Research on E-commerce product demand forecast and inventory optimization based on improved network optimization model

Xiangjun Meng, Qian Zhao, Yingxi Bi*

North China Institute of Science and Technology, Heibei, China

ABSTRACT

This study aims to accurately predict the future demand for goods on e-commerce platforms and optimize inventory through an improved network optimization model. First, take merchants, warehouses, and commodities as dimensions and build a probability density cluster analysis model based on time series to determine the classification of commodities under the most similar characteristics. Then, using the shipment volume as the dependent variable and the clustering correlation factors as the independent variables, an SBO-LSTM neural network model was constructed to optimize the shipment volume prediction. Through training and testing, the model shows high accuracy and stability. Finally, the trained model is applied to the data of each merchant in each warehouse, and the commodity demand forecast value from 2023-05-16 to 2023-05-30 is output, providing optimal demand forecast and inventory strategy for each merchant. The research results show that when the merchant number is 6 and the warehouse number is 43, the forecast value of product demand in the first three days is 2, 2, and 3.

Keywords: E-Commerce Platform; Demand Forecast; Inventory Optimization; Probability Density Clustering; SBO-LSTM Neural Network

1 INTRODUCTION

With the rapid development of the e-commerce industry, uncertainty in product demand and inventory management challenges have become one of the key factors affecting the economic development of e-commerce platforms. In order to better promote the economic growth of e-commerce platforms, demand forecasting and inventory optimization have become the focus of research. This study is based on an improved network optimization model and is dedicated to solving the problems of e-commerce product demand forecasting and inventory optimization to improve the efficiency and flexibility of the commodity supply chain.

In the e-commerce world, accurately forecasting product demand is critical to optimizing inventory, improving customer satisfaction, and reducing costs [1]. Traditional forecasting methods often have difficulty handling the complex relationships between different merchants, warehouses and goods, so a more sophisticated and flexible model is needed to address this challenge.

This article first builds a probability density cluster analysis model based on time series to classify merchants, warehouses, commodities and other multi-dimensional dimensions to find the most similar commodity characteristic conditions. Subsequently, we introduced the SBO-LSTM neural network model, using shipment volume as the dependent variable and clustering related factors as independent variables to build a shipment optimization prediction model. By training and testing the model, we verified its advantages in accuracy and stability.

In the specific study, we used 85% of the data as the training set and 15% of the data as the test set. The training accuracy and 1-wmape dimension obtained by the model were 0.945 and 0.947, indicating that the model has good performance [2]. Finally, we applied the model to the data of each merchant in each warehouse and output the forecast value of commodity demand from May 16, 2023 to May 30, 2023, providing the e-commerce platform with the optimal demand forecast and Inventory strategy.

The goal of this research is to provide an effective method for e-commerce platforms to improve the efficiency of the commodity supply chain, reduce costs, and promote the sustainable development of e-commerce platforms through more accurate demand forecasting and inventory optimization.

2 RELEATED WORK

E-commerce product demand forecasting and inventory optimization are important research directions in the field of supply chain management, and many scholars and researchers have made important contributions in this field. The following are some works and research directions related to this study: Time series analysis. For demand forecasting, many studies have adopted time series analysis methods, such as ARIMA (Autoregressive Integrated Moving Average) model and exponential smoothing method [3]. These methods solve the problem of time correlation to a certain extent, but their performance may be limited for the complex multi-dimensional data relationships of e-commerce platforms. Cluster Analysis, Some studies use cluster analysis methods to classify, goods or customers to better understand and predict demand,patterns. However, traditional clustering methods are difficult to cope with multi-dimensional, non-linear and dynamic relationships, so more advanced clustering techniques are needed [4]. Neural Network Model, Deep learning and neural networks have made significant, progress in the field of demand forecasting in recent years. Models such as Recurrent Neural Network (RNN) and Long Short-Term Memory Network (LSTM) are widely used for modeling time series data [5]. However, in a multi-dimensional e-commerce environment, further network optimization and improvement are necessary. Inventory optimization, research on inventory optimization mainly focuses on reducing inventory holding costs, improving service levels and reducing excess inventory [6]. Some research focuses on how to achieve inventory level optimization through smarter forecasting. Network optimization model. Network optimization model is widely used in supply chain management, including path optimization, resource allocation, etc [7]. However, research on integrating network optimization with demand forecasting and inventory management is still relatively limited.

In summary, although there are many studies on demand forecasting and inventory optimization, further improvements and innovations are still needed in processing multi-dimensional data on e-commerce platforms and improving the accuracy of models. This study aims to more comprehensively solve the problems in supply chain management of e-commerce platforms by introducing an improved network optimization model.

3 BASED ON PROBABILITY DENSITY CLUSTER ANALYSIS MODEL

3.1 Construction of DM model

The DM cluster analysis model is also called a cluster analysis model based on probability density and has a lot of attachment data. In order to avoid the cross-clustering phenomenon caused by noise interference, we proposed a DM cluster analysis model. The specific process of the model is shown below. Suppose there is a set Y with n combination sets, expressed as: $y = \{y_1, \dots, y_i, \dots, y_n\}$ [8]. The length of each combination time is n_i, and the corresponding observation point is x_i. There is a p-order polynomial regression relationship between y_i and x_i. The error satisfies the Gaussian superposition form, and the expression is as follows:

$$y_i = X_i\beta + \varepsilon_i\varepsilon_i \sim N(0, \sigma^2 I) \tag{1}$$

Among them, X_i is the Vandermond matrix corresponding to $n_i \times p$, β is the p-order regression coefficient, and the expression of X_i is as follows:

$$X_i = \begin{pmatrix} 1 & x_{i1} & x_{i1}^2 & \ell & x_{i1}^p \\ \mp & \mp & \mp & \ell & \mp \\ 1 & x_{in_1} & x_{in_2}^2 & \ell & x_{in_i}^p \end{pmatrix} \tag{2}$$

From this, the conditional probability density distribution function of y_i and x_i can be given as: $N(y_i \mid X_i\beta_k, \sigma_k^2 I)$. We can express the probability density function as a joint density distribution form that depends on k, so the expression of the polynomial regression model is as follows:

$$p(y_i \mid x_i, \Theta) = \sum_k^K \alpha_k p_k(y_i \mid x_i\theta_k) \sum_k^K \alpha_k N(y_i \mid X_i\beta_k, \sigma_k^2 I) \tag{3}$$

Use Z_i to represent the classification cluster of curve i, and the expression of the joint density functions y_i and Z_i is as follows:

$$\begin{aligned} p(y_i, z_i \mid x_i) &= \alpha_{zi} p_{zi}(y_i \mid x_i) \\ &= \alpha_{zi} N(y_i \mid X_i\beta_{zi}, \sigma_{zi}^2 I) \end{aligned} \tag{4}$$

$\{Z_i\}$ is hidden, and the related posterior probability is $p(z_i \mid y_i, x_i)$. The likelihood function L_c of the complete data can be expressed as the logarithmic joint density sum of n curves:

$$L_c = \sum_t \log \alpha_{zi} N\big(\gamma_i \mid X_i\beta_{zi}, \sigma_{zi}^2 I\big) \tag{5}$$

By calculating the posterior probability $p(z_i \mid y_i, x_i)$, we obtain the probability that the i-th curve belongs to class Z_i. The probability cluster about Z_i can be expressed as:

$$\begin{aligned} w_{ik} = p(z_i = k \mid y_i, x_i) &\propto \alpha_k p_k(y_i \mid x_i) \\ &= \alpha_k N(y_i \mid X_i\beta_k, \sigma_k^2 I) \end{aligned} \tag{6}$$

The Q function can represent the posterior expectation of the likelihood function, and the calculation expression is as follows:

$$Q = E[L_c \mid \nu_i, x_i] = \sum_i \sum_k \nu_{ik} \log \alpha_k N(\gamma_i \mid X_i\beta_k, \sigma_k^2 I) \tag{7}$$

The corresponding parameters when Q obtains the maximum value are $\{\beta_k, \sigma_k^2, \alpha_k\}$, and the direct expression of the solution is:

$$\hat{\beta}_k = \left[\sum_i w_{ik} X_i' X_i\right]^{-1} \sum_i w_{ik} X_i' y_i \tag{8}$$

$$\hat{\sigma}_k^2 = \frac{1}{\sum_i w_{ik}} \sum_i w_{ik} \parallel y_i - X_i \beta_k \parallel^2 \tag{9}$$

$$\hat{\sigma}_k^2 = \frac{1}{n} \sum_i w_{ik} \tag{10}$$

Iterate the data according to the above formula to obtain the DM cluster set.

3.2 Model solution

Import the preprocessed attachment data into the constructed DM model for iterative calculation. The DM clustering results of the model are shown in Table 1 below.

Table 1: DM clustering results

Cluster Category	Frequency	Percentage%
Cluster Category_1	33037	9.694
Cluster Category_2	42290	12.409
Cluster Category_3	37984	11.146
Cluster Category_4	54108	15.877
Cluster Category_5	55604	16.316
Cluster Category_6	38328	11.247
Cluster Category_7	37797	11.091
Cluster Category_8	662	0.194
Cluster Category_9	2	0.001
Cluster Category_10	40986	12.026
Total	340798	100.0

According to Table 1, it can be concluded that DM clustering is divided into 10 clusters. The detailed center points of merchants, warehouses, and commodities in these 10 cluster center points are shown in Table 2 below.

Table 2: Clustering center point

Cluster Type	Merchant Code	Product Code	Warehouse Code	First-Level Product Classification	Product Secondary Classification
1	2	70	7	3	4
2	23	762	12	6	18
3	21	464	14	13	35
4	21	1103	38	11	32
5	9	916	20	5	14
6	6	189	10	5	13
7	28	605	13	12	41
8	15	406	6	9	31
9	25	547	1	12	44
10	15	331	11	10	26

Cluster Type	Merchant Classification	Inventory Classification	Business Scale	Warehouse Category	Warehouse Area
1	2	1	1	1	2
2	6	2	2	2	3
3	9	2	2	1	2

4	9	2	1	2	3
5	4	2	1	2	3
6	3	2	1	1	2
7	10	2	1	1	2
8	6	2	1	1	1
9	9	1	1	1	1
10	8	2	2	1	2

According to the clustering center point in Table 2, the optimal classification result of the time series formed by merchants, warehouses, and commodities can be obtained. For example, when the merchant code is 2, the warehouse code is 7, the warehouse area is 2, and the commodity code is 70, The demand for the most similar product in the same category is 8 [9]. Then use the obtained DM clustering results as the import parameters of the following forecast model to perform forecast analysis of demand.

4 SBO-LSTM DEMAND FORECAST MODEL

4.1 LSTM neural network demand forecast model

The long short-term memory neural network model (LSTM) is an extension and improvement of the traditional RNN model. Its unique structural design enables it to exhibit superior performance when processing long-term data. The internal structure of LSTM is shown in Figure 1, which mainly includes three core components: forgetting gate, input gate and output gate. The forget gate is responsible for deciding what information is retained or forgotten from the cell state [10]. It works by first receiving the hidden layer output of the previous time step and the input of the current time step, and then using this information to calculate a value between 0 and 1. Specifically, the forgetting gate processes these inputs through a sigmoid function and outputs a corresponding "retention probability". This probability value determines which information from the previous time step will be retained. In this way, through the forgetting gate, LSTM can effectively filter out historical information that is useful for the current task and discard information that is no longer important. The calculation formula for the sum is:

$$f_t = \sigma\left(W_f \cdot (h_{t-1}, x_i) + b_f\right) \tag{11}$$

$$\sigma(x) = \frac{1}{1 - e^{-x}} \tag{12}$$

In the formula, W_f is the weight matrix of the forgetting gate; b_f is the bias; σ is the sigmoid activation function. The function of the input gate is to filter which new information will be added to the unit state, which is completed by the sigmoid function layer and the tanh function layer [11]. Among them, the sigmoid function layer is responsible for determining the selectivity of new information, while the tanh function layer is responsible for generating a candidate value $\widetilde{a_t}$ representing the learned new information. The calculation formula is:

$$i_t = \sigma(W_i \cdot (h_{t-1}, x_i) + b_i) \tag{13}$$

$$\widetilde{a_t} = \tanh(W_c \cdot (h_{t-1}, x_i) + b_c) \tag{14}$$

$$\tanh x = \frac{1 - e^{-x}}{1 + e^{-x}} \tag{15}$$

In the formula, W_i and W_c weight matrices correspond to the sigmoid function layer and tanh function layer respectively; b_i and b_c bias terms are also related to the sigmoid function and tanh function, where tanh refers to the hyperbolic tangent activation function. At the same time, based on the unit state at the previous moment and the input information at the current moment, the unit state will be updated accordingly through the forget gate and input gate. The update formula is:

$$C_t = f_t C_{t-1} + i_t a_t \tag{16}$$

The function of the output gate is to filter important information from the current unit state to form a new hidden layer. It first uses the sigmoid function to determine the part of the current unit state that should be output, then adjusts this state through the tanh function, and finally obtains the new hidden state h_t:

$$h_t = \sigma(W_0 \cdot (h_{t-1}, x_i) + b_0) \tag{17}$$

In the formula, W_0 is the weight matrix of the output gate; b_0 is the offset. To sum up, the hidden layer output h_t and cell state (C_t) of LSTM at the current moment are jointly composed of the hidden layer output (h_{t-1}), cell state (C_{t-1}) and the current moment at the previous moment. The input (x_t) is collaboratively determined. At the same time, according to relevant research, when exploring the complex nonlinear relationship between dam deformation and influencing factors, an LSTM model containing more than two hidden layer stacks can be constructed for training.

4.2 SBO model optimization LSTM

(1) Initialize a random population. This algorithm generates a starting population containing NB individuals, where each individual has a position in the D-dimensional space, and t represents the current evolutionary generation.

(2) Given the fitness of an individual, first determine its relative position in the total fitness of the entire population. This relative position reveals the probability of this individual being selected during the selection phase.

$$prob_i = \frac{fit_i}{\sum_{n=1}^{NB} fit_n} \tag{18}$$

In the formula, fit_i is the fitness value of the i-th courting booth; $f(x_i)$ is the cost function of the i-th courting booth, which can be calculated.

$$fit_i = \begin{cases} 1/1 + f(x_i), f(x_i) \geq 0 \\ 1 + |f(x_i)|, f(x_i) < 0 \end{cases} \tag{19}$$

(3) Update the population. The male bird will constantly adjust the position of the courting booth during the courtship process. The position update formula is as follows:

$$x_{ik}^{t+1} = x_{ik}^{t} + \lambda_k \left(\frac{x_{jk} + x_{elite,k}}{2} - x_{ik}^{t} \right) \tag{20}$$

In the formula, for the i-th individual in the t-th generation, its k-th dimension component is represented by x_{ik}^{t+1}; the component of the currently found optimal position in the k-th dimension is represented by x_{ik}^{t}; and in the entire population , the component of the current optimal position on the k-th dimension is $x_{elite,k}$. λ_k is the step size factor, which can be calculated by the following formula:

$$\lambda_k = \frac{\alpha}{1} + P_j \tag{21}$$

In the formula, α is the maximum threshold of the step size; P_j is the probability of being selected for the target courtship booth.

(4) Individual variation. Since the place where a mate is sought may be destroyed, random mutations are required with a certain probability at the end of the iterative process of the algorithm. In this mutation, x_{ik} follows a normal distribution with the following formula:

$$x_{ik}^{t+1} \sim N(x_{ik}^{t}, \sigma^2) \tag{22}$$

$$N(x_{ik}^{t}, \sigma^2) = x_{ik}^{t} + \left(\sigma * N(0,1)\right) \tag{23}$$

In the above formula, the standard deviation is calculated as follows:

$$\sigma = z * (\mathrm{var}_{\max} - \mathrm{var}_{\min}) \tag{24}$$

In the formula, z is the scaling factor, $\mathrm{var}_{\max}$ and $\mathrm{var}_{\min}$ are the upper and lower limits of the variable respectively.

4.3 SBO-LSTM optimization prediction

To sum up the mathematical process mentioned above, our proposed SBO-LSTM prediction process is as follows.

(1) The original data was denoised by outliers through two algorithms: Gaussian filtering and Symlet wavelet filtering, which complemented the data features and noise types. The reconstructed feature components after noise removal were used for prediction models. training, and divide the reconstructed data into a training set and a test set.

(2) Determine the hyperparameters through SBO, use the above formula to determine the step factor, mutation probability, maximum number of iterations and calculate the fitness value at the beginning of the iteration, and finally output the global optimal mating booth position, and obtain the optimal LSTM network parameter.

(3) Import the obtained parameters into the LSTM model, use the obtained model to train the training set, and predict the test set. Finally, calculate the root mean square error between the predicted value and the measured value and use it as an evaluation index to verify the accuracy of the model's prediction. sex.

(4) Compare the calculation result with the given error value. If the result is less than the given value, the result will be output directly, and then the cycle will end; otherwise, the parameters will be updated and the cycle will continue from step (3).

(5) If the error value is still greater than the given value when iterating to the maximum number of iterations, the optimal value in the iteration process (that is, the result closest to the given value) will be directly output as the final result, and then the loop will end.

4.4 Model solution

Use 85% of the data as training data and 15% of the data as the test set, then import relevant data into the model for training and testing. The training accuracy and test results of the model are as shown in the table below.

Table 3: Test and training results

Index	Accuracy	Recall	Accuracy	F1	Rmse	1-Wmape
Train	0.946	0.945	0.945	0.947	0.945	0.945
Test	0.942	0.945	0.945	0.945	0.947	0.946

According to Table 3, it can be concluded that the training accuracy of the model is 0.945, and the model test 1-wmape and RMSE values are 0.945 and 0.947, indicating that the prediction accuracy of the model is high. Then import the data of each merchant in each warehouse into the trained model for prediction. The prediction results of the model are shown in Figure 1 below.

Figure 1: Part of the prediction results

According to Figure 1, the shipment forecast values of each merchant in each warehouse from 2023-5-16 to 2023-5-30 can be obtained. The merchant number is 6, the warehouse number is 43, and the forecast values for the first three days are 2, 2, 3. The shipment forecast values of the remaining merchants in each warehouse from 2023-5-16 to 2023-5-30 are output to the corresponding positions in Table 1.

5 CONCLUSION

First, based on time series, we constructed a probability density cluster analysis model in multiple dimensions such as merchants, warehouses and commodities. The goal of this model is to output the most similar product classification by considering multi-dimensional feature conditions. Taking the scenario where the merchant code is 2, the warehouse code is 7, the warehouse area is 2, and the product code is 70, for example, the demand for goods in the same category is optimized to 8, achieving the optimal classification conditions.

Secondly, we use the SBO-LSTM neural network model, taking shipment volume as the dependent variable and clustering correlation factors as independent variables to achieve optimized prediction of shipment volume. By training and testing the model with 85% of the data as the training set and 15% of the data as the test set, the obtained model training accuracy and 1-wmape dimension reached 0.945 and 0.947 respectively, verifying the efficiency and accuracy of the model. .

Finally, after the model training was completed, we applied it to the data of each merchant in each warehouse and output the forecast value of commodity demand from May 16, 2023 to May 30, 2023. Specifically, the predicted values for the first three days of the merchant number 6 and the warehouse number 43 are 2, 2, and 3 respectively. These predicted values are stored in the corresponding locations in Table 1, providing a powerful decision-making basis for the e-commerce platform.

Through the above methods, we have achieved significant research results in e-commerce product demand forecasting and inventory optimization, and have made positive contributions to improving supply chain management efficiency and achieving economic development.

6 DISCUSSION

Future research directions can focus on further deepening the methods and applications of e-commerce product demand forecasting and inventory optimization to meet changing market demands and improve the economic benefits of e-commerce platforms. The following are some possible future research directions: model fusion and optimization, exploring the fusion of different prediction models, such as the combination of deep learning and traditional time series models, to improve the prediction accuracy of the model. Further optimize the network structure and parameters to adapt to the diversity of different merchants, warehouses and commodities. Real-time and dynamic research: Aiming at the real-time and dynamic nature of e-commerce platform demand and inventory management, we study how to adapt to market changes in a more timely manner and improve forecast accuracy through real-time data updates and dynamic adjustment of model parameters. Consider more influencing factors and consider more factors that may affect demand, such as market promotions, competitor behavior, etc., to improve the comprehensiveness and adaptability of the model. Multi-scale forecasting studies demand forecasting at different time scales, such as hourly level, daily level, and quarterly level, to better meet the different operational needs of e-commerce platforms.

Risk management: introduce a risk management model to consider possible risk factors in the supply chain, such as emergencies, market fluctuations, etc., to reduce potential economic losses. Empirical research: Conduct empirical research on a real e-commerce platform to verify the actual effect and operability of the model, and consider the feasibility and sustainability of the model in actual operations. Interpretability and visualization, improve the interpretability of the model, enable decision-makers to better understand the prediction results of the model, and further develop visualization tools to more intuitively display model output and influencing factors. Global adaptability: For e-commerce platforms in different regions and countries, study how to adjust the model to adapt to global market differences.

These future research directions are expected to deepen our understanding of e-commerce product demand forecasting and inventory optimization issues, and provide more intelligent, flexible and sustainable solutions for the e-commerce industry.

REFERENCES

[1] Zhou, J. (2023). Logistics inventory optimization method for agricultural e-commerce platforms based on a multilayer feedforward neural network. *Pakistan Journal of Agricultural Sciences*, *60*, 487-496.

[2] Salamai, A. A., Ageeli, A. A., & El-kenawy, E. S. M. (2022). Forecasting E-commerce adoption based on bidirectional recurrent neural networks. *Computers, Materials & Continua*, *70*(3), 5091-5106.

[3] Lin, H. L., Cho, C. C., Ma, Y. Y., Hu, Y. Q., & Yang, Z. H. (2019). Optimization plan for excess warehouse storage in e-commerce–based plant shops: A case study for Chinese plant industrial. *Journal of Business Economics and Management*, *20*(5), 897-919.

[4] Zhu, L. (2020). Optimization and simulation for e-commerce supply chain in the internet of things environment. *Complexity*, *2020*, 1-11.

[5] Li, J., Wang, T., Chen, Z., & Luo, G. (2019). Machine learning algorithm generated sales prediction for inventory optimization in cross-border E-commerce. *International Journal of Frontiers in Engineering Technology*, *1*(1), 62-74.

[6] Guo, L. (2022). Prediction Method of Short-Term Demand for e-Commerce Goods Based on Deep Neural Network. *Advances in Multimedia*, *2022*.

[7] Deng, Y., Zhang, X., Wang, T., Wang, L., Zhang, Y., Wang, X., ... & Peng, X. (2023). Alibaba realizes millions in cost savings through integrated demand forecasting, inventory management, price optimization, and product recommendations. *INFORMS Journal on Applied Analytics*, *53*(1), 32-46.

[8] Ekren, B. Y., Mangla, S. K., Turhanlar, E. E., Kazancoglu, Y., & Li, G. (2021). Lateral inventory share-based models for IoT-enabled E-commerce sustainable food supply networks. *Computers & operations research*, *130*, 105237.

[9] Li, Z., & Zhang, N. (2022). Short-Term Demand Forecast of E-Commerce Platform Based on ConvLSTM Network. *Computational Intelligence and Neuroscience*, *2022*.

[10] Suchánek, P., & Bucki, R. (2011). Method of supply chain optimization in E-commerce. *Journal of Applied Economic Sciences*, *6*(3), 58-64.

[11] Guo, X., Liu, C., Xu, W., Yuan, H., & Wang, M. (2014, July). A prediction-based inventory optimization using data mining models. In *2014 Seventh International Joint Conference on Computational Sciences and Optimization* (pp. 611-615). IEEE.