WEP Warwick Evans Publishing

Academic Journal of Management Science and Engineering

ISSN: 3080-9827 (Print), ISSN: 3080-9835 (Online) | Volume 1, Number 2, Year 2026
DOI: https://doi.org/10.62051/ajmse.v1n2.14
Journal homepage: https://ajomse.com



Research on Chain Supermarket Inventory Optimization based on Multi-dimensional Spatio-temporal Data

Hanlin Zhou

Shanghai University, Shanghai, China

ABSTRACT

With the increasing scale and complexity of chain supermarkets, inventory management in multi-dimensional spatio-temporal data environments becomes more challenging, as traditional methods fail to handle regional demand heterogeneity. This paper proposes an integrated inventory management approach that leverages spatio-temporal data by combining the DeepAR probabilistic forecasting model with the newsvendor model for region-specific optimization. Sales regions are divided into five areas, and four demand-influencing features are selected. The DeepAR model, with a Gaussian likelihood, estimates demand distributions for each region. Experimental results show stable training convergence and high prediction interval coverage (PICP above 0.85 for most regions), while adaptively widening intervals in data-scarce or highly seasonal regions. These forecasts are then used in the newsvendor model to determine optimal order quantities by minimizing overage and underage costs. The proposed framework enables differentiated inventory strategies, improving service levels and operational efficiency.

KEYWORDS

Multi-dimensional Spatio-temporal Data; Inventory; Optimization; DeepAR.

1. INTRODUCTION

With the in-depth development of economic globalization, market competition has become increasingly fierce, especially in the retail industry. Chain supermarkets have become an important force in the retail industry due to their advantages in economies of scale, diversified commodities and high-qualityservices, but inventory management has always been a bottleneck restricting their development, facing challenges such as excessive or insufficient inventory, commodity loss and capital constraints. Existing studies mostly focus on specific factors such as customer demand and sales data, ignoring the impact of spatio-temporal data on inventory management. In fact, supermarket inventory management needs to take into account multi-dimensional data such as time, space and commodity categories, and customer demand varies in different time periods and regions, all of which will affect inventory management.

With the development of the Internet of Things and cloud computing, supermarkets can collect a large amount of multi-dimensional spatio-temporal data. Although deep learning technology is widely used in demand prediction, it tends to ignore the spatio-temporal differences, fluctuations and nonlinear problems of demand. To solve this problem, this paper comprehensively considers multi-dimensional influencing factors such as time, space and cost, and constructs an inventory optimization model under multi-dimensional spatio-temporal data to achieve accurate demand prediction and refined inventory management. This research has both theoretical and practical significance. It can not only promote the interdisciplinary development of operations research, data mining and other disciplines,

Content from this work may be used under the terms of CC BY-NC 4.0 licence (https://creativecommons.org/licenses/by-nc/4.0/).
Published by Warwick Evans Publishing.

but also provide scientific inventory strategies for chain supermarkets and improve their competitiveness and profitability.

2. LITERATURE REVIEW

(1) Research Status of Machine Learning-based Demand Forecasting

With the in-depth development of economic globalization, market competition has become increasingly fierce, especially in the retail industry. Chain supermarkets have become an important force in the retail industry due to their advantages in economies of scale, diversified commodities and high-quality services, but inventory management has always been a bottleneck restricting their development, facing challenges such as excessive or insufficient inventory, commodity loss and capital constraints. Existing studies mostly focus on specific factors such as customer demand and sales data, ignoring the impact of spatio-temporal data on inventory management. In fact, supermarket inventory management needs to take into account multi-dimensional data such as time, space and commodity categories, and customer demand varies in different time periods and regions, all of which will affect inventory management.

With the development of the Internet of Things and cloud computing, supermarkets can collect a large amount of multi-dimensional spatio-temporal data. Although deep learning technology is widely used in demand prediction, it tends to ignore the spatio-temporal differences, fluctuations and nonlinear problems of demand. To solve this problem, this paper comprehensively considers multi-dimensional influencing factors such as time, space and cost, and constructs an inventory optimization model under multi-dimensional spatio-temporal data to achieve accurate demand prediction and refined inventory management. This research has both theoretical and practical significance. It can not only promote the interdisciplinary development of operations research, data mining and other disciplines, but also provide scientific inventory strategies for chain supermarkets and improve their competitiveness and profitability.

With the increase in data availability in the big data environment, the demand for predicting product sales has also increased. To address this challenge, researchers have widely applied machine learning methods for prediction. Researchers such as Liu Zhi[1] determined the main characteristic indicators affecting demand through the collection and analysis of actual e-commerce data, and predicted the future total sales of e-commerce enterprises based on three machine learning models: neural network model, random forest model (RF) and support vector machine model (SVM). Their experimental results show that the optimal prediction accuracy can reach 93%. Wang Changqiong and Cao Nieqing[2] used the adaptive boosting algorithm in machine learning and integrated Baidu Index in the data processing stage. The final results show that this method has high accuracy in demand prediction and can effectively improve customer satisfaction. Wang Yuanming [3] established a neural network model with long short-term memory by analyzing the sales data of e-commerce enterprises in the past three years, and used this model to predict future demand. In addition, scholars such as Baardman[4] used a research method for new product demand prediction, performed cluster analysis on existing similar products and predicted the demand for new products, which can effectively improve the accuracy of demand prediction.

It can be seen from the above literature that the current research focus on demand prediction is mainly on improving prediction accuracy. At the same time, with the diversification and increasing availability of data, the method of using machine learning for demand prediction will be more widely applied. Some research literatures directly apply the demand prediction results to enterprise operation decisions, while others use them as input for subsequent optimization models. However, in these studies, most scholars only considered historical customer demand data, and few researchers introduced environmental characteristic data. The important contribution of this paper is that it

considers spatial characteristics, data of different varieties and product prices, thereby expanding the scope of data analysis and improving the accuracy of analysis.

(2)Data-driven Inventory Optimization: A Literature Review

In recent years, stochastic inventory optimization has become a key research direction in supply chain management, and scholars have conducted in-depth studies in this area. By collecting and analyzing inbound and outbound data from e-commerce enterprises to solve practical problems in inventory management, Li et al.[5] gathered a large amount of inbound and outbound information from Cainiao, a division of Alibaba Group. Through data analysis, they proposed a method for forecasting demand for high-volume products, determined the optimal shipment quantity, and resolved issues of unreasonable inventory allocation. Li Changchun[6] analyzed sales data from the Tianchi Big Data Platform and tackled the problem in two stages. In the first stage, machine learning models such as neural networks, random forests, and linear regression were used to predict future demand. In the second stage, they assumed demand follows a Gaussian distribution and derived the optimal inventory decision by fitting the parameters of the demand distribution. Ban et al. [7]proposed an approach for real-time multi-stage procurement of new products. By analyzing historical sales and feature indicators of similar existing products, they used linear regression to forecast future sales of new products, and then applied an improved scenario tree method to solve a multi-stage stochastic programming problem. Comparisons with real data from ZARA showed that this method can effectively reduce costs by 6% to 15%.

The above-mentioned literature all adopts a sequential decision-making approach [8], dividing the problem into two separate sub-problems, demand forecasting and inventory optimization, and processing them independently. In the demand forecasting stage, features are selected and input through data analysis to predict sales. Then, in the inventory optimization stage, the demand forecast results are used as intermediate data and input into the inventory optimization model to determine appropriate inventory decisions. However, the research objective of this paper is to consider the heterogeneity of demand across sales regions of chain supermarkets, analyze the demand characteristics of each sales region from the spatial dimension, and develop corresponding inventory optimization strategies tailored to the regional demand characteristics.

3. DEMAND DISTRIBUTION FORECASTING UNDER MULTI-DIMENSIONAL SPATIO-TEMPORAL DATA

The accuracy of demand forecasting critically impacts production planning. Overproduction causes inventory buildup, while underproduction leads to shortages, both preventing profit maximization. This chapter therefore trains the DeepAR algorithm on historical sales data from each region of Chain Supermarket A to forecast next-week sales, providing a sound basis for adjusting ordering strategies and reducing inventory costs.

(1) Feature Indicator Selection

Given the significant disparities in sales volume across different provinces and regions of Chain Supermarket A, this forecast spatially groups the various provinces and regions into five major areas to enable the DeepAR machine learning algorithm to accurately predict future sales based on each region's sales characteristics.

The spatial division of provinces is as follows: The Northern Region includes Liaoning, Jilin, Heilongjiang, Shandong, Shaanxi, Hebei, Tianjin, Ningxia Hui Autonomous Region, Beijing, Inner Mongolia Autonomous Region, Xinjiang Uygur Autonomous Region, Qinghai, and Gansu. The Southern Region includes the Hong Kong Special Administrative Region, Macao Special Administrative Region, Yunnan, Guangxi Zhuang Autonomous Region, and Guizhou. The Eastern Region includes Jiangsu, Zhejiang, Fujian, Guangdong, Hainan, and Taiwan. The Western Region

includes Chongqing, Sichuan, and the Tibet Autonomous Region. The Central Region includes Henan, Anhui, Hunan, Hubei, Jiangxi, and Shanxi.

By analyzing the chain supermarket's multi-dimensional dataset and synthesizing features related to demand variation mentioned in existing research, this paper selects four features—previous period sales, product category, purchase cost, and sales revenue—as demand influencing factors, as described below:

First, previous period sales: The time of the last sale of the product is taken as one period. Since previous period sales significantly influence current period sales, it is included as a feature.

Second, product category: People have different demands for different products. For example, sales of daily necessities are stable and less affected by seasonal changes, whereas non-essential goods are more influenced by promotional activities and seasons. Therefore, product category is included as a feature.

Third, product purchase cost: Different retail products have different purchase costs. The purchase volume of high-cost products affects the procurement quantity of other products, thereby impacting sales. Hence, purchase cost is included as a feature.

Fourth, product sales revenue: Retail products with high sales revenue often influence their standing among all products, affecting their purchase volume and shelf placement, which in turn impacts their sales. Therefore, sales revenue is included as a feature.

(2) DeepAR Model Construction

DeepAR is a time series model based on recurrent neural networks, widely applied in forecasting tasks. This model adopts an autoregressive approach and features hidden states. By learning periodic patterns in time series along with time-based covariates, DeepAR can make accurate predictions and plays an important role in the field of time series forecasting. Compared to manual processing, the DeepAR model can reduce the data processing workload required to handle highly complex and cross-correlated groups [9].

Let the value of the -th sequence at time step be denoted, which in this paper represents the feature indicators affecting sales volume, and let denote the forecast start time. DeepAR predicts the probability distribution based on an autoregressive recurrent neural network, expressed using a likelihood function.

The model is illustrated in the figure below, with the training process on the left and the prediction process on the right.

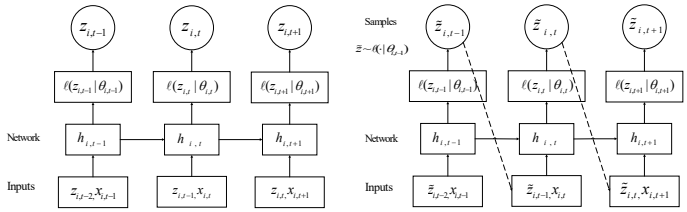


Figure 1. Training Process and Prediction Process

During training, at each time step, the model's input includes four feature indicators, the sales volume from the previous time step, and the state from the previous time step. The current state is first computed, then the parameters of the likelihood are calculated, and finally the network parameters are learned by maximizing the log-likelihood.

$$\ell = \sum_i \sum_t \log l(z_{i,t} \mid \theta(\vec{h}_{i,t})$$

t_0 represents the starting point for our prediction of demand. [1, t_{0-1}] is the training interval, and [t_0, T] is the prediction interval. In this case, the training dataset starts on January 1, 2018, and ends on

August 24, 2018. After training is completed, historical data earlier than August 24, 2018, is fed into the network to obtain the initial state . Based on the ancestral sampling method, the prediction results are obtained as follows: for t_0, $t_{\{0+1\}}$, at each time step, a sample is obtained through random sampling and used as the input value for the next time step. By repeating this process continuously, sampling values from t_0 to T can be obtained. Calculating these sampling values yields the data we need, such as expectation, variance, and quantiles.

The likelihood function determines θ, while statistical features determine the likelihood function we choose. Common likelihood function models include the Gaussian distribution and the negative binomial distribution. The Gaussian distribution is suitable for real-valued data, whereas the negative binomial distribution is suitable for non-negative integer data. There are also some other likelihood function models, such as the t-distribution, Bernoulli distribution, and beta distribution. When using the Gaussian distribution, it is parameterized using the mean and standard deviation, i.e., θ = (μ, σ). The affine transformation function of the neural network outputs the value of the mean \mu, and the affine transformation of the softplus activation function can obtain the value of σ[10]. Among them,

$$l_G(z \mid \mu, \sigma) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left(\frac{-(z-\mu)^2}{2\sigma^2}\right)$$

$$\mu(h_{i,t}) = \vec{w}_{\mu}^{T}\vec{h}_{i,t} + b_{\mu}$$

$$\sigma(\vec{h}_{i,t}) = \log(1 + \exp(\vec{w}_{\sigma}^{T}\vec{h}_{i,t} + b_{\sigma}))$$

Since daily sales volume fluctuates randomly and the goods sold vary with seasonality, with sales quantities reaching peaks under certain conditions, this variation trend conforms to a Gaussian distribution. Therefore, this paper adopts the Gaussian distribution as the likelihood model.

(3) Evaluation Metrics

PICP——Prediction Interval Coverage Probability

$$PICP = (1-a) * 100\%$$

It represents the proportion of actual observations contained within the prediction intervals generated by the model. For a given dataset, a PICP value closer to 1 indicates higher prediction accuracy of the model. Here, α denotes the confidence level of the prediction interval, typically ranging from 0 to 1. Specifically, α represents the proportion of actual observations that are expected to fall within the prediction interval. For instance, for a prediction interval with a 90% confidence level, α= 0.1.

PINAW — Prediction Interval Normalized Averaged Width

$$PINAW = \left(\frac{1}{n}\right) * \sum\left(\frac{w_i}{U-L}\right)$$

where n is the number of prediction intervals, w_i is the width of the i-th prediction interval, U is the upper bound of the i-th prediction interval, and L is the lower bound of the i-th prediction interval. For each prediction interval, the width w_i can be calculated by subtracting the lower bound L from the upper bound U. Then, the width of each prediction interval is standardized by the difference between its upper and lower bounds to eliminate dimensional differences among them. Finally, averaging all the standardized widths yields PINAW. A smaller PINAW value indicates that the average width of the prediction intervals generated by the model is narrower, meaning the prediction uncertainty is lower. Such prediction intervals are generally more reliable and accurate. It represents the average width of the prediction intervals generated by the model after standardization. A smaller PINAW value indicates that the model's prediction intervals are more compact and accurate.

4. ANALYSIS OF PREDICTION RESULTS

(1) Analysis of the Model Training Process

At the beginning of training, the model's initial parameters are set randomly, so the prediction results may be inaccurate. However, through the iterative process of backpropagation and optimization algorithms, the model gradually adjusts its parameters based on feedback from the loss function, bringing the prediction results progressively closer to the true values. The decrease in the loss function value reflects the improvement in the model's predictions and the increase in accuracy.

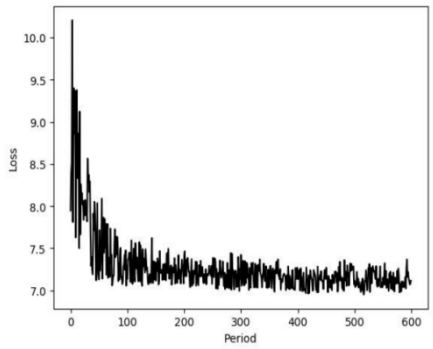


Figure 2. Training Loss Curve of the DeepAR Model

As the number of iterations increases, the loss function value of the model's prediction results gradually decreases and eventually stabilizes around 6.5, indicating that the accuracy of the DeepAR model training is continuously improving. A smaller loss function value suggests better performance, with relatively ideal accuracy on both the validation and test sets.

(2) Demand Forecasting Accuracy Analysis

This paper employs the DeepAR model, a time series forecasting framework, to predict sales volumes across different regions of a chain supermarket, and analyzes the prediction accuracy using the PICP and PINAW evaluation metrics, as shown in Table 1 below:

Table 1. Sales Price Information for the Western Region

Region	PICP	PINAW
Northern Region Prediction Results	0.864	1.555
Southern Region Prediction Results	0.907	1.451
Eastern Region Prediction Results	0.867	2.506
Western Region Prediction Results	0.857	1.627
Central Region Prediction Results	0.952	4.149

Judging from the calculation results of the two metrics, except for the Southern and Central regions, the prediction intervals of the RNN-based autoregressive recurrent neural network model (DeepAR) are able to cover the vast majority of actual prediction values, with PICP values remaining above 0.85. In the Southern region, which is characterized by relatively weaker consumption capacity, fewer sales data are available, resulting in a limited amount of obtainable data. To ensure that actual sales volumes are predicted as accurately as possible, DeepAR forecasts a relatively wide range of sales volumes, thereby safeguarding inventory under conditions of data scarcity. As for the Central region, the sales data are sparse and seasonal demand fluctuations are significant, leading to sudden increases

in sales during certain months. To ensure that the actual values fall within the prediction intervals, the DeepAR model generates wider prediction intervals.

Comparative Analysis of Demand Forecasting Results

In this section, the DeepAR model is used to forecast the sales volume for the upcoming week in each of the five regions, where the upper and lower bounds of the predictions correspond to the predicted values at the 95% and 5% confidence levels, respectively.



Figure 3. Forecasting Performance Charts for the Five Regions

Judging from the forecasting performance charts, the 90% prediction interval output by the DeepAR model is able to cover the majority of actual observed values. This indicates that the model effectively captures the uncertainty in future sales and provides a reliable range within which the true sales figures are likely to fall. Consequently, the DeepAR model can be used to predict the sales volume of each sales region of the chain supermarket for the upcoming week with a relatively high degree of accuracy, offering a dependable foundation for subsequent inventory management and ordering decisions.

5. CONCLUSION

Based on the comprehensive analysis presented above, the following conclusions can be drawn regarding the application of the DeepAR model for multi-regional demand forecasting in chain supermarkets.

This study employed the DeepAR model, a probabilistic time series forecasting framework built upon autoregressive recurrent neural networks, to predict weekly sales volumes across five distinct sales regions (North, South, East, West, and Central). The model training process demonstrated a clear convergence, with the loss function value declining steadily and eventually stabilizing around 6.5,

indicating that the iterative optimization of parameters effectively improved the model's predictive accuracy and generalization on both validation and test sets.

The performance of the prediction intervals was quantitatively evaluated using two metrics: the Prediction Interval Coverage Probability (PICP) and the Prediction Interval Normalized Averaged Width (PINAW). The results revealed that for the Northern, Eastern, and Western regions, the DeepAR model achieved PICP values above 0.85, meaning that the prediction intervals successfully covered the majority of actual sales observations while maintaining relatively narrow interval widths. This confirms the model's ability to produce well-calibrated uncertainty estimates for regions with stable demand patterns and sufficient historical data. In contrast, the Southern region exhibited a lower PICP but a narrower interval, which can be attributed to limited data availability due to weaker consumption capacity in that area. To compensate for data scarcity and ensure inventory protection, the model appropriately generated wider predictive intervals, prioritizing coverage over precision. The Central region, characterized by sparse sales records and pronounced seasonal demand fluctuations, also necessitated wider prediction intervals to capture sudden demand spikes during certain months, as reflected by the highest PINAW value (4.149) and a PICP of 0.952. These regional differences highlight the model's adaptability: DeepAR automatically learns region-specific temporal patterns and adjusts its predictive distributions accordingly, thereby effectively handling the inherent heterogeneity across different sales territories.

A comparative analysis of the demand forecasts further validated the model's practical utility. By setting the upper and lower bounds at the 95% and 5% confidence levels, respectively, the 90% prediction intervals generated by DeepAR were able to envelope the vast majority of actual sales observations for all five regions. The visual inspection of the forecasting charts confirmed that the model captures both the central tendency and the variability of future demand. This ability to quantify uncertainty is particularly valuable for inventory management, as it allows decision-makers to balance the risks of overstocking and stockouts.

In summary, the DeepAR model proves to be an effective tool for multi-dimensional spatio-temporal demand forecasting in chain supermarket contexts. It successfully identifies and adapts to region-specific demand characteristics, providing reliable probabilistic forecasts that serve as a scientific basis for subsequent inventory optimization. The integration of such data-driven demand predictions into the newsvendor framework can enable differentiated, region-tailored ordering strategies, ultimately enhancing inventory satisfaction and service levels. Future work could explore the incorporation of additional covariates—such as promotional activities, macroeconomic indicators, and local events—to further refine prediction accuracy, especially in data-sparse or highly volatile regions.

REFERENCES

[1] Liu Zhi. Research on E-commerce Sales Forecasting Based on Data Mining [D]. Beijing Jiaotong University, 2018.

[2] Wang Changqiong, Cao Mieqing, Wang Yanli, et al. Research on Combined Forecasting of E-commerce Order Quantity Based on Fusing Baidu Index [J]. Computer Engineering and Applications, 2018, 54(12): 21-225.

[3] Wang Yuanming. Research on E-commerce Demand Forecasting Based on LSTM Neural Network [D]. Shandong University, 2018.

[4] Baardman L, Levin I, Perakis G, et al. Leveraging comparables for new product sales forecasting [J]. Available at SSRN 3086237, 2017.

[5] Li X, Zheng Y, Zhou Z, et al. Demand Prediction, Predictive Shipping, and Product Allocation for Large-scale E-commerce [J]. Predictive Shipping, and Product Allocation for Large-Scale E-Commerce, 2019.

[6] Li Changchun. Commodity Demand Forecasting and Warehouse Allocation Planning under Big Data Background [J]. Mathematics in Practice and Theory, 2017, 47(07): 70-79.

[7] Ban G Y, Gallien J, Mersereau A J. Dynamic procurement of new products with covariate information: The residual tree method [J]. Manufacturing & Service Operations Management, 2019, 21(4): 798-815.

[8] Huber J, Muller S, Fleischmann M. A Data-Driven Newsvendor Problem: From Data to Decision [J]. European Journal of Operational Research, 2019, 278(3): 904-15.

[9] Bai Yuntao. Analysis of optimal allocation of nuclear power based on newsboy model [J]. Earth and Environmental Science, 2020, 510: 1315-1755.

[10] Zhu Gang, Li Wen, Du Shouguo, et al. Time Series Forecasting Based on Deep Learning Model DeepAR and Its Application Examples [J]. E-Business Journal, 2020, No.247(07): 83-86.