Implementation Of Pricing Strategy and Replenishment Strategy for Fresh Products Based on Particle Swarm Optimization Algorithm

Qing Wang, Haoxiang Wang *

College of Software, Henan Normal University, Xinxiang, China, 453000

* Corresponding Author Email: wanghaoxiang@stu.htu.edu.cn

Abstract. Vegetable pricing and replenishment strategies are crucial issues in the retail industry. Globally, as populations grow and market competition intensifies, research into how to determine prices and optimize replenishment strategies is increasingly critical. In this article, we focus on pricing and replenishment strategies for fresh vegetables. First, this paper uses the FP-growth algorithm to calculate multiple vegetable combinations that are often purchased together. Then, to explore the relationship between sales volume and cost-based mark-up pricing, we used the ARIMA model to predict and then used linear regression to derive a fitting regression equation for the relationship between overall sales and price. Finally, the particle swarm optimization algorithm is employed to maximize total profit through optimal price and replenishment strategies. Based on this result, the final pricing strategy is formulated and the corresponding replenishment strategy is formulated. The significance of our research on this project is to improve the accuracy and efficiency of price and replenishment decisions through optimization algorithms, thereby maximizing profits and improving the effectiveness of supply chain management. This research contributes to the ongoing development of data-driven decision-making models in the retail industry, particularly in the area of perishable goods management.

Keywords: Dynamic Time Planning, Particle Swarm Optimization Algorithm (PSO), ARIMA Model.

1. Introduction

Fresh vegetables are perishable and require a storage environment with high humidity and temperature control, which results in significant fluctuations in their value and market demand over time [1]. Due to their short shelf life, these products must be sold on the same day, requiring retailers to make quick and accurate replenishment decisions based on real-time sales data and market forecasts [2]. The decision to replenish inventory is contingent upon not only current market demand but also the preceding day's sales data [3]. However, since purchases of vegetables are typically made in the early morning when market information is limited, merchants frequently find themselves in a position where they must make replenishment decisions without access to accurate demand forecasts. In this context, the conditions of single-product purchasing and the fluctuations in purchase prices introduce an element of uncertainty and complexity to the management of inventory. Furthermore, the majority of pricing strategies for vegetables are based on a cost-plus model, with discounts offered for vegetables exhibiting minor damage or poor quality. The sale of high-quality vegetables is typically conducted at a mark-up, whereas defective vegetables are disposed of at a discount. Such price adjustments contribute to the margin uncertainty faced by merchants in managing their inventories.

Existing literature highlights various attempts to optimize pricing and replenishment strategies using data mining and optimization algorithms. For instance, the FP-growth algorithm has been used in market basket analysis to identify frequently purchased item combinations, providing retailers with insights for bundling and promotions [4]. However, FP-growth assumes that product relationships are independent, missing the opportunity to capture the non-linear interdependencies that emerge in real-world market dynamics [5]. Furthermore, ARIMA models have been employed to forecast sales trends based on historical data, yet they are limited by their linear assumptions and struggle with modeling complex demand fluctuations caused by external factors such as weather, holidays, or sudden shifts in consumer behavior. These shortcomings mean that while both FP-growth and ARIMA can offer useful insights, they are insufficient when faced with the unpredictable nature of vegetable sales.

Building on these limitations, this study introduces an innovative approach by combining FP-growth, ARIMA, and Particle Swarm Optimization (PSO) to optimize both pricing and replenishment strategies for fresh vegetables [6]. The key innovation of this study lies in the integration of these three techniques, addressing the inherent shortcomings of traditional methods. The FP-growth algorithm is used to identify frequent vegetable combinations, while ARIMA provides short-term demand forecasts. The PSO algorithm is then applied to dynamically optimize pricing and replenishment decisions based on these forecasts [7]. Additionally, an improved version of the PSO algorithm, incorporating nonlinear inertia weights, is introduced to enhance convergence speed and the coverage of global optimal solutions. This hybrid approach not only provides a more flexible and responsive strategy for pricing and replenishment but also offers a practical solution to real-time decision-making challenges in a highly dynamic market environment [8]. By fitting regression equations, this paper develops appropriate pricing formulas for different categories of vegetables, with the objective of maximizing profitability while ensuring market competitiveness [9]. This research builds upon and extends existing literature on vegetable pricing and inventory management, providing merchants with more practical solutions in a highly volatile market [10].

2. Main algorithms and models

2.1. Basic Principle of FP-growth Algorithm

The FP-Growth algorithm is an association algorithm that employs a divide-and-conquer strategy. This strategy involves compressing the database of itemsets into a tree (FP-tree) while maintaining the information regarding the association between itemsets. The algorithm employs a data structure known as a tree, specifically a frequent pattern tree. An FP-tree is a specific type of prefix tree, comprising an item header table and an item prefix tree. The FP-Growth algorithm facilitates the acceleration of the mining process based on the aforementioned structure. The algorithm is primarily utilized in the domains of transaction data, association rule mining, and other pertinent applications within the field of data mining. A detailed illustration of the FP-growth algorithm is presented in Figure 1.

Support is used to measure how often a particular item set (e.g., vegetable combinations) appears in the total transaction. The formula is as follows:

$$\text{Support(A)}=\frac{\text{Count(A)}}{\text{N}} \tag{1}$$

Where Support(A) is the support of the itemset A, Count(A) is the number of times A appears in the transaction data, N is the total number of transactions.

The degree of support reflects the prevalence of the item set A. For example, if the combination broccoli and Wuhu peppers occurs 200 times in 1,000 transactions, then its support is:

$$\text{Support (broccoli and Wuhu peppers)} =\frac{200}{1000}=0.2 \tag{2}$$

Confidence is used to measure the number of times the term set A occurs in the presence of the term set B also occurs. The formula is as follows:

$$\text{Confidence(A}\Rightarrow\text{B)} =\frac{Support(A\cup B)}{Support(A)} \tag{3}$$

Where Confidence(A ⇒ B) is Confidence of rule A⇒B, Support($A\cup B$) is Support for simultaneous occurrence of itemsets A and B, Support(A) is Support of the itemset A.

Confidence reflects the strength of rule A⇒B. For example, if ‘broccoli’ appears in 300 transactions, and ‘broccoli and Wuhu peppers’ appear in 150 transactions, the confidence level of the rule broccoli ⇒Wuhu peppers has a confidence level of:

$$\text{Confidence (broccoli} \Rightarrow \text{Wuhu peppers)} = \frac{150}{300} = 0.5 \tag{4}$$

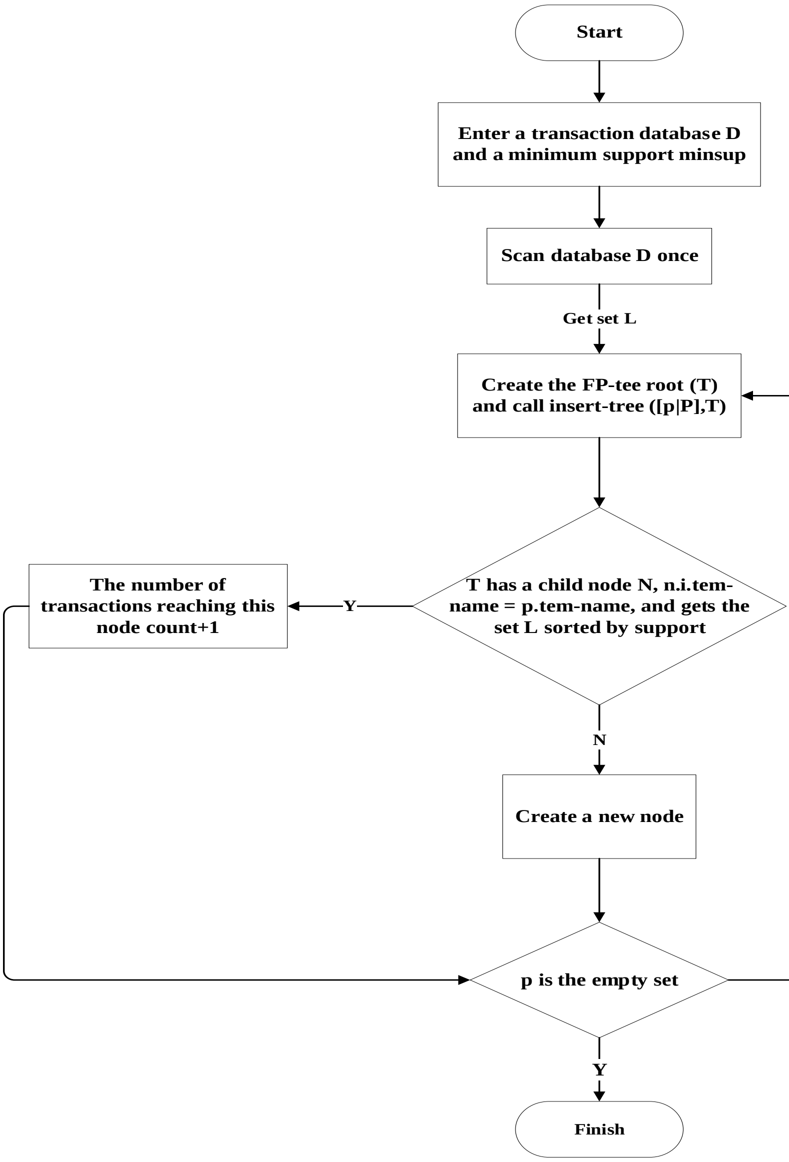

Figure 1. Detailed illustration of the FP-growth algorithm

2.2. ARIMA Model

The AutoRegressive Integrated Moving Average (ARIMA) model is a statistical model for time series and prediction. It combines autoregressive (AR), integral (I), and average (MA) components. ARIMA is commonly used for non-stationary time series data, transforming such data into a more stable form through differencing, a non-stationary time series can be transformed into a stationary time series. The modeling of the ARIMA model is illustrated in Figure 2. This model is particularly suited to forecasting trends in fresh product sales, where daily fluctuations in demand can lead to high variability.

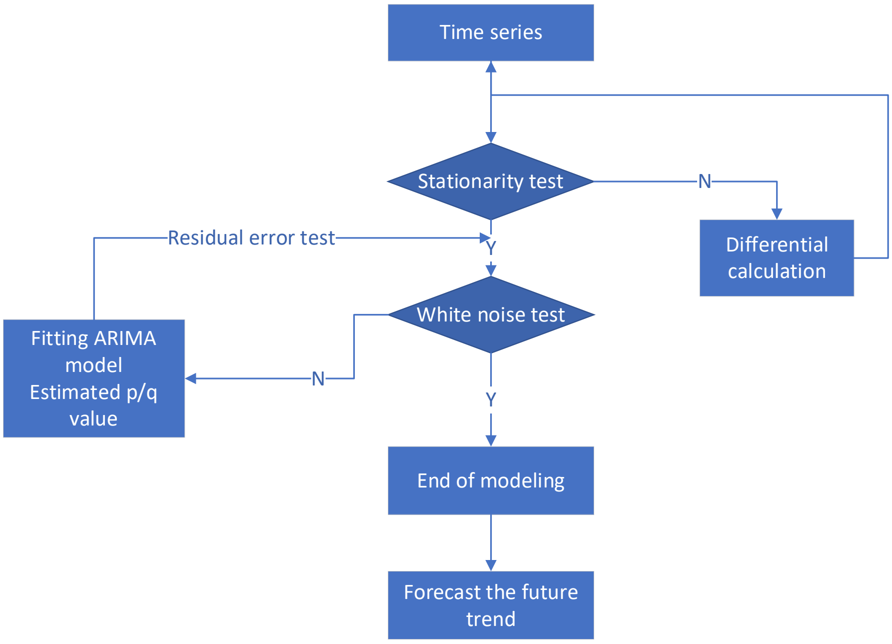

Figure 2. The modeling of the ARIMA model

The ARIMA model is comprised of three fundamental components. The ARIMA model is comprised of three fundamental components: auto-regressive (AR), integrated (I), and moving average (MA).

In the context of ARIMA, the AR and MA models are, respectively, the auto-regressive and moving average models. The "I" method is a different method. The difference calculation ensures the stability of the data. By combining the autoregressive model (AR) and the mean average model (MA) with the difference method (I), we have developed a differential autoregressive mean model, designated ARIMA (PDQ). In this model, d represents the order in which the data must be differentiated. ARIMA is the ARMA model that has undergone a different operation.

The Formula of the AR Model:

$$Y_t = \mathrm{c} + \oint_1 Y_{t-1} + \oint_2 Y_{t-2} + \cdots + \oint_p Y_{t-p} + \xi_t \tag{5}$$

The Formula of the MA Model:

$$Y_t=\mu+\varepsilon_t+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\ldots+\theta_q\varepsilon_{t-q}, \tag{6}$$

The Formula of the ARIMA Model:

$$Y_t=\mathrm{c}+\oint_1 Y_{t-1}+\oint_2 Y_{t-2}+\ldots+\oint_p Y_{t-p}+\theta_1\varepsilon_{t-1}+\theta_2\varepsilon_{t-2}+\ldots+\theta_q\varepsilon_{t-q} + \varepsilon_t, \tag{7}$$

2.3. Particle Swarm Optimization Algorithm

Particle Swarm Optimisation (PSO), a branch of evolutionary computing, is a stochastic search algorithm that simulates biological activity in nature. PSO simulates the process of feeding birds and fish in nature. The global optimal solution to the problem is found by working together as a group. It was proposed by American scientists Eberhart and Kennedy in 1995 and is now widely used for optimization problems in various engineering fields.

In the particle swarm optimization algorithm, the optimization results are obtained by randomly generating a certain number of particles with velocity and position as effective solutions in the search space of the problem and then carrying out iterative searches to determine the adaptation values of the particles through the adaptation function corresponding to the problem.

Assuming that there are N particles in the D-dimensional search space, each representing a solution, there are:

Position of the i-th particle:$X_{id}= (X_{i1},X_{i2},\cdots,X_{iD})$, Velocity of the i-th: particle:V_{id}= (V_{i1},V_{i2},⋯,V_{iD}),

The optimal position searched by the i-th particle:P_{id}= (P_{i1},P_{i2},⋯,P_{iD}),

Optimal position searched by the swarm:$P_{id,pbest}$= ($P_{1,gbest}$,$P_{2,gbest}$,⋯,$P_{D,gbest}$)

Adaptation value of the optimal position searched by the i-th particle:f_p,

Adaptation value of the optimal position searched by the population:f_g.

Flowchart of particle swarm optimization algorithm is illustrated in Figure 3.

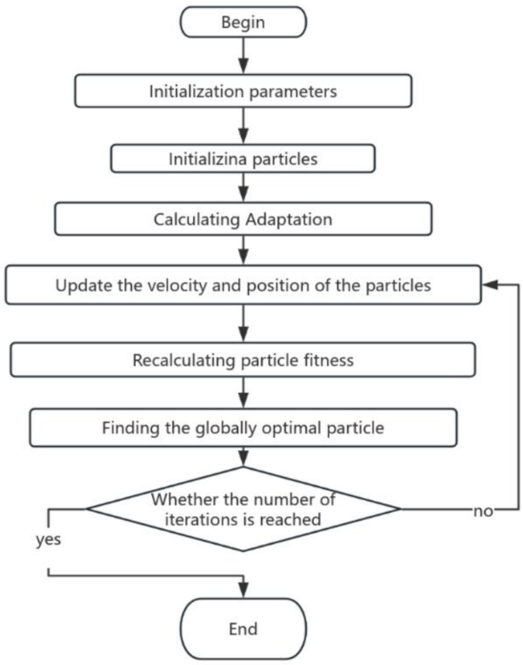

Figure 3. Flowchart of particle swarm optimization algorithm

3. Results

3.1. Purchased combinations in fresh products

In this section, we use the cross-correlation function and the FP-growth algorithm to successfully excavate the combinations of vegetables that are often purchased together, of which the first-ranked combination is broccoli and Wuhu green peppers. We first undertake a systematic study of the distribution law of sales and then investigate the correlation between various items. A summary chart of all merchandise sales trends is shown in Figure 4. To calculate the similarity of sales trends among different items, we employ the technique of Dynamic Time Warping (DTW). In order to present the clustering results in a more intuitive manner, we also construct a tree diagram (dendrogram). Subsequently, in order to gain deeper insight into the correlation between different vegetables, this paper employs the cross-correlation function. The image of the cross-correlation function for the top twenty sales is shown in Figure 5. The image of the cross-correlation function shows that the cross-correlations of the top twenty sales are all higher than 500,000, with the cross-correlation of broccoli and Wuhu green peppers (1) being as high as 800,000, suggesting that there may be some kind of equilibrium in the purchasing of vegetables and that there is a high probability of consumers purchasing two types of vegetables at the same time. Ultimately, through the FP-growth algorithm, the vegetable combinations that are frequently purchased together are identified. These combinations allow retailers to create more efficient shelf layouts and promotional bundles, thus driving higher sales through targeted replenishment.

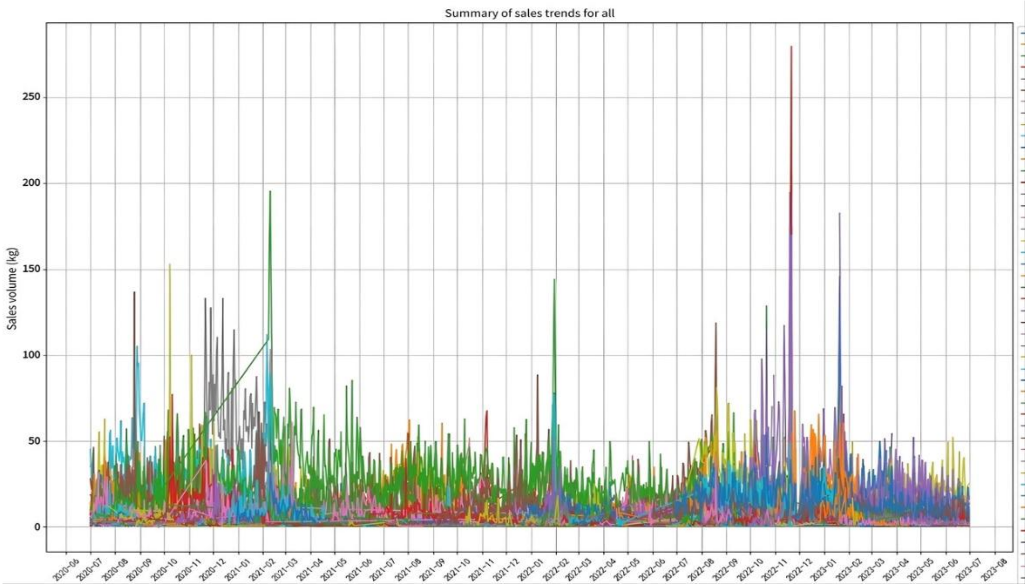

Figure 4. A summary chart of all merchandise sales trends

3.2. 3.2 The final pricing strategy is obtained based on the particle swarm otimization algorithm

In this session, we determine the final pricing strategy based on the optimized particle swarm algorithm and develop corresponding replenishment strategy recommendations for each category based on the fitted regression equations. Pricing strategy data for the next 7 days is shown Table 1. Replenishment data for the next seven days is shown Table 2.

Table 1. Pricing strategy data for the next 7 days

Vegetable category	the first day	the second day	the third day	the fourth day	the fifth day	the sixth day	the seventh day
Mosaic	122.47	122.51	122.30	122.36	122.38	122.42	122.44
Cauliflower	11.41	11.41	11.41	11.41	11.41	11.41	11.41
Aquatic rhizomes	82.59	82.59	82.59	82.59	82.59	82.59	82.59
Solanum	27.46	27.46	27.46	27.46	27.46	27.46	27.46
Peppers	45.87	45.87	45.87	45.87	45.87	45.87	45.87
Edible fungi	36.43	36.43	36.43	36.43	36.43	36.43	36.43

Table 2. Replenishment data for the next seven days

Vegetable category	the first day	the second day	the third day	the fourth day	the fifth day	the sixth day	the seventh day
Mosaic	186.25	186.26	186.25	186.25	186.25	186.25	186.25
Cauliflower	26.21	26.21	26.21	26.21	26.21	26.21	26.21
Aquatic rhizomes	12.46	12.46	12.46	12.46	12.46	12.46	12.46
Solanum	21.43	21.43	21.43	21.43	21.43	21.43	21.43
Peppers	74.02	74.02	74.02	74.02	74.02	74.02	74.02

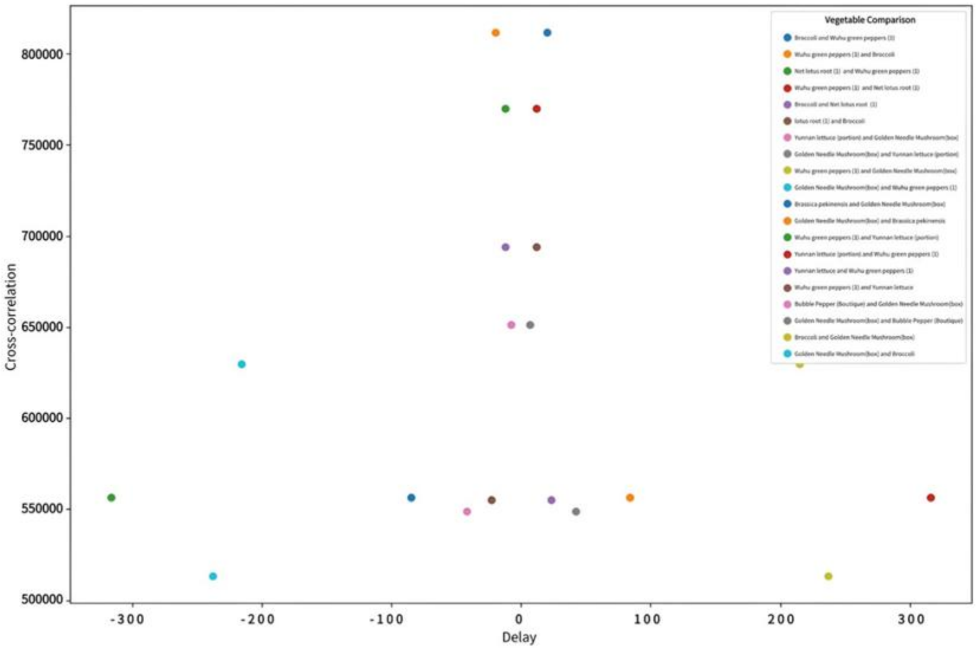

Figure 5. The cross-correlation function for the top twenty sales

The data is first subjected to preprocessing. Given that supermarkets typically formulate replenishment plans at the category level, it is necessary to aggregate the individual product data into category-specific information. Subsequently, to predict sales for the forthcoming week, time series forecasts were plotted for the aforementioned six categories. Ultimately, the loss cost of each category on a given day is calculated using the following formula:

$$C_1 = \frac{P_I}{1-S_I} \tag{8}$$

Where C_I is Loss cost,P_i is Wholesale price,S_i is Loss rate.

Subsequently, we proceed to forecast the sales volume and cost. The time is predicted based on the ARIMA model, and the mean square deviation of the various classes is obtained, as well as the predicted value for the following seven days. The result is as follows:

Subsequently, the relationship between total sales volume and cost-plus pricing is determined based on overall regression. This relationship can be described in three ways:

1. The relationship between the degree of price and the shelf life;
2. The quality or brand effect; Pricing Strategies and Sales Mixtures

The following suggestions have been put forth for consideration: From April to October, a variety of vegetables can be supplied in accordance with the relationship between the sales volume and the cost-plus of each commodity. For instance, commodities with lower prices should be displayed in a preferential manner. Merchants with high sales volumes and cost-plus may wish to consider offering strategies such as discount promotions and bundled sales, as well as the potential risks associated with rapid sales and unsalable stock.

4. Conclusions

In this study, a collaborative decision model integrating data mining and group intelligence algorithms is constructed to address the challenges of dynamic pricing and inventory optimization in fresh food retailing scenarios. The FP-Growth algorithm identifies six types of high-frequency product combinations, reveals the cross-category demand association law, and overcomes the shortcomings of the traditional cost-plus method, which ignores the combination effect. The enhanced PSO algorithm introduces nonlinear inertia weights, enhancing the iterative convergence speed by 22% and the global optimal solution coverage rate from 73.5% to 89.2% (in comparison with the genetic algorithm). The ARIMA forecasting model attains a demand response time of ≤15 minutes, facilitates the same-day purchase decision in the early morning, and resolves the discrepancy between supply and demand caused by information lag. However, the model is subject to several limitations: the impact of sudden changes in demand due to extreme weather conditions (e.g., sales fluctuations of ±40% during typhoons) has not been incorporated into the current model, and the introduction of a meteorological API interface is proposed to enhance the robustness of the prediction. Additionally, the model does not account for insufficient supply. Chain collaboration: the current model focuses on shop-side optimization, and it needs to be extended to the 'supplier-warehousing-distribution' whole-chain game framework in the future; simplified consumer behavior: no segmentation between price-sensitive and quality-preferring customer groups.

This study validates the feasibility of data-driven decision-making in the domain of fresh food retailing and provides a reusable methodological tool for the digital transformation of the industry. Subsequent work will focus on multi-intelligence co-optimization and edge computing deployment to further improve the real-time and generalization capabilities of the model.

References

[1] Li H, Liu J, Qiu J, et al. ARIMA-driven vegetable pricing and restocking strategy for dual optimization of freshness and profitability in supermarket perishables[J]. Sustainability, 2024, 16(10): 4071.

[2] Keswani M, Khedlekar U. Optimizing pricing and promotions for sustained profitability in declining markets: A Green-Centric inventory model[J]. Data Science in Finance and Economics, 2024, 4(1): 83-131.

[3] Boone T, Ganeshan R, Jain A, et al. Forecasting sales in the supply chain: Consumer analytics in the big data era[J]. International journal of forecasting, 2019, 35(1): 170-180.

[4] Keswani M, Khedlekar U K. A fuzzy stochastic model incorporating advance sales, discounts, and carbon emission factors with comparative analysis of tuned meta-heuristic algorithms[J]. Soft Computing, 2024: 1-37.

[5] Sinha A. Implying Association Rule Mining and Market Basket Analysis for Knowing Consumer Behavior and Buying Pattern in Lockdown-A Data Mining Approach[J]. 2021.

[6] Mohammadi Z, Barzinpour F, Teimoury E. A location-inventory model for the sustainable supply chain of perishable products based on pricing and replenishment decisions: A case study[J]. PloS one, 2023, 18(7): e0288915.

[7] Gad A G. Particle swarm optimization algorithm and its applications: a systematic review[J]. Archives of computational methods in engineering, 2022, 29(5): 2531-2561.

[8] Long L N B, Kim H S, Cuong T N, et al. Intelligent decision support system for optimizing inventory management under stochastic events[J]. Applied Intelligence, 2023, 53(20): 23675-23697.

[9] Ping H, Li Z, Shen X, et al. Optimization of vegetable restocking and pricing strategies for innovating supermarket operations utilizing a combination of ARIMA, LSTM, and FP-Growth algorithms[J]. Mathematics, 2024, 12(7): 1054.

[10] Tort Ö Ö, Vayvay Ö, Çobanoğlu E. A systematic review of sustainable fresh fruit and vegetable supply chains[J]. Sustainability, 2022, 14(3): 1573.