the LSTM-ELM weighted combination prediction model proves to be effective in exchange rate forecasting.

Keywords: Exchange rate prediction; Long Short-Term Memory neural network; Extreme Learning Machine

1. Introduction

In recent years, China has continuously pushed forward with the reform of its exchange rate marketization. As the status of the Renminbi (RMB) has risen in the international market, its exchange rate fluctuations have become more pronounced than before. These fluctuations not only affect investors' investment decisions but also have significant implications for enterprises' cross-border investments, arbitrage hedging, risk management, and other crucial determinations. Furthermore, they are factors that demand particular consideration when the government formulates economic policies and manages exchange rate risks. Particularly, the escalation of China-US trade tensions since 2018 and the outbreak of the COVID-19 pandemic in 2019, followed by its global spread, have further intensified the risks in

DOI: 10.54254/2755-2721/32/20230177

the RMB foreign exchange market [1]. As a result, research related to the prediction of exchange rate fluctuations has garnered extensive attention from various sectors.

Currently, scholars have conducted substantial research on the RMB exchange rate, and the methods for predicting RMB exchange rates are continuously being updated and optimized. Due to the influence of numerous intricate factors on exchange rates, predicting them remains a challenging issue. An analysis of existing literature reveals that exchange rate predictions often focus on research related to the driving forces of economic fundamentals [2], as well as technical studies based on the temporal characteristics of exchange rates themselves [3][4][5][6].

2. Theoretical Foundations

2.1. LSTM Network

The Long Short-Term Memory (LSTM) network is a specialized type of recurrent neural network that relies on three "gates" to selectively process input information. The structure of a single LSTM neuron is depicted in Figure 2.1.

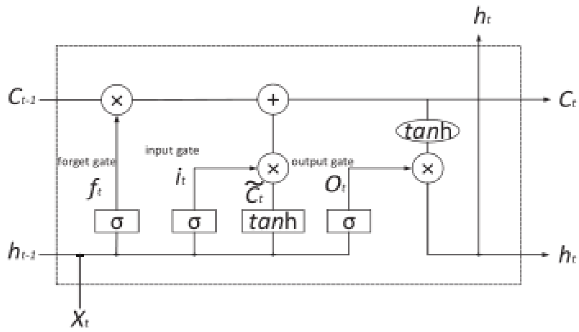

Figure 2.1. the structure of LSTM

In this figure, X_t represents input data entering the LSTM unit from the external environment, h_t denotes the output of this LSTM unit, C_{t-1} signifies the state of the previous LSTM unit at the preceding time step, h_{t-1} represents the output of the previous LSTM unit, and i_t, O_t, f_t denote the input gate, output gate, and forget gate respectively. The LSTM unit computes the current state and output based on these input data.

The specific calculation formulas are as follows:

$$i_t = s(W_{iX}X_t + W_{iM}M_{t-1} + W_{ic}C_{t\text{-}1} + b_i), \tag{2.1}$$

$$f_t = s\left(W_{fX}X_t + W_{fM}M_{t-1} + W_{fC}C_{t-1} + b_f\right), \tag{2.2}$$

$$C_t = f \otimes C_{t-1} + i_t \otimes g(W_{CX}X_t + W_{CM}M_{t-1} + b_C), \tag{2.3}$$

$$O_t = \sigma(W_{OX}X_t + W_{OM}M_{t-1} + W_{OC}C_t + b_O), \tag{2.4}$$

$$M_t = O_t \otimes h(C_t), \tag{2.5}$$

$$h_t = W_{yM}M_t + b_y. \tag{2.6}$$

Where σ is a $sigmoid$ function, $W_{iX}, W_{iM}, W_{iC}, W_{fX}, W_{fM}, W_{fC}, W_{CX}, W_{CM}, W_{OX}, W_{OM}, W_{OC}, W_{yM}$ are the weight coefficients for the forget gate, and b_i, b_f, b_C, b_O, b_y are biases term in the calculations.

2.2. Extreme Learning Machine (ELM)

Extreme Learning Machine (ELM) is a novel neural network with unique characteristics and excellent performance. It generates all hidden layer parameters randomly and balances recognition accuracy with algorithm extensibility. It has found widespread applications in various research fields. Figure 2.2 illustrates the structure of ELM.

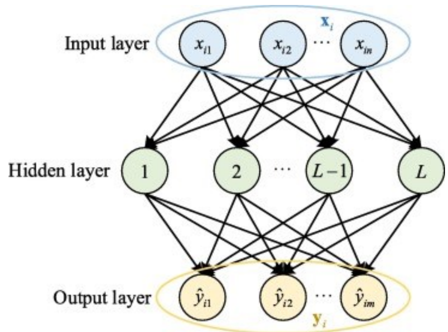

Figure 2.2. the structure of ELM

Considering the sample matrix $\{x_i, t_i\}$, where $i = 1, \cdots, N$ and N is the number of samples, $x_i = (x_{i1}.x_{i2}, \cdots, x_{in})^T \in {R_n}^m$ is the network input vector, $t_i = (t_{i1}.t_{i2}, \cdots, t_{im})^T \in R^m$ is the network output vector, $n, \overset{\Lambda}{N}, m$ are the dimensions of the input layer, hidden layer, and output layer respectively, and the activation function $g(x)$ is typically a $Sigmoid$ type. Then, the mathematical expression of ELM is given by:

$$o_j = \sum_{i=1}^{\overset{\Lambda}{N}} \beta_i g(w_i \cdot x_j + b_i), j = 1, \cdots, N. \tag{2.7}$$

Where β_i represents the connection weights between the ith hidden layer node and the output layer, w_i denotes the connection weights between the input layer and the i -th hidden layer node, and b_i is the bias of the i -th hidden layer node.

The loss function [7] of ELM is as follows:

$$E = \sum_{j=1}^{L}(\varepsilon_j)\text{，} \varepsilon_j = \sum_{j=1}^{\overset{\Lambda}{N}} \beta_i g(w_i \cdot x_j + b_i) - o_j. \tag{2.8}$$

Where $\varepsilon_j = [\varepsilon_{j1}, \varepsilon_{j2}, \cdots, \varepsilon_{jm}]$ is the error for the j -th sample. Achieving zero error approximation to t_i leads to the ideal expectation: $\sum_{i=1}^{N}\|o_i - t_i\| = 0$, which means that there exists β_i, w_i, and b_i that makes $\sum_{j=1}^{\overset{\Lambda}{N}} \beta_i g(w_i \cdot x_j + b_i) = t_i$.

2.3. *Marine Predators Algorithm (MPA) and Optimization*

The Marine Predators Algorithm (MPA) is a new type of intelligent optimization algorithm proposed by Faramarzi et al. [8]. In MPA, each predator acts as a searching individual, and its position represents a candidate solution. Predators update their positions using predation and individual dispersion operators to ultimately obtain prey (optimal solutions). Compared to existing intelligent optimization algorithms, MPA possesses a unique search mechanism and demonstrates significant advantages in solving various classical optimization problems. However, during the optimization process involving alternating $Brownian$ and $Lévy$ motions, large step lengths may lead to the intersection of optimal solutions. To address this, an adaptive parameter controlling step length, originally expressed as Equation (2.9):

$$CF = (1 - \frac{t}{t_{\max}})^{(2\frac{t}{t_{\max}})} \tag{2.9}$$

is replaced with Equation (2.10):

$$CF = \frac{(1+\cos(\frac{\pi \times t}{t_{\max}}))}{2} \tag{2.10}$$

Furthermore, ideas are presented for addressing the issues of a limited initial population and bypassing local optima, as well as providing extensive exploration of the search space by introducing the Opposite-Based Learning strategy (OBL) [9]. OBL mitigates the shortcomings of a random population and enhances the convergence of the Marine Predators Algorithm. For OBL, assuming $Opp = (X_{\min} + X_{\max}) - X$ is the inverse function of a real number $X \in [X_{\min}, X_{\max}]$, with Opp being the inverse variable, the above formula can be written as:

$$\vec{Opp}_i = \left(\vec{X}_{\min} + \vec{X}_{\max}\right) - \vec{X}_i \tag{2.11}$$

Where $\vec{X_i}$ is the component of the i -th solution, and $\vec{Opp_i}$ is the inverse solution corresponding to $\vec{X_i}$.

3. Empirical Analysis

3.1. Data Source

The daily average price data used in this study is sourced from the S&P Capital IQ database. All other data, including daily trading data for USD/CNY exchange rates and indices such as NASDAQ Composite Index, Dow Jones Industrial Average, Shanghai Composite Index, and Hang Seng Index, are obtained from the Wind database. The daily average price data is used to predict the next trading day's USD/CNY exchange rate, while the eight aforementioned variables are used to predict the USD/CNY closing price for the following day. Specific data details are as follows:

(1) USD/CNY Price Data

This study selects the daily trading data for the USD/CNY exchange rate between January 1, 2015, and January 1, 2020. The data includes daily average price, opening price, highest price, lowest price, and closing price.

(2) Stock Price Data

Stock price data covers the daily trading data of the NASDAQ Composite Index (Code: IXIC.GI), Dow Jones Industrial Average (Code: DJI.GI), Shanghai Composite Index (Code: 000001.SH), and Hang Seng Index (Code: HSI.HI) between January 1, 2015, and January 1, 2020.

Excluding weekends, a total of 1305 daily average price data points are used for univariate predictions of the USD/CNY exchange rate. For multivariate predictions, a total of 1221 daily data points including opening price, highest price, lowest price, closing price, NASDAQ Composite Index, Dow Jones Industrial Average, Shanghai Composite Index, and Hang Seng Index are used for the same period.

3.2. Evaluation Metrics

Commonly used evaluation metrics to assess the performance of prediction models are R^2、MAPE、MSE, introduced as follows:

(1)R^2

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2/n}{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2/n}, \tag{3.1}$$

DOI: 10.54254/2755-2721/32/20230177

where $\hat{y}_i$ is the predicted value, y_i is the true value, $\bar{y}_i$ is the mean of y, and R^2 ranges between 0 and 1.

(2)MSE

$$\text{MSE} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2, \tag{3.2}$$

where $\hat{y}_i$ is the predicted value, y_i is the true value, and a lower MSE indicates better predictive performance.

(3)MAPE

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right|. \tag{3.3}$$

where $\hat{y}_i$ is the predicted value, y_i is the true value, and a lower MAPE indicates more accurate predictions.

3.3. LSTM-ELM Weighted Combination Exchange Rate Prediction Model

3.3.1. Model Construction

In this study, the predicted value from the LSTM model is denoted as Y_1, and the predicted value from the ELM model is denoted as Y_2. A combined prediction model is established by multiplying the two prediction results by their respective weights and then adding them together to obtain the final result of the combined method:

$$Y = W_1Y_1 + W_2Y_2. \tag{3.4}$$

Where W_1, W_2 represents the weighting coefficients ($W_1 + W_2 = 1$).

The improved Marine Predators Algorithm (MPA) is employed to determine the optimal ratio of the two models in the combination model. Using MPA, with MAPE as the fitness function during the optimization process, the weights of the combination are optimized. These optimized weights are assigned to the prediction values of each model to obtain the prediction values of the combined model. The optimization process is detailed in Figure 3.1.

DOI: 10.54254/2755-2721/32/20230177

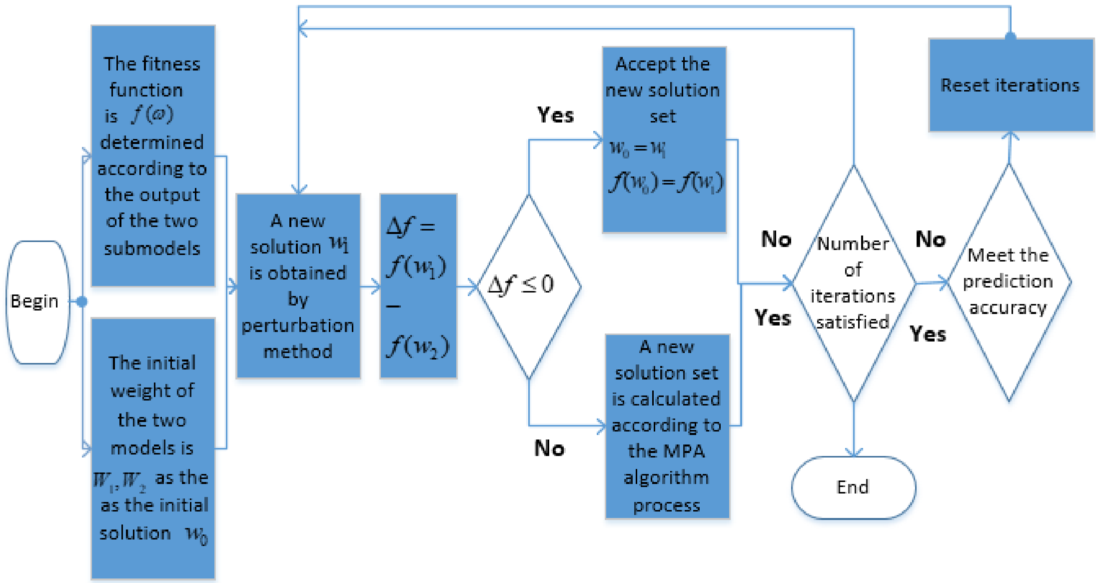

Figure 3.1. Flowchart of the Marine Predators Algorithm (MPA) for Determining Optimal Weights of Two Exchange Rate Prediction Models

3.3.2. Experimental Results

Table 3-1 presents the evaluation metrics of the univariate LSTM-ELM weighted combination exchange rate prediction model using different training datasets (80% and 90% of the data).

Table 3-1 Evaluation Metrics of Univariate LSTM-ELM Weighted Combination Exchange Rate Prediction Model with Different Training Sets

Evaluation Metric	80% Training Set Value	90% Training Set Value
MAPE	0.00164	0.00174
MSE	0.00035	0.00042
R^2	0.99898	0.99687

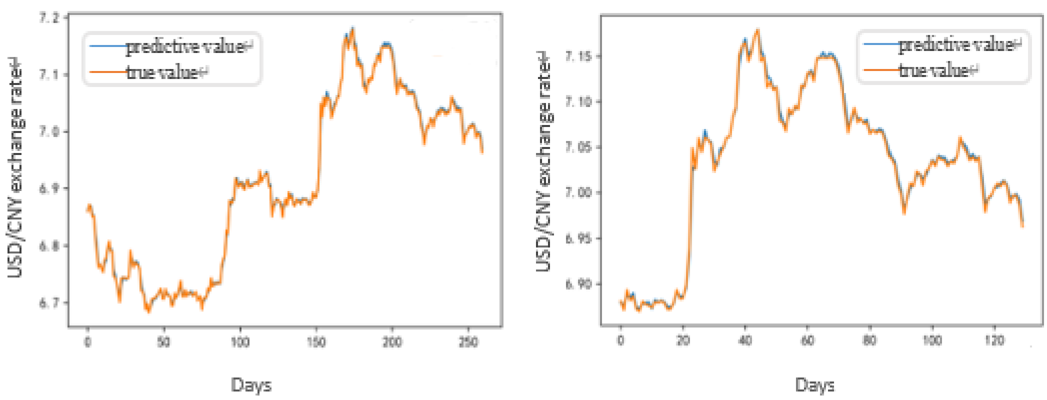

Figure 3.2. Comparison Chart of Real and Predicted Values of USD/CNY Exchange Rate by the LSTM-ELM Weighted Combination Model for 80% and 90% Training Sets

From the tables and figures, it is evident that in both training dataset scenarios, the 80% training dataset performs better in terms of prediction accuracy, as indicated by the lower MAPE and MSE values.

DOI: 10.54254/2755-2721/32/20230177

Additionally, the R^2 value is closer to 1 in the 80% training dataset scenario, indicating better fitting performance.

Table 3-2 and Figure 3.3 show the evaluation metrics and comparison of real and predicted values for the multivariate LSTM-ELM weighted combination exchange rate prediction model using different training datasets (80% and 90% of the data). Similar to the univariate scenario, the 80% training dataset outperforms the 90% training dataset in terms of prediction accuracy, as evidenced by the lower MAPE and MSE values. The R^2 value is also closer to 1 in the 80% training dataset scenario, indicating better fitting performance.

Table 3-2. Evaluation Metrics of Multivariate LSTM-ELM Weighted Combination Exchange Rate Prediction Model with Different Training Sets

Evaluation Metric	80% Training Set Value	90% Training Set Value
MAPE	0.00226	0.00262
MSE	0.00022	0.00031
R^2	0.98939	0.98897

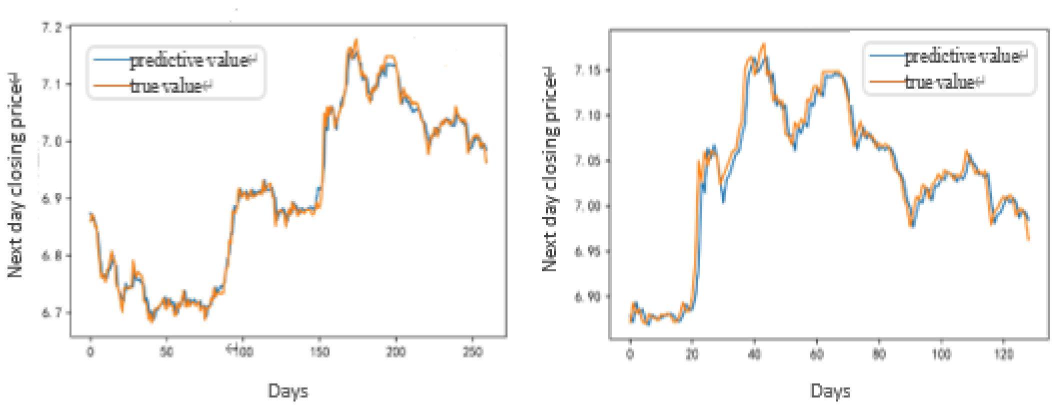

Figure 3.3. Comparison Chart of Real and Predicted Values of Multivariate LSTM-ELM Weighted Combination Exchange Rate Prediction Model for 80% and 90% Training Sets

3.4. Comparison of Prediction Results for Different Models

Based on the predicted values of the six methods, the evaluation metrics for each method are calculated for both univariate and multivariate predictions using both 80% training and 20% testing datasets, as well as 90% training and 10% testing datasets. The evaluation results are summarized in Tables 3-3, 3-4, 3-5, and 3-6.

Table 3-3. Comparison of Evaluation Metrics for Univariate Six Different Prediction Models with 80% Training Set and 20% Testing Set

Model / Evaluation Metric	LSTM-ELM Weighted	LSTM-ELM Average	LSTM	ELM	Random Forest	SVM (Linear Kernel)
MAPE	0.00164	0.11264	0.00494	0.00377	0.00214	0.00273
MSE	0.00035	0.01023	0.00041	0.00101	0.00061	0.00039
R^2	0.99898	0.99491	0.97997	0.98592	0.99323	0.98595

DOI: 10.54254/2755-2721/32/20230177

Table 3-4. Comparison of Evaluation Metrics for Univariate Six Different Prediction Models with 90% Training Set and 10% Testing Set

Evaluation Metric \ Model	LSTM-ELM Weighted	LSTM-ELM Average	LSTM	ELM	Random Forest	SVM (Linear Kernel)
MAPE	0.00174	0.12679	0.00683	0.00232	0.00182	0.00273
MSE	0.00042	0.01155	0.00076	0.00054	0.00051	0.00047
R^2	0.99687	0.97998	0.89034	0.99219	0.99503	0.95400

Table 3-5. Comparison of Evaluation Metrics for Multivariate Six Different Prediction Models with 80% Training Set and 20% Testing Set

Evaluation Metric \ Model	LSTM-ELM Weighted	LSTM-ELM Average	LSTM	ELM	Random Forest	SVM (Linear Kernel)
MAPE	0.00226	0.11634	0.00342	0.00167	0.00263	0.00281
MSE	0.00022	0.00038	0.00056	0.00029	0.00061	0.00032
R^2	0.98939	0.98179	0.97441	0.98629	0.98144	0.98464

Table 3-6. Comparison of Evaluation Metrics for Multivariate Six Different Prediction Models with 90% Training Set and 10% Testing Set

Evaluation Metric \ Model	LSTM-ELM Weighted	LSTM-ELM Average	LSTM	ELM	Random Forest	SVM (Linear Kernel)
MAPE	0.00262	0.12992	0.00297	0.00169	0.00284	0.00272
MSE	0.00031	0.00057	0.00046	0.00032	0.00091	0.00034
R^2	0.98897	0.90936	0.86368	0.95337	0.98356	0.94890

From the tables, it is clear that among the six methods, the LSTM-ELM weighted combination exchange rate prediction model outperforms the others. This model exhibits the lowest MAPE and MSE values, indicating superior prediction accuracy. In terms of predictive performance, the R^2 value for the LSTM-ELM weighted combination model is closest to 1. Thus, the proposed LSTM-ELM weighted combination exchange rate prediction model, whether in terms of fitting performance or error values, outperforms the other five comparison models. It demonstrates excellent capability in predicting both the average USD/CNY exchange rate and the closing price for the following day.

4. Conclusion

In this study, a weighted combination exchange rate prediction model using LSTM and ELM was proposed. Through comparison with five other models, it was found that the proposed prediction model exhibited superior predictive performance and achieved favorable results in exchange rate forecasting. The primary focus of this paper was to investigate the LSTM-ELM weighted combination exchange rate prediction model in the context of both univariate and multivariate exchange rate predictions. Two testing set allocation schemes were adopted: 90% as training set and 10% as testing set, and 80% as training set and 20% as testing set, to explore the impact of different training and testing set distributions on exchange rate prediction. Concerning the LSTM-ELM weighted combination exchange rate prediction model, the allocation scheme of 80% training set and 20% testing set yielded higher prediction accuracy and better fitting results. In the future, the model proposed in this study could also be applied to address other complex forecasting problems, such as crude oil price prediction, traffic flow prediction, stock index prediction, among others.

DOI: 10.54254/2755-2721/32/20230177

References

[1] Wang, P. P. (2021). Sino-U.S. Trade Frictions, U.S. Economic Policy Uncertainty, and RMB Exchange Rate Fluctuations. *World Economy Studies*, (329), 75-92.

[2] Tatsuma, W. (2022). Out-of-sample forecasting of foreign exchange rates: The band spectral regression and LASSO. *Journal of International Money and Finance*, 128.

[3] Pengfei, L., Ze, W., Daoqun, L., et al. (2023). A CNN-STLSTM-AM model for forecasting USD/RMB exchange rate. *Journal of Engineering Research*, 11(2).

[4] Yujie, Z., Qian, Y., Zhicai, L., et al. (2023). A Prediction Model Based on Gated Nonlinear Spiking Neural Systems. *International Journal of Neural Systems*.

[5] Xueling, L., Xiong, X., Baojun, G. (2023). Increasing the prediction performance of temporal convolution network using multimodal combination input: Evidence from the study on exchange rates. *Frontiers in Physics*.

[6] Zhang, L., Sun, S., Wang, Y. (2021). Exchange Rate Forecasting Based on Deep Learning LSTM Model. *Statistics & Decision*, 37(13), 158-162.

[7] Mi, X., Liu, H., Li, Y. (2017). Wind speed forecasting method using wavelet, extreme learning machine and outlier correction algorithm. *Energy Conversion and Management*, 151, 709-722.

[8] Faramarzi, A., Heidarinejad, M., et al. (2020). Marine predators algorithm: A nature-inspired metaheuristic. *Expert Systems with Applications*, 152, 113377.

[9] Aarts, E., Aarts, E. H., Lenstra, J. K. (2003). Local search in combinatorial optimization. Princeton University Press.