Supply Chain Demand Forecast Based on SSA-XGBoost Model

Shifeng Ni, Yan Peng, Ke Peng, Zijian Liu

School of Computer Science and Engineering, Sichuan University of Science and Engineering, Yibin, China
Email: nsf_96@163.com

Abstract

Supply chain management usually faces problems such as high empty rate of transportation, unreasonable inventory management, and large material consumption caused by inaccurate market demand forecasts. To solve these problems, using artificial intelligence and big data technology to achieve market demand forecasting and intelligent decision-making is becoming a strategic technology trend of supply chain management in the future. Firstly, this paper makes a visual analysis of the historical data of the Stock Keeping Unit (SKU); Then, the characteristic factors affecting the future demand are constructed from the storage level, product level, historical usage of SKU, etc; Finally, a supply chain demand forecasting algorithm based on SSA-XGBoost model has proposed around three aspects of feature engineering, parameter optimization and model integration, and is compared with other machine learning models. The experiment shows that the forecasting result of SSA-XGBoost forecasting model is highly consistent with the actual value, so it is of practical significance to adopt this forecasting model to solve the supply chain demand forecasting problem.

Keywords

Data Visualization Analysis, SSA-XGBoost, Supply Chain, Demand Forecast

1. Introduction

The development of supply chain management has experienced a transition from a relatively simple labor-intensive model to a relatively complex global functional network model. With the development of manufacturing globalization, the enterprise resource planning (ERP) system has greatly improved the availability and accuracy of data, and the term "supply chain" has been widely recognized [1] [2] [3]. With the maturity of the supply chain management model, the application of computer technology to supply chain management ac-

tivities is conducive to enhancing the trust and cooperation between upstream and downstream nodes, and promoting the digital transformation of the supply chain.

The research on supply chain demand forecasting can be mainly conducted from the following perspectives: The first is from the perspective of long-term forecast and short-term forecast. The former usually forecasts the demand by year or quarter, while the latter usually forecasts the demand by month or day [4] [5]. The second is from the perspective of single method and integrated method. The former only uses one algorithm model [6] [7] [8], while the latter uses two or more algorithms to build an integrated model for demand forecasting [9]. The third is to use different model methods. It mainly includes: mathematical statistics methods [10] [11] [12], such as grey prediction model, exponential smoothing method [13], ARIMA model [14], etc; Prediction models based on machine learning [15], such as support vector machine, decision tree, random forest regression, etc; Prediction models based on deep learning [16], such as BP neural network [17], LSTM neural network, etc.

In general, mathematical statistics method is still the mainstream method of supply chain demand forecasting, while machine learning and deep learning methods based on big data mining are relatively few. The traditional mathematical statistics methods are mostly used for forecasting data series with obvious time series law and stable trend. In real scenarios, the changing trend and fluctuation of supply chain demand are unstable, and the changes of various factors have different effects on demand. Therefore, this paper proposes a SSA-XGBoost model for supply chain demand forecasting from three aspects of feature analysis, optimization and demand forecasting. The experiment takes the historical consumption data of SKU as the demand, analyzes the influence degree of each influencing factor on the demand, combines the demand and influencing factors to build a data set as the input of SSA-XGBoost forecasting model, obtains the demand forecasting results, and compares them with different models.

2. Theoretical Basis of Related Research

2.1. XGBoost Model

XGBoost is an integrated learning algorithm in machine learning that takes decision tree as the base classifier and proposes optimization based on GBDT principle [18]. Integrated learning is a technical framework, which trains several weak models to complete machine learning tasks with a certain combination strategy. XGBoost proposes optimization on the basis of GBDT, which has greatly improved accuracy and speed, enabling the model to perform better in big data processing projects. The optimization of XGBoost mainly includes:

The regularization term is added to the cost function to control the complexity of the model and reduce the possibility of over fitting.

Different from the way that GBDT uses Gini coefficients, XGBoost obtains the node splitting mode after optimization and derivation. In this way, XGBoost can automatically learn its splitting direction even if there are missing values in the

samples of the feature set.

To determine the best segmentation point, the decision tree needs to sort the eigenvalues in the training process, which is a time-consuming step. Before training, XGBoost will sort the eigenvalues and save them as block structures. This structure will be reused in subsequent iterations to reduce the amount of computation. At the same time, this block structure makes it possible to calculate the gain of each feature by multithreading.

Table 1 shows several important parameters of XGBoost algorithm. The values of these parameters determine the size and prediction accuracy of XGBoost model. For example, the smaller the value of n_estimators, the easier the model is to be under fitted, while the larger the value, the more over fitted the model will be, and the longer the training time of the model will be. Each parameter has a different value range, and its combination will produce a large number of combinations. XGBoost's traditional method of finding the optimal parameter combination is to first set several values for each parameter according to experience, and then select the combination with the highest accuracy by calculating the accuracy of each combination. Then, within a certain range of the combination, grid search is carried out to select the optimal combination. The traditional method is easy to understand and implement, but its disadvantage is that it has a large amount of computation and is easy to fall into the local optimal value.

2.2. Sparrow Search Algorithm

Sparrow search algorithm (SSA) is an algorithm that simulates the process of sparrow group foraging. The traditional SSA algorithm focuses on the update of a single parameter. This research will improve the SSA algorithm, so that the model parameters can be updated in different ranges in the form of combination. In this way, SSA algorithm can be combined with XGBoost algorithm. In formula (1), X represents the current position of the population composed of n sparrows, and m is the number of parameters to be optimized.

$$X = \begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \\ x_{2,1} & x_{2,2} & \cdots & x_{2,m} \\ \vdots & \vdots & \ddots & \vdots \\ x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix} \tag{1}$$

In SSA, individual fitness value is used to evaluate the level of sparrows' energy reserves. The fitness value matrix of the whole sparrow population is shown in Formula (2):

$$F_X = \begin{bmatrix} f\left(\begin{bmatrix} x_{1,1} & x_{1,2} & \cdots & x_{1,m} \end{bmatrix}\right) \\ \vdots \\ f\left(\begin{bmatrix} x_{i,1} & x_{i,2} & \cdots & x_{i,m} \end{bmatrix}\right) \\ \vdots \\ f\left(\begin{bmatrix} x_{n,1} & x_{n,2} & \cdots & x_{n,m} \end{bmatrix}\right) \end{bmatrix} \tag{2}$$

Table 1. Meaning of important parameters of XGBoost algorithm.

Parameter	Meaning
n_estimators	Number of basic models
max_depth	Maximum depth of decision tree
learning_rate	Weight of the model generated by each iteration (learning rate)
objective	Given loss function
reg_alpha	Weight of L1 regular term
reg_lambda	Weight of L2 regular term

In Equation (2), $f\left(\left[x_{i,1} \; x_{i,2} \; \cdots \; x_{i,m}\right]\right)$ is used to calculate the fitness value of the ith sparrow's current position, where $i \in [1,n]$. F_X' can be obtained by sorting fitness values. Let p be the discoverer ratio ($p \in (0,0.5]$), then the first f sparrows in F_X' are discoverers, where $f = n \times p$.

In the process of foraging, there are two roles of sparrows: discoverer and joiner [19]. The discoverer refers to the sparrow with high energy storage level in the population. The responsibility of the discoverer is to find areas rich in food, and provide the scope and direction of foraging for the joiners of the group. When the finder finds the predator, it will immediately send a warning to the group and fly to other safe areas for feeding. Equation (3) describes the location update rule of discoverer in the improved SSA algorithm:

$$X_i^{t+1} = \begin{cases} GridSearchCV\left(X_i^t, R\right) & \text{if } w < s \\ X_i^t + Q * L & \text{if } w > s \end{cases} \tag{3}$$

In Formula (3), t represents the number of current iterations, X_i^t represents the position information of the ith sparrow at time t, and R represents a matrix of $1 \times m$. Meanwhile, w represents the optimal individual fitness value among the discoverers at time t ($w \in [0,1]$), and s represents the safety value ($s \in [0.5,1]$). When $w < s$, it means that there are no predators around the foraging environment at time t, and the finder performs grid search in $\left[X_i^t - R, X_i^t + R\right]$ to find the location of the local optimal fitness value; Q and L both represent a matrix of $1 \times m$, where each element in L is −1 or 1, and the operation symbol $*$ represents Hadamard product; When $w > s$, it means that there are predators around at time t. Some discoverers have found predators and sent alarm signals to other individuals in the group. At this time, all sparrows in the group need to fly to other safe areas for feeding.

The remaining individuals of the group are the joiners. Some of the joiners with high reserve energy are mainly responsible for monitoring the discoverer. Once they know that the discoverer has found a better place to look for food, they will immediately fly near the discoverer to compete with it. If the energy reserve of the joiner is higher than that of the discoverer after the location update, the roles of the two will be exchanged to keep the proportion of the two in the population unchanged. The remaining joiners are very hungry due to their

low energy reserves, and they will fly to other places to find food in order to obtain more energy. Equation (4) describes the rules for updating the joiners' location in the improved SSA algorithm:

$$X_i^{t+1} = \begin{cases} X_i^t + I * \dfrac{X_{worst}^t - X_i^t}{\sqrt[2]{i}} & \text{if } i > n/2 \\ X_{best}^t + Q * L & \text{otherwise} \end{cases} \tag{4}$$

In formula (4), I represents a matrix of $1 \times m$, and the elements in matrix I conform to the normal distribution with the distribution center of 0 and the standard deviation of 1; X_{worst}^t represents the worst position occupied by individual population at time t; X_{best}^t represents the optimal position occupied by the discoverer at time t; When $i > n/2$, the i-th joiner with low fitness is hungry and must fly to other places for feeding. Other joiners monitored and robbed food around the discoverer.

After the location update of the discoverer and the joiner is completed, 20% of the sparrows in the group will realize that they are on the edge of the group and are easy to be attacked by predators. This part of the sparrow will immediately move to the safe area to get a safer position. In the improved SSA algorithm, the position update of sparrows at the edge of the group is as shown in Formula (5):

$$X_i^{t+1} = X_i^t + \frac{\beta \cdot (f_i - f_{best})}{f_{worst} - f_{best}} \cdot \left(X_{best}^t - X_i^t \right) \text{ if } i \geq \theta \cdot n \tag{5}$$

In Equation (5), f_i represents the fitness value of the ith sparrow at time t; f_{best} and f_{worst} represents the best and worst fitness values of the current population respectively; β is a constant ($\beta \in (0,1)$), used to avoid that the fraction is equal to 1; θ is a constant ($\theta \in (0,1)$), set to 0.8 here.

3. Experimental Data Set

The experiment shows the temporal distribution of historical demand data by visualizing the data, and confirms the influence of product level and geographical level of the unit on demand through one-way ANOVA, providing decision support for the subsequent supply chain demand forecasting link.

3.1. Dataset Introduction

The dataset used in the experiment is from the 2021 Alibaba Cloud Infrastructure Supply Chain Competition, including training set, test set, geographic level information dataset and product level information dataset. The data set contains 632 product unit historical demand information. The timing range of the training set is from June 6, 2018 to March 1, 2021, with a total of 284,832 pieces of data. The time series range of the test set is from March 2, 2021 to June 7, 2021, with a total of 61,936 pieces of data. None of the data sets have missing values. Table 2 shows the data field details of training set and test set. Table 3 and Table 4 respectively show the geographic level information dataset and product level information dataset. It can be seen that the unit name, geographic level information and product level information of the dataset are desensitized.

During the experiment, the geographic level information and product level information need to be labeled with numbers for subsequent model training and prediction.

3.2. Data Visualization and Analysis

Figure 1 shows the historical data change curve of 6 units randomly selected. It can be seen that the historical demand change of most units shows an overall rise over time, while that of a few units shows an overall rise first and then a decline, accompanied by a sharp increase or decrease in demand. In general, between different units, the time sequence of demand change is different.

Figure 2 shows the histogram of historical average demand of different categories at GL1 and PL1 levels on March 1, 2021. At the same time, the historical demand values of different categories at the geographical level and product level are quite different. Table 5 shows the ANOVA results of GL1 and PL1 level categories and historical demand data on March 1, 2021. The PR values of both are greater than 0.05, which indicates that at the significance level of 0.05, there is a significant difference in the average unit historical demand between different categories of geographic and product levels, and there is a correlation between characteristics (geographic and product levels) and forecast variables (unit demand).

Table 2. Data field details of training set and test set.

Column name	Description	First row data
Unit	Cargo unit	9b8f48bacb1a63612f3a210ccc6286cc
ts	Date	2018/6/4
qty	Resource usage	11926.8286
Geography	Geographic information of geography_level_3	36ab7b000da26b0547bfc3c3fdf143dc
Product	Product information of product_level_2	5cc8015f03554313900f069182bdaf9c

Table 3. Geographic level information dataset.

Column name	Abbreviation	First row data
Geography_level_1	GL1	860b874ad79c7f2072b4dd24a952f027
Geography_level_2	GL2	50fcfda7a3aeb86c6dabc2bd467067b0
Geography_level_3	GL3	0b41e006d801550cba4c38157729bf87

Table 4. Product level information dataset.

Column name	Abbreviation	First row data
Product_level_1	PL1	2eb2930111864beeb409e946751215b1
Product_level_2	PL2	ba569a9ecbaa99645f9baf060d5061e7

Table 5. ANOVA Results of GL1 and PL1 level categories and historical demand data.

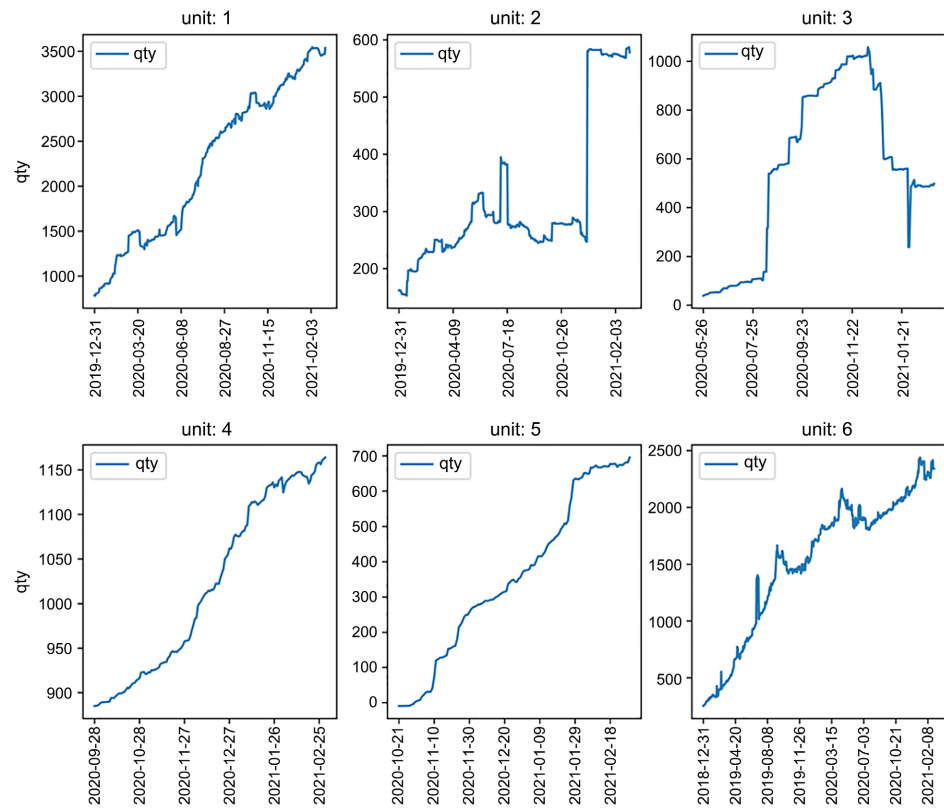| Name	PR (>F)
GL1	0.071461
PL1	0.316654

Figure 1. Historical demand data curve of 6 units randomly selected.
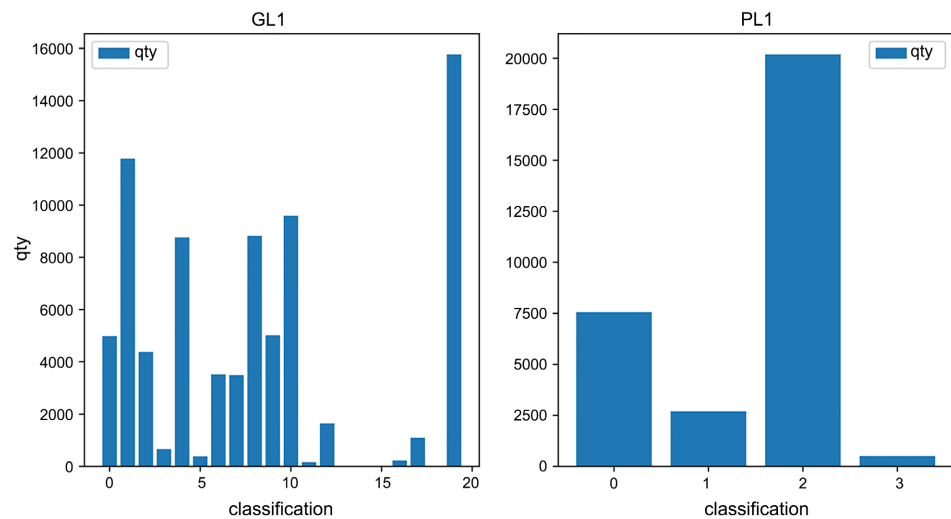

Figure 2. Histogram of average historical demand of different categories on March 1, 2021.

4. Experiment of Fresh Food E-Commerce Logistics Demand Forecasting

4.1. Preliminary Work

The experiment was carried out on Windows 10 system, using PyCharm2022.2.3 and Python 3.9 as experimental tools. According to the geography and product columns of Table 2, merge the data of Table 2 with Table 3 and Table 4. At the

same time, according to the time information, the characteristics "holiday", "month", "week" and "weekday" are constructed to represent whether the current time is a holiday, the month, the week of the year and the day of the week. Finally, the sliding window statistics is used to construct the features "LW", "data_smooth", "mean", "max", "min", and "std", which represent the historical value, exponential smoothing value, mean value, maximum value, minimum value, and standard deviation of the previous seven days, respectively. To sum up, 15 influencing factors, namely "GL3", "GL2", "GL1", "PL2", "PL1", "holiday", "month", "week", "weekday", "LW", "data_smooth", "mean", "max", "min" and "std", are selected as the characteristics of the data set, and "qty" is used as the label data set.

4.2. Establishment of Evaluation Index

The coefficient of determination (R^2), mean square error (MSE), root mean square error (RMSE) and mean absolute error (MAE) were selected as evaluation indicators to evaluate the prediction performance of the model. Coefficients of determination can be used to determine the goodness of fit between the predicted value and the actual value. The closer the result is to 1, the higher the goodness of fit is. The specific formula is as follows:

$$R^2 = 1 - \frac{\sum_{i=1}^{k}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{k}(y_i - \bar{y}_i)^2} \tag{6}$$

In the formula, y_i is the true value, $\bar{y}_i$ is the average of the true value, $\hat{y}_i$ is the predicted value, and k is the number of data items in the dataset. When $R^2 = 1$, it means that the predicted value of the model is equal to the true value, and the prediction accuracy of the model is high; When $R^2 = 0$, it means that the predicted values of the model are equal to the mean value, and the prediction accuracy of the model is low; When $R^2 < 0$, it means that the model cannot predict accurately.

The mean square error refers to the mean value of the square sum of the corresponding point errors of the prediction data and the original data. The closer the result is to 0, the smaller the model prediction error is. The specific formula is as follows:

$$\text{MSE} = \frac{1}{k}\sum_{i=1}^{k}(y_i - \hat{y}_i)^2 \tag{7}$$

The root mean square error is the arithmetic square root of the mean square error, which is very sensitive to the reflection of outliers in measurement and can reflect the dispersion of samples.

$$\text{RMSE} = \sqrt[2]{\text{MSE}} \tag{8}$$

The average absolute error refers to the average value of the distance between the predicted value of the model and the true value of the sample, which can better reflect the actual situation of the predicted value error.

$$\text{MAE} = \frac{1}{m}\sum_{i=1}^{m}|y_i - \hat{y}_i| \tag{9}$$

4.3. Construction of SSA-XGBoost Model

By integrating SSA model and XGBoost model, a SSA-XGBoost model is proposed. SSA-XGBoost model can automatically find the global optimal parameter combination of XGBoost model to improve the accuracy of XGBoost model. Table 6 shows the specific flow of SSA-XGBoost algorithm, where the function $F(X)$ is to calculate the fitness value of the position of the sparrow individual. Substitute the position information of the sparrow individual at the current time into the XGBoost model to obtain the prediction data of the verification set. Using the prediction data and the real data, the fitness value of the sparrow individual's position at the current time can be obtained. The formula is as follows:

$$FV_j = \frac{\sum_{i=1}^{k}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{k}(y_i - \overline{y}_i)^2}, j \in [0, n] \tag{10}$$

Table 6. SSA-XGBoost algorithm.

Input: p(population quantity), M(iterations), dim(dimension), d(proportion of discoverers)

1. Define the objective function $F(X)$, where the variable $X = (X_1, X_2, \ldots, X_n)$, the variable $X_i = (x_1, x_2, \ldots, x_m)$
2. Randomly initialize the positions of n sparrows X and define their relevant parameters
3. Use function $F(X)$ to obtain the population fitness value matrix
4. for $i \leftarrow 0$ to M do
5.   Sort the fitness value matrix from small to large to obtain a matrix named fit
6. $R = \min(\text{fit})$
7.   if $R <$ safety value then
8.   for $i \leftarrow 0$ to p^*d do
9.   Update sparrow position using $GridSearchCV\left(X_i^t, R\right)$ in formula (3)
10.   end for
11.   else
12.   for $i \leftarrow 0$ to p^*d do
13.   Update sparrow position using $X_i^t + Q * L$ in formula (3)
14.   end for
15.   for $i \leftarrow 0$ to $p^*(1\text{-}d)$ do
16.   if $(i + p^*d) > p/2$ then
17.   Update sparrow position using $X_i^t + I * \dfrac{X_{worst}^t - X_i^t}{\sqrt[2]{i}}$ in formula (4)
18.   else
19.   Update sparrow position using $X_{best}^t + Q * L$ in formula (4)
20.   end for
21. $\theta \leftarrow 0.8$
22.   for $i \leftarrow 0$ to p do
23.   if $i >= \theta^*p$ then Update sparrow position using $X_i^t + \dfrac{\beta \cdot (f_i - f_{best})}{f_{worst} - f_{best}} \cdot \left(X_{best}^t - X_i^t\right)$ in formula (5)
24.   end for
25.   Update the global optimal fitness value and the individual optimal position of the population
26. end for

Outout:

BestX: location of global optimal fitness value

Convergence_Curve: global optimal fitness value for each iteration

4.4. Model Training

Parameters n_estimators, max_depth and learning_rate are selected as optimization objectives in the experiment, and the corresponding upper and lower limits of each parameter are shown in Table 7. Each parameter is randomly initialized between the upper and lower limits.

The experiment takes the training set as the input of SSA-XGBoost model, extracts 0.3 as the verification set, sets the number of training rounds as 15, and outputs the minimum fitness value of each round of training and its corresponding parameter combination value. As can be seen from Figure 3, with the increase of iteration times, the fitness value of SSA-XGBoost model on the training set is declining, which proves that the model can automatically find a better value and is not easy to fall into the local optimal value. The best parameter combination of the final output of the model is [387, 16, 0.03]. That is, n_estimators is 387, max_depth is 16, and learning_rate is 0.03.

4.5. Demand Forecast and Evaluation Analysis

The experiment uses the trained SSA-XGBoost model (SSAX) to forecast the demand of the test set, and selects ARIMA, exponential smoothing (ES), decision tree (DT), GBDT, XGBoost (XGB) models as the comparison models. Randomly select the prediction results of a unit. The comparison between the predicted value and the real value of the unit is shown in Figure 4. It can be seen that the SSA-XGBoost model has a good prediction effect on the test set. It can not only better predict the trend of data changes, but also has a small prediction error.

According to the model prediction results, the determination coefficient, mean square error, root mean square error and mean absolute error of different models are calculated. According to the evaluation index results of each model in Table 8, the SSA-XGBoost model has the best fitting result among the six models, with the highest R^2 value of 0.988, indicating that SSA-XGBoost model has the best prediction effect on the experimental data set compared with the other five models.

Table 7. List of parameters to be optimized.

Parameter	Upper limit	Lower limit	Common difference
n_estimators	400	100	1
max_depth	30	5	1
learning_rate	0.3	0.01	0.01

Table 8. Comparison of evaluation indicators of different models.

Model	ARIMA	ES	SSAX	DT	GBDT	XGB
R^2	0.9279	0.7647	0.988	0.7224	0.9254	0.9524
MSE	7319.2118	7362.5743	703.5428	1083.3392	772.3229	964.4053
MAE	20.162	20.718	2.4518	2.1164	3.4001	4.936
RMSE	85.5524	85.8054	26.5243	32.9141	27.7906	31.0548
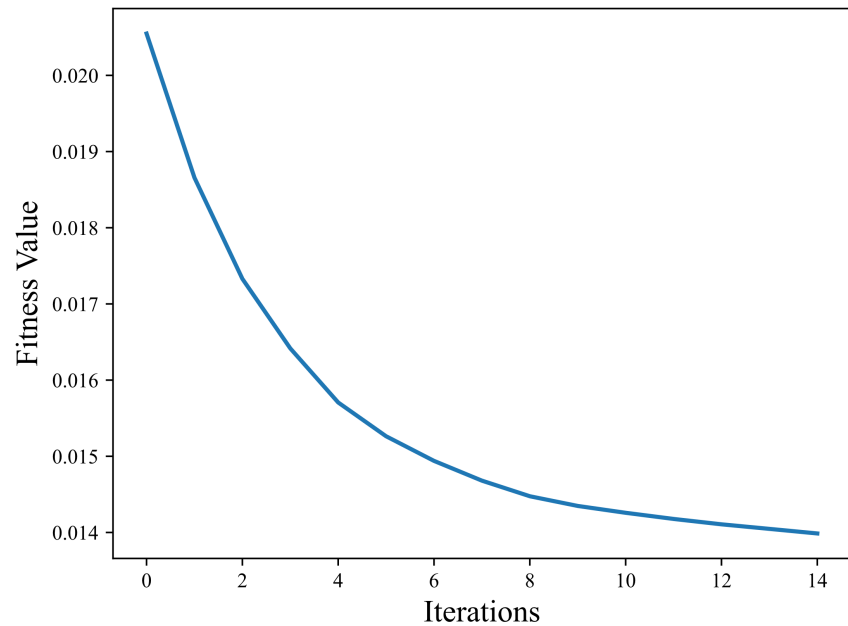
Figure 3. Minimum fitness value of SSA-XGBost model in each iteration.
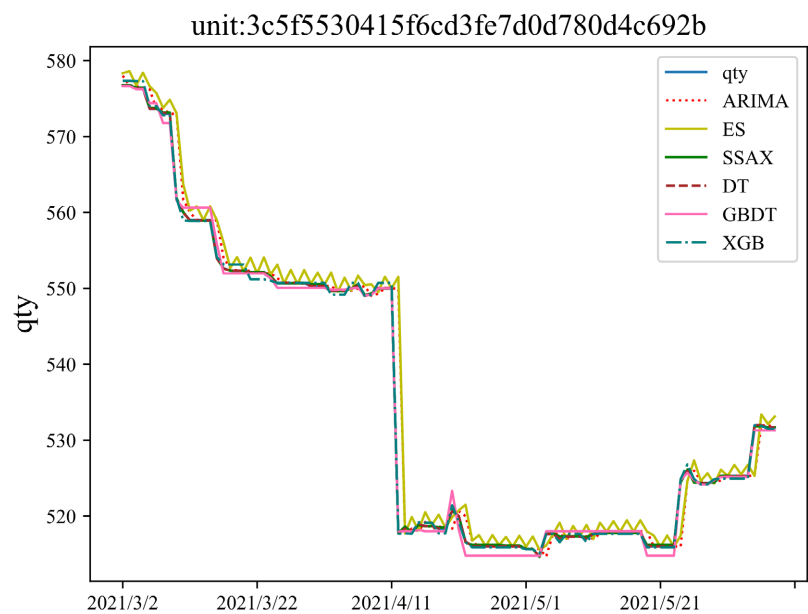
unit:3c5f5530415f6cd3fe7d0d780d4c692b

Figure 4. Comparison chart of model predicted value and real value.

5. Conclusions

Aiming at the problem of difficult supply chain demand forecasting, this paper uses enterprise historical data as a data set to analyze the influence of various influencing factors on demand, and proposes a supply chain demand forecasting model based on SSA-XGBoost model.

The future demand of this category can be effectively predicted by using the geographical level information, product level information, time information and historical consumption information of SKU. The experiment proves that in the

face of no obvious time sequence law for the change of single SKU demand and the overall change trend of each SKU demand is different, the SSA-XGBoost model can not only realize automatic parameter searching, but also accurately predict the daily demand of each SKU by improving the SSA parameter combination update method.

Automatic parameter searching can improve the problem that manual parameter searching is easy to fall into local optimal value. The supply chain demand forecasting model based on SSA-XGBoost can take into account the overall change trend of different SKU demand and the impact of changes in various factors on demand. It is applicable to the supply chain demand forecasting problem. It can help enterprises quickly and accurately feedback the demand information to the production side, reduce the information gap, reduce the transportation and inventory costs, and thus improve the efficiency of the entire supply chain.

In future research, the SKU demand forecast results and inventory situation can be used to build an inventory control and ordering model suitable for supply chain management to achieve efficient operation and reduce inventory costs.

Acknowledgements

Conflicts of Interest

The authors declare no conflicts of interest regarding the publication of this paper.

References

[1] Lummus, R.R., Krumwiede, D.W. and Vokurka, R.J. (2001) The Relationship of Logistics to Supply Chain Management: Developing a Common Industry Definition. *Industrial Management & Data Systems*, **101**, 426-431. https://doi.org/10.1108/02635570110406730

[2] Liu, J., Zhang, S. and Cao, W.J. (2002) A Case Study of an Inter-Enterprise Workflow-Supported Supply Chain Management System. *Operational Research*, **2**, 17-34. https://doi.org/10.1007/BF02940119

[3] Shah, R., Goldstein, S.M. and Ward, P.T. (2002) Aligning Supply Chain Management Characteristics and Interorganizational Information System Types: An Exploratory Study. *IEEE Transactions on Engineering Management*, **49**, 282-292. https://doi.org/10.1109/TEM.2002.803382

[4] Li, G.X., Ma, W.B. and Xia, G.E. (2021) Research on Logistics Demand Forecasting Model Based on Deep Learning. *Chinese Journal of Systems Science*, **29**, 85-89.

[5] Giri, C. and Chen, Y. (2022) Deep Learning for Demand Forecasting in the Fashion and Apparel Retail Industry. *Forecasting*, **4**, 565-581. https://doi.org/10.3390/forecast4020031

[6] Hamie, H., Hoayek, A. and Auer, H. (2021) Modeling Post-Liberalized European

Gas Market Concentration—A Game Theory Perspective. *Forecasting*, **3**, 1-16. https://doi.org/10.3390/forecast3010001

[7] Chen, S.M. (2021) Online Forecasting Model of Supply Chain Demand Based on Incomplete Sales Information. Ph.D. Thesis, South China University of Technology, Guangzhou.

[8] Liu, J.Y. (2020) Research on Key Technologies of LASSO Time Series Prediction and Recommendation System and Its Application in Supply Chain Management. Ph.D. Thesis, Shanghai Jiao Tong University, Shanghai.

[9] Wu, W.D. (2021) Research and Implementation of Household Appliance Demand Forecast Based on Multi Model Fusion. Ph.D. Thesis, Southwest University, Chongqing.

[10] Deng, Q. (2021) Research on Supply Chain Demand Forecast of L Company. Ph.D. Thesis, University of International Business and Economics, Beijing.

[11] Xie, F. (2020) Research on Demand Forecast and Comprehensive Production Plan of BD Shanghai Company. Ph.D. Thesis, Shanghai University of Finance and Economics, Shanghai.

[12] Xu, Z. (2020) Research on Demand Forecast and Inventory Management of P Company's Clothing Products. Ph.D. Thesis, Donghua University, Shanghai.

[13] Sheng, Z.G. (2014) The Application of Exponential Smoothing Method in the Forecast of the Demand for Product Oil Distribution in Sinopec Kunming Area. Ph.D. Thesis, Yunnan University, Kunming.

[14] Chen, L.P. (2020) Research on the Application of Data Mining Technology in the Supply Chain Management of Operators. Ph.D. Thesis, Jilin University, Changchun.

[15] Fang, K. (2020) Research on Demand Forecasting of Food Supply Chain Based on Machine Learning. Ph.D. Thesis, North China Electric Power University, Beijing.

[16] Zhu, D.Q. (2020) Research on Supply Chain Demand Forecasting Model Based on Data Mining. Ph.D. Thesis, Huazhong University of Science and Technology, Wuhan.

[17] Xiao, S.T. (2021) Research on Demand Forecast and Ordering Strategy of M Garment Enterprise. Ph.D. Thesis, Beijing Jiaotong University, Beijing.

[18] Chen, T.Q. and Guestrin, C. (2016) XGBoost: A Scalable Tree Boosting System. *Proceedings of the* 22*nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*, San Francisco, 13-17 August 2016, 785-794. https://doi.org/10.1145/2939672.2939785

[19] Xue, J.K. (2020) Research and Application of a New Type of Swarm Intelligence Optimization Technology. Ph.D. Thesis, Donghua University, Shanghai.