Study on the Key Factors Influencing Consumer Loyalty in Fresh Food E-commerce Platforms Based on SVM-RF Model

Yaoyi Huang [1, a], Jiaxun Chen [2, b], Bobin Zhang [1, c], Feiyu Jin [3, d],

Yanmin Lu [4, e], Changxiao Ma [5, f]

[1] Wuhan University of Science and Technology, Wuhan, Hubei, China

[2] Xiamen University, Xiamen, Fujian, China

[3] Zhengzhou University of Science and Technology, Zhengzhou, Henan, China

[4] Hebei Normal University of Science and Technology, Qinhuangdao, Hebei, China

[5] Anhui University, Hefei, Anhui, China

[a] 2311364473@qq.com, [b] 35820222202112@stu.xmu.edu.cn, [c] 3382311941@qq.com,

[d] 2432220407@qq.com, [e] 2410948603@qq.com, [f] 1364216704@qq.com

All six authors are Co-first authors of this article

Abstract. With the rapid development of internet technology, fresh food e-commerce platforms represented by companies such as Hema Fresh have quickly risen to prominence, becoming a new force in the modern retail sector. These platforms meet consumer demands for high-quality fresh products as well as convenient and efficient shopping experiences through their unique business models and service advantages. Therefore, in the context of the highly competitive market environment of fresh food e-commerce platforms, exploring how to enhance user loyalty to the brand of these platforms has become an urgent task. This study aims to investigate the key factors influencing consumer loyalty in fresh food retail e-commerce platforms by constructing an SVM-RF combined regression model, and to offer practical suggestions for improving platform management and market competitiveness.

Keywords: Fresh Food E-commerce Platform; Consumer Loyalty; Support Vector Machine; Random Forest; Weighted Fusion Model.

1. Introduction

With the rapid development of internet technology and the increasing demand for convenient shopping experiences, fresh food e-commerce supermarkets have gradually emerged as a prominent retail model in the market. Leveraging cold chain logistics, information management, and big data technologies, these platforms offer a convenient and efficient shopping experience [1].

The primary advantages of fresh food e-commerce lie in its robust supply chain management capabilities. Firstly, consumers can easily purchase affordable and diverse fresh products on these platforms [2]. Secondly, with the application of big data and artificial intelligence, e-commerce platforms can use data analysis to provide precise product recommendations, thereby enhancing the personalization and accuracy of consumer shopping [3]. Additionally, the integration of online platforms with offline stores through the "O2O" model enables consumers to enjoy convenient shopping while benefiting from more reliable delivery services [4].

Despite these advantages, fresh food e-commerce platforms still face significant challenges. Due to the perishable nature of fresh products, the timeliness and temperature control requirements for logistics and delivery are high. How to ensure product quality while reducing logistics costs and improving delivery efficiency remains a core issue for the development of fresh food e-commerce [5]. Furthermore, consumers have high expectations regarding the freshness and quality of products, and how platforms ensure product quality and enhance customer satisfaction is a key factor for their success [6].

In recent years, with the intensification of market competition, fresh food e-commerce platforms have gradually realized the importance of consumer loyalty for their long-term development. Studies have shown that consumer loyalty is closely related not only to factors such as product quality and logistics services but also to user experience and brand recognition [7]. To improve consumer loyalty, e-commerce platforms need to focus on optimizing personalized services, enhancing customer support quality, and continuously improving product delivery and preservation services [8]. For example, by utilizing social media and user review systems, platforms can strengthen their brand influence through word-of-mouth marketing and social marketing [9].

2. Construction of the SVM-RF-Based Model for Factors Influencing Consumer Loyalty

2.1 Indicator Construction

Table 1. Variable Explanation

Variable Name	**Variable Name**	**Variable Name**	**Variable Name**
Product Quality	Categorical	$Q = \begin{cases} 1, & \text{Low} \\ 2, & Medium \\ 3, & \text{High} \end{cases}$	Represents the quality level of the product
Product Price	Numeric	$P =$ Actual Payment Amount	The actual transaction price of the product
Delivery Speed	Numeric	$D = T_{\text{received}} - T_{\text{orderd}}$	The number of days from placing the order to receiving the product. T represents time
Advertising Intensity	Categorical	$A = \begin{cases} 1, & \text{Low} \\ 2, & Medium \\ 3, & \text{High} \end{cases}$	The level of advertising intensity
Social Attributes	Numeric	$S = \frac{\text{Number of Friends } + 2 \times \text{ Interaction Frequency}}{\text{Number of Interactions}}$	A weighted score based on the user's social network density and interaction frequency
Personal Decision	Numeric	$I = \frac{\sum_{i=1}^{n} \text{ Preference Score }_i}{n}$	The average preference score based on the user's past purchasing behavior
Repeat Purchase Rate	Numeric	$R = \frac{\text{Repeat Purchase Frequency of Specific Produ}}{\text{Total Purchase Frequency}}$	The proportion of repeat purchases of specific products by the user
Average Shopping Frequency	Numeric	$F = \frac{\text{Number of Purchases in a Specific Period}}{\text{Period Length}}$	The average shopping frequency of the user within a specified period
User Retention Rate	Numeric	$L = \frac{\text{Number of Active Users After a Specific Time}}{\text{Total Users at the Start}}$	The proportion of users who remain active on the platform after a certain period
User Engagement	Numeric	$E = \frac{\text{Number of User Reviews+Numer of Shares}}{\text{Number of User Interactions}}$	A weighted average indicator of user engagement, retention, repeat purchase rate, and average shopping frequency
User Stickiness	Numeric	$\text{K} = \frac{\text{E +L+R+F}}{4}$	

To gain deeper insights into the underlying driving factors of user loyalty, this study introduces a series of key independent variables: product quality, product price, delivery speed, advertising intensity, social attributes, and personal decision-making. These variables encompass multiple dimensions, such as product characteristics, marketing strategies, and user individuality, providing a comprehensive perspective for analyzing consumer loyalty. The detailed explanation of these variables is shown in **Table 1**.

2.2 Construction of the SVM-RF Combined Regression Model

In this study, a weighted combined model based on Support Vector Machine (SVM) and Random Forest (RF) is constructed to predict consumer loyalty on fresh food e-commerce platforms. This combined model enhances predictive accuracy and robustness by leveraging the strengths of both SVM and RF.

Specifically, SVM excels in capturing nonlinear relationships in data and is effective at handling complex patterns, while RF, by integrating multiple decision trees, is robust against noise and performs well in terms of generalization. The model separately trains the SVM and RF to obtain their respective predictions, and assigns corresponding weights based on the model's performance on the training set. Finally, the model combines the SVM and RF predictions through a weighted average method, integrating the advantages of both approaches and improving prediction accuracy.

Table 2. Algorithm Description of the Weighted Fusion Model

Algorithm	SVM-RF Weighted Fusion
1	**Input:** Training data $D = \{(x_i, y_i)\}_{i=1}^{n}$, Validation data D_{val}
2	**Output:** Final prediction y_{pred}
3	**Step 1:** Train base models
4	Train SVM model h_1 on D
5	Train RF model h_2 on D
6	**Step 2**: Evaluate model errors
7	Predict on D_{val} using h_1 , obtain $y_{val,1}$
8	Predict on D_{val} using h_2 , obtain $y_{val,2}$
9	Compute validation errors: $e_1 = \text{MAE}(y_{val}, y_{val,1}), e_2 = \text{MAE}(y_{val}, y_{val,2})$
10	**Step 3:** Determine fusion strategy
11	If (
12	Select model with smaller error: $y_{\text{pred}} = y_{val,1}$ if $e_1 < e_2$, else $y_{\text{pred}} = y_{val,2}$
13	Else: Perform weighted fusion
14	Compute weights: $w_1 = e_2/(e_1 + e_2), w_2 = e_1/(e_1 + e_2)$
15	Fuse predictions: $y_{\text{pred}} = w_1 \cdot y_{val,1} + w_2 \cdot y_{val,2}$
16	**Step 4:** Return final predictions
17	**Return:** y_{pred}

3. Analysis of Results

3.1 The Result of SVM-RF Model

Based on the results in **Table 3**, the MAE (Mean Absolute Error) values of the test set for different consumer groups are all below 9, indicating that the prediction results are relatively stable. Specifically, the MAE values for price-sensitive shoppers and loyal brand fans are relatively low, suggesting that the model's predictions are quite accurate. On the other hand, the MAE values for quality-seeking customers, convenience-first consumers, and socially influenced shoppers are slightly higher but still within an acceptable range. This demonstrates the model's stability and its reasonable generalization ability. Figures 1 through 5 show the test set results for each consumer type,

further validating the model's predictive performance across various consumer groups. Overall, the SVM-RF weighted fusion model performs well in predicting consumer loyalty, showcasing strong robustness.

Table 3. MAE of the Test Set for Five Consumer Types

Consumer Type	Train Set MAE	Test Set MAE
Price-Sensitive Customers	2.2	4.06
Quality-Seeking Customers	4.49	7.74
Convenience-First Customers	3.25	6.44
Loyal Brand Customers	1.26	4.86
Socially Influenced Consumers	3.89	8.34

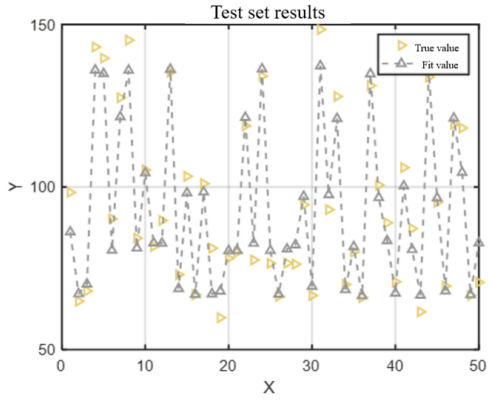

Figure 1. Price sensitive shopper test set

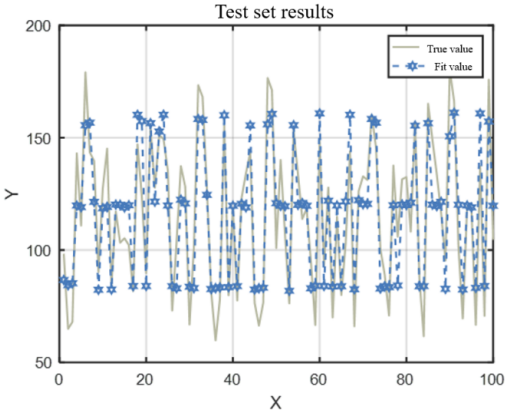

Figure 2. Quality Pursuit Customer Test Set

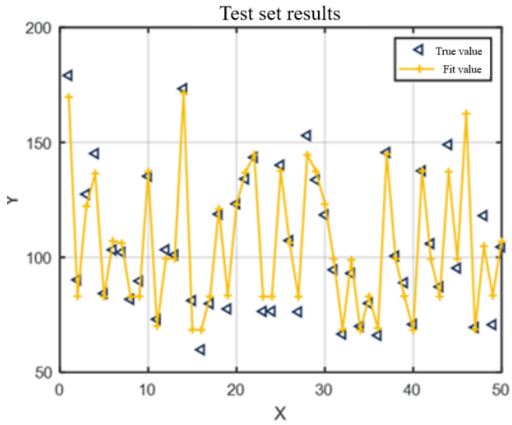

Figure 3. Convenient Priority Customer Test Set

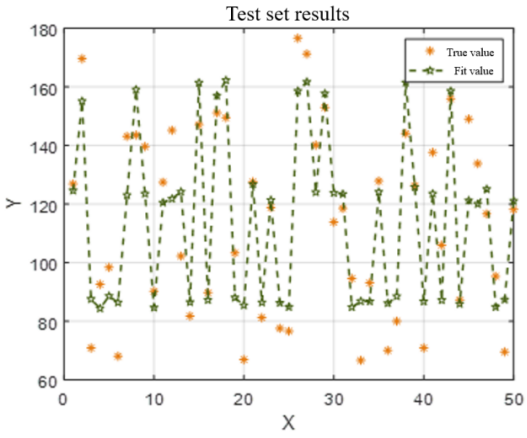

Figure 4. Test set of loyal brand Customer

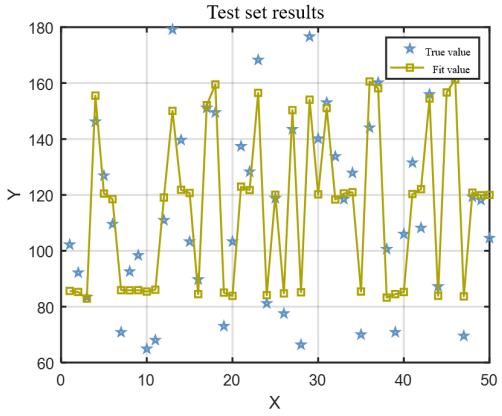

Figure 5. Social Impact Consumer Test Set

3.2 Identification and Analysis of Key Influencing Factors

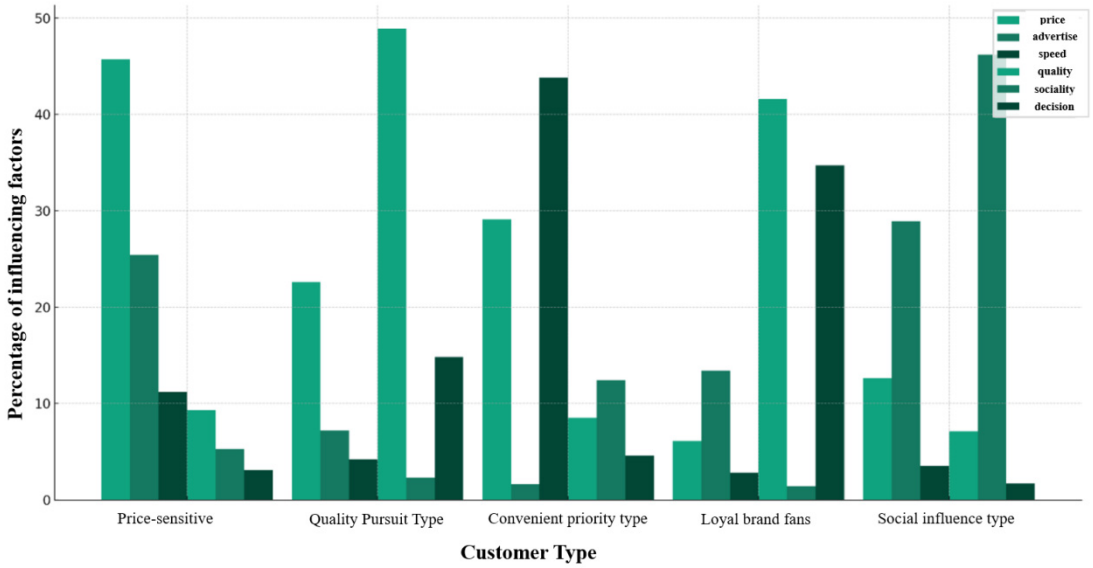

Figure 6. Proportion of influencing factors of different user types

In this study, the factors influencing consumer loyalty vary across different consumer groups. For price-sensitive shoppers, product price is the most significant factor, accounting for 45.7%, reflecting their emphasis on cost-effectiveness, especially during promotional events. Advertising also plays a notable role, contributing 25.4%, particularly when ads highlight price advantages. Although delivery speed, product quality, and social attributes have smaller effects, they still contribute to loyalty, suggesting that improving service quality and strengthening social media marketing can enhance loyalty.

Quality-seeking customers are primarily influenced by product quality, which accounts for 48.9%, indicating their strong focus on quality. Product price contributes 22.6%, suggesting that price still plays a role in their decision-making. Personal decision-making, with a share of 14.8%, shows that brand loyalty and personalized marketing have a significant impact on them.For convenience-first consumers, loyalty is mainly driven by delivery speed, accounting for 43.8%, followed by price at 29.1%. Social attributes also have an impact at 12.4%, indicating that the influence of social circles cannot be ignored. To cater to this group, e-commerce platforms need to optimize logistics systems, set reasonable prices, and enhance social interactions.

Loyal brand fans' loyalty is most strongly influenced by product quality (41.6%) and personal decision-making (34.7%), highlighting the alignment of brand belief and values. Advertising

contributes 13.4%, showing that brand marketing plays a role in attracting them. Maintaining a strong brand image and telling a compelling brand story is crucial for enhancing loyalty.

Socially influenced consumers are most impacted by social attributes, which account for 46.2%, with advertising contributing 28.9%. Product price has a smaller impact, contributing 12.6%, showing that social circles and social media play a significant role in their decision-making. E-commerce platforms can increase loyalty by enhancing social sharing, encouraging user reviews, and using targeted advertising.

4. Conclusion

In this paper, by constructing an SVM-RF weighted fusion model, we investigated the key factors influencing consumer loyalty on fresh food e-commerce platforms. The model analyzed various factors, including product quality, price, delivery speed, advertising intensity, social attributes, and personal decision-making, and identified the critical drivers of loyalty for different types of consumers. The SVM-RF weighted fusion model, by combining the strengths of Support Vector Machine (SVM) and Random Forest (RF), improved the prediction accuracy and model robustness. The test results across multiple consumer types demonstrate that the model performed well for all groups, confirming the reliability of the SVM-RF fusion model in predicting consumer loyalty.The findings of this study provide valuable insights for fresh food e-commerce platforms to optimize user stickiness and enhance market competitiveness.

References

[1] Cheng, Q., Sun, C., Zhang, X., et al. (2004). Short-Term load forecasting model and method for power system based on complementation of neural network and fuzzy logic. Transactions of China Electrotechnical Society, 19(10), 53-58.

[2] Fangfang. (2011). Research on power load forecasting based on Improved BP neural network. (Master's thesis). Harbin Institute of Technology.

[3] Amjady, N. (2001). Short-term hourly load forecasting using time series modeling with peak load estimation capability. IEEE Transactions on Power Systems, 16(4), 798-805.

[4] Ma, K. (2014). Short term distributed load forecasting method based on big data. (Master's thesis). Changsha: Hunan University.

[5] Shi, B., Li, Y., Yu, X., & Yan, W. (2010). Short-term load forecasting based on modified particle swarm optimizer and fuzzy neural network model. Systems Engineering-Theory and Practice, 30(1), 158-160.

[6] Zhang, L., & Chen, H. (2022). Research on logistics efficiency of fresh e-commerce: Challenges and solutions. Journal of E-Commerce Research, 18(3), 104-112.

[7] Liu, M., & Zhao, Y. (2021). Exploring consumer loyalty in online fresh food retailing. E-commerce and Digital Marketing, 12(5), 51-59.

[8] Zhou, W., & Wang, F. (2023). Enhancing customer satisfaction and loyalty in online fresh food supermarkets. Journal of Retail and Consumer Services, 30(8), 192-200.

[9] Li, X., & Wang, H. (2020). Social media influence on consumer behavior in online fresh food shopping. Social Media Studies, 9(4), 67-74.