Big Data and Inventory Management

Guanghua Tang*, Yun Xie

Guangdong Neusoft University, Foshan 528225, China

*Email: tangguanghua@nuit.edu.cn

Abstract. This paper primarily explores the application of big data technology in inventory management, focusing on three main aspects: product demand forecasting, inventory optimization, and real-time inventory monitoring. Big data technology enables precise product demand predictions and inventory optimization, helping companies reduce inventory costs, improve inventory efficiency, mitigate inventory risks, and enhance customer satisfaction. This, in turn, boosts the company's competitiveness and profitability.

Keywords: Big Data, Inventory Management, Supply Chain, Demand Forecasting

1 Introduction

Inventory management plays a crucial role in a company's operations, especially in supply chain management, and is essential for business success. Excessive inventory leads to high holding costs and increased warehousing costs, occupying a significant amount of a company's working capital, reducing liquidity, and lowering capital utilization efficiency. Conversely, insufficient inventory can result in supply shortages, affecting customer satisfaction and sales opportunities, leading to sales losses and customer attrition. Effective inventory management helps companies balance inventory costs and sales risks, improve operational efficiency and cost-effectiveness, enhance customer satisfaction, and optimize the supply chain, thereby boosting market competitiveness.

As globalization and digitalization accelerate, the market environment companies face is becoming increasingly complex and dynamic. Inventory management, a core link in supply chain management, is becoming more significant. Traditional inventory management methods often rely on experience and simple forecasting models, which are inadequate in addressing the rapid changes and uncertainties of modern market demands. The rise of big data technology presents new opportunities and challenges for inventory management. Big data can integrate data from different channels, such as sales data, market trends, and customer behavior data. Through big data analysis, companies can extract valuable information from massive data sets and, using advanced algorithms and models, gain deep insights and predictive capabilities for accurate de-

C. W. K. Chen et al. (eds.), *Proceedings of the 9th International Conference on Engineering Management and the 2nd Forum on Modern Logistics and Supply Chain Management (ICEM-MLSCM 2024)*,
Advances in Engineering Research 243,
https://doi.org/10.2991/978-94-6463-531-7_18

mand forecasting, real-time inventory monitoring, and intelligent replenishment decisions. This optimizes inventory levels, reduces operational costs, and enhances customer satisfaction.

This paper examines the application of big data technology in inventory management, analyzing its role in improving inventory management efficiency, reducing costs, and enhancing customer satisfaction. Specifically, the study will cover the following aspects: improving the accuracy of demand forecasting, optimizing inventory levels, real-time monitoring and intelligent replenishment, and enhancing supply chain visibility.

2 Research Objectives and Significance

The main objectives of this study include:

2.1 Improving the Accuracy of Demand Forecasting:

Big data analysis can process vast amounts of data from multiple internal and external sources, including historical sales data, market trends, and seasonal changes. Demand forecasting model algorithms can predict product demand more accurately, helping companies avoid overstock or stockouts, minimize inventory backlogs and capital occupation, and reduce warehousing costs and risks.

2.2 Optimizing Inventory Levels

The study will explore how to use big data to optimize inventory levels, distributing inventory reasonably based on demand to reduce excess inventory while avoiding stockouts. Big data-driven decisions can manage inventory more precisely, lowering holding and warehousing costs and improving the utilization rate of working capital[3].

3 Research on the Application of Big Data in Inventory Management

3.1 Big Data-Based Demand Forecasting

Big data-based demand forecasting utilizes vast amounts of data and advanced analytical techniques to predict future market demand for products. Specifically, it involves constructing demand forecasting algorithm models based on big data technology. These models analyze historical sales data, market trends, consumer behavior, and seasonal variations to achieve more accurate predictions of future product demand[1]. The key steps in demand forecasting are illustrated in Figure 1.

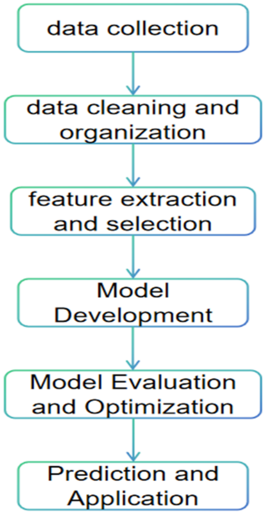

Fig. 1. Modeling steps for demand forecasting.

1) Data Collection.

Data collection is one of the key steps in demand forecasting based on big data. It involves gathering diverse data through multiple channels[6], including historical sales data of products, market data, customer data, macroeconomic data, competitor data, and data on major events, to obtain comprehensive market information.

2) Data Cleaning and Preparation.

Data cleaning and organization involve cleaning, transforming, and organizing the collected data. This includes removing duplicate data, handling missing values, standardizing data formats, and assessing data quality to ensure data quality, consistency, and suitability.

3) Feature Extraction and Selection.

Feature extraction is the process of extracting and selecting useful features from the collected raw data that can aid in prediction. This step is crucial for improving the performance of the model. Different types of data require different processing methods to establish a demand forecasting model. Below are some common feature extraction methods:

(1) Time Series Feature Extraction: Lag Features: These are demand values from a previous period of time. Lag features help the model capture trends and seasonal patterns in the time series.Rolling Window Features: These involve calculating statistics within a specific time window, such as mean, maximum, minimum, etc.Time Features: These involve extracting date and time information, such as year, month, day, quarter, day of the week, and whether it is a holiday or not (as shown in Figure 2).

```
data['year'] = pd.to_datetime(data['date']).dt.year
data['month'] = pd.to_datetime(data['date']).dt.month
data['day'] = pd.to_datetime(data['date']).dt.day
data['day_of_week'] = pd.to_datetime(data['date']).dt.dayofweek
data['is_weekend'] = data['day_of_week'].isin([5, 6]).astype(int)
```

Fig. 2. Extract time series features.

(2) Numerical Feature Extraction: For numerical data, statistical methods can be used to extract features such as mean, variance, median, etc. Data transformation methods such as polynomial feature transformation and dimensionality reduction can also be applied to extract higher-order numerical features.

Feature Selection: The primary method used is embedded feature selection, which combines feature selection with the model training process, allowing the model to learn the importance of features directly during training. Common methods include Lasso regression, decision trees, and support vector machines.

4) Building the Prediction Model.

Based on the collected data and selected features, future sales revenue can be predicted using statistical and machine learning algorithms to establish a demand forecasting model. To forecast future product demand, time series models such as ARIMA or Prophet can be utilized. These models leverage trends and seasonal information from historical data, employing it for training and optimization purposes. The specific modeling process proceeds as follows:

First, historical sales data is aggregated and analyzed annually to form a sales growth rate by comparing the annual sales volume with that of the previous year. The basic algorithm calculates the annual growth rate as the difference between the current year's sales volume and the previous year's sales volume divided by the previous year's sales volume. This process is repeated over multiple years to compute average growth rates and acceleration.

Next, the historical sales data for each year is further aggregated and analyzed monthly to determine the month-on-month growth rate of product sales for each month. The month-on-month growth rate is calculated by subtracting the sales volume of the same month in the previous year from the current month's sales volume, divided by the sales volume of the same month in the previous year.

Based on parameters such as annual and monthly average growth rates and acceleration, along with machine learning algorithms, a basic forecasting model can be developed. Two potential algorithms for modeling include:

(1) Moving Average Method (Moving Average)

The moving average method is a technique used to smooth time series data by calculating the average of the data over a specified number of periods. Common types include Simple Moving Average (SMA) and Weighted Moving Average (WMA).

Simple Moving Average Formula: $\mathrm{SMA}_t = \frac{1}{n}\sum_{i=0}^{n-1} Y_{t-i}$.

Where:

- SMAt is the moving average value at time t.
- Yt−i is the product demand at time t−i.
- n is the window size of the moving average.

(2) ARIMA Model (AutoRegressive Integrated Moving Average).

The ARIMA model combines both Autoregressive (AR) and Moving Average (MA) methods and includes differencing operations to make the time series data stationary.

The ARIMA(p,d,q) model formula is relatively complex, involving Autoregressive terms (AR), Differencing terms (I), and Moving Average terms (MA):

The general form of the ARIMA model is:

$$Y_t = c + \phi_1 Y_{t-1} + \phi_2 Y_{t-2} + \cdots + \phi_p Y_{t-p} + \epsilon_t + \theta_1 \epsilon_{t-1} + \theta_2 \epsilon_{t-2} + \cdots + \theta_q \epsilon_{t-q}$$

where:

- Yt is the product demand at time t.
- c is the constant term.
- Φ is the coefficient for the AR part.
- θ is the coefficient for the MA part.
- ϵt is the white noise error term.

Secondly, analyze the development trends of the product industry in recent years, the trends of competitors, and major events. Establish adjustment parameters for product forecasts and further optimize the basic prediction model using machine learning algorithms to refine the forecast results. Adjustment parameters are an essential part of the prediction model, fully accounting for the uniqueness of the future market, which can make predictions more accurate. There are two main types of adjustment parameters: weighted parameters (for weighted adjustment) and constant parameters (for incremental adjustment in the forecast results).

F（year，month）=f(year，month）* S1*S2*...*Sn +K1+K2+...+Kn

S：Weighted Parameters　K：Constant Parameters

Fig. 3. Product demand forecasting function.

5) Model Evaluation and Adjustment.

Evaluate and validate the established model by splitting the data, typically using 70% for training, 15% for validation, and 15% for testing. Assess the model's accuracy by comparing the prediction results with the actual sales data, and adjust and optimize the prediction model based on machine learning methods.

6) Prediction and Application.
Use the established demand prediction model to forecast future market demand. Based on the prediction results, develop corresponding product production plans, inventory strategies, and marketing strategies[8].

3.2 Inventory Optimization

Inventory optimization is based on the results of demand forecasting, aiming to improve inventory efficiency, reduce costs, and ensure supply chain stability and customer satisfaction through accurate demand prediction, intelligent inventory management, and optimization strategies[2]. The following are specific steps and methods:

1) Determining Service Levels and Safety Stock.
Establishing service level targets for different products to ensure meeting customer demand at specified service levels[5]. Initially, categorize products based on sales volume, profit margins, etc., into categories such as ABC, and determine the service levels for each category:

A-category products: Set a 95% service level, such as popular electronic products or current/future flagship products. B-category products: Set a 90% service level, including common household items with good sales. C-category products: Set an 85% service level, such as low-sales small commodities.

Based on product service levels, production or supply cycles, and demand standard deviations, calculate safety stock levels using the normal distribution. For example, if the demand standard deviation (S) for a specific A-category product is known, with a supply cycle of N days, and a 95% service level corresponding to a standard normal distribution value of 1.65, the formula for calculating safety stock for this product would be:

$$\text{Safety Stock} = 1.65 * S * \sqrt{N}$$

2) Determining Reorder Point (ROP) and Economic Order Quantity (EOQ).
Reorder Point (ROP) is the inventory level at which an order should be placed to replenish stock just as it reaches the safety stock level. Assuming the average daily demand for a product is X units and the supply cycle is N days, the formula for calculating ROP is:

$$ROP = X * N + \text{Safety Stock (units)}$$

Each product has associated costs such as ordering costs, logistics costs, and storage costs. When placing orders, these costs must be considered together to optimize overall costs. These costs are divided into two parts: ordering costs per order and holding costs per unit. Generally, larger order quantities reduce the average ordering cost per unit, but excessive ordering can increase overall holding costs. Therefore, the optimal balance is achieved by considering both factors. Assuming the annual demand is L units, the ordering cost per order is C dollars, and the holding cost per unit is P dollars, the formula for Economic Order Quantity (EOQ) is:

$$EOQ = \sqrt{(2 * L * C / P)}$$

Ordering according to EOQ minimizes total inventory costs (including ordering and holding costs).

3) Implementation of Inventory Optimization Strategy and Dynamic Inventory Optimization.

Inventory strategy encompasses safety stock levels, procurement cycles, order quantities, and more for each product. Based on these parameters, machine learning algorithms and optimization techniques (such as dynamic programming, genetic algorithms) are used to formulate and continuously optimize inventory strategies in real-time. Multi-level inventory management and supply chain coordination are considered to ensure efficiency and stability across the entire supply chain. Automated inventory management systems can be developed and deployed based on inventory optimization strategies, integrating functions such as demand forecasting, inventory optimization, and replenishment management. The system updates inventory data in real-time, generates replenishment orders, and optimizes strategies automatically[7].

4) Inventory Optimization Evaluation and Adjustment.

Regular evaluations assess the effectiveness of inventory optimization, analyzing key metrics (such as inventory turnover, stockout rates, holding costs, etc.). Based on evaluation results, adjustments are made to demand forecasting models and inventory optimization strategies to ensure continuous improvement[4]. Feedback from users and management is collected promptly to identify and resolve issues. Continuous improvement and optimization of inventory management strategies are conducted in response to market changes and supply chain conditions, enhancing system flexibility and adaptability.

4 Challenges and Responses of Big Data in Inventory Management

Big data involves a large amount of sensitive information, such as customer information, supplier information, and competitor information, making data breaches and security risks a significant challenge. Companies should strengthen data security and privacy protection measures by establishing strict data security management systems, such as data encryption, access control, and audit logs, to ensure data security and compliance. Additionally, they should increase investment in data security and privacy protection technologies, employing measures such as encrypted data storage, access permission control, access auditing, and data anonymization to safeguard data security and privacy.

5 Conclusion

The development of big data technology has brought new opportunities and challenges to inventory management for businesses. Through reasonable application and effective management, big data can accurately predict product demand, optimize inventory, improve inventory management efficiency, reduce inventory costs, and enhance supply chain operations, thereby boosting corporate competitiveness and promoting sustainable development. However, the application of big data in inventory management also faces challenges such as data security and privacy protection. Therefore, companies need to strengthen data governance and data management efforts, and enhance data security technologies and management measures to address these challenges and maximize the value of big data in inventory management.

References

1. Zhou Zhongwu, Tang Junpeng. Application of Big Data Technology in Enterprise Inventory Management [J]. Enterprise Economy, 2020, 39(5): 88-93.
2. Li Zhifeng, Wang Chunyan. Discussion on Optimization Strategy of Inventory Management Based on Big Data [J]. China Management Informatization, 2019, 22(6): 45-48.
3. Zhang Tao, Liu Ming. Inventory Management and Control in the Era of Big Data [J]. Commercial Economy Research, 2018(12): 102-104.
4. Wu Xiaowei, Chen Xiaodong. Optimization Model of Inventory Management in Big Data Environment [J]. Operations Research and Management, 2017, 26(3): 123-129.
5. Wang Ying, Li Na. Research on Inventory Management Based on Big Data Technology [J]. Science and Technology Enterprise, 2019(2): 56-59.
6. Liu Yang, Zhang Jing. Application of Internet of Things and Big Data Integration in Inventory Management [J]. Information Communication, 2018(7): 109-111.
7. Li Wei, Han Mei. Big Data-Driven Inventory Management Optimization Strategy [J]. Science and Technology Economic Guide, 2018, 26(36): 57-59.
8. Liu Xiaodong. Research and Application of Intelligent Inventory Management System [J]. Computer Applications Research, 2017, 34(5): 1547-1551.