Frontiers in Business, Economics and Management
ISSN: 2766-824X | Vol. 10, No. 3, 2023

Research on Product Selection Model for Small E-Commerce Stores Based on Improved RFM Model

Faqun Cai*

College of Economics and Management, Nanjing Polytechnic Institute, Nanjing 210048, P. R. China

* Corresponding author. E-mail address: caifaqun@njpi.edu.cn

Abstract: Traditionally, merchants determine whether to enter or exit the market based on the product life cycle. However in the era of e-commerce, the needs of customers are constantly changing, and the life cycle of products is very short. If a merchants accidentally enters the market during the recession period, it may lead to business failure. Therefore, stores must evaluate the value of their products in real time, grasp the trend of product value, and adjust their business direction in a timely manner to ensure the sustainability of their operations. The RFM model is a customer value evaluation model, which includes three important indicators: purchase proximity R, purchase frequency F, and purchase amount M. Products also have indicators such as proximity of purchase, frequency of purchase, and amount of purchase. Therefore, in this article, we propose a products value evaluation model based on the RFM model. In the model, we use the entropy weight method to assign weights to various indicators. Through empirical calculation, analysis, and testing of the business data of sample stores, we confirm that the model is effective in Product Selection for small e-commerce stores.

Keywords: Product selection; Product Value; RFM model; Standardization; Entropy Weight Method.

1. Introduction

In traditional marketing, the Product Life Cycle Theory defines the product life cycle as four periods: introduction, growth, maturity, and decline(Dean, 1950; Levirt, 1965), most merchants choose to intervene in the market during their growth or maturity period. Managers often determine which products should be retained, replaced, and eliminated by analyzing their sales revenue and profitability. Therefore, the main methods for predicting product sales include time series analysis, causal modeling (Stadtler, Kilger & Meyr, 2015), and monitoring product market share over time (Bendle, Farris, Pfeifer & Reibstein, 2016). Time series prediction methods also include exponential smoothing method and seasonal variation method. These methods are based on industry market data to make overall judgments. However, for small e-commerce merchants, it is almost impossible to obtain these data in a timely manner, let alone use the analysis results of these data to quickly make business adjustments.

In 1959, Duncan Luce proposed the discrete choice model MNL (Multinomial Logit Model) to predict customer choices (Duncan,1959). Assuming that there are J types of products that customers can choose from; The corresponding utility of each types of product is U, which is composed of the sum of fixed and random parts; The fixed utility V can be explained by a certain observable factor X, while the random part represents the influence of unobserved utility and error. In 1994, Hughes, a database research institute in the United States, first proposed the RFM model, which consists of three indicators: purchase progress (R), purchase frequency (F), and purchase limit (M) (Hughes,1994). This model categorizes customer value based on the three indicator values of customers in the enterprise's own business data. In 2008, Cheng C.H et al designed a user classification model that quantifies user RFM data, then uses k-means clustering algorithm to classify user RFM indicator data, and analyzes the characteristics of each type (Cheng C.H et al., 2008). Over the last few decades, firms have become more customer-centric, adding a customer perspective to the analysis of expected revenues, which had been previously predicted solely from expected product sales. Although this new perspective is very relevant, the previous perspective of product-orientation should not be forgotten (R. Heldt, et al.2019; Kumar & Reinartz, 2016), a successful firm has to create or co-create (Vargo & Lusch, 2004) perceived value for customers through the development of products and brands. How to evaluate and predict the value of products? Similar to the indicators in the customer value evaluation system, there are also indicators such as the purchase amount, recent purchase time, and purchase frequency during the sales process of products. Therefore, we can construct an RFM model for evaluating the value of products

2. Models and Methods

When operating on the internet, the business data of stores is retained on the network platform, allowing merchants to easily access this data and make reasonable use of it to achieve market prediction. In the data, the amount of products purchased, the latest time of purchase, and the frequency of purchase provide the basis for constructing a product value evaluation model based on RFM model.

2.1. Indicator Selection

Similar to customer data, product data also includes three indicators: R, F, and M, including purchase proximity (R), purchase frequency (F), and purchase amount (M). R represents the interval between the last purchase time and the statistical deadline. Therefore, the smaller the better, this indicator can be used to determine the recent activity of a product. F represents the frequency of product purchases， it refers to the number of times a product has been purchased within a certain period of time. Generally, The higher the frequency of purchase, the higher the popularity of the product, and the greater its potential value. M represents the amount purchased, It refers to the total amount of products purchased within a certain period of time. The higher the amount purchased, the greater the value that the product can create for the store. Indicators R and F can be used to infer the popularity of product types. The value of indicator M can be used to measure the contribution rate of products to merchants.

2.2. Weighting methods for indicators

In traditional RFM model, the weights of indicators are the same. However, in reality, the impact of R, F, and M on the product value is different among different industries or different products within the same industry. The repeat purchase rate of consumables is generally high, so the frequency of purchase can better reflect the sustainability of product sales. The purchasing frequency of durable product is often low, so the purchase amount M may be more important. Therefore, the R, F, and M indicators of different products should be assigned different weights.

The indicator weight assignment method used in this article is the entropy weight method, which is an objective weighting method. This method calculates the weights of each indicator based on the degree of dispersion of data, which is the information entropy of each indicator, and then makes certain modifications to obtain more objective indicator weights.

The specific steps of the entropy weight method are as follows:

(1) Data standardization

Due to the non-uniform measurement units of various indicators, it is necessary to standardize them before calculating their comprehensive weights, that is, convert the absolute values of the indicators into relative values to eliminate the influence of measurement units. Different standardized processing algorithms are required for positive, negative, and moderate indicators. The standardized calculation formulas for positive indicators, negative indicators, and moderate indicators are formulas (1), (2), and (3), respectively.

$$y_{ij} = 0.998 \cdot \frac{x_{ij} - \min\{x_{i1}, \cdots x_{in}\}}{\max\{x_{i1}, \cdots x_{in}\} - \min\{x_{i1}, \cdots x_{in}\}} + 0.002 \quad (1)$$

$$y_{ij} = 0.998 \cdot \frac{\max\{x_{i1}, \cdots x_{in}\} - x_{ij}}{\max\{x_{i1}, \cdots x_{in}\} - \min\{x_{i1}, \cdots x_{in}\}} + 0.002 \quad (2)$$

$$y_{ij} = \begin{cases} \dfrac{x_{ij} - \min\{x_{i1}, \cdots x_{in}\}}{\max\{x_{i1}, \cdots x_{in}\} - \min\{x_{i1}, \cdots x_{in}\}} & x_{ij} < A \\ \dfrac{x_{ij} - \min\{x_{i1}, \cdots x_{in}\}}{\max\{x_{i1}, \cdots x_{in}\} - \min\{x_{i1}, \cdots x_{in}\}} & x_{ij} \geq A \end{cases} \quad (3)$$

(2) Calculation of variance value of indicators

Calculate the proportion of each object's indicator value in each indicator according to formula (4), that is, the variation indicator value, and construct a probability matrix.

The variation value of the indicator can represent the degree of dispersion of the indicator to a certain extent, preparing for the calculation of information entropy value.

$$p_{ij} = \frac{y_{ij}}{\sum_{j=1}^{n} y_{ij}}, (i = 1, 2, \cdots, m) \quad (4)$$

(3) Calculate information entropy value

The formula for calculating the information entropy value e_i of the i-th indicator is (5).

$$e_i = -k \cdot \sum_{j=1}^{n} p_{ij} \cdot \ln(p_{ij}), k = 1/\ln(n), (i = 1, 2, \cdots, m) \quad (5)$$

(4) Calculate information entropy redundancy value

Calculate the information entropy redundancy value g_i using formula (6). It's 1 minus the information entropy value.

$$g_i = 1 - e_i, (i = 1, 2, \cdots, m) \quad (6)$$

(5) Calculate the weight of each indicator

When the information entropy redundancy value of each indicator is calculated, the weight of each indicator ω_i can be calculated according to formula (7).

$$\omega_i = \frac{g_i}{\sum_{i=1}^{m} g_i} \quad (7)$$

3. Data Description and Analysis

This article selects the data of an online store operating consumable small goods as the sample data. Using the transaction information of the store from January 3, 2022 to August 10, 2022 as the data source, there are nearly 6000 pieces of data in total. According to research needs, we divided the data into two periods, with a cut-off time of 24:00 on May 31, 2022.

. In the phase I, there are over 3500 pieces of data used to calculate comprehensive product ratings, predict future sales of products, and for model calibration. In the phase II, there are over 2500 pieces of data in total for model validation.

3.1. Data Introduction

The sample store operates over 50 types of products, with nearly 6000 recorded sales data. There are three main types of product purchase data. The first type is the transaction amount data of the product, including the sales amount, postage, refunds, and other information of the product. The second type of data is used to describe the transaction time of products, including order creation time, payment time, confirmation of receipt time, etc. The third type is transaction frequency data, including order numbers, logistics tracking numbers, payment tracking numbers, etc.

3.2. Data processing

Before data analysis and model establishment, we must clean the data, including missing data, incorrect data, abnormal data, duplicate data, and inconsistent data. Firstly, for incomplete or missing data, we use the Ismissing function in Matlab software to search, and fill it in with the Fillmissing function. Then, we detect and eliminate erroneous data, abnormal data, and duplicate records. Finally, 5559 useful data were obtained, including 3221 data from the phase I and 2338 data from the phase II.

There are 52 types of products sold in this sample store, of which 9 types do not have any sales records. For the convenience of quantitative analysis, only 43 types of products with actual sales are examined.

3.3. Data Descriptive Analysis

The sales amount is an important indicator of the best-selling level of a product and one of the important factors considered by stores when selecting products. By summarizing and sorting 3221 pieces of data from the merchant from January 3, 2022 to May 31, 2022, we can clearly see that products C10, C21, C45, C08, and C35 occupy the top 5 in sales, as shown in Figure 1.

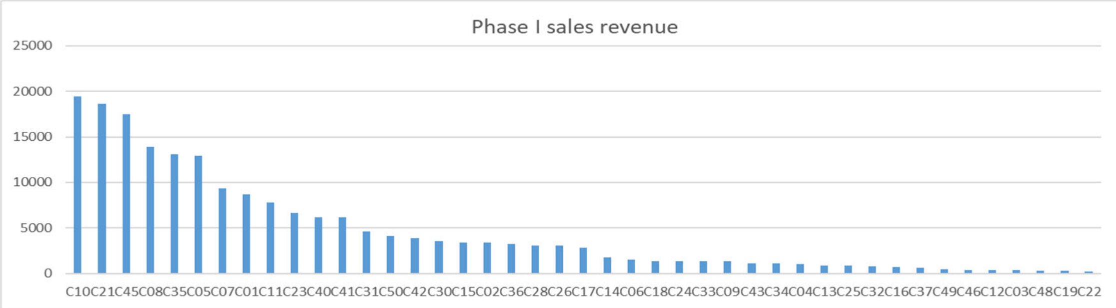

Figure 1. Ranking Chart of Product Sales from January 1, 2022 to May 31, 2022

If we speculate the sales trend based on the current sales volume, these five products should increase their purchase quantity. They should perform well in the later stage and can bring more profits to the enterprise. Meanwhile, the sales of 10 products, including C22, C19, C48, C03, C12, C46, C49, C37, C16, and C32, are all very low and should be eliminated.

However, this is not the case. As shown in Figure 2, in the phase II, the sales of product C10, with its highest sales in the phase I, decreased significantly. C25, C16, and C14, which performed well in the phase I, had zero sales in the phase II. C04, C06, C09, C17, C24, C26, C34, and C43, which were originally good in sales, performed poorly in the phase II. while C22 and C49, which should have been eliminated, have a significant increase in sales. Therefore, inferring the sales trend in the later stage based on the previous sales volume may lead to significant errors.

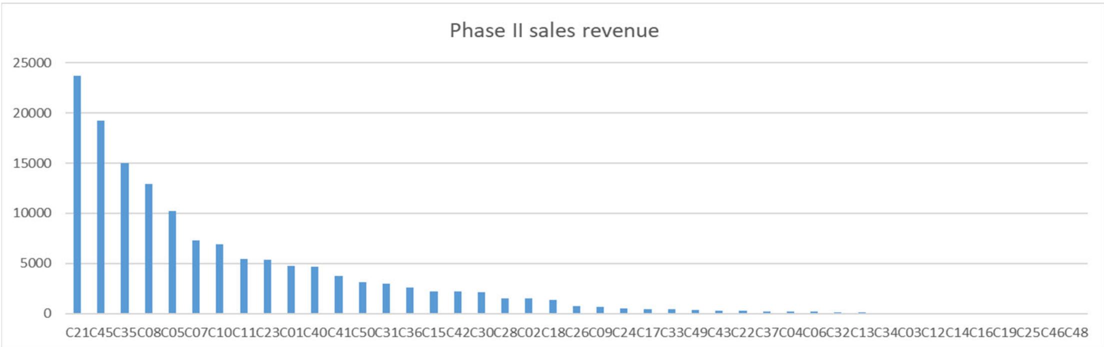

Figure 2. Ranking Chart of Product Sales from June 1, 2022 to August 10, 2022

For the convenience of observation and comparison, we have standardized the sales of the previous and subsequent periods. The higher the sales volume, the higher the value of the product. Therefore, the Positive standardization formula (2) can be used for processing. All standardized values are between 0 and 1. When we put the standardized values of the two periods in the same graph, it can be seen that there are significant differences between them, as shown in Figure 3. What makes them different? How can we accurately infer the later performance of a product?

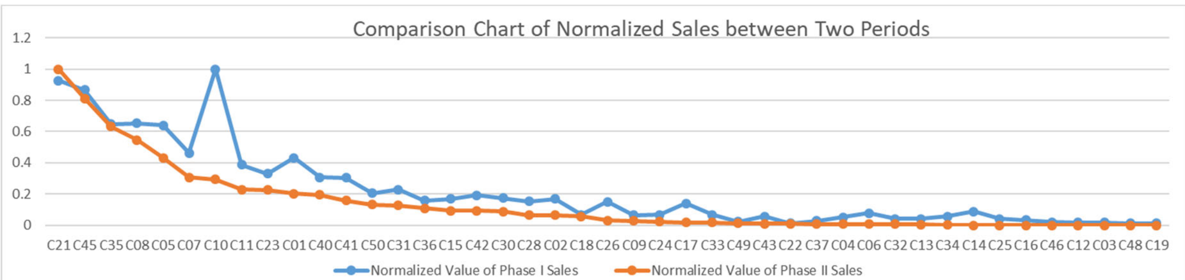

Figure 3. Comparison Chart of Standardized Values of Sales volume in Two Periods

4. Model Construction and Calibration

4.1. Statistical caliber of indicators

There are three time values in the data, which are the order placement time, payment time, and confirmed receipt time. In general, the transaction ends only after the customer confirms receipt, so the data deadline used in this model is based on the customer's confirmed receipt time.

The statistical caliber of the R indicator: the time interval between the latest purchase and confirmed receipt of various types of products and the statistical deadline. The statistical deadline specified here is 24:00 on May 31, 2022;

Statistical caliber of indicator F: the total number of purchases of various types of products between January 3, 2022 and May 31, 2022;

Statistical caliber of M indicator: the sales volume of various types of products from January 3, 2022 to May 31, 2022;

For the convenience of expressing all numerical data to a maximum of 4 decimal places.

According to the calculation method of indicator values in the RFM model, we classified and summarized the data of the phase I, and calculated the R, F, and M values of 43 types of products, which have sales records. The results are shown in Table 1.

Table 1. Values of indicators R, F, M

Product types	R	F	M
C01	5.3561	42	8704.76
C02	3.1381	8	3393.54
C03	32.5275	1	345.31
C04	17.3643	8	1036.65
C05	0.2280	351	12887
……	……	……	……
C45	0.2361	390	13811.37
C46	24.2081	2	396.56
C48	86.9719	1	262.7
C49	3.4027	1	444.44
C50	1.1612	39	4163.78

The indicators R, F and M are not conducive to observation and measurement due to differences in measurement units. In order to eliminate the impact of indicator units on the evaluation of products value, we first standardize the data. The smaller the R value, the higher the value of the product, so R is a negative indicator. Therefore, the R indicator is treated with a negative standardization formula, as shown in formula (8); The larger the value of the F and M dimension indicators, the higher the value of the product, so they are all positive indicators. Therefore, the F and M indicators should be processed using a positive standardization formula, as shown in formula (9) and (10).

$$\boldsymbol{R}_i^{'} = 0.998 \cdot \frac{\boldsymbol{R}_{\max} - \boldsymbol{R}_i}{\boldsymbol{R}_{\max} - \boldsymbol{R}_{\min}} + 0.002 \tag{8}$$

$$\boldsymbol{F}_i^{'} = 0.998 \cdot \frac{\boldsymbol{F}_i - \boldsymbol{F}_{\min}}{\boldsymbol{F}_{\max} - \boldsymbol{F}_{\min}} + 0.002 \tag{9}$$

$$\boldsymbol{M}_i^{'} = 0.998 \cdot \frac{\boldsymbol{M}_i - \boldsymbol{M}_{\min}}{\boldsymbol{M}_{\max} - \boldsymbol{M}_{\min}} + 0.002 \tag{10}$$

$\boldsymbol{R}_i^{'}, \boldsymbol{F}_i^{'}, \boldsymbol{M}_i^{'}$ respectively represent the standardized values of each dimension indicator for the i-th type of product, and $\boldsymbol{R}_i, \boldsymbol{F}_i, \boldsymbol{M}_i$ represents the actual value of the product sales data correspondingly.

$\boldsymbol{R}_{\max}$, $\boldsymbol{R}_{\min}$, $\boldsymbol{F}_{\max}$, $\boldsymbol{F}_{\min}$, $\boldsymbol{M}_{\max}$, $\boldsymbol{M}_{\min}$ represent the maximum and minimum values of the product's sales proximity (indicator R), frequency (indicator F), and quota (indicator M), respectively.

After standardizing the sales data of products, standardized data that is not affected by indicator units is ultimately obtained, as shown in Table 2. These data eliminate the influence of the original sample data units on the evaluation of product value.

Table 2. Standardized values of indicators R, F, M

Product types	R	F	M	R'	F'	M'
C01	5.3561	42	8704.76	0.9553	0.0679	0.3619
C02	3.1381	8	3393.54	0.9738	0.0132	0.1362
C03	32.5275	1	345.31	0.7284	0.0020	0.0066
C04	17.3643	8	1036.65	0.8550	0.0132	0.0360
C05	0.2280	351	12887	0.9981	0.5645	0.5396
……	……	……	……	……	……	……
C45	0.2361	390	13811.37	0.9980	0.6272	0.5789
C46	24.2081	2	396.56	0.7979	0.0036	0.0088
C48	86.9719	1	262.7	0.2738	0.0020	0.0031
C49	3.4027	1	444.44	0.9716	0.0020	0.0108
C50	1.1612	39	4163.78	0.9903	0.0631	0.1689

4.2. Calculate the weight of indicators

Firstly, for internet stores, maintaining a sustained growth in order volume is a necessary condition for their survival. Unlike traditional markets, customers' orders on the internet are often in small batches, with multiple batches being processed. Therefore, the long tail theory is more effective than the 80-20 Rule, especially for consumable small goods. Therefore, for online stores, although the sales amount is important, it may not necessarily be the most important, and the purchase frequency may be more important. Secondly, on online platforms, customers have a higher degree of autonomy in choosing. If a product does not have any orders for a long period of time, it is likely that the product is not popular with customers. stores should attach great importance to it and consider replacing or eliminating it, which once again demonstrates the importance of purchasing frequency.

There are significant differences in the impact of proximity R, frequency F, and quota M on the value of different types of products. Therefore, we cannot use the traditional RMF model's weighting method to weight each indicator. The research object of this article is stores operating consumables on the internet, therefore, the entropy weight method is used to assign weights to each indicator. The entropy weight method algorithm can consider the interrelationships between various indicators and comprehensively consider the role of multiple indicators; It can adapt well to the characteristics and needs of the evaluated object, and can improve the overall objectivity and credibility of the evaluation results; In situations where information is uncertain or incomplete, the entropy weight method algorithm can effectively avoid the influence of subjective factors.

According to the steps of the entropy weight method, after standardizing the values of each indicator, we should calculate the corresponding variability values p_R , p_F , p_M for each indicator. The calculation method is shown in formula (11), and the results are shown in Table 3.

$$p_{ij} = \frac{i_j^{'}}{\sum_{j=1}^{43} i_j^{'}} \;,\; i = R, F, M \tag{11}$$

Table 3. The variation values of indicators R, F, M

Product types	R	F	M	R'	F'	M'	p_R	p_F	p_M
C01	5.3561	42	8704.76	0.9553	0.0679	0.4259	0.0249	0.0131	0.0459
C02	3.1381	8	3393.54	0.9738	0.0132	0.1600	0.0254	0.0026	0.0173
C03	32.5275	1	345.31	0.7284	0.0020	0.0074	0.0190	0.0004	0.0008
C04	17.3643	8	1036.65	0.8550	0.0132	0.0420	0.0223	0.0026	0.0046
C05	0.2280	351	12887	0.9981	0.5645	0.5396	0.0260	0.1087	0.0685
……	……	……	……	……	……	……	……	……	……
C45	0.2361	390	13811.37	0.9980	0.6272	0.5789	0.0260	0.1208	0.0735
C46	24.2081	2	396.56	0.7979	0.0036	0.0088	0.0208	0.0007	0.0011
C48	86.9719	1	262.7	0.2738	0.0020	0.0031	0.0071	0.0004	0.0004
C49	3.4027	1	444.44	0.9716	0.0020	0.0108	0.0253	0.0004	0.0014
C50	1.1612	39	4163.78	0.9903	0.0631	0.1689	0.0258	0.0121	0.0214

Then we use formula (12) to calculate the information entropy values of each indicator, and the information entropy values e_R , e_F , e_M of indicators R, F, M are: 0.9857, 0.7189, 0.8311

$$e_i = -k \cdot \sum_{j=1}^{43} p_{ij} \cdot \ln(p_{ij}) \;\;, k = 1/\ln(43) \;\;,\; i = R, F, M \tag{12}$$

The information entropy value represents the degree of dispersion of indicator values, which is the degree of uncertainty or confusion. The larger the information entropy value, the greater the degree of chaos in the system, indicating that the role of the indicator in the evaluation system is smaller, and the corresponding indicator weight is smaller; On the contrary, the smaller the entropy value, the greater the role and weight of the indicator in the evaluation system.

Therefore, we can infer the impact degree of each indicator based on its information entropy value previously. The information entropy value of F is the smallest, so its role in the evaluation system should be the greatest. The information entropy value of R is the largest, so its role in the evaluation system should be the smallest.

Next, we need to calculate the information entropy redundancy of each indicator based on the information entropy values. The calculation method is shown in formula (13), and the information quotient redundancy of each indicator, g_R , g_F , g_M is calculated as follows: 0.0143,0.2811,0.1689. Contrary to the information entropy value, the smaller the information entropy redundancy of an indicator, the greater its role in the indicator system, and vice versa.

$$g_i = 1 - e_i \;\;, i = R, F, M \tag{13}$$

Finally, we calculate the weights of each indicator according to formula (14). The weights of indicators R, F, and M are represented by ω_R , ω_F and ω_M respectively, and the calculated results are: 0.0308, 0.6054, 0.3638

$$\omega_i = \frac{e_i}{\sum_i e_i} \;, i = R, F, M \tag{14}$$

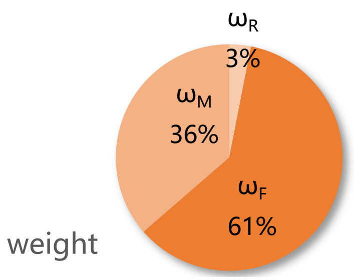

Figure 4. Schematic diagram of indicator weights

As shown in Figure 4, F has the highest weight, nearly 61%, and has the greatest impact on the evaluation of commodity value. M is second in weight, and R has the smallest weight, which also has the smallest impact on the evaluation of commodity value. This result is quite consistent with the characteristics of consumables.

4.3. Calculation of Comprehensive Scores

With both the weights and standardized values of the indicators known, we can calculate the comprehensive scores of various types of products based on the comprehensive score calculation formula (15). The results are shown in Tables 4.

$$\Omega_j = \sum_i \omega_i \cdot i'_j \quad \boldsymbol{i = R, F, M} \tag{15}$$

Table 4. Comprehensive scores of various types of products

Product types	C01	C02	C03	C04	……	C46	C48	C49	C50
Comprehensive score	0.2022	0.0876	0.0261	0.0475	……	0.0300	0.0108	0.0351	0.1301

Rank the comprehensive scores of each product, as shown in Figure 5. Product C21 has the highest score of 0.9217, followed by C35, C45, C05, and C08. Although C10 had the highest sales volume in the phase I, its comprehensive score was only 0.5272, far lower than C21, which was ranked behind C10 in the phase I. Through data observation, we found that although the sales volume of C10 are highest, but the frequency of purchase is not high, resulting in a large number of orders with large amounts .

After communicating with the merchant, we found that there was a promotion for C10 in the early stage, and the price was much lower than the market price. This price attracts some retailer to come for wholesale. This is an occasional event, so the comprehensive score of C10 is not high, which is in line with the actual situation and also proves the effectiveness of our evaluation model.

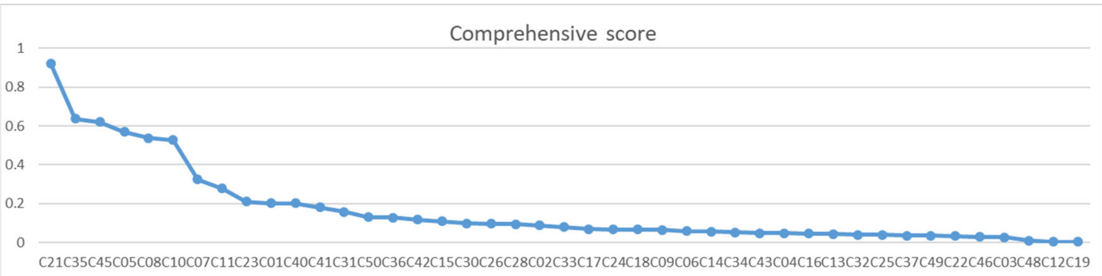

Figure 5. Comprehensive score ranking chart for various types of products

5. Model Validation

Can the comprehensive scores of the value of various types of products predict the future value direction of the products? To verify the effectiveness of the prediction model, we can compare the predicted results with the actual results.

For the sake of comparison, we normalize the sales volume of various products in Phase II based on the forward standardization formula (2), first, so that the processed data values are all between 0 and 1.

By placing the comprehensive scores of various types of products in the same image as the standardized values of sales volume in the phase II, we can see that there is a high degree of consistency between the comprehensive scores of the phase I and their future sales, as shown in Figure 6.

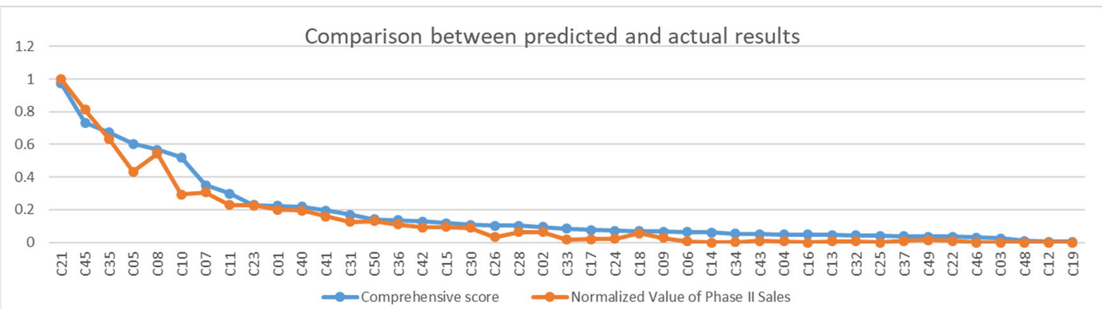

Figure 6. Comparison between predicted results and actual results

We conducted Paired-Samples T Test on the comprehensive scores of the phase I (Abbreviation:CS1)and the standardized values of sales volume in the Phase II (Abbreviation: SV2), and the results are shown in Tables 5, 6, and 7.

Table 5. Paired Samples Statistics

		Mean	N	Std. Deviation	Std. Error Mean
Pair 1	CS1	.1673	43	.20941	.03193
	SV2	.1408	43	.22703	.03462

Table 5 is a paired sample statistical table. From the table, it can be seen that the average of comprehensive scores of various types of products is 0.1673, and the average of standardized value of the actual sales volume of the phase II is 0.1408. The predicted result is slightly higher than the actual result.

Table 6. Paired Samples Correlations

		N	Correlation	Sig.
Pair 1	SV2 & CS1	43	.971	.000

To further compare whether the predicted results and actual results have reached a statistical difference level, paired sample t-tests are also required. According to Table 6, the correlation between the comprehensive score and actual sales performance is 0.971, with a probability P-value of 0.000, which is less than the significance level of 0.05. Therefore, they have a significant correlation and are suitable for paired sample T-test.

Table 7. Paired Samples Test

		Paired Differences					t	df	Sig. (2-tailed)
		Mean	Std. Deviation	Std. Error Mean	95% Confidence Interval of the Difference				
					Lower	Upper			
Pair 1	SV2 - CS1	-.02651	.05530	.00843	-.04353	-.00949	-3.144	42	.003

From Table 7, it can be seen that the paired sample t-test between the comprehensive score and the actual sales volume of the product has a T-value of -3.144, and a probability P-value of 0.003, which is less than 0.05, at the significance level. Therefore, there is no significant difference between the comprehensive score and the actual sales volume of the product.

6. Conclusion

In this article, we apply the RFM model to product value evaluation, which is commonly used for customer value evaluation. Additionally, we have improved the RFM model based on the characteristics of the products in the sample store.The traditional RFM model often assume that the weights of each indicator are the same, leading to bias in the evaluation results. Therefore, we use the entropy weight method to assign weights to each indicator. Through testing the prediction results, the improved RFM model has higher accuracy in predicting the value of products. The application of this model can help Small merchants use the internal data of the store to predict market trends, thereby get rid of dependence on external data.

The sample data of one certain online store for several months has limitations, which may lead to errors in model validation results. Therefore, we need to collect more data from more stores for a longer period of time to ensure the accuracy and universality of the conclusions in future research,.

Finally, researchers have discovered the impact of online comments. This includes both the impact of comments on user decisions and Including its impact on corporate sales (Zhao Meng & Qi Jiayin, 2014). Customer perceived value is directly reflected in the satisfaction evaluation of consumers towards the enterprise(Wu Jun, 2022)，therefore, customer evaluation should be included in the indicator system. Due to the low proportion of customers participating in the evaluation of the sample stores, customer evaluation indicators were not included in the evaluation system of this article. To improve the accuracy and comprehensiveness of the evaluation model, customer evaluation indicators should be added in later research.

The development of the modern retail industry, whether online or offline, cannot ignore the user experience. Only by grasping the real needs of users can enterprises achieve maximum profits (Evgeni & Gergana, 2017).

Acknowledgment

This work is supported by the Humanities and Social Sciences Research Project of Nanjing Polytechnic Institute (Grant No: NJPI-2021-ZD-06).

References

[1] Ching-Hsue Cheng, You-Shyang Chen. Classifying the segmentation of customer valuevia RFM model and RS theory[J].Expert Systems With Applications,2008,36(3).

[2] Dean J. Pricing policies for new products[J]. Harvard Business Review, 1950, 28(6) : 45 - 53.

[3] Bendle, N. T., Farris, P. W., Pfeifer, P. E., & Reibstein, D. J. (2016). Marketing metrics: The manager's guide to measuring marketing performance. Upper Saddle River: Pearson.

[4] Evgeni Genchev, Gergana Todorova. Sales Promotion Activities – Effective Tool of Marketing Communication Mix[J]. Social Science Electronic Publishing, 2017:181-185.

[5] Heldt Rodrigo, Silveira Cleo Schmitt, Luce Fernando Bins. Predicting customer value per product: From RFM to RFM/P[J]. Journal of Business Research,2019 (prepublish).

[6] Hughes Arthur M. Strategic Database Marketing[M]. Chicago: Probus publishing,1994.

[7] Kumar, V., & Reinartz, W. (2016). Creating enduring customer value. Journal of Marketing, 80, 36–68. https://doi.org/10.1509/jm.15.0414.

[8] Levirt T. Exploit the product life cycle [J]. Harvard Business Review, 1965, 43(6) : 81 - 94.

[9] Luce, R. Duncan Individual choice behavior: A theoretical analysis. John Wiley & Sons, Inc., New York; Chapman & Hall, Ltd., London 1959.

[10] Stadtler, H., Kilger, C., & Meyr, H. (2015). Supply chain management and advanced planning: Concepts, models, software, and case studies. New York: Springer.

[11] Vargo, S. L., & Lusch, R. F. (2004). Evolving to a new dominant logic for marketing. Journal of Marketing, 68, 1–17. https://doi.org/10.1509/jmkg.68.1.1.24036.

[12] Wu Jun. Research on Customer Value Segmentation and Customer Relationship Management Enhancement Strategies Based on Improved RFM model [D]. Northeastern University of Finance and Economics, 2023. DOI: 10.27006/d. cnki. gdbcu.2022.000650.

[13] Zhao Meng, Qi Jiayin. Research on Customer Lifetime Value Based on Purchasing Behavior RFM and Comment Behavior RFMP Models [J]. Statistics and Information Forum, 2014,29 (09): 91-98.