A Study on Promotional Shipment Forecasting for E-Commerce Merchants Based on ARIMA Time Series and LSTM Models

Weijie Zhang [1, #], Weicheng Hu [2, #], Junjie Zhu [1, #], Wendi Wu [1, #], Rusheng Zhou [2, *]

[1] School of Mechanical and Energy Engineering, Guangdong Ocean University, Yangjiang, China, 529500

[2] School of Computer Science and Engineering, Guangdong Ocean University, Yangjiang, China, 529500

* Corresponding Author Email: zrs@gdou.edu.cn

[#]These authors contributed equally

Abstract. E-commerce merchants usually want to bear low inventory costs while ensuring that the goods meet the demand, in order to sell a large number of goods, merchants will also carry out a variety of promotional activities, so e-commerce merchants need to accurately predict the promotional shipments of goods. In order to accurately predict the promotional shipments of goods, taking the double eleven data as an example, this paper adopts ARIMA time series analysis model and LSTM long and short-term memory model to predict the shipments of June promotions; the accuracy rate of 1-wmape of ARIMA time series model is 0.752, and the accuracy rate of 1-wmape of LSTM is 0.834, and the accuracy rate of LSTM is higher, the results are predicted by LSTM model. As a result, one of the merchants' promotional sales in June predicted by using LSTM model is 18, 19, 15, 12, 15, 7, 16, 22, 11, 32, 22, 25, 18, 16, 16, 19, 20, 22, 15, 13. However, the degree of recognition between the ARIMA time series model and the LSTM model prediction results is 80%, which can prove that the LSTM model reliability and rationality of the LSTM model.

Keywords: ARIMA, LSTM, Shipment Forecasting.

1. Introduction

With the rapid development of Internet technology, the e-commerce industry is rapidly emerging. In order to better understand market changes and provide better services and lower prices to consumers, e-commerce companies are constantly striving to improve their demand forecasting capabilities. Accurate demand forecasting of goods can effectively reduce out-of-stocks and lower inventory backlogs, and help companies make reasonable ordering strategies and inventory decisions. However, the volatility of the market and the instability of customers, as well as the increase in the degree of informationization and intelligence, have greatly affected the accuracy of demand forecasting. In order to accurately predict the promotional shipments of goods, Taghizadeh [1] used an artificial neural network (ANN) model to make predictions of future demand for extreme weather-sensitive retail goods, and the study showed that the ANN was more effective than the decision tree with the addition of the weather as an influencing factor. Bao Jixiang et al [2] took an enterprise's paper-based goods as the object and used LSTM neural network demand to do prediction, and the experiment achieved better prediction accuracy. Wang Yuanming [3] optimized the long short-term memory neural network (LSTM) into a neural network with adaptive training ability for predicting the monthly sales of 12 SKUs of a famous e-commerce food retailer, and the experimental results showed that the method is more effective than manually adjusting the parameters.

In this paper, in order to accurately predict the demand for goods during the large-scale promotion in June, we use the demand data of the Double Eleven promotion to make similar predictions, import the collected dataset of the Double Eleven into the original dataset, and redo the transcoding and clustering process, and make predictions from 2023-06-01 to 2023-06-20 based on the established

ARIMA time-series model and the LSTM (Long Short-Term Memory) model, confirming the applicability of the model by comparing the 1-wmap.

2. Analysis of the general idea

In order to accurately predict the shipment of goods, the collected double eleven dataset is imported into the original dataset using the inner join function in Excel, and the transcoding clustering process is re-conducted to predict the shipment of goods from 2023-06-01 to 2023-06-20 on the basis of the established ARIMA time-series model and the LSTM (Long Short-Term Memory) model, confirming the applicability of the model by comparing the 1-wmap. The Analysis of ideas is shown in Figure 1.

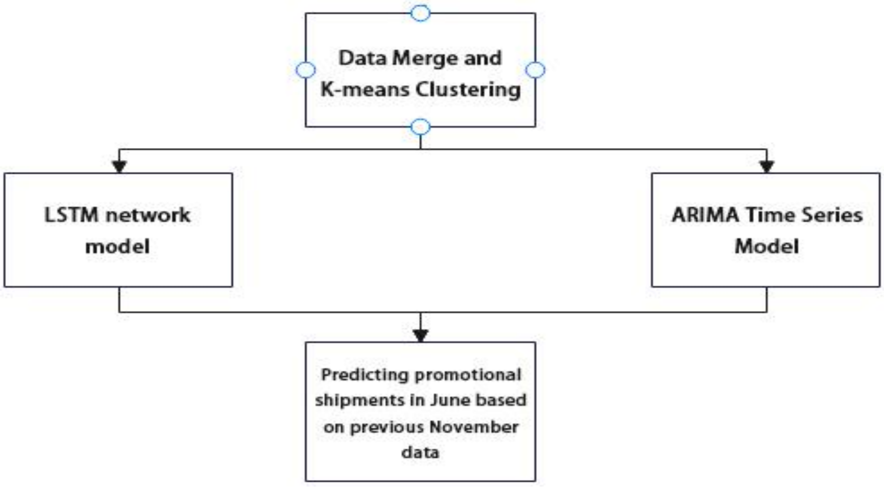

Figure 1. Analysis of ideas

2.1. Data integration and pre-processing

Due to the very large data collected, as well as the need for multiple datasets to be analyzed jointly, in order to make the analysis and prediction more convenient and efficient, matlab is used to merge multiple datasets and import the double eleven data to get a summarized dataset.

2.2. K-means clustering model based on the elbow method

Considering that the data samples provided in the problem are only data for which the indicators are not specifically delineated as a result, the unsupervised learning model is applicable.

Unlike tasks such as classification and sequence labeling, clustering is the process of dividing the samples into classes by the intrinsic relationships between the data without prior knowledge of any sample labels, resulting in high similarity between samples of the same class and low similarity between samples of different classes (i.e., increasing the intraclass cohesion and decreasing the class spacing)[4]

The elbow method is a way to determine the optimal K-value by using the relationship graph between the sum of squares of errors (SSE) and the K-value, and its algorithmic idea is that the SSE decreases gradually with the gradual increase of the clustering center (K-value). When the K value is smaller than the true number of clusters, the SSE changes more as the K value increases; when the K value is larger than the true number of clusters, the SSE changes less as the K value increases [5]. The expression of SSE, the core index of its elbow method, is shown in formula:

$$\mathrm{SSE} = \sum_{j=1}^{k} \sum_{p \in C_j} |p - m_i| \tag{1}$$

The k-clustering elbow method is shown in Figure 2.

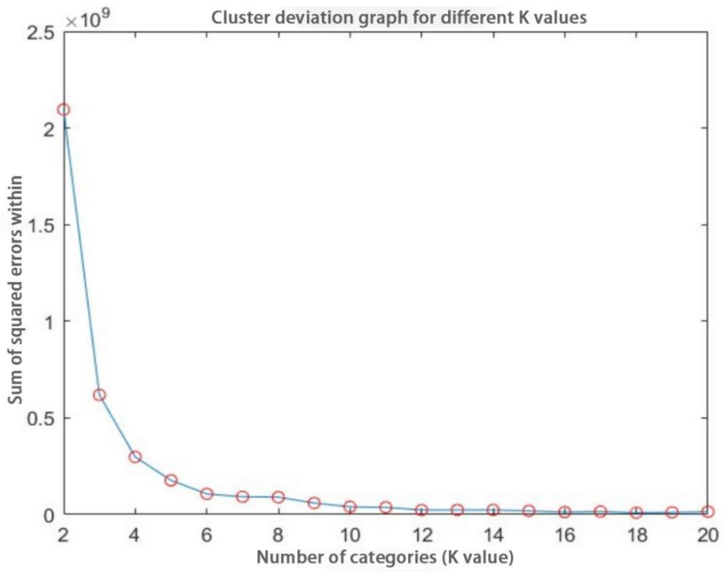

Figure 2. k-clustering elbow method

It can be observed that the k-value tends to stabilize at 10, so in this paper, the new merchants are merged together, and the table is divided into 10 classes. In this paper, the following clustering results are obtained by using matlab.Cluster situation as shown in Table.1.

Table 1. Cluster situation.

Cluster categories	frequent and continuous	Percentage(%)
Category 1	89078	25
Category 2	9732	3
Category 3	25249	7
Category 4	57582	16
Category 5	30434	8
Category 6	34252	10
Category 7	34161	9
Category 8	26920	7
Category 9	43774	12
Category 10	11578	3

2.3. Introduction to the LSTM network model

LSTM belongs to the unsupervised learning algorithm, i.e., it analyzes the attributes and characteristics of the data to make predictions based on the fact that it is not clear what factors are constraining the dependent variable[6].The structure of the recurrent unit of the LSTM network is shown in Figure 3.

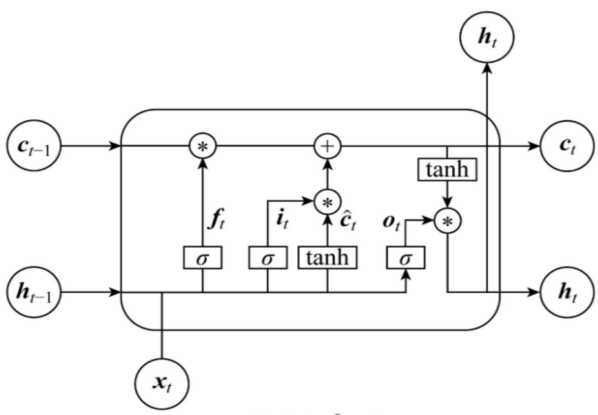

Figure 3. LSTM cycle unit structure

The improvement of LSTM on traditional RNN can be summarized in two aspects: on the one hand, it introduces a new internal state c_t in addition to the implicit state h, and c_t and h_t together serve as the network's memory content passed between cyclic units at adjacent moments.

On the other hand, a gating mechanism is introduced to selectively add new information to the memory content and forget the old information, thus controlling the rate of information accumulation The gating mechanism is three vectors controlling the information transfer, i.e., the forgetting gate f_t, the input gate i_t, and the output gate o. The gate vectors are obtained from the external input vector $x_t$at a certain time step and the unitary implicit vector h_{t-1} at the previous moment, through the fully-connected layer, and the activation function of the fully-connected layer is chosen as a Logistic function, also called Sigmoid function, denoted as, and its output value is (0, 1). The activation function is chosen as Logistic function, also known as Sigmoid function, denoted as $\sigma(.)$, and its output value is (0, 1).The LSTM forward propagation computation process is as follows [6-7].

(1) First calculate the 3 gate vectors as well as the candidate internal state vector c_t using the input vector x_t at the current moment and the implied vector h_{t-1} at the previous moment calculated as

$$\widehat{c_t} = tanh(W_c x_t + U_c h_{t-1} + b_c) \tag{2}$$

In the formula: W_c and U_c are the weight parameters; b_c is the bias parameter.

(2) Combining forgetting gate, input gate and candidate internal state to update the internal state First, the internal state of the previous moment is selected by the forgetting gate, and then the candidate internal state selected by the input gate is added to the internal state:

$$c_t = f_t * c_{t-1} + i_t * \widehat{c_t} \tag{3}$$

(3) Finally, the output gate is utilized to compute the implied vector of outputs at the current moment based on the internal state vectors after the activation of the tanh function:

$$h_t = o_t * tanh(c_t) \tag{4}$$

2.4. Implementation of the LSTM network model

In this paper, in order to explore the accuracy of different models in predicting the problem, here the LSTM method is used for prediction to ensure the uniqueness of the variables, and the same objects as in Problem 1 are selected for analysis and prediction as follows. LSTM Prediction Objects as shown in Table.2.

Table 2. LSTM Prediction Objects

Indicator Name	selected object
Merchant code	Seller_28
product code	Product_741
Warehouse code	Wh_1
Product Level 1 Classification	Pet life
Merchant classification	Pet Health
Inventory classification	A
Merchant scale	Large
Warehouse category	Central warehouse
Warehouse area	East China

The LSTM time model prediction of the above objects using matlab yields the following results, only some of which are shown due to space constraints.The Overall prediction results of the LSTM model is shown in Figure 4.

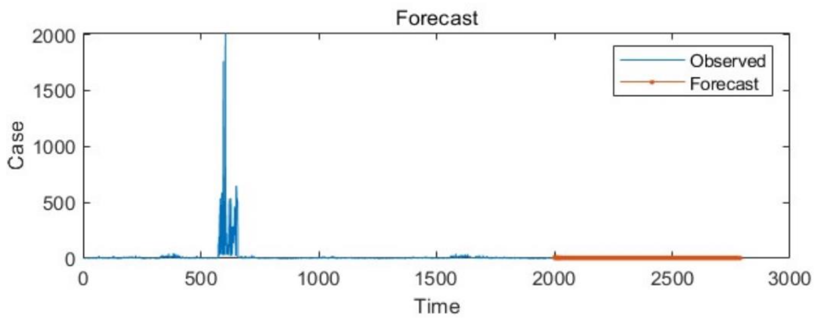

Figure 4. Overall prediction results of the LSTM model

The LSTM model prediction results is shown in Figure 5.

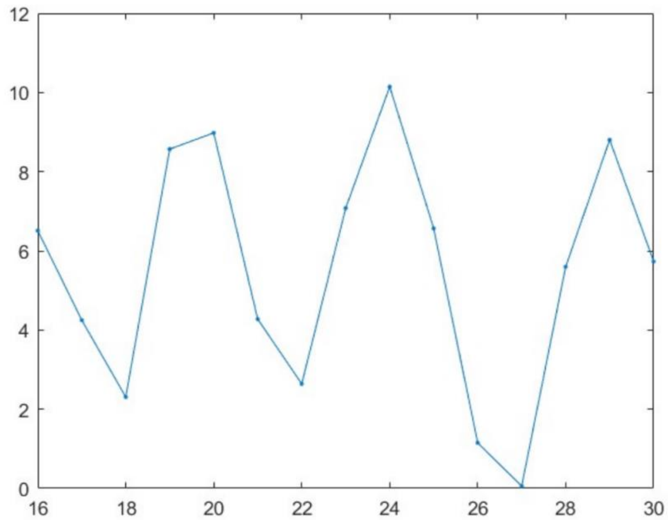

Figure 5. LSTM model prediction results

LSTM prediction results Objects as shown in Table.3.

Table 3. LSTM prediction results

seller_no	product_no	warehouse_no	date	forecast_qty
seller_28	product_741	wh_1	2023/6/01	18
seller_28	product_741	wh_1	2023/6/02	19
seller_28	product_741	wh_1	2023/6/03	15
seller_28	product_741	wh_1	2023/6/04	12
seller_28	product_741	wh_1	2023/6/05	15

2.5. ARIMA time series modeling

2.5.1 Principles of ARIMA time series modeling

The full name of the ARMA model is Autoregressive Moving Average Model and it is the most commonly used model for fitting smooth series. It is based on a combination of autoregression (AR) and moving average (MA) and can be used to analyze trends, seasonality and stochasticity in time series. It studies a series of data representing a phenomenon that is time-varying yet correlated, thus describing and exploring the regularity of the development of the phenomenon over time [8].

2.5.2 Determining the ARIMA (p,d,q) parameters

First draw the time series plot, in order to determine whether it is smooth or not, the ADF root test is introduced, and it can be found that t is closest to 0 when the difference of order 0, so d is 0;

Then the autocorrelation plot (ACF) and partial correlation plot (PACF) are used to determine the specific values of parameter p (autoregressive order) and parameter q (moving average order) [9].

Find p as 2 and q as 1;

2.5.3 Building an ARIMA (2, 0, 1) time series

The ARIMA time series prediction results is shown in Figure 6.

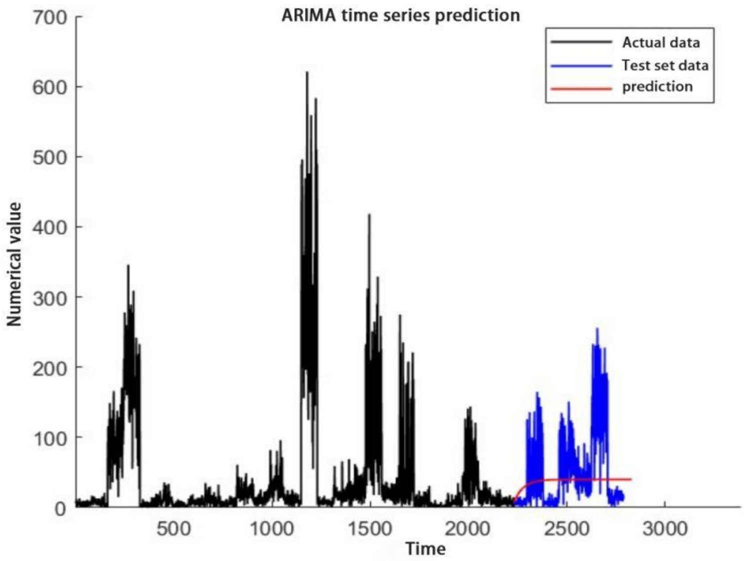

Figure 6. ARIMA time series prediction results

The goodness of fit is 0.714, and the results are well predicted with a high degree of confidence, as shown in the table below, with only some of the results shown due to space constraints. ARIMA forecasting results as shown in Table.4.

Table 4. ARIMA forecasting results.

seller_no	product_no	warehouse_no	Date	forecast_qty
seller_28	product_741	wh_1	2023/6/01	18
seller_28	product_741	wh_1	2023/6/02	16
seller_28	product_741	wh_1	2023/6/03	15
seller_28	product_741	wh_1	2023/6/04	12
seller_28	product_741	wh_1	2023/6/05	15

2.6. Comparing LSTM and ARIMA model results

There are many evaluation metrics for prediction accuracy, the more commonly used metric is 1-wmape, defined as follows:

$$1 - \mathrm{wmape} = 1 - \frac{|\Sigma\, y_i - \widehat{y_i}|}{\Sigma\, y_i} \tag{5}$$

Where y_i is the true demand for the ith sequence(the number of various commodities stored by the merchant in each warehouse on a daily basis) and $\widehat{y_i}$ is the predicted demand for the ith sequence [10].

By calculating the 1-wmape values, LSTM: 1-wampe = 0.834; ARIMA: 1-wmape = 0.752.

For comparison, it can be obtained that the LSTM model results are predicted more accurately, and the two corroborate each other, which strengthens the reliability, rationality and correctness of this paper[11].

3. Conclusions

In order to allow various e-commerce companies to serve consumers with lower prices and more adequate storage of goods during the June large-scale promotion, it is necessary to predict more accurately the demand for goods in the June large-scale promotion.This paper analyzes the data of the Double Eleven promotion, and based on the ARIMA time-analysis model and the LSTM model, it is possible to get the prediction results with a higher accuracy rate.

The experimental results show that the accuracy rate 1-wmape of ARIMA time series model is 0.752, and the accuracy rate 1-wmape of LSTM is 0.834, by comparing the two 1-wmape, it is concluded that the accuracy rate of LSTM is higher, which indicates that LSTM is more suitable for the e-commerce demand prediction to a certain extent, and one of the merchants predicted by using the LSTM model for the June promotional sales volume results are 18, 19, 15, 12, 15, 7, 16, 22, 11, 32, 22, 25, 18, 16, 16, 19, 20, 22, 15, 13, at the same time, the ARIMA time-series model and the LSTM model prediction results have 80% identity, which can prove the reliability and rationality of the LSTM model. Therefore, the use of LSTM can more accurately predict the demand for goods as well as effectively reduce out-of-stocks and reduce inventory backlogs, help enterprises to develop reasonable ordering strategies and inventory decisions, and reduce the waste of resources for goods.

References

[1] TAGHIZADEH E. Utilizing Artificial Neural Networks to Predict Demand for Weather-sensitive Products at Retail Stores[C]//International annual conference of the American Society for Engineering Management, 2017: 718-727.

[2] Bao Jixiang, Li Lin, ZHAO Mengge. E-commerce Demand forecasting based on LSTM model considering lag [J/OL]. Computer Engineering and Applications: 1-13 [2022-03-17].

[3] Wang Yuanming. Research on E-commerce demand forecasting based on LSTM neural network [D]. Jinan: Shandong University, 2018

[4] Chen Chao, WU Hui, Wan Qing et al. General aircraft after-sales support method based on K-means algorithm after optimal clustering [J]. Trainer Aircraft, 2023, (04): 52-56.

[5] Xie Jiaxin. Composition analysis of ancient glass based on decision tree and K-means clustering [J]. New technology and new products in China, 2023, (6): 66-69.

[6] Chen Siyu, Xu Aidi, Wang Zhenxu et al. Corn grain storage temperature prediction based on LSTM algorithm [J/OL]. Experimental technology and management, 1-7 [2024-01-31].

[7] Deng Chenhui, Zhang Jihan. Research on leakage prediction method of offshore LNG transfer system based on LSTM [J/OL]. Mechanics and practice, 1-15 [2024-01-31].

[8] Zheng Yan, Huang Xing, XIAO Yujie. Research on commodity demand forecasting model based on time series [J]. Journal of Chongqing University of Technology (Natural Science), 2019, 33(09): 217-222. (In Chinese)

[9] Zhai Xiaowei, Zhou Xiang, Song Bobo et al. Mine disaster prediction method based on the ARIMA model research [J/OL]. Coal technology, 1-6 [2024-01-31].

[10] Li Z Y. Research on product demand forecasting based on correlation clustering and product attribute classification [D]. Zhejiang university, 2019.

[11] Zheng Liyuan, Wu Xin, Li Jun et al. Electricity commodity demand forecasting based on TCN [J]. Journal of digital technology and applications, 2023, 9 (01): 110-113.