ISSN 1330-3651 (Print), ISSN 1848-6339 (Online)
https://doi.org/10.17559/TV-20240314001398
Original scientific paper

# SSA-CNN-LSTM Fusion of Multi-Source Heterogeneous Data for Order Performance Evaluation

Wei QIANG*, Yihan FANG

**Abstract:** With the continuous development of the flexible supply chain in the manufacturing industry of the Industrial Internet of Things and the widespread promotion of the order-oriented production mode, the quantity and types of data involved in order performance evaluation tasks are constantly increasing. The applicability of traditional evaluation methods is significantly weakened, leading to additional investment of manpower and time resources. In response to this issue, this paper proposes a SSA-CNN-LSTM multi-source heterogeneous data fusion model aimed at achieving precise order performance evaluation. By integrating and learning data of different structures from various sources, the model fully explores the correlation of data features to obtain precise fusion results, thereby enabling the evaluation of order performance. Simulation experiments conducted on a dataset from a certain intelligent collection customization company demonstrate that the RMSE, MAE, and MAPE of the SSA-CNN-LSTM model results are reduced by 58.71%, 62.94%, and 63.29% respectively compared to the LSTM model, validating the superior accuracy of the proposed model. It also indicates that the model proposed in this study provides new ideas and methods for completing performance evaluation tasks, offering reliable basis and reference for enterprise decision-making, and enriching the research content of the deep learning multi-source heterogeneous data fusion field.

**Keywords:** CNN neural network; LSTM neural network; multi-source heterogeneous data fusion; SSA algorithm

## 1 INTRODUCTION

In modern manufacturing, supply chain flexibility has become an indispensable advantage in coping with market fluctuations and diverse customer demands. However, the data accumulated in manufacturing exhibit a trend of vast quantity and high complexity. These data contain rich information, encompassing various types such as digitized production indicators, sensor-captured environmental parameters, and textual description data [1, 2], as illustrated in Fig. 1.

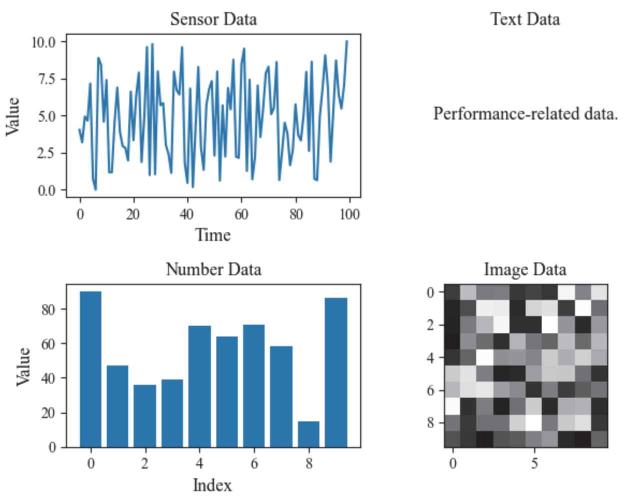


**Figure 1** Examples of heterogeneous data types

In such a rich and complex data environment, accurately evaluating order performance faces increasingly severe challenges. Achieving a flexible supply chain and ensuring the assessability of order performance remain challenging issues. Traditional methods for evaluating order performance often overlook the complex correlations and spatiotemporal dynamic changes among heterogeneous data sources, no longer suitable for heterogeneous data scenarios. Benchmarking [3] typically assesses performance by comparing the gap between targets and benchmarks. However, it struggles to integrate different data formats and sources since it primarily focuses on comparing indicators rather than the structure or type of data itself. Therefore, benchmarking may overlook data diversity for handling heterogeneous data, leading to less accurate assessment results. The Balanced Scorecard method [4] evaluates performance through four different perspectives (financial, customer, internal business processes, and learning and growth). However, it is difficult to quantify multiple heterogeneous data sources and incorporate them into the Balanced Scorecard system. Since the Balanced Scorecard mainly focuses on balancing specific indicators and performance evaluation, its ability to handle and synthesize heterogeneous data is limited. The Fuzzy Comprehensive Evaluation method [5] attempts to deal with the fuzziness and uncertainty of data. However, it struggles to accurately define fuzzy sets for multiple heterogeneous data sources. Heterogeneous data may have different data types, structures, and qualities, making it difficult to effectively process and balance data in fuzzy comprehensive evaluation. Data Envelopment Analysis [6] aims to evaluate the efficiency between multiple input and output indicators. However, it requires high consistency and homogeneity of data. Heterogeneous data typically have different scales and units of measurement, making it difficult to effectively compare and evaluate data in data envelopment analysis. Analytic Hierarchy Process [7] is commonly used to determine the weights of different indicators for performance evaluation or decision support. However, it is difficult to allocate weights to heterogeneous data in the analytic hierarchy process. Because heterogeneous data may involve different data types and measurement units, establishing weights in the analytic hierarchy process may result in inconsistency and inaccuracy. In the actual production environment, enterprises face influences from various aspects such as supply chains, production processes, and market demands, necessitating more comprehensive and accurate evaluation models to guide decision-making and improve overall operational efficiency.

To address these issues, this paper proposes a manufacturing supply chain order performance evaluation model based on a CNN-LSTM model with multi-feature input and single-feature output. The aim is to fully explore the feature correlations among heterogeneous data sources in the evaluation task to achieve intelligent assessment of order performance. Additionally, it introduces the Sparrow Search Algorithm, which possesses stronger global search capability, faster convergence speed, and fewer parameters, to enhance model performance, thus better adapting to the variability and uncertainty of order performance evaluation tasks. Simulation experiments are conducted on an order dataset from a certain intelligent smartphone customization company using MATLAB software. The dataset includes numerical and textual data from sources such as sensors, manual monitoring, and online platforms. The simulation results demonstrate that the proposed model, by combining the Sparrow Search Algorithm, convolutional neural network, and long short-term memory neural network, can achieve fusion analysis of heterogeneous data sources, applicable for supply chain performance evaluation while maintaining good evaluation accuracy. The remaining parts of this paper are organized as follows: Section 2 reviews the current research status. Section 3 introduces the research methodology and the structure of the model. Section 4 presents the experimental results and discussion. Section 5 concludes and provides prospects for future research.

## 2 LITERATURE REVIEW

The technology of multi-source heterogeneous data fusion is a method of comprehensively utilizing data from different sources, structures, and types, aiming to enhance the comprehensiveness and accuracy of information by integrating heterogeneous data [8, 9]. This technology covers the entire process of data collection, integration, processing, and analysis, and it integrates data from multiple sources into a unified framework to produce more comprehensive results with higher information density. In recent years, deep learning has been widely applied in various fields and has accumulated rich research results. In order to enable machines to learn more comprehensively and efficiently, many scholars have been committed to endowing deep learning models with the understanding and reasoning ability of multi-source heterogeneous data, enabling the models to acquire more information during the decision-making process and improve the overall accuracy of decisions. The goal of this research direction is to establish models capable of handling and correlating information from multiple modalities [10]. Maimaitijiang M. et al. [11] proposed a soybean yield assessment model based on DNN neural networks, which integrates field spectra, thermal data, soybean texture features, etc. Experimental results show that this method can provide relatively accurate and stable crop yield estimates. Chen F. C. et al. [12] proposed a deep learning framework for crack detection in nuclear power plants based on CNN and naive Bayes data fusion scheme, aggregating multi-source image data extracted from different video frames to enhance the overall performance and robustness of the system. Simulation results demonstrate good evaluation performance. Khan S. et al. [13] proposed a CNN-LSTM-based multi-source heterogeneous data fusion model for region-based traffic flow assessment in smart cities. Simulation experiments show that the proposed model has strong applicability. Hatami N. et al. [14] proposed an integrated model based on CNN-LSTM for clinical outcome assessment, which integrates patient admission data and clinical test data to obtain decision results.

Experimental results show that this method is helpful for clinical management. The above research results indicate that deep learning algorithms can explore feature correlations in multi-source heterogeneous data, fully utilize data from various sources for integrated processing, make the fusion results closer to the real data needed, and LSTM performs well in various algorithms. The advantage of the CNN-LSTM integrated model over LSTM lies in its ability to effectively extract spatial features and capture temporal relationships, achieve more comprehensive integration of local and global information, and improve result accuracy, but model parameters still need to be manually set. To address the CNN-LSTM parameter optimization problem, Haoran Zhang et al. [15] proposed the VMD-CNN-LSTM fusion model, which uses VMD for data decomposition and denoising to improve data continuity and stability, enhancing prediction results. However, algorithms like VMD are more suitable for signal domains, with lower applicability to performance evaluation data characteristics. Zhiwei Li et al. [16] proposed a WOA-CNN-LSTM model to optimize the parameters of five feature values of input data, improving the accuracy of evaluation results, but WOA has a slow convergence speed for complex problems, resulting in low model scalability. Lu W. et al. [17] established a GA-CNN-LSTM feature fusion model for traffic prediction, using GA to optimize the number of neurons to achieve more accurate fusion results. However, there are many parameter settings and difficulty in adjustment, making it difficult to deploy in practical applications. The current research hotspot is to use more effective optimization algorithms to optimize the performance of CNN-LSTM models.

## 3 RESEARCH METHODOLOGY

### 3.1 Supply Chain Order Performance Analysis

With the diversification of market demands and the rapid development of science and technology, an increasing number of enterprises are transitioning from inventory-based production to order-based production. However, under the premise of diversified order demands, there still exist significant challenges in effectively controlling orders throughout their entire lifecycle. Employing order performance evaluation schemes to assess the performance of order execution, rapidly gauging order completion status, can optimize production cycles and enhance decision-making capabilities for enterprises [18, 19]. While some enterprises can effectively evaluate order performance, the evaluation process involves a large amount of multi-source heterogeneous data, leading to substantial human resource consumption for calculations and posing challenges in ensuring the accuracy and timeliness of performance assessments. The order performance analysis is illustrated in Fig. 2.

This research focuses primarily on four aspects influencing the performance of orders:

(1) Economic Benefits: The profit margin brought to the enterprise by the order, calculated from relevant data such as order revenue, production costs, and logistics costs.

(2) Production Efficiency: The degree of utilization of production resources during the order production process, typically calculated by production monitoring sensor equipment.

(3) Time Efficiency: From the time the order is placed to the time it is completed, time primarily includes production time and logistics time.

(4) Customer Satisfaction: User comments on the order after completion, mostly in the form of textual information.

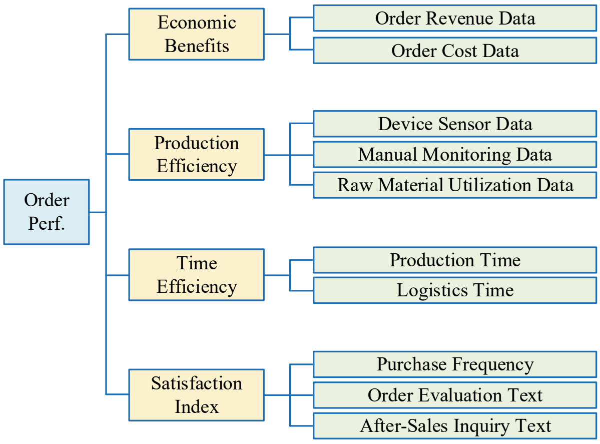


**Figure 2** Order performance analysis diagram

## 3.2 Sentiment Analysis Tool

Text sentiment analysis tools are widely used in the data text analysis of post-order evaluations in the manufacturing industry. These evaluations involve multiple aspects such as product quality, service attitude, etc., with complex sentiment expressions and varying text lengths, and may also involve domain expertise. Through these tools, companies can quickly and accurately understand customers' sentiment tendencies, adjust product and service strategies in a timely manner, and enhance customer satisfaction and brand reputation.

The performance comparison of the most common text sentiment analysis tools currently available is shown in Table 1. In the context of this study, order review data often appear in the form of short texts, which do not require the use of computationally intensive methods for calculation. Considering that the data fusion process requires sentiment index data rather than simple negative, neutral, and positive data, while ensuring the reliability of the results, we comprehensively consider using VADER for text sentiment analysis to achieve feature extraction.

**Table 1** Comparison of sentiment analysis tools

| Tools | Short text accuracy | Speed | Difficulty |
|---|---|---|---|
| NLTK | Middle | Fast | Low |
| TextBlob | Middle | Fast | Low |
| VADER | High | Fast | Middle |
| LSTM | High | Low | High |
| CNN | High | Low | High |
| BERT | High | Low | High |

## 3.3 Model Construction

### 3.3.1 SSA-CNN-LSTM Model

This paper combines the SSA optimization algorithm with CNN and LSTM models. The preprocessing stage of the model utilizes CNN, a combination of convolutional layers, batch normalization layers, activation function layers, and pooling layers. This combination is effective in reducing data dimensionality, extracting key features, and reducing training time. CNN processes a large amount of high-dimensional heterogeneous data, providing more refined and targeted input for the subsequent LSTM neural network. In the fusion stage, an LSTM model with multiple feature inputs and a single feature output is used for feature integration [20, 21]. LSTM has the ability to handle time series data with long short-term memory, effectively addressing the issues of gradient explosion and vanishing gradient through the control of forget gates, input gates, and output gates. This allows the neural network to better capture information changes in time series. However, parameter tuning is required.

SSA simulates the collective behavior of sparrows in activities such as foraging, socializing, and evading predators to find the optimal solution. The core idea of the algorithm includes local search and global search. Local search explores local optimal solutions through the cooperation of individuals, while global search seeks global optimal solutions through random movements of individuals. This algorithm is novel and has strong optimization capabilities with a fast convergence rate. It can be used for parameter tuning of the LSTM model [22,23]. The computational process and structure of the model are shown in Figs. 3 and 4 respectively. The following is the computational process of the SSA-CNN-LSTM model:

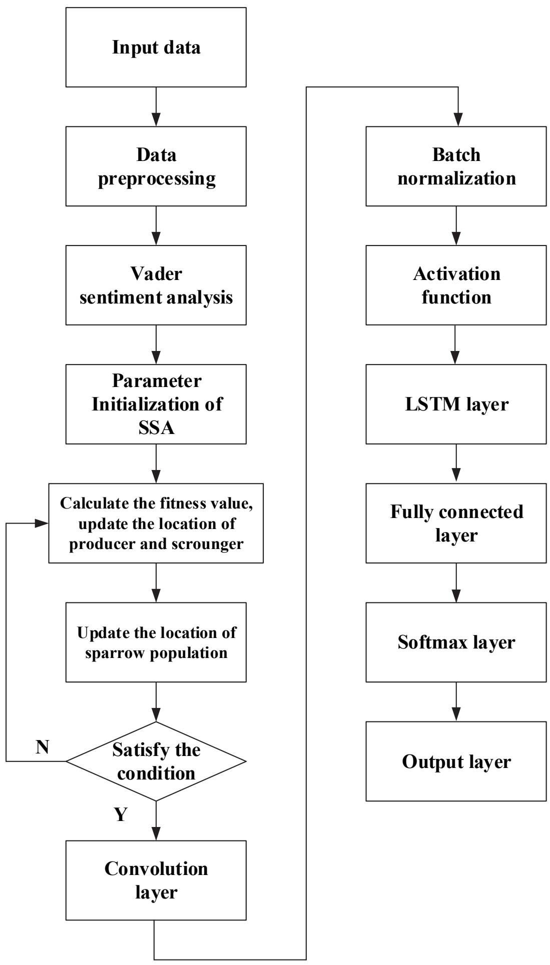


**Figure 3** Model computation process diagram

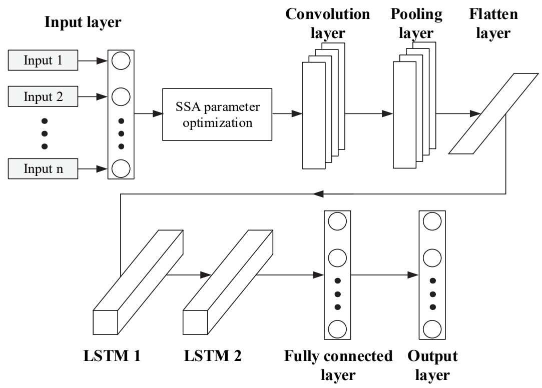

**Figure 4** Model structure diagram

(1) Data Processing

First, the raw data undergoes labeling, partitioning, normalization, and format conversion. Subsequently, the VADER sentiment analysis tool [24] is utilized to recognize the positive sentiment in textual data. It provides a sentiment index ranging between –1 and 1 for each text. A score closer to 1 indicates a higher level of user satisfaction for the respective order, while a score closer to -1 suggests a lower level of user satisfaction. Compared to methods like Bayesian and LSTM text sentiment analysis, VADER operates using a pre-trained model, offering the advantages of reduced computational burden and faster processing speed. Additionally, it is more suitable for short texts.

(2) SSA parameter tuning

1)SSA parameter initialization settings: The number of sparrows (*n*) corresponds to the number of samples in the dataset; the number of producers (*PD*) corresponds to the number of CNN convolutional layers, representing the complexity of feature extraction; the number of sparrows perceiving danger (*SD*) corresponds to the part of the model used for detecting anomalies or noise, reflecting the quantity of abnormal situations in the data; the safety threshold (*ST*) corresponds to the sensitivity setting of the model to anomalies during training or prediction, serving as a reference value to determine whether to trigger an alarm; the alarm value ($R_2$) corresponds to the value in the model used to determine whether to trigger an alarm, activating the alarm when the perceived anomalies exceed this value.

2) Updater Position $X_{i,j}^{t+1}$ : Update the weights of the convolutional neural network to further optimize feature extraction performance.

$$X_{i,j}^{t+1} = \begin{cases} X_{i,j}^{t} \cdot \exp\left(\dfrac{-i}{\alpha \cdot iter_{\max}}\right) R_2 < ST \\ X_{i,j}^{t} + Q \cdot L \end{cases} \tag{1}$$

3) Updater's Position Update: Adjusting the weights of the LSTM neural network to fine-tune the network parameters and enhance learning performance on time series data.

$$X_{i,j}^{t+1} = \begin{cases} Q \cdot \exp\left(\dfrac{X_{worst}^{t} - X_{i,j}^{t}}{i^2}\right) i > n/2 \\ X_P^{t+1} + \left|X_{i,j}^{t} - X_P^{t+1}\right| \cdot A^{+} \cdot L \end{cases} \tag{2}$$

4) Updater's Awareness of Hazardous Sparrow Position: Adjusting the anomaly detection segment of the CNN-LSTM model to enhance its ability to handle anomalies and noise.

$$X_{i,j}^{t+1} = \begin{cases} X_{best}^{t} + \beta \cdot \left|X_{i,j}^{t} - X_{best}^{t}\right| f_i > f_g \\ X_{i,j}^{t} + K \cdot \left(\dfrac{\left|X_{i,j}^{t} - X_{worst}^{t}\right|}{\left(f_i - f_w\right) + \varepsilon}\right) f_i = f_g \end{cases} \tag{3}$$

5) Termination Condition: If satisfied, output the sparrow position; if not satisfied, return to step (3) where $X_{i,j}^{t+1}$ represents the parameter information of the j-th dimension in the model parameters for the *i*-th sample at the *t*-th iteration; $iter_{\max}$ is the maximum number of iterations; $\alpha$ is a random number in the range [0, 1]; $R_2$ is a random number in the range [0, 1] representing the alert value; *ST* is a constant in the range [0.5, 1] representing the safety value; *Q* is a random number following a normal distribution; *L* is a 1 × *d* matrix of all ones; $X_{worst}^{t}$ is the worst parameter in the *t*-th iteration; $X_p$ is the current optimal parameter; *A* is a 1 × *d* matrix with elements randomly assigned as 1 or –1; $\beta$ is a random number following a normal distribution with mean 0 and variance 1; *K* is a random number in the range [–1, 1], where the sign indicates the direction for finding the optimal solution, and the magnitude represents the step size control parameter; $f_i$ is the fitness value of the parameter; $f_g$ is the current maximum fitness value; $f_w$ is the current minimum fitness value.

(3) Convolutional Neural Network Model

Through convolutional operations, the CNN model extracts local features using different convolutional kernels, capturing crucial information from input data. The application of the ReLU activation function enhances the model's expressive power. Pooling layers are used to downsample the feature maps, reducing spatial dimensions while retaining essential features. This process is effective in extracting high-level abstract representations from multiple data sources, providing more meaningful inputs for subsequent model layers [25, 26]. The CNN architecture is illustrated in Fig. 5.

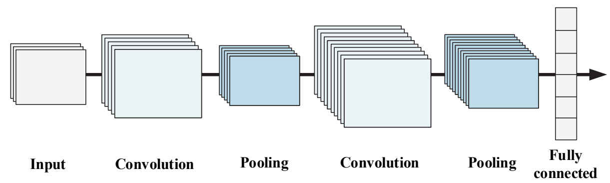

**Figure 5** CNN structure diagram

(4) Long Short-Term Memory Model

The LSTM neural network controls the degree of memory and forgetting of previous and current data through forget gates, input gates, and output gates, giving the neural network the capability of long short-term memory. Upon entering the LSTM, the forget gate $f_t$ calculates by applying the Sigmoid activation function layer to the previous hidden state $h_{t-1}$ and the current input value $x_t$. The previous hidden state $h_{t-1}$ and the current input value $x_t$ generate the input gate $i_t$ through a Sigmoid activation function layer and a candidate cell state $\hat{c}_t$ through an activation function layer. The current cell state $c_t$ is calculated from the previous cell state $c_{t-1}$, the new cell state $\hat{c}_t$, and the input gate $i_t$. The output gate $o_t$ is computed by applying the Sigmoid activation function to the previous hidden state $h_{t-1}$ and the current input value $x_t$. Finally, the output gate $o_t$ and the current cell state $c_t$ are calculated to obtain the current hidden state [27, 28]. The LSTM architecture is depicted in Fig. 6.

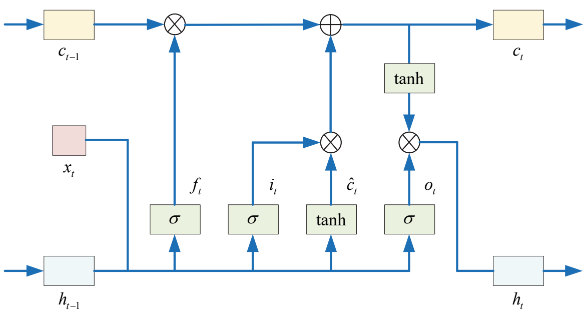


**Figure 6** LSTM structure diagram

1) The input gate is used to control the extent to which the current computational state is updated to the memory cell. The calculation formula is as follows:

$$\hat{c}_t = \tanh\left(W_c x_t + U_c h_{t-1} + b_c\right) \tag{4}$$

$$i_t = \sigma\left(W_i x_t + U_i h_{t-1} + b_i\right) \tag{5}$$

where $x_t$ represents the input value; $h_{t-1}$ represents the hidden state; $W_c$, $U_c$, $W_i$, $U_i$ are weight matrices; $b_c$, $b_i$ are biases.

2) The forget gate is used to control the extent to which input and the current computational state are updated to the memory cell. The calculation formula is as follows:

$$f_t = \sigma\left(W_f x_t + U_f h_{t-1} + b_f\right) \tag{6}$$

where $x_t$ represents the input value; $h_{t-1}$ represents the hidden state; $W_f$, $U_f$ are weight matrices; $b_f$ is the bias.

3) The state is deleted and updated by the forget gate and input gate. The calculation formula is as follows:

$$c_t = f_t c_{t-1} + i_c \hat{c}_t \tag{7}$$

4) The output gate is used to control the input and current output based on the current memory cell's degree, and the formula is as follows:

$$o_t = \sigma\left(W_o x_t + U_o h_{t-1} + b_o\right) \tag{8}$$

$$h_t = o_t \cdot \mathrm{tanc}_t \tag{9}$$

where $x_t$ represents the input value; $h_{t-1}$ represents the hidden state; $W_o$, $U_o$ are weight matrices; $b_o$ is the bias.

### 3.3.2 Model Parameter Configuration

The learning rate and the number of hidden layer nodes are crucial parameters in neural networks, significantly impacting training outcomes. The learning rate regulates the model's learning progress, with excessively high values making convergence difficult and causing the network to oscillate around the optimal values. Conversely, a too-low learning rate results in slow convergence, prolonging the time needed to find optimal values. In this study, the initial learning rate for the LSTM neural network is set to 0.0035.

The number of hidden layer nodes influences the model's performance. Too many nodes increase training time and the risk of overfitting, while too few nodes may hinder successful learning, requiring more training cycles and impacting training accuracy. In this paper, the LSTM model's two hidden layers are configured with 128 and 30 nodes, respectively. As epochs increase, network parameters continuously update to find optimal values. Tab. 2 shows five sets of optimal parameter combinations obtained through repeated parameter tuning and the key parameters for the SSA-CNN-LSTM model are presented in Tabs. 3, 4, and 5.

**Table 2** Optimal hyperparameter combinations

| Group | Convolutional kernel | LSTM 1 hidden layer | LSTM 2 hidden layer | Learning rate | RMSE |
|---|---|---|---|---|---|
| 1 | 32 | 128 | 32 | 0.0035 | 0.3352 |
| 2 | 32 | 112 | 48 | 0.0035 | 0.3355 |
| 3 | 40 | 192 | 32 | 0.0035 | 0.3355 |
| 4 | 40 | 112 | 48 | 0.0035 | 0.3356 |
| 5 | 40 | 96 | 96 | 0.0001 | 0.3961 |

**Table 3** CNN parameters

| Parameters | Settings |
|---|---|
| CNN layers | 1 |
| Convolutional kernel size | 10 × 1 |
| Number of convolution kernels | 32 |
| Stride | 1 |
| Activation function | ReLU |
| Dropout | 0.3 |

**Table 4** LSTM parameters

| Parameters | Settings |
|---|---|
| LSTM layers | 2 |
| Learning rate | 0.0035 |
| Number of hidden layer nodes | 128, 30 |
| Activation function | Tanh and $\sigma$ |
| Epochs | ReLU |

**Table 5** SSA parameters

| Parameters | Settings |
|---|---|
| The numbers of polulations | 2 |
| The maximum iterations | 0.0035 |
| The number of producers | 128, 30 |
| The number of scroungers | Tanh and $\sigma$ |
| The sparrow that senses danger | ReLU |
| Safety value | 0.8 |
| Lower limit range of parameters | $(1 \times 10^{-10}, 10)$ |
| Upper limit range of parameters | $(1 \times 10^{-2}, 200)$ |

# 4 EXPERIMENTAL RESULTS AND DISCUSSION

## 4.1 Dataset

This paper utilizes a custom dataset collected from a specific type of smart mobile devices, consisting of 972 entries. The data is sorted based on the order generation time, with the output for evaluation being the order fulfillment, i.e., order performance data. The input data includes four sets of information: order profit margin, order production utilization, order delivery cycle, and user review text. The profit margin is calculated and presented as a percentage. The production utilization is calculated from sensor data collected during the production process, also presented as a percentage. The order delivery cycle represents the time required for an order from generation to completion, measured in hours. The user review text is composed of English characters, creating a heterogeneous dataset. Some sample data is presented in Tab. 6.

**Table 6** Partial dataset examples

| Order ID | Profit Margin | Production Utilization Rate | Delivery Cycle | Review | Order Fulfillment Rate |
|---|---|---|---|---|---|
| 1 | 23.3 | 89.06 | 55 | Terrible product | 83.3132 |
| 2 | 20.9 | 95 | 37 | Fabulous | 89.1057 |
| 3 | 22.62 | 90.72 | 47 | Slightly disappointed | 85.9269 |
| 4 | 24 | 94.94 | 32 | Good choice | 91.9022 |
| 5 | 25 | 88.96 | 28 | Perfect product! | 92.3814 |

Preprocessing the dataset involves removing missing and outlier data to ensure the reliability of each entry. The user review texts are subjected to sentiment analysis using the VADER tool, generating corresponding sentiment scores within the range of [–1, 1]. A sentiment score closer to -1 indicates lower satisfaction, while a score closer to 1 indicates higher satisfaction.

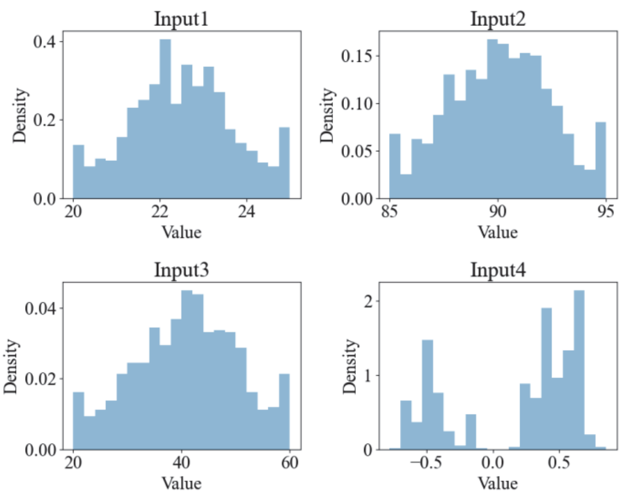


**Figure 7** Feature distribution

**Table 7** Partial processed data

| Order ID | Profit Margin | Production Utilization Rate | Delivery Cycle | Customer Satisfaction | Order Fulfillment Rate |
|---|---|---|---|---|---|
| 1 | 23.3 | 89.06 | 55 | –0.4767 | 83.3132 |
| 2 | 20.9 | 95 | 37 | 0.5707 | 89.1057 |
| 3 | 22.62 | 90.72 | 47 | –0.4228 | 85.9269 |
| 4 | 24 | 94.94 | 32 | 0.4404 | 91.9022 |
| 5 | 25 | 88.96 | 28 | 0.6114 | 92.3814 |

Ultimately, 800 processed entries are selected as model input, with 720 entries used for training and 80 for testing. A portion of the processed data is illustrated in Tab. 7. The distribution of dataset features is shown in Fig. 7.

## 4.2 Evaluation Metrics

This paper selects the statistical metrics Root Mean Square Error (*RMSE*), Mean Absolute Percentage Error (*MAPE*), and Mean Absolute Error (*MAE*) to assess and compare evaluation accuracy. Smaller values for these metrics indicate higher accuracy. The formulas are as follows:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\hat{y}_i - y_i\right)^2} \tag{10}$$

$$MAPE = \frac{100\%}{n}\sum_{i=1}^{n}\left|\frac{\hat{y}_i - y_i}{y_i}\right| \tag{11}$$

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left|\hat{y}_i - y_i\right| \tag{12}$$

where $\hat{y}_i$ represents the assessment value of the fusion result, and $y_i$ represents the ground truth.

## 4.3 Ablation Experiment

Test LSTM, CNN-LSTM, and SSA-CNN-LSTM models separately using datasets. The experimental results show that all three models can relatively accurately assess order performance. The evaluation performance indicators for each method are shown in Tab. 6. Compared to the LSTM model, the CNN-LSTM model and the SSA-CNN-LSTM model reduce RMSE by 21.93% and 58.71%, MAE by 18.26% and 62.94%, and MAPE by 17.72% and 63.29%, respectively. From Tab. 8, it can be observed that the proposed method in this paper achieves optimal results in the comparison of the three evaluation metrics compared to other assessment methods. This indicates the superiority of the proposed method.

**Table 8** Ablation experiment performance comparison

| Model | RMSE | MAE | MAPE |
|---|---|---|---|
| LSTM | 0.8107 | 0.6938 | 0.0079 |
| CNN-LSTM | 0.6410 | 0.5670 | 0.0065 |
| SSA-CNN-LSTM | 0.3352 | 0.2573 | 0.0029 |

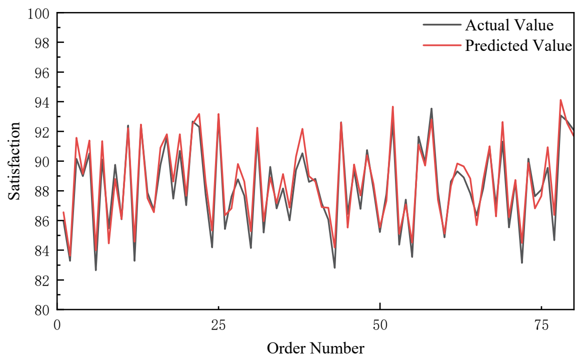


**Figure 8** LSTM assessment result line chart

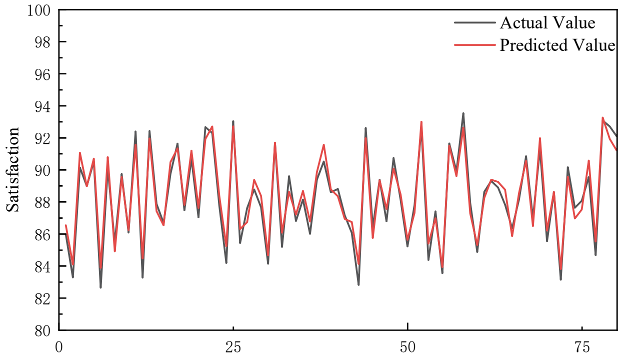


**Figure 9** CNN-LSTM assessment result line chart

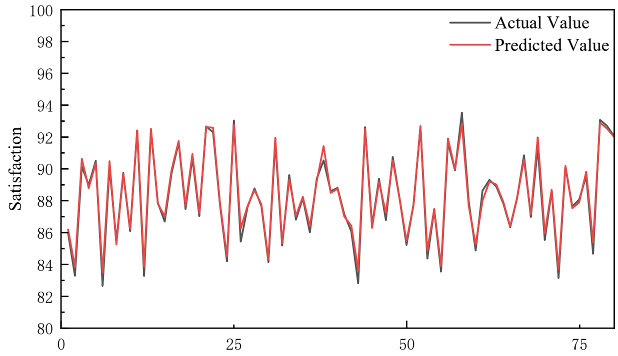


**Figure 10** SSA-CNN-LSTM assessment result line chart

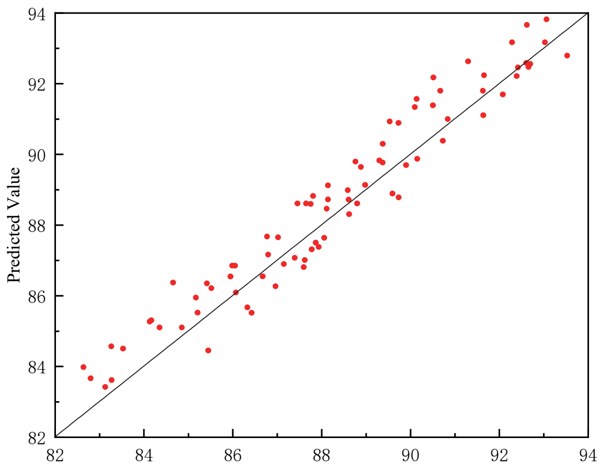


**Figure 11** LSTM assessment result scatter plot

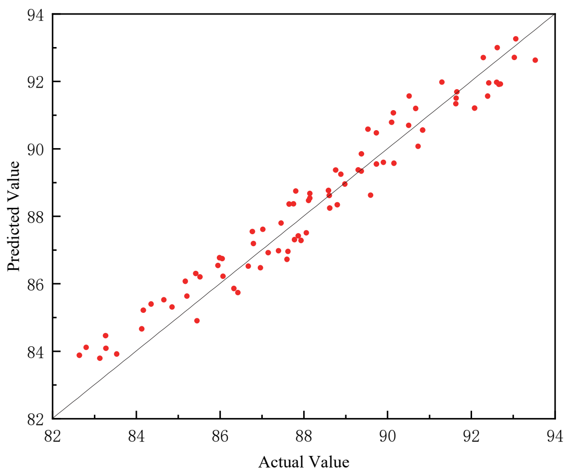


**Figure 12** CNN-LSTM assessment result scatter plot

The evaluation result line charts are shown in Figs. 8, 9, 10, and the scatter plots are displayed in Figs. 11, 12, 13. It can be seen that the assessment curves of the proposed method fit well with the real curves, exhibiting smaller overall errors and superior assessment results compared to other methods. The error comparison chart for the three models is illustrated in Fig. 14.

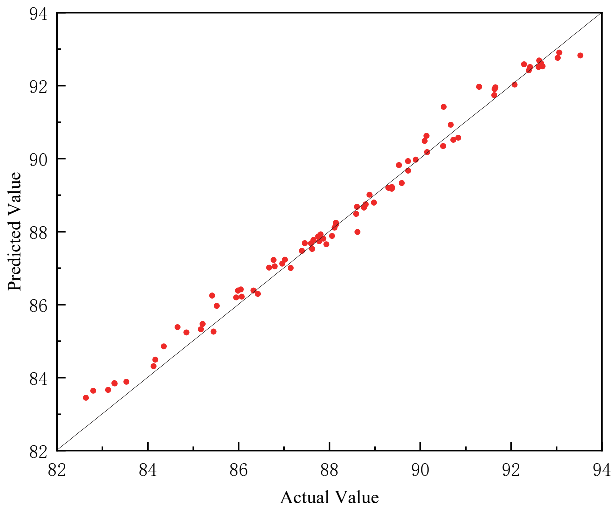


**Figure 13** SSA-CNN-LSTM assessment result scatter plot

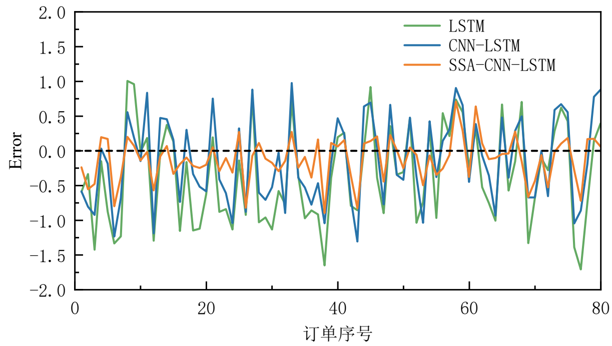


**Figure 14** Error comparison chart for three models

## 4.4 Comparative Experiment

To assess the robustness of the model against missing values and noise in the input data, robustness evaluation experiments were conducted. To highlight the superiority of the SSA-CNN-LSTM model, it was compared against several commonly used models for multi-source heterogeneous data fusion in experimental trials. The experimental results, as shown in Tab. 9, demonstrate that the fusion accuracy of the SSA-CNN-LSTM model is higher compared to MLP, RNN, and CNN-RNN models.

**Table 9** Comparative experiment performance comparison

| Model | RMSE | MAE | MAPE |
|---|---|---|---|
| MLP | 0.9183 | 0.8098 | 0.0087 |
| RNN | 0.7812 | 0.7027 | 0.0071 |
| CNN-RNN | 0.5136 | 0.4830 | 0.0055 |
| SSA-CNN-LSTM | 0.3352 | 0.2573 | 0.0029 |

The number of parameters (Params) and floating-point operations (FLOPs) are considered important indicators for evaluating model complexity. Params represent the number of parameters in the model, while FLOPs indicate the computational complexity of the model. The complexity of the three networks is shown in Tab. 10. The proposed model exhibits the lowest Params and FLOPs, highlighting the effectiveness of its parameter optimization functionality. The lower Params and FLOPs also indicate that its computational and storage requirements are relatively low when dealing with larger-scale data, thus effectively adapting to larger datasets.

**Table 10** Model complexity comparison

| Model | Params (M) | Flops (G) |
|---|---|---|
| MLP | 22.32 | 10.29 |
| RNN | 25.98 | 11.67 |
| CNN-RNN | 28.12 | 4.23 |
| SSA-CNN-LSTM | 11.74 | 3.88 |

### 4.5 Robustness Experiment

To assess the robustness of the model against missing values and noise in the input data, robustness evaluation experiments were conducted. The input data was randomly augmented with missing values and noise, divided into two groups. In the first group, 40 missing values and 80 noise data points were added to each of the four sets of inputs, with the noise fluctuating within ±20% range of the original values. In the second group, 80 missing values and 160 noise data points were added to each of the four sets of inputs, with the noise fluctuating within ±50% range of the original values. These two sets of data were separately used to evaluate the SSA-CNN-LSTM model, and the experimental results are shown in Fig. 15 and Fig. 16.

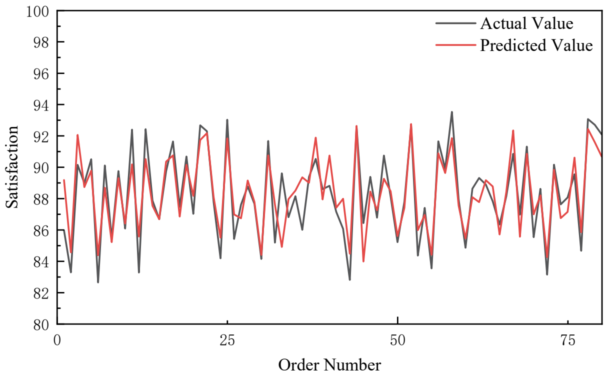


**Figure 15** Experiment 1 results for noise

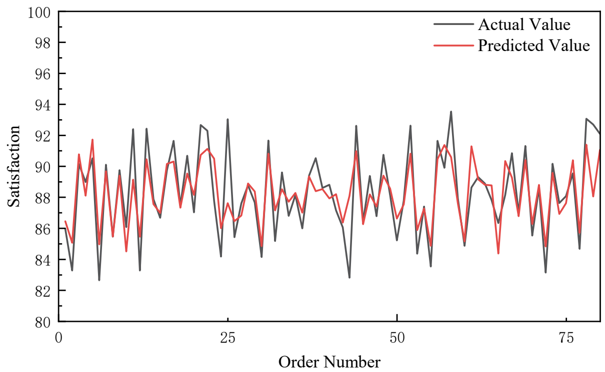


**Figure 16** Experiment 2 results for noise

The results of the robustness experiment indicate that, under the influence of a small to moderate amount of missing values and noise, the model can still roughly evaluate performance indicators. However, it imposes high demands on data quality, as its fusion calculation capability relies on the integrity of multiple data sources. Each input has an indispensable impact on the results; thus, the presence of missing values and noise significantly reduces the accuracy of the evaluation results. Similarly, its adaptability to changes in data characteristics is relatively low; the model may fail when encountering concept drift. Therefore, it is necessary to regularly update the training model based on the application scenario to adapt to the continuously changing data characteristics.

## 5 CONCLUSION

This paper presents an SSA-CNN-LSTM model for the fusion of multi-source heterogeneous data in order performance evaluation. The proposed model integrates CNN for feature extraction and LSTM for learning data feature associations, with SSA employed to optimize the model parameters. Experimental results using an order dataset from a custom smart mobile devices manufacturer demonstrate the superiority of the SSA-CNN-LSTM model compared to the LSTM and CNN-LSTM models in terms of evaluation accuracy. The proposed model's ability to handle complex multidimensional data and achieve higher accuracy highlights its potential for improving operational efficiency in enterprise performance evaluation tasks. However, the study has some limitations, such as the need for a larger sample size to fully learn data features and the limited types of data considered in the multi-source heterogeneous dataset. Future research directions include exploring methods to enhance the model's accuracy for unknown events, employing more advanced deep learning models for sentiment analysis on complex texts, expanding the dataset to include higher-dimensional data, and optimizing the model structure to reduce training time.

## 6 REFERENCES

[1] Merschmann, U. & Thonemann, U. W. (2011). Supply chain flexibility, uncertainty and firm performance: An empirical analysis of German manufacturing firms. *International Journal of Production Economics*, *130*(1), 43-53. https://doi.org/10.1016/j.ijpe.2010.10.013

[2] Feizabadi, J. (2022). Machine learning demand forecasting and supply chain performance. *International Journal of Logistics Research and Applications*, *25*(2), 119-142. https://doi.org/10.1080/13675567.2020.1803246

[3] Stewart, G. (1995). Supply chain performance benchmarking study reveals keys to supply chain excellence. *Logistics information management*, *8*(2), 38-44. https://doi.org/10.1108/09576059510085000

[4] Kaplan, R. S. (2009). Conceptual foundations of the balanced scorecard. *Handbooks of management accounting research*, *3*, 1253-1269. https://doi.org/10.1016/S1751-3243(07)03003-9

[5] Zhu, L. (2022). Research and application of AHP-fuzzy comprehensive evaluation model. *Evolutionary Intelligence*, *15*(4), 2403-2409. https://doi.org/10.1007/s12065-020-00415-7

[6] Ji, Y. B. & Lee, C. (2010). Data envelopment analysis. *The Stata Journal*, *10*(2), 267-280. https://doi.org/10.1177/1536867X1001000207

[7] Vaidya, O. S. & Kumar, S. (2006). Analytic hierarchy process: An overview of applications. *European Journal of operational research*, *169(*1), 1-29. https://doi.org/10.1016/j.ejor.2004.04.028

[8] Zhang, L., Xie, Y., Xidao, L., & Zhang, X. (2018, May). Multi-source heterogeneous data fusion. *2018 International*

*conference on artificial intelligence and big data (ICAIBD)*, 47-51. https://doi.org/10.1109/ICAIBD.2018.8396165

[9] Gao, J., Li, P., Chen, Z., & Zhang, J. (2020). A survey on deep learning for multimodal data fusion. *Neural Computation*, *32*(5), 829-864. https://doi.org/10.1162/neco_a_01273

[10] Zhang, H., Cheng, C., Xu, Z., & Li, J. (2020). A Comprehensive Review of Data Fusion Methods Based on Deep Learning. *Computer Engineering and Application*, (24), 1-11. https://doi.org/10.3778/j.issn.1002-8331.2007-0475

[11] Maimaitijiang, M., Sagan, V., Sidike, P., Hartling, S., Esposito, F., & Fritschi, F. B. (2020). Soybean yield prediction from UAV using multimodal data fusion and deep learning. *Remote sensing of environment*, *237*, 111599. https://doi.org/10.1016/j.rse.2019.111599

[12] Chen, F. C. & Jahanshahi, M. R. (2017). NB-CNN: Deep learning-based crack detection using convolutional neural network and Naïve Bayes data fusion. *IEEE Transactions on Industrial Electronics*, *65*(5), 4392-4400. https://doi.org/10.1109/TIE.2017.2764844

[13] Khan, S., Nazir, S., García-Magariño, I., & Hussain, A. (2021). Deep learning-based urban big data fusion in smart cities: Towards traffic monitoring and flow-preserving fusion. *Computers & Electrical Engineering*, *89*, 106906. https://doi.org/10.1016/j.compeleceng.2020.106906

[14] Hatami, N., Cho, T. H., Mechtouff, L., Eker, O. F., Rousseau, D., & Frindel, C. (2022, July). CNN-LSTM based multimodal MRI and clinical data fusion for predicting functional outcome in stroke patients. *2022 44th Annual International Conference of the IEEE Engineering in Medicine & Biology Society (EMBC)*, 3430-3434. https://doi.org/10.1109/EMBC48229.2022.9871735

[15] Zhang, H., Zhang, Y., & Xu, Z. (2022, January). Thermal Load Forecasting of an Ultra-short-term Integrated Energy System Based on VMD-CNN-LSTM. *2022 International Conference on Big Data, Information and Computer Network (BDICN)*, 264-269. https://doi.org/10.1109/BDICN55575.2022.00058

[16] Li, Z., Li, Y., Liao, C., Zhang, C., Wang, L., & Wang, L. (2023, April). State-of-Health Estimation of Lithium-ion Batteries Based on WOA-CNN-LSTM-Attention. *2023 8th International Conference on Cloud Computing and Big Data Analytics (ICCCBDA)*, 572-578. https://doi.org/10.1109/ICCCBDA56900.2023.10154659

[17] Lu, W., Rui, H., Liang, C., Jiang, L., Zhao, S., & Li, K. (2020). A method based on GA-CNN-LSTM for daily tourist flow prediction at scenic spots. *Entropy*, *22*(3), 261. https://doi.org/10.3390/e22030261

[18] Lima-Junior, F. R. & Carpinetti, L. C. R. (2017). Quantitative models for supply chain performance evaluation: A literature review. *Computers & Industrial Engineering*, *113*, 333-346. https://doi.org/10.1016/j.cie.2017.09.022

[19] Gopal, P. R. C. & Thakkar, J. (2012). A review on supply chain performance measures and metrics: 2000-2011. *International journal of productivity and performance management*, *61*(5), 518-547. https://doi.org/10.1108/17410401211232957

[20] Kim, T. Y. & Cho, S. B. (2019). Predicting residential energy consumption using CNN-LSTM neural networks. *Energy*, *182*, 72-81. https://doi.org/10.1016/j.energy.2019.05.230

[21] Shao, X., Kim, C. S., & Sontakke, P. (2020). Accurate deep model for electricity consumption forecasting using multi-channel and multi-scale feature fusion CNN-LSTM. *Energies*, *13*(8), 1881. https://doi.org/10.3390/en13081881

[22] Xue, J. & Shen, B. (2020). A novel swarm intelligence optimization approach: sparrow search algorithm. *Systems science & control engineering*, *8*(1), 22-34. https://doi.org/10.1080/21642583.2019.1708830

[23] Yue, Y., Cao, L., Lu, D., Hu, Z., Xu, M., Wang, S., ... & Ding, H. (2023). Review and empirical analysis of sparrow search algorithm. *Artificial Intelligence Review*, 1-53. https://doi.org/10.1007/s10462-023-10435-1

[24] Hutto, C. & Gilbert, E. (2014, May). Vader: A parsimonious rule-based model for sentiment analysis of social media text. *Proceedings of the international AAAI conference on web and social media*, *8*(1), 216-225. https://doi.org/10.1609/icwsm.v8i1.14550

[25] Li, Z., Liu, F., Yang, W., Peng, S., & Zhou, J. (2021). A survey of convolutional neural networks: analysis, applications, and prospects. *IEEE transactions on neural networks and learning systems.* https://doi.org/10.1109/TNNLS.2021.3084827

[26] Gu, J., Wang, Z., Kuen, J., Ma, L., Shahroudy, A., Shuai, B., ... & Chen, T. (2018). Recent advances in convolutional neural networks. *Pattern recognition*, *77*, 354-377. https://doi.org/10.1016/j.patcog.2017.10.013

[27] Yu, Y., Si, X., Hu, C., & Zhang, J. (2019). A review of recurrent neural networks: LSTM cells and network architectures. *Neural computation*, *31*(7), 1235-1270. https://doi.org/10.1162/neco_a_01199

[28] Abbasimehr, H., Shabani, M., & Yousefi, M. (2020). An optimized model using LSTM network for demand forecasting. *Computers & industrial engineering*, *143*, 106435. https://doi.org/10.1016/j.cie.2020.106435

**Contact information:**

**Wei QIANG**
(Corresponding author)
Northwestern Polytechnical University school of computing (CN),
127 Youyi West Road, Xi'an City
E-mail: 13909212121@163.com

**Yihan FANG**
School of Mechanical and Electrical Engineering,
Beijing Institute of Printing and Technology (CN),
No. 1, Section 2, Xinghua Street, Daxing District, Beijing
E-mail: fyh006@126.com