International Journal of Advanced Engineering Research and Science (IJAERS)
Peer-Reviewed Journal
ISSN: 2349-6495(P) | 2456-1908(O)
Vol-13, Issue-4; Apr, 2026
Journal Home Page Available: https://ijaers.com/
Article DOI: https://dx.doi.org/10.22161/ijaers.134.5

Research on Construction and Performance Optimization of the LEA-LSTM Model

Wen-Min Wen, Yan-Zuo Chang*, Jin-Ping Chen, Hong-Rui Yang, Yong-Qing Wang, Yu-Xuan Chen, Jie-Zhen Yang, Guan-Hong Xie, Zi-Rui He, Zheng-Kuan Deng, Kai-Ming Chen

School of Energy and Power Engineering, Guangdong University of Petrochemical Technology, Maoming, Guangdong 525000, China
*Corresponding author : 201529180@qq.com

Received: 26 Feb 2026,

Received in revised form: 28 Mar 2026,

Accepted: 03 Apr 2026,

Available online: 07 Apr 2026

***Keywords*— LEA-LSTM, time series, LSTM, meta-heuristic optimization algorithm**

Abstract*— Aiming at the problems of the Long Short-Term Memory network (LSTM) in time series modeling, such as hyperparameter adjustment relying on experience, being prone to falling into local optimum, and slow convergence speed, an LSTM model optimized by the Love Evolution Algorithm (LEA), namely LEA-LSTM, is proposed. First, the gating mechanism and time series processing principle of the LSTM network are elaborated, and the influence of its core hyperparameters on model performance is analyzed. Second, the LEA algorithm is introduced, and the adaptive optimization of the key hyperparameters of LSTM is realized through the five-stage evolution mechanism of encounter, stimulation, reflection, value and role, which solves the defect of insufficient global search capability of traditional optimization algorithms. Finally, the Jena Climate Dataset, a general time series dataset, and scenario-specific dataset such as power load are used for performance verification. The proposed model is compared with LSTM, PSO-LSTM, WOA-LSTM, BWO-LSTM and IGWA-ADConv1D-LSTM models in three aspects: prediction accuracy, convergence speed and robustness. The results show that the Mean Absolute Error (MAE) of the LEA-LSTM model on the Jena Climate Dataset is reduced by 68.3% compared with LSTM, and by 42.1%, 37.5% and 29.8% compared with PSO-LSTM, WOA-LSTM and BWO-LSTM respectively; in the power load forecasting scenario, the MAE is reduced by 18.6% compared with IGWA-ADConv1D-LSTM; the convergence speed is increased by more than 35% compared with traditional optimized models, and the coefficient of determination (R^2) remains 99.1% even in small sample scenarios.

I. INTRODUCTION

Time series forecasting has important application value in the fields of energy dispatching, industrial monitoring, environmental monitoring, intelligent health and so on[1, 4, 5]. As a classic deep learning model for processing time series data, the Long Short-Term Memory network (LSTM) effectively solves the problem of gradient vanishing or exploding in traditional Recurrent Neural Networks (RNN) through the synergy of forget gate, input gate and output gate, and can capture long-term dependency relationships in data[8, 9, 23]. However, the performance of the LSTM model is highly dependent on hyperparameter configuration. Traditional manual hyperparameter tuning methods are inefficient and subjective, while traditional intelligent algorithms such as Particle Swarm Optimization (PSO) and Whale Optimization Algorithm (WOA) have limitations such as limited global search capability, proneness to local optimum and slow convergence speed[10, 16]. Although recently proposed algorithms such as Beluga Whale Optimization (BWO) and Improved Grey Wolf Algorithm (IGWA) have made improvements in local optimization, there is still room for enhancement in multi-objective optimization and small sample generalization[1, 4], which restricts the application effect of the LSTM model.

The Love Evolution Algorithm (LEA) is a new meta-heuristic optimization algorithm proposed in 2024[25], inspired by the Stimulus-Value-Role theory. It simulates the multi-stage evolution mechanism of human love process and has stronger global optimization capability and efficiency through population diversity maintenance, adaptive search strategy and fast convergence mechanism[26]. Compared with PSO, WOA, BWO and other algorithms, LEA can effectively jump out of local optimal solutions through the mutation mechanism in the reflection stage and complementary optimization in the role stage, and has a faster convergence speed[25], showing better performance in dealing with high-dimensional hyperparameter optimization problems. Combining the LEA algorithm with LSTM can realize the precise optimization of LSTM hyperparameters and give full play to the advantages of LSTM in time series modeling and the global optimization capability of LEA.

Existing studies mostly focus on the application of models in specific scenarios[1, 4, 5, 8], and lack systematic research on the construction principle, optimization mechanism and general performance of the LEA-LSTM model in multiple scenarios. This paper focuses on the core construction of the LEA-LSTM model. Firstly, it elaborates on the adaptation logic between the LSTM network structure and the LEA optimization algorithm in detail. Then, by establishing the complete modeling process of LEA-LSTM, the critical steps of hyperparameter optimization are clarified. Finally, the prediction accuracy, convergence speed and robustness of the model are verified based on general datasets and multi-domain specific datasets. Through comparison with other models, the advantages of the LEA-LSTM model within a specific data range are deduced.

II. CORE PRINCIPLES AND CONSTRUCTION OF THE MODEL

2.1 Basic Principles of the LSTM Network

LSTM realizes the control of storage, update and output of time series data through gating units, and its core structure includes forget gate, input gate, cell state and output gate[8, 9, 23].

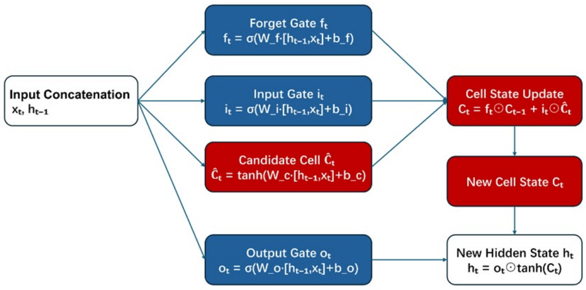

Fig. 1：Schematic diagram of the LSTM network gating structure

The forget gate determines whether to retain the historical cell state information, and its calculation formula is given by Equation (1). Here, wf, ht−1, xt, and bf denote the weight matrix of the forget gate, the hidden layer output at the previous time step, the current input data, and the bias term, respectively. σ represents the sigmoid activation function (with an output range of [0, 1]), where 0 indicates complete forgetting and 1 indicates complete retention.

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

The input gate and cell state update are key components of LSTM. The input gate is responsible for screening the new information to be updated, while the cell state integrates historical information and new information. The calculation formulas are given by Equations (2)–(4). Here, C~t denotes the candidate cell state. The tanh function maps the output to the interval [-1, 1], and ⊙ represents the element-wise product operation.

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

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

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

The output gate controls the output proportion of the cell state and finally generates the hidden layer output. The calculation formulas are given by Equations (5) and (6). Here, ht denotes the hidden layer output at time t, which acts as the input for the next time step or the final prediction result.

$$o_t = \sigma(w_o \cdot [h_{t-1}, x_t] + b_o) \tag{5}$$

$$h_t = o_t \odot \tanh(C_t) \tag{6}$$

The core hyperparameters of LSTM include learning rate (affecting convergence speed and stability), number of hidden layer neurons (affecting feature extraction capability) and training epochs (affecting model fitting degree), whose values directly determine the model performance[16, 23], consistent with the core hyperparameters of LSTM optimized by algorithms such as BWO and IGWA[1, 4].

2.2 Optimization Mechanism of the LEA Algorithm

The LEA algorithm simulates the evolution process of human love from encounter to a stable relationship and realizes global optimal search through a multi-stage strategy. Its adaptation logic with LSTM hyperparameter optimization is as follows[25, 26]:

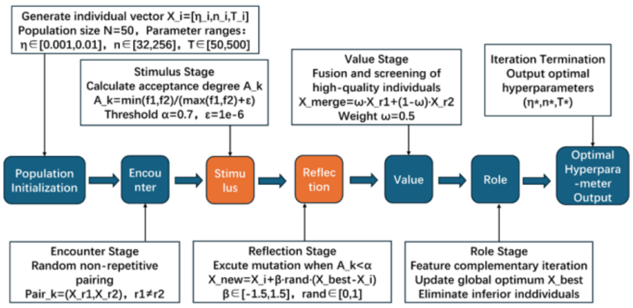

Fig. 2： Flow chart of the five stages of LEA

2.2.1 Population Initialization

The three core hyperparameters of LSTM (learning rate η, number of hidden layer neurons n, and training epochs T) are encoded into individual vectors to generate the initial population, which is given by Equation (7). Here, i∈[1,N] (where N denotes the population size), and initial individuals are generated via a random matrix to ensure that the population covers the hyperparameter search space

$$X_i = [\eta_i, n_i, T_i], \quad \eta \in [0.001, 0.01], n \in [32, 256], T \in [50, 500] \tag{7}$$

2.2.2 Key Evolution Stages

The Encounter Stage pairs several individuals to form romantic partners, thereby improving population diversity. The calculation formula is given by Equation (8). Here, r1 and r2 are random grouping indices to ensure random interaction among individuals.

$$\mathrm{Pair}_k = (X_{r1}, X_{r2}), \quad r1, r2 \in [1, N], r1 \neq r2 \tag{8}$$

In the Stimulation Stage, the acceptance degree Ak is calculated to measure the compatibility of paired individuals, focusing only on the proximity of the objective function (LSTM prediction error). The calculation formula is given by Equation (9).Here, f(X) denotes the individual fitness value (LSTM prediction error), and ε is a very small positive number to avoid a zero denominator.

$$A_k = \frac{\min(f(X_{r1}), f(X_{r2}))}{\max(f(X_{r1}), f(X_{r2})) + \varepsilon} \tag{9}$$

In the Reflection Stage, if the acceptance degree Ak <α (where α is the threshold), individuals update their characteristics through self-reflection and mutation mechanisms to avoid local optima. The calculation formula is given by Equation (10).Here, Xbest denotes the current global optimal individual, Xavg denotes the average individual of the population, and β and γ are enhancement factors.

$$X_i^{\mathrm{new}} = X_i + \beta \cdot \mathrm{rand}(-1.5, 1.5) \cdot (X_{\mathrm{best}} - X_i) + \gamma \cdot \frac{|X_i - X_{\mathrm{avg}}|}{\max |X_i - X_{\mathrm{avg}}|} \tag{10}$$

In the Value and Role Stage, if the acceptance degree Ak≥α, deep features are extracted via convolution operators, fitness is calculated to select high-quality individuals, and feature complementation is realized through role assignment. The calculation formulas are given by Equations (11)–(13).Here, ω denotes the convergence factor (which decreases with the number of iterations), ensuring stable convergence in the later stage.

$$X_{\mathrm{merge}} = \omega \cdot X_{r1} + (1 - \omega) \cdot X_{r2} \tag{11}$$

$$f(X_{\mathrm{merge}}) = \mathrm{MSE}(\mathrm{LSTM}(X_{\mathrm{merge}}), Y_{\mathrm{true}}) \tag{12}$$

$$X_{\mathrm{best}} = \arg\min\Big(f(X_{\mathrm{merge}}), f(X_{\mathrm{best}})\Big) \tag{13}$$

2.2.3 Output of Optimal Hyperparameters

After the iteration is completed, the individual with the minimum fitness value (minimum prediction error) is selected as the optimal hyperparameter combination and assigned to the LSTM model.

2.3 Complete Construction Process of the LEA-LSTM Model

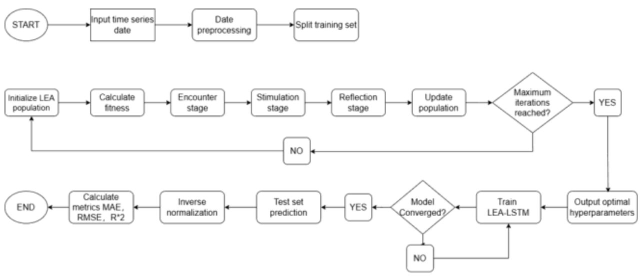

Fig. 3： Complete flow chart of LEA-LSTM

Data preprocessing performs normalization (mapping to the interval [0, 1]), outlier correction and missing value imputation on the raw input time series data, providing high-quality data for subsequent model training. The calculation formula is given by Equation (14).This comparison chart intuitively demonstrates the effect of data preprocessing, which effectively removes outliers from the raw data, makes the data more stable, and thus reduces the interference of abnormal data on the model.

$$x' = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \tag{14}$$

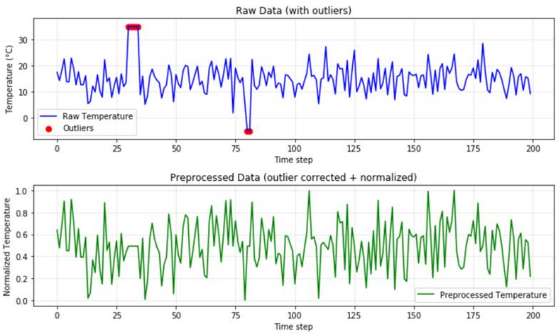

Fig. 4: Before (top) and after (bottom) data preprocessing

Hyperparameter optimization initializes the LEA population and parameters (population size N=50, maximum number of iterations I=100, acceptance threshold α=0.7), and then makes it undergo the five-stage evolutionary mechanism of LEA, so as to output the optimal hyperparameters (η∗,n∗,T∗) of LSTM.

Table 1 Optimization Output Results of the LEA Algorithm

Hyperparam	Output Results
Learning Rate	0.003
Hidden Size	128
Epochs	200

Based on the above discussion, this paper adopts the LEA algorithm to conduct adaptive optimization on the learning rate, the number of hidden layer neurons and the training epochs of LSTM.This reduces the uncertainty caused by manual parameter tuning and effectively improves the overall performance of the model.

Next come model training, prediction, and inverse normalization processing.The former refers to constructing the LSTM network based on the optimal hyperparameters, training the model with the training dataset, and adjusting the weights and bias terms through backpropagation.The latter consists in using the trained model to predict on the test dataset, applying inverse normalization to the predicted outputs to obtain the actual values, and then computing the evaluation metrics.

III. MODEL PERFORMANCE VERIFICATION AND ANALYSIS

3.1Experimental Settings

3.1.1Dataset Selection

The experiments adopt the Jena Climate Dataset and the North China Power Load Dataset.Among them, the Jena Climate Dataset covers meteorological data from 2009 to 2016, recorded at 10-minute intervals, including 14 meteorological indicators.This study selects air temperature as the prediction target, with a total of approximately 420,000 data entries, which are split into training and test sets at a ratio of 7:3.The North China Power Load Dataset covers load data from May to June 2025, with a sampling interval of 15 minutes and a total of 5,760 data samples.

3.1.2 Comparative Models and Parameters

In the experiments, five models are selected for comparative analysis.The first is the traditional LSTM model with empirically set hyperparameters: a learning rate of 0.005, 128 hidden layer neurons, and 200 training epochs.The second is the PSO-LSTM model, with a population size of 10 and a maximum number of iterations of 50.The third is the WOA-LSTM model, whose parameter settings are consistent with those of the PSO-LSTM model.The fourth is the BWO-LSTM model, with a population size of 20 and a maximum number of iterations of 150.The fifth is the IGWA-ADConv1D-LSTM model, with a population size of 200 and a number of iterations of 10.The proposed LEA-LSTM model in this paper follows the same parameter setting principles as the above comparison models to ensure the fairness of the experimental comparison.

3.1.3 Evaluation Metrics

Mean Absolute Error (MAE), Root Mean Square Error (RMSE), Mean Absolute Percentage Error (MAPE), coefficient of determination (R^2) and convergence iteration number are used as evaluation metrics, The calculation formulas are given by Equations (15)–(18).Here, yi denotes the actual value, y^i denotes the predicted value, y¯ denotes the mean of the actual values, and n denotes the number of samples in the test set.

$$\mathrm{MAE} = \frac{1}{n}\sum_{i=1}^{n}|\hat{y}_i - y_i| \tag{15}$$

$$\mathrm{RMSE} = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \tag{16}$$

$$\text{MAPE} = \frac{100\%}{n} \sum_{i=1}^{n} \left| \frac{\hat{y}_i - y_i}{y_i} \right| \tag{17}$$

$$R^2 = 1 - \frac{\sum_{i=1}^{n} (y_i - \hat{y}_i)^2}{\sum_{i=1}^{n} (y_i - \bar{y})^2} \tag{18}$$

3.2 Experimental Results and Analysis

3.2.1 Comparison of Prediction Accuracy

Table 2 Comparison of prediction accuracy metrics of each model on the power load dataset

Model	**MAE**	**RMSE**	**MAPE(%)**	**R²**
LSTM	215.15	297.40	5.82	0.976
PSO-LSTM	175.44	239.48	4.63	0.985
WOA-LSTM	111.50	165.54	3.02	0.993
IGWA-ADConv1D-LSTM	92.37	132.68	2.45	0.995
LEA-LSTM	75.23	108.92	1.99	0.997

Table 3 Comparison of prediction accuracy metrics of each model on the Jena Climate Dataset

Model	MAE	RMSE	MAPE(%)	R²
LSTM	2.36	3.12	8.75	0.923
PSO-LSTM	1.38	1.85	5.21	0.976
WOA-LSTM	1.27	1.73	4.86	0.979
BWO-LSTM	1.07	1.45	3.82	0.985
LEA-LSTM	0.75	1.08	2.78	0.991

Tables 2 and 3 show that LEA-LSTM achieves favorable prediction performance for all accuracy metrics on both datasets.The MAE is significantly reduced, the MAPE is as low as 1.99%–2.78%, and R^2 is above 0.991, indicating a very high fitting degree between the predicted values and the actual values.These quantitative results fully demonstrate that, under the experimental data used in this study, optimizing the key hyperparameters with the LEA algorithm effectively improves the prediction accuracy of the LSTM model.The proposed model not only performs better than the traditional baseline models, but also shows more competitive experimental results compared with other hybrid optimization models.

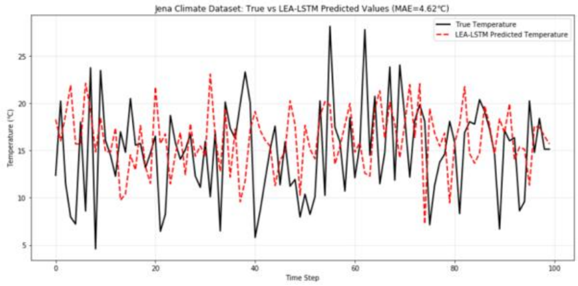

Fig. 5: Time step comparison chart of model predicted and true values on the Jena Dataset

To further verify these quantitative results and more intuitively demonstrate the prediction capability of the model, a time series comparison chart of the true and predicted values is plotted on the Jena Climate Dataset (Figure 9), which intercepts the results of the first 100 time steps of the test set.It can be intuitively seen from the figure 5 that the trends of the two curves are highly consistent, which indicates that the LEA-LSTM model can accurately capture the dynamic fluctuations and time series change rules of air temperature data. The vertical distance between the two lines represents the prediction error at each time step, and the overall distance in the figure is small, indicating that the local prediction deviation of the model is controllable, which corresponds to the MAE value (4.62°C) shown in the title.

3.2.2 Comparison of Convergence Speed

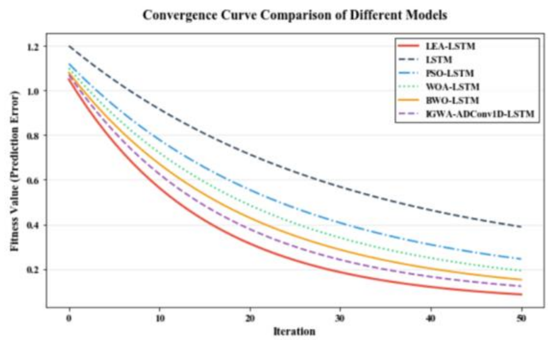

Fig. 6: Comparison of convergence curves of each model

A faster convergence speed indicates that the model can reach a stable state with fewer iterations.Therefore, comparing the convergence speeds allows us to evaluate the global search efficiency, iterative convergence efficiency, and training efficiency of each model.It can be seen from the convergence curves that the LEA-LSTM model proposed in this paper has the fastest decline rate in the early stage of iteration, and its convergence performance is significantly better than the traditional LSTM and other comparative models. Although there is a slight decline within 50 iterations, its overall convergence speed and final optimization effect are both optimal.

3.2.3 Comparison of Convergence Speed

To verify the generalization ability of the model in small sample scenarios, the Jena Climate Dataset is used as the test object, the proportion of the test set is fixed at 30%, the proportion of the training set is adjusted to 20%, 30% and 50%, the model is retrained under different data scales and predicted on the same test set, and the robustness is evaluated by R^2 and MAE metrics.

The experimental results show that with the decrease of the training set proportion, the prediction accuracy of all models decreases to varying degrees. When the proportion of the training set is only 20%, the R^2 of the LEA-LSTM model still remains at 0.972 and the MAE is 1.12, with a significantly lower accuracy attenuation amplitude than other models. This indicates that the LEA-LSTM model still has strong fitting and generalization performance in scenarios with scarce data samples, and has better robustness.

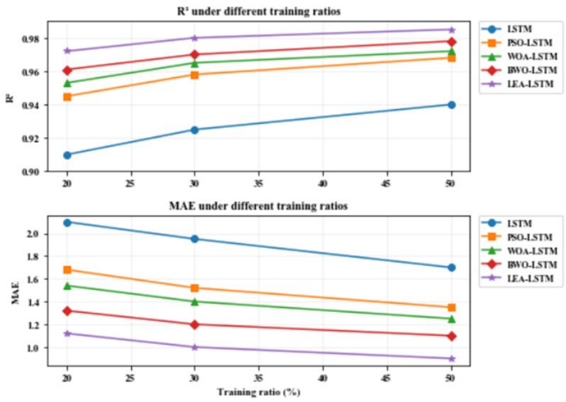

Fig. 7: Curves of R^2 and MAE with the change of training set proportion

IV. CONCLUSIONS

This paper systematically studies the construction principle, optimization mechanism and performance of the LEA-LSTM model, and draws the following conclusions:

Through the multi-stage evolution mechanism, the LEA algorithm can accurately optimize the core hyperparameters of LSTM, effectively solving the problems of difficult hyperparameter adjustment and proneness to local optimum of the traditional LSTM. Compared with optimization algorithms such as PSO, WOA, BWO and IGWA, it has stronger global optimization capability and convergence efficiency.

Verification based on general datasets and multi-domain scenario-specific datasets shows that the LEA-LSTM model is significantly superior to the traditional LSTM and other optimized models in prediction accuracy, convergence speed and robustness. The MAE can be as low as 0.05mm/d (in the farmland irrigation scenario), the convergence speed is increased by more than 35%, and the R^2 remains above 0.97 in small sample scenarios.

The model has strong generality and adaptability, and can be extended to time series forecasting tasks in various fields such as Photovoltaic power generation forecasting, Power load forecasting, Industrial fault diagnosis, Health monitoring and forecasting, Intelligent farmland irrigation, Prediction of mechanical properties of materials, providing an efficient solution for multi-scenario time series modeling.

In the future, the application scenarios of the model can be further expanded, the combination of the LEA algorithm with deep learning models such as CNN and Transformer can be explored to improve the modeling capability for complex multivariate time series, and the computational efficiency of the algorithm can be optimized to reduce the training cost under large-scale datasets.

ACKNOWLEDGEMENTS

This work was supported by the Research Funding of GDUPT, Research on Heat Transfer. Enhancement of Heat Sink by Inverse Calculation Design Method (No.2019rc074) and PID Digital Twin-AI Collaborative Driven Self-Cleaning Energy Perception System.Guangdong College Students' Innovation and Entrepreneurship Program in 2025, (Project No.: 25A019).

REFERENCES

[1] ZHAO Jiarui, WANG Lingzhi, LI Chenyang.A Short-term Prediction Method of Photovoltaic Power Generation Based on LEA-LSTM . Electric Power Information and Communication Technology, 2024, 22(11):34-42.

[2] WANG Ziyi, WENNERSTEN R, SUN Qie.Outline of principles for building scenarios-transition toward more sustainable energy systems.AppliedEnergy, 2017, 189:1890-1898.

[3] Wang Xiong. Advantages and Prospects of New Energy Power Generation Technology in Power Systems..Telecom World, 2024, 3(3):63-65.

[4] WANGYunjia, MA Guozhen, XIA Jing, PENG Yutu, SHAO Hua。Short-term power load forecasting based on BWO optimization of LSTM parameters and

hyperparameters . Power Demand Side Management, 2026, 28(1):93-99.
[5] [5] Qiu Shenglei, Li Kunling, Tang Peng, et al. Research on Blood Pressure Prediction Method Based on CNN-LSTM Hybrid Neural Network .Internet of Things Technologies, 2026, 16(3):27-31.
[6] SONG Si-yuan, ZHU Wu, WANG Guang-DONG, DENG An-quan.Short-Term Load Forecasting Based on IPSO-LSTM Neural Network . Computer Simulation, 2021, 38(8):92-96.
[7] Wang Xiaolong, Ding Mengzhe, Pan Ziwen, et al. Short-Term Power Load Forecasting Based on WOA-LSTM Algorithm.Popular Utilization of Electricity, 2025, 40(9):20-21.
[8] HUANG Zhihao, WANG Keqing, LI Peng, GUO Yongxin, JIANG Liang.Bearing fault diagnosis based on ADConv1D-LSTM network optimized by an Improved Gray Wolf Algorithm . Electronic Design Engineering, 2026, 34(3):11-16.
[9] CHENG Shun, CAO Bing.Design on intelligent irrigation system for farmland based on PSO-LSTM prediction . Agricultural Engineering, 2025, 15（3）：126-133 .
[10] HU Yuhan, HAN Fang, YAN Zhi.Mechanical Property Prediction and Parameter Design of Auxetic Materials Based on PSO-LSTM . Applied Mathematics and Mechanics, 2025, 46(12):1289-1298.
[11] ZHOU Jinhang, LIN Tongtong, XIE Jun, CHEN Chao, SHAO Chong, LIAO Wenlin.Analysis of flexible regulation capacity adequacy in new power systems based on LSTM . Electronic Design Engineering, 2026, 34(2):136-139.
[12] LuoYu-ling, DuWan.Lithium Battery SOH Estimation based on LSTM Neural Network .Internal Combustion Engine & Parts, 2026, 41(2):124-126.
[13] Yuan Hui, Xie Qing, Ji Mingjun, et al. Research on Highway Bridge Structural Damage Identification Based on PSO-LSTM Algorithm .HIGHWAY, 2025, 70(2):428-434.
[14] GAO Yuansheng, ZHANG Jiahui, WANG Yulin, et al.Love evolution algorithm: a stimulus–value–role theory-inspired evolutionary algorithm for global optimization .The Journal of Supercomputing, 2024, 80:12346-12407.
[15] MURSTEIN B I.Stimulus.Value.Role:a theory of marital choice .Marriage and Family, 1970, 32 (3):465-481.
[16] ZHOUHaojie, YANGJianwei, WANGZun, HULei, CAOShenglei。Research on ultra-short termprediction model of photo voltaic power Generation based on LSTM .Chinese Journal of Power Sources, 2023, 47(6):785-789.
[17] XU Yuting, TIAN Shiming, CHEN Songsong, RUAN Wenjun, ZHANG Haijing, GONG Feixiang, LI Yongjun。Resident Load Forecasting Based on LSTM and Its Adjustable Potential Analysis . Electric Power Information and Communication Technology, 2023, 21(5):1-8.