World Journal of Engineering Research
Print ISSN: 2959-9865
Online ISSN: 2959-9873
DOI: https://doi.org/10.61784/wjer3039

LOGISTICS SORTING OPTIMIZATION BASED ON MACHINE LEARNING

Yao Xue*, WenJing Huang, WangYi Xu
Southwest University of Science and Technology, Mianyang 621010, Sichuan, China.
Corresponding Author: Yao Xue, Email: 18382469606@163.com

Abstract: The aim of this research is to achieve accurate prediction of the cargo volume in a logistics sorting center and efficient allocation of personnel. To this end, various machine learning algorithms such as Random Forest, Multilayer Perceptron Regressor (MLPRegressor), Support Vector Machine, and Multiple Linear Regression are comprehensively applied to deeply mine the historical cargo volume data of the sorting center, and features such as lagged cargo volume and moving average cargo volume are constructed. The grid search method is used for parameter optimization, and the optimal model is selected for future cargo volume prediction. On this basis, based on the prediction results, methods such as Linear Programming (LP), graph theory models, and queuing theory are utilized to construct a model for the optimal allocation of personnel. Taking into account factors such as employees' work efficiency and attendance rules, the rational deployment of personnel in the sorting center is realized. The experimental results show that the constructed cargo volume prediction model has a high accuracy, and the optimized personnel allocation scheme is reasonable and feasible. The conclusion indicates that the multi-algorithm fusion strategy can effectively improve the accuracy of cargo volume prediction and the efficiency of personnel allocation in the logistics sorting center, providing strong support for the intelligent management of the sorting center.
Keywords: Cargo volume prediction; Staffing allocation; Machine learning; Multi-objective optimization

1 INTRODUCTION

1.1 Research Background

In the modern logistics system, as a crucial hub for the rapid circulation of goods, the sorting center's operational efficiency directly affects the smoothness of the entire logistics chain. With the vigorous development of e-commerce, the volume of logistics orders has experienced explosive growth, which places higher demands on the accuracy of cargo volume prediction and the rationality of personnel allocation in the sorting center. If there is a large deviation in cargo volume prediction, it may lead to either idle resources or resource shortages. In addition, unreasonable personnel allocation will reduce the sorting efficiency.

Early cargo volume predictions mostly relied on simple statistical methods, such as the moving average method and linear regression. These methods were difficult to capture the complex and changeable patterns of cargo volume fluctuations, and thus the prediction accuracy was limited. In recent years, machine learning algorithms have gradually been applied in this field. Xu et al. proposed a time series prediction model based on deep learning and achieved good results in logistics order prediction, demonstrating the advantages of deep neural networks in handling complex nonlinear relationships[1-2]. Additionally, Ghosh et al. developed a simulation-based optimization algorithm for scheduling the operation of the sorting center, which utilized real-time prediction and dynamic scheduling to maximize the throughput[3]. However, a single algorithm is restricted by the characteristics of the data and the limitations of the model itself, making it difficult to fully adapt to complex logistics scenarios.

Regarding personnel allocation, traditional methods mostly rely on experience and lack scientific planning. Some studies have introduced mathematical models. For example, Zhang et al. used the ARIMA model to predict the cargo volume in the logistics network center and achieved good results[4]. Aiming at the above problems of cargo volume prediction and personnel allocation, the objective is to overcome the deficiencies of existing studies. By integrating the advantages of multiple machine learning algorithms, a more accurate cargo volume prediction model and a more scientific personnel allocation optimization model are constructed to improve the operation efficiency of the logistics sorting center and achieve the efficient utilization of resources.

1.2 Research Objectives and Technical Route

Aiming at the above problems, the objective is to integrate multiple machine learning algorithms to improve the accuracy of cargo volume prediction in logistics sorting centers, and construct an optimized personnel allocation scheme based on the prediction results. The main objectives include: Accurate Cargo Volume Prediction:

Comprehensively utilize machine learning algorithms such as Random Forest (RF), Multilayer Perceptron Regressor (MLPRegressor), Support Vector Machine (SVM), and Multiple Linear Regression (MLR) to mine the historical cargo volume data of the sorting center, and construct a series of prediction features (such as lagged cargo volume, moving average cargo volume, time-series features, etc.). Employ Grid Search to optimize hyperparameters, and evaluate the accuracy of the prediction model using Root Mean Squared Error (RMSE)[5].

Intelligent Personnel Optimization and Allocation: Based on the cargo volume prediction results, construct optimization models of Linear Programming (LP), Graph Theory, and Queuing Theory. On the basis of considering factors such as employees' work efficiency, attendance rules, and shift constraints, optimize the personnel allocation scheme to achieve scientific scheduling.

Model Validation and Performance Evaluation: Validate the effectiveness of the proposed cargo volume prediction model and personnel optimization scheme through actual datasets, and evaluate their applicability in different logistics scenarios.

The contribution lies in integrating machine learning and optimization methods to achieve accurate prediction of cargo volume and intelligent allocation of human resources in the sorting center, providing a feasible technical solution for intelligent logistics scheduling. The research results can offer data-driven decision support for logistics enterprises, further enhancing the operational efficiency of the logistics system and the utilization rate of resources.

2 DATA AND RESEARCH METHODS

2.1 Data Acquisition and Preprocessing

The research data for this study are sourced from the open-source website (http://mathorcup.org). It has collected data from 57 sorting centers within the logistics network. Specifically, the daily cargo volume of each sorting center over the past four months, the hourly cargo volume over the past 30 days, the average cargo volume of each transportation route between the sorting centers over the past 90 days, and data assuming that changes will occur in the transportation routes between the sorting centers in the coming 30 days have been gathered.

Preliminary inspection reveals that the collected data suffer from format chaos and lack of temporal dependence in the recording order. To ensure the temporal consistency of the data and the accuracy of subsequent modeling, it is essential to conduct strict temporal sorting of the data. Further data exploration shows that there are significant differences in the total cargo volumes of different sorting centers, and their fluctuations within different dates are particularly evident. Especially in November, affected by promotional activities, the amplitude of fluctuations increases significantly, resulting in large differences in the mean and variance of each center. At the same time, the collected data also have the problem of missing values. To ensure data continuity and integrity and improve the accuracy of the prediction model, an interpolation algorithm is employed to reasonably estimate and complete the missing data. After strict data preprocessing, the data quality is significantly improved, laying a solid foundation for the subsequent construction of a prediction model and optimization algorithm with universality and robustness.

2.2 Introduction of the Methods

2.2.1 Problem analysis of cargo volume prediction in logistics sorting centers

In the field of cargo volume prediction, commonly used models include neural networks, random forests, support vector machines, multiple linear regression, ARIMA, and mean regression and other time prediction models. These models utilize historical cargo volume data and relevant characteristic variables, and based on their respective unique algorithms, conduct learning and fitting operations on the data, thereby achieving the prediction of future cargo volumes. By leveraging the advantages of different models and comparing their degrees of fitting on the same dataset, a more accurate prediction model can be screened out, effectively improving the accuracy and reliability of the prediction.

2.2.2 Problem analysis of the interconnection prediction of logistics centers considering transportation routes

The general model of Linear Programming (LP), a widely adopted quantitative analysis method when logistics enterprises formulate transportation plans, solves the corresponding decision variables by constructing an objective function and setting a series of constraint conditions. Under the established framework of constraint conditions and objectives, this model is capable of optimizing the system to achieve the goal of obtaining the maximum output with the minimum input, providing a scientific quantitative basis for logistics transportation decision-making.

2.2.3 Strategies for personnel allocation in logistics sorting centers

In the algorithms related to graph theory, the Dijkstra algorithm can efficiently search for the shortest paths from a given starting point to all other points in the graph by assigning permanent labels and temporary labels to each vertex during the calculation process and continuously updating these labels. When applied to the personnel allocation in logistics sorting centers, it can save a large amount of human and material resources[6].

2.2.4 Optimization of attendance plans in logistics sorting centers by integrating artificial networks

Artificial neural networks are constructed based on the structure and operation principles of the human neural network in biology. Their basic processing units, artificial neurons, are composed of connection weights, summation units, and activation functions[7]. Numerous neurons are interconnected to form a network structure, which processes input signals and outputs results. By simulating the learning, memory, and problem-solving methods of the human brain, this network can learn from data and acquire knowledge, store the knowledge in the connection weights, and handle new problems according to the acquired knowledge, demonstrating powerful capabilities in tasks such as complex pattern recognition and prediction.

3 MODEL ESTABLISHMENT AND SOLUTION

3.1 Modeling Analysis of Precise Prediction of Logistics Sorting Volume Based on Multiple Models

3.1.1 ARIMA model

Combining the Autoregressive (AR) model equation (1) and the Moving Average (MA) model equation (2) results in the Autoregressive Moving Average (ARMA) model equation (3), which incorporates the characteristics of both AR and MA models. This allows it to capture both the autocorrelation and random fluctuations in time series data, providing a more comprehensive description and prediction of time series data.

$$y_t=\mu+\sum_{i=1}^{p}\gamma_i y_{t-i}+\varepsilon_t \tag{1}$$

$$y_t=\mu+\sum_{i=1}^{q}\theta_i \varepsilon_{t-i}+\varepsilon_t \tag{2}$$

$$y_t=\mu+\sum_{i=1}^{p}\gamma_i y_{t-i}+\varepsilon_t+\sum_{i=1}^{q}\theta_i \varepsilon_{t-i} \tag{3}$$

Where, y_t represents the forecasted value at time t, μ denotes the mean of the series, γ_i are the autoregressive coefficients, indicating the relationship between the current value and the values from the past i periods, p is the order of the autoregressive part, signifying the number of past observations considered, ε_t is the white noise error term, $\theta_i$are the moving average coefficients, reflecting the relationship between the current value and the errors from the past i periods, q is the order of the moving average part, indicating the number of past error terms considered.

3.1.2 Random forest

Random Forest is a classification algorithm that combines multiple weak classifiers into a strong classifier. It is a collection that contains multiple unpruned decision trees, in which the parameter set is an independent and identically distributed random vector. Under the given independent variables, each decision tree classification model determines the optimal classification result through voting[8]. The principle of Random Forest is shown in Figure 1 below.

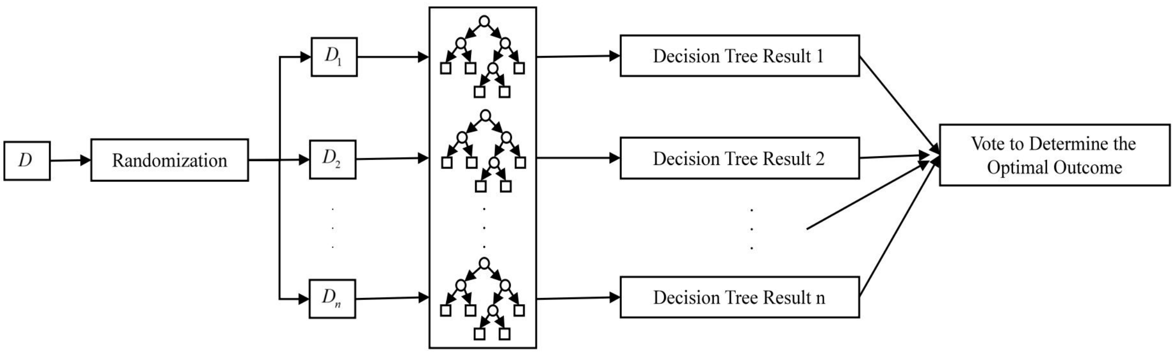

Figure 1 The Principle of the Random Forest Model

The Random Forest increases the differences among classification models by constructing different training sets, thereby improving the predictive ability of the combined classification model. Through k rounds of training,k classifiers are obtained,$\{h_1(x),h_2(x),\cdots,h_k(x)\}$,and then they are used to construct a multi-classification model system that adopts the voting method. Its decision function is shown in Equation (4).

$$H(x)=\operatorname{argmax}\sum_{i=1}^{k} I(h_i(x)=Y) \tag{4}$$

In the equation: $H(x)$ is the classification combination model; $h_i(x)$ is the decision tree classification model;Y is the target variable;$I(\cdot)$is the indicator function.

3.1.3 Linear ε-support vector regression machine

The architecture of the Support Vector Machine is shown in Figure 2 below, where x_i represents the input data and $K(x,x_i)$represents the kernel function., k=1,2,…,K.

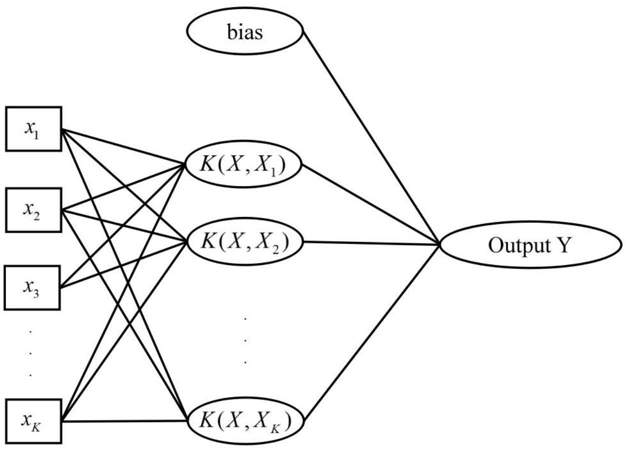

Figure 2 The Architecture of the Support Vector Machine

The following algorithm can be established for the Linear ε-Support Vector Regression Machine.
Step 1: Given the training sample data

$$T=\{(x_1,y_1),\cdots,(x_i,y_i),\cdots,(x_k,y_k)\} \tag{5}$$

Where $x_i \in R^n, y \in R$, and select the penalty factor c and the loss parameter ε;
Step 2: Construct a convex programming problem

$$\min\frac{1}{2}\sum_{i,j}^{K}(\lambda_i^{'}-\lambda_i)(\lambda_j^{'}-\lambda_j)(x_i x_j)+\varepsilon\sum_{i,j}^{K}(\lambda_i^{'}-\lambda_i)-\sum_{i,j}^{K}y_i(\lambda_j^{'}-\lambda_j)$$

$$\text{s.t.}\ \sum_{i,j}^{K}(\lambda_i^{'}-\lambda_i) \tag{6}$$

The first term of the objective function is a function of the inner product of data points x_i and x_j, multiplied by the difference in Lagrange multipliers, representing the maximization of the margin. The second term is a penalty term used to control classification errors, where ε is a very small positive number that allows for some classification errors. The third term is another penalty term, which may be related to specific constraints or regularization conditions.
It is determined that $\overline{\lambda}=(\overline{\lambda_1},\overline{\lambda_1},\cdots,\overline{\lambda_k},\overline{\lambda_k})$
Step 3: Calculate the value of $\overline{b}$
Select either $\overline{\lambda_k}$ or $\overline{\lambda_j}$ that is located in (0, c)
If the selected one is $\overline{\lambda_j}$,

$$\overline{b}=y_k-\sum_{i=1}^{k}(\overline{\lambda_i}-\overline{\lambda_i^{'}})(x_i x_j)+\varepsilon \tag{7}$$

If the selected one is $\overline{\lambda_k}$,

$$\overline{b}=y_k-\sum_{i=1}^{k}(\overline{\lambda_i}-\overline{\lambda_i^{'}})(x_i x_k)+\varepsilon \tag{8}$$

Step 4: The decision function is obtained.

$$y=h(x)=\sum_{i=1}^{k}(\overline{\lambda_i}-\overline{\lambda_i^{'}})(x_i x)+\overline{b} \tag{9}$$

3.1.4 Establishment of a multiple linear regression model

The multiple regression model mainly collects data, analyzes the data and establishes a model. Through this model, the relationship between the undetermined variables and the predictive variables can be analyzed, and the data can be accurately predicted through mathematical expressions. Equation (10) is the general form of the multiple linear regression model, and the specific establishment process of the model is shown in Figure 3.

$$y=\beta_0+\beta_1 x_1+\beta_2 x_2+\cdots+\beta_k x_k+\varepsilon \tag{10}$$

In the equation, $x_1,x_2,\cdots,x_k$ is the independent variable; $\beta_1,\beta_2,\cdots,\beta_k$ is the regression coefficient; y is the dependent variable;ε is the random error term.
In this problem, a series of features are also constructed for prediction, including the cargo volume with a lag of 7 days, the cargo volume with a 3-day and 7-day moving average, as well as the time series features such as the month, the day, whether it is a working day, and the trend index. A corresponding feature set is constructed, and a hybrid model combining multiple linear regression and time series features is established as shown in Equation (11) below.

$$Y=\beta_0+\beta_1\alpha+\beta_2 M_2+\beta_3 M_7+\beta_4 \text{mont}h+\beta_5 \text{day}+\beta_6 \text{is}_\text{w}\text{orkday}+\beta_7 \text{trend}+\varepsilon \tag{11}$$

This model can better capture the complex patterns of the target variable, and it is also capable of making longer-term predictions. At the same time, it can make use of more information from relevant characteristic variables.

3.1.5 Model implementation

First, import the data and visualize it. Sort the data by time, and supplement the missing data and the data with a value of 0 using the mean method. Draw graphs based on the data. Select multiple models for fitting according to the graph trends, and construct a feature set that includes lagged cargo volume, time series features, etc. Use models such as neural networks and random forests for prediction and fitting. Obtain the optimal parameters through this method, compare the results of the test set, and combine with the model with the highest score. Fit different models for different centers. Statistically analyze the predicted results of the cargo volume for the next 30 days of 57 sorting centers.

3.2 Construction of the Cargo Volume Prediction Model for Sorting Centers Considering Changes in the Transportation Network

3.2.1 Model specification

Similar to the logistics sorting cargo volume model, neural networks, random forests, ARIMA, support vector machines, and linear regression are adopted. Based on the average cargo volume of each transportation route at each sorting center over the past 90 days and the updated transportation routes, Figure 3 and Figure 4 can be drawn.

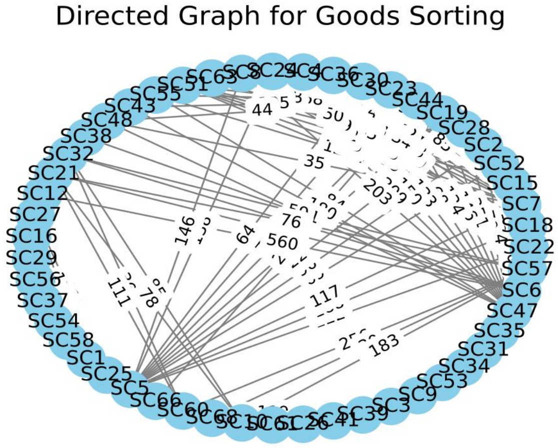

Figure 3 Directed Graph of Cargo Sorting

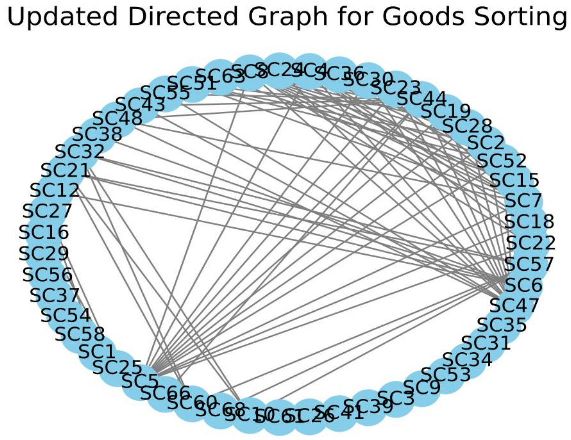

Figure 4 Update the Directed Graph of Sorting

From Figure 3 and Figure 4, one can intuitively observe the flow of cargo volume and the changes in transportation routes. However, it is unknown how the cargo volumes at different sorting centers will change after the transportation network has been altered. Therefore, the quantities of received and dispatched goods are taken into consideration. Then, an average received quantity and an average dispatched quantity are constructed as new features, and subsequently, the cargo volumes at 57 sorting centers are predicted.

3.2.2 Model implementation

First, conduct data import and pre - processing. After importing the relevant data, clean and integrate it. In terms of feature construction, based on the relevant features of the logistics sorting cargo volume without considering the transportation network in Part A, further construct the cargo volume features of the origin and destination stations according to the average cargo volume data of the transportation routes. During the model selection and training phase, select multiple models such as ARIMA and random forests. Combine historical cargo volume data and the comprehensive feature set, and fully consider the impact of changes in transportation routes to make adaptive adjustments to the models. Then, use the grid search method to adjust the parameters and evaluate the accuracy of each model in cargo volume measurement. Based on the evaluation results, adopt a combination of multiple models and use different models for fitting for different sorting centers[9]. Finally, calculate the predicted results of the cargo volume at each moment within the next 30 days for 57 sorting centers.

3.3 Linear Programming and Multi - theory Solving Model for the Number of Attendees in Sorting Center Shifts

3.3.1 Model specification

Based on the changes in the transportation network, the predicted cargo volume data for each sorting center in the next 30 days are obtained. The goals are to achieve a reasonable division of shifts, reduce the total number of days that employees are on duty, and control the attendance rate.

According to the problem description, define the decision variables: x_{ij1} represents the number of full - time employees attending the j- th shift on the i- th day, x_{ij0} represents the number of temporary employees attending the j- th shift on the i - th day. Construct the following objective function to determine the number of attendees for each shift at each sorting center, and then write the constraint conditions according to the problem description. The model is shown in Equation (12).

$$\min \sum_{i=1}^{30} \sum_{j=1}^{6} x_{ij1} \text{ s.t.} \begin{cases} \sum_{j=1}^{6} x_{ij1} \leq 60 \\ \sum_{j=1}^{n} (20x_{ij1}+16x_{ij0}) \geq w_{ik} \end{cases} \tag{12}$$

$$\text{Where } j=\begin{cases} 1,0<k<8 \\ 2,5<k<13 \\ 3,8<k<16 \\ 4,12<k<20' \\ 5,14<k<24 \\ 6,16<k<24 \end{cases}$$

$$n=\begin{cases} 1,0<k<5,22<k<24 \\ 2，5<k<12,13<k<14,20<k<22 \\ 3,12<k<13,14<k<20 \end{cases}$$

3.3.2 Personnel planning based on queuing theory

Queuing theory conducts statistical research on the arrival of service objects and service time, and derives the statistical laws of these quantitative indicators (such as waiting time, queue length, etc.). Then, based on these laws, the structure of the service system is improved or the serviced objects are reorganized, so that the service system can meet the needs of the service objects[10].

$$\min \sum_{i=1}^{n} \sum_{j=1}^{m} X_{ij}$$

$$\text{s.t.} \begin{cases} X_ij \leq L \\ \sum_{i=1}^{n} \sum_{j=1}^{m} x_{ij} c_{ij} \geq D_{ij} \\ X_{ij} \geq 0 \end{cases} \tag{13}$$

Through the model in Equation (13), the optimal configuration of the number of attendees for each shift at each sorting center in the next 30 days can be solved to achieve as much balance in labor efficiency as possible. In actual operation, it may be necessary to make fine adjustments to the model according to specific circumstances to meet the actual needs.

3.4 Optimization Modeling of Staff Scheduling in Sorting Centers Based on Artificial Neural Network

To optimize the shift attendance plans of full-time employees and temporary workers in the coming 30 days, an artificial neural network model is established. Firstly, the historical shift scheduling data, cargo volume processing data, and personnel attendance data are collected and used as the training set. In terms of model construction, the input is defined as the employees' shift scheduling situations, and the output is set as the completion rate of cargo volume processing, attendance rate, hourly labor efficiency, and labor cost. Subsequently, a multi-layer neural network including an input layer, a hidden layer, and an output layer is designed, and various activation functions and optimizers are employed to enhance the performance of the model. During the model training and evaluation stage, according to the output of the model and the actual situation, the mean squared error or cross-entropy loss function is selected. The prepared training set is used to train the model, and the parameters are continuously adjusted to minimize the loss function. After the training is completed, the validation set is utilized to evaluate the performance of the model, checking whether it can accurately predict the attendance plan and meet various constraints. Finally, the model is optimized based on the evaluation results, such as adjusting the network structure and optimizing hyperparameters. The trained and optimized model, as shown in Equation (14), can be used to predict the employees' shift attendance plans, so as to minimize the labor cost while meeting the cargo volume requirements.

$$\min \sum_{i=1}^{30} \sum_{j=1}^{6} \sum_{L=0}^{1} x_{ijL}$$

$$\text{s.t.} \begin{cases} \sum_{j=1}^{6} x_{ij1} \leq 60 \\ \sum_{j=1}^{n} (20x_{ij1}+16x_{ij0}) \geq w_{ik} \end{cases} \tag{14}$$

$$\text{Where } j=\begin{cases} 1,0<k<8 \\ 2,5<k<13 \\ 3,8<k<16 \\ 4,12<k<20' \\ 5,14<k<24 \\ 6,16<k<24 \end{cases}$$

$$n=\begin{cases} 1,0<k<5,22<k<24 \\ 2，5<k<12,13<k<14,20<k<22 \\ 3,12<k<13,14<k<20 \end{cases}$$

4 MODEL VALIDATION

4.1 Validation of the Logistics Sorting Cargo Volume Model

In order to validate the logistics sorting cargo volume model of A., the predicted cargo volumes for the past four months are compared with the cargo volumes of the past four months given in the original data. The comparison diagram is shown in Figure 5 below.

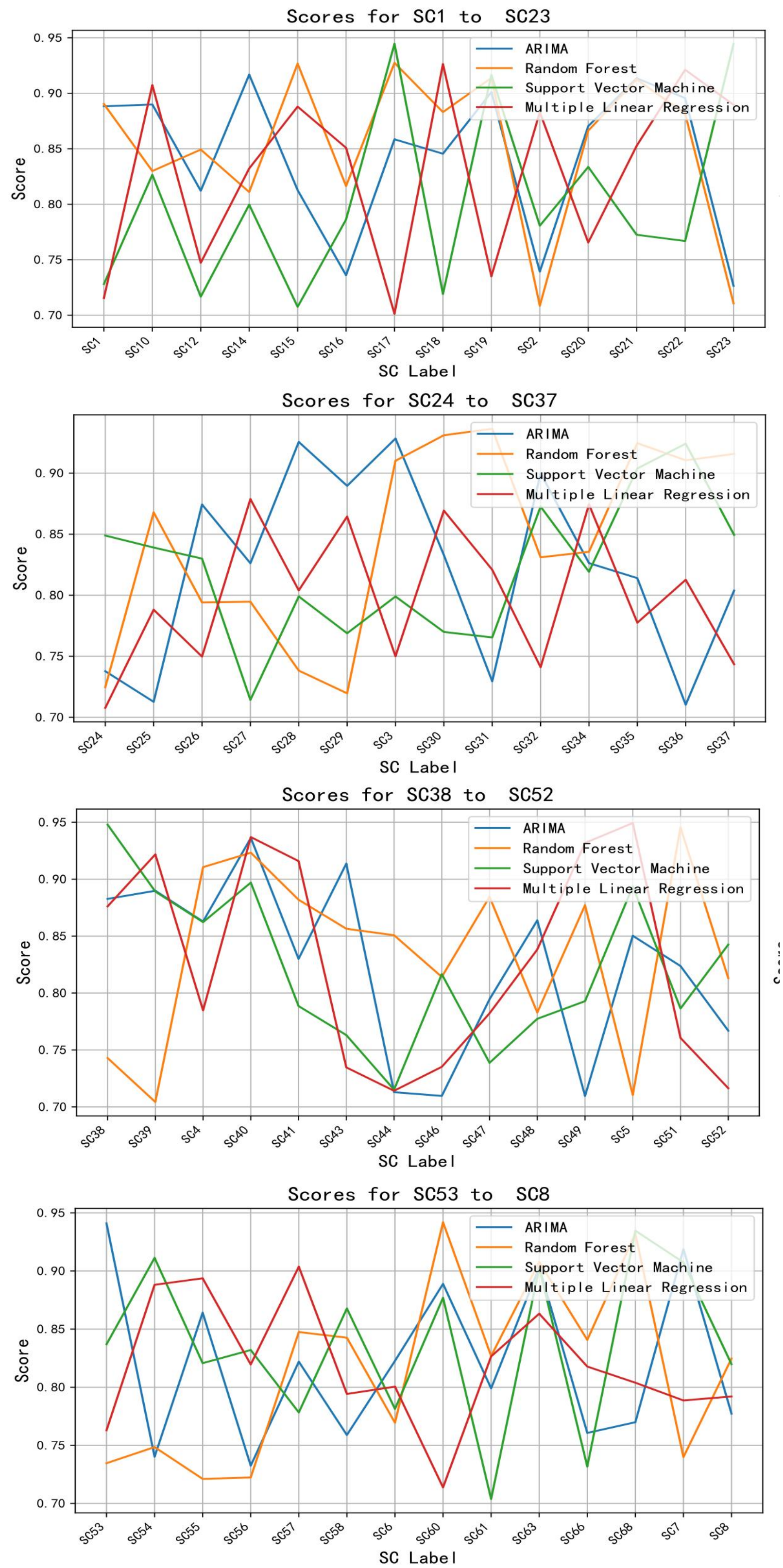

Figure 5 Model Scoring Diagram

It is found that the two curves roughly coincide. Using the Root Mean Squared Error (RMSE) as the evaluation index, it is obtained that most of the scores are above 0.7, and the vast majority are between 0.85 and 0.95. This indicates that the artificial neural network model used has a very good fitting effect, and the predicted daily cargo volume is relatively accurate. Subsequently, the hourly cargo volumes predicted by the prediction model are summed up on a daily basis, and the obtained values are compared with the predicted daily cargo volumes. The comparison diagram is shown in Figure 6. The overall ratio range is between 0.9 and 1.1, with the majority falling between 0.95 and 1.05, indicating good forecasting performance.

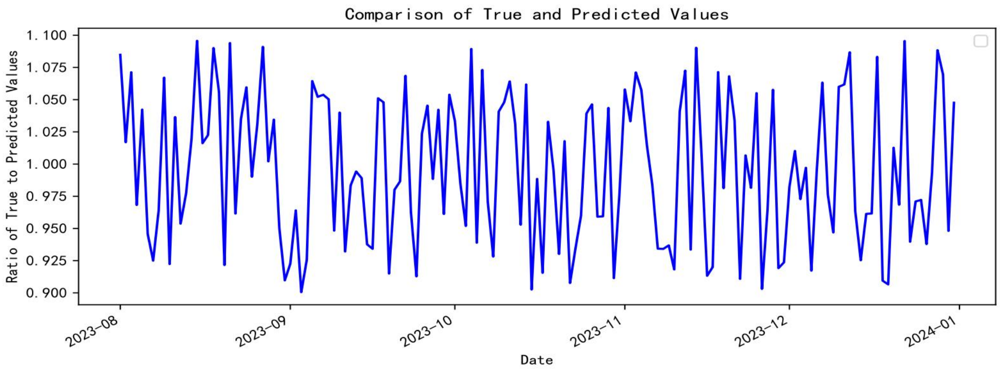

Figure 6 Comparison Diagram between the True Values and the Predicted Values

It is found that the two curves roughly coincide. By using the Root Mean Squared Error (RMSE) as the evaluation index, the scores are obtained to be between 0.97 and 1.03. This demonstrates that the artificial neural network model employed has an excellent fitting effect, and the predicted hourly cargo volume is relatively accurate.

4.2 Validation of the Transportation Network Change Model

In order to validate the transportation network change model of B., the predicted hourly cargo volumes after the route changes are summed up to obtain the predicted daily cargo volumes, which are then compared with the predicted daily cargo volumes. The comparison diagram is shown in Figure 7.

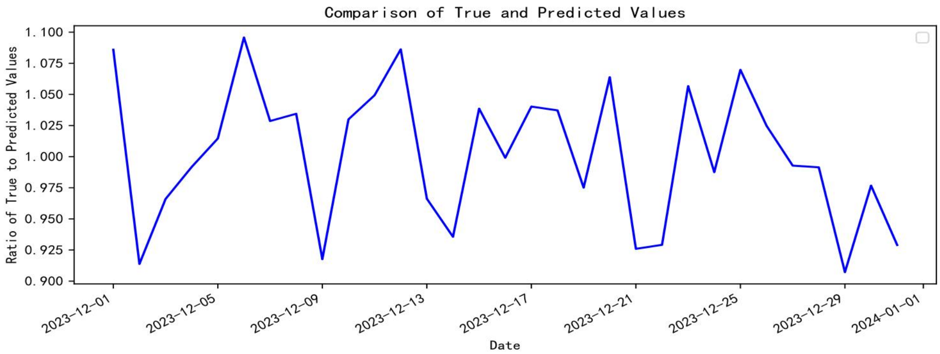

Figure 7 Comparison Chart between the Actual Values and the Predicted Values

4.3 Validation of the Personnel Scheduling Optimization Model

In order to validate the personnel scheduling optimization model, the obtained daily cargo volume is compared with the previous predicted value. The comparison diagram is shown in Figure 8.

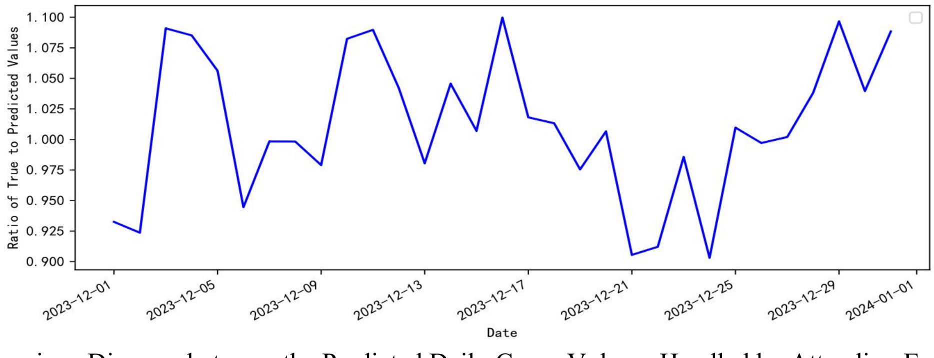

Figure 8 Comparison Diagram between the Predicted Daily Cargo Volume Handled by Attending Employees and the Previously Predicted Daily Cargo Volume

It is found that the two curves roughly match. By using the Root Mean Squared Error (RMSE) as the evaluation index, the scores are obtained to be between 0.97 and 1.03. This indicates that the artificial neural network model used has a high degree of accuracy, and the attendance arrangements derived from the personnel optimization model (ANN) are relatively reasonable.

5 CONCLUSION

Focusing on the accurate prediction of cargo volume and the efficient allocation of personnel in logistics sorting centers, data from multiple sorting centers are collected, and problems such as data format disorder and missing values are resolved. During the research process, a multi-algorithm fusion strategy is innovatively adopted. When predicting cargo volume, multiple machine learning models including ARIMA and Random Forest are comprehensively applied, and multi-dimensional features are constructed. When optimizing personnel allocation, models based on Linear Programming, Queuing Theory, and Artificial Neural Networks are integrated, taking various practical factors fully into account. Experimental results indicate that the constructed models have high prediction accuracy, and the personnel allocation schemes are reasonable. Compared with previous studies that employed a single algorithm, the fusion strategy significantly improves the efficiency of prediction and allocation, providing a brand-new and effective technical solution for the intelligent management of logistics sorting centers.

Looking ahead, the research could further incorporate more advanced deep learning methods, such as Recurrent Neural Networks (RNN), Long Short-Term Memory Networks (LSTM), and the like. These methods can be employed to explore and extract the complex features and long-term dependency relationships within the data, thereby further enhancing the accuracy of cargo volume prediction. Meanwhile, in terms of personnel allocation, it is advisable to consider incorporating more practical factors, such as the actual working status of employees and the dynamic response mechanism. This approach can enhance the robustness and universality of the models. In addition, integrating the combination of multiple models with a real-time scheduling system to achieve a closed-loop feedback between prediction and scheduling will provide more comprehensive and robust technical support for the intelligent upgrading of the logistics industry.

COMPETING INTERESTS

The authors have no relevant financial or non-financial interests to disclose.

REFERENCES

[1] Xu Y, Liu Y, Montreuil B. Dynamic workforce scheduling and relocation in hyperconnected parcel logistic hubs. arXiv preprint, 2024. https://doi.org/10.48550/arXiv.2405.04785

[2] Xu J, Wu H, Cheng Y, et al. Optimization of worker scheduling at logistics depots using genetic algorithms and simulated annealing. arXiv preprint, 2024. https://doi.org/10.48550/arXiv.2405.11729

[3] Ghosh S, Pal A, Kumar P, et al. A simulation driven optimization algorithm for scheduling sorting center operations. arXiv preprint, 2021. https://doi.org/10.48550/arXiv.2112.04567

[4] Tianyu Zhang, Miaoxin Wang, Liwei Guo, et al. Research on the Prediction of Cargo Volume in Logistics Network Centers Based on the ARIMA Model. Journal of Liaoning Technical University, 2023, 57(8): 1-10.

[5] Bin Nie, Jianqiang Du, Yue Yu, et al. Method for Screening Active Substances of Traditional Chinese Medicine by Partial Least Squares Based on the Criterion of Minimum Root Mean Square Error. Journal of Medical Informatics, 2024, 45(12): 29-36.

[6] Huiling Jiang, Wei Fang, Tianfeng Xu, et al. Indoor Evacuation Optimal Path Planning Model Based on Dijkstra Algorithm. Journal of Tsinghua University(Science and Technology), 2025, (04), 742-749.

[7] Cheng Huang, Jian Wu, Weiwei Ding, et al. Research on the Prediction Model of Chlorophyll-a Concentration in the Beijiang River Water Body Based on the BP Artificial Neural Network. Sichuan Environment, 2025, 44(01): 109-115.

[8] Zhiming Yuan, Qi Zhang, Kang Huang, et al. Research on the Prediction Method of Train Arrival Time Based on Random Forest. Railway Transport and Economy, 2016, 38 (05): 60-63+79.

[9] Haiyang Xu, WenwenDeng, Tong Li. A Hybrid ARIMA Time Series Data Prediction Model Based on BPNN. Information Technology & Informatization, 2025, (01): 102-105.

[10] Xihai Zhang. Evaluation of the Anchorage Capacity of Large Oil Tankers in Lanshan Port Based on Queuing Theory. Journal of Qingdao Ocean Shipping Mariners College, 2012, 33(04): 5-8.