ISSN: 2688-9323

# Customer Loyalty Evaluation of Internet Catering based on Customer Segmentation

Jiaxue Wang[1], Chengying Niu[2, *]

[1] School of Statistics and Applied Mathematics, Anhui University of Finance and Economics, Bengbu 233030, China

[2] School of Statistics, Lanzhou University of Finance and Economics, Lanzhou 730101, China

## Abstract

**Under the background of the Internet economy, Internet catering has become the main mode of catering, and accurately mining the catering needs of different customers, conducting customer loyalty assessment, and maintaining the competitiveness of enterprises are urgent problems for Internet catering businesses to solve. This paper collects data through online questionnaire surveys, uses the random forest method to screen out the important influencing factors affecting customers' Internet catering consumption, and on this basis, uses the PAM clustering method to segment Internet catering customers, and comprehensively evaluates customer loyalty on the basis of customer segmentation. The research results show that the general customer group accounts for the largest proportion, the low-value customer group accounts for the smallest proportion, and the recommendation of the five types of customers to Internet catering is not high.**

## Keywords

**Customer Segmentation; Customer Loyalty Assessment; Pam Clustering; Multi-level Fuzzy Comprehensive Evaluation.**

## 1. Introduction

With the rapid development of the Internet economy, Internet catering, as a product of the close integration of the modern Internet and the traditional catering industry, is also showing a strong upward trend. Due to the popularization and rapid development of the Internet, the competition among various catering companies has become increasingly fierce. As a high-frequency consumption industry, Internet catering can increase customer loyalty more conducive to increasing corporate profits.

Customer loyalty represents the customer's repeated and continuous purchase intention, which shows the identification and following of the product both psychologically and behaviorally. According to the 80/20 law proposed by the Italian economist Pareto, 20% of highly loyal customers account for 80% of the company's sales and profits. In this regard, ensuring the high loyalty of old customers can enable enterprises to achieve high-quality sustainable development. In the study of Internet catering marketing model, according to the purchase behavior and value of different types of customers, catering customers can be subdivided into five categories: important retention customers, important development customers, important retention customers, general customers, and low-value customers. The research on Internet catering customer loyalty based on customer segmentation provides a reference for catering companies to provide targeted products and services for different customers.

In recent years, in terms of enterprise customer value analysis and customer segmentation, researchers have established RFM models based on enterprise customer data and customer

ISSN: 2688-9323

value theory, combined with Recency, Frequency and Monetary, and used data mining to segment customers on this basis. Most of the commonly used methods are K-means clustering [1]. Harish and Malathyused historical retail transaction data combined with K-means clustering and Markov model to segment custodians [5]. Anitha and Patil analyzed real-time transaction and retail datasets based on the RFM model and using the K-Means algorithm [6]. However, K-means clustering is sensitive to outliers, irregular data is not easy to process, the effect is not good when the distance between classes is short, and the processing efficiency is low in large data sets, so the algorithm effect is not ideal. At the same time, traditional clustering methods including systematic clustering will greatly affect the clustering effect when there are many outliers, and are not suitable for complex data types. Due to the large number of Internet catering customers, high consumption frequency and high data dimension, the traditional clustering method is not the best customer segmentation method.

In terms of research on the relationship between customer loyalty and customer satisfaction, most scholars at home and abroad believe that customer loyalty is the emotional and behavioral trend of continuous purchase after the customer has a high satisfaction evaluation of the product after the purchase behavior. Hwang et al. believe that product service quality, satisfaction, trust and cost are significant determinants of brand loyalty [7]. The empirical research of Sasono et al. shows that customer satisfaction has a significant positive impact on customer loyalty [8]. Hoq and Amin also proposed that customer satisfaction is the most important driver of customer loyalty [9]. Since customer satisfaction is an important and significant positive factor affecting customer loyalty, this paper uses customer satisfaction of Internet catering customers to evaluate their loyalty. In terms of method research, Safari et al. used AHP to determine the weights of indicators at each level and sorted them [10]. Chan used the customer lifetime value (LTV) model to evaluate segmented customers [11]. Chan et al. developed a consumer segmentation model using the PSO algorithm [12]. However, the above method is too simple and traditional to calculate the specific score of customer loyalty according to the situation, and the index of customer satisfaction of Internet catering has more than one layer. In order to reasonably evaluate customer loyalty, traditional evaluation methods such as principal component analysis and other single-level methods are no longer applicable. It is necessary to determine the weights of each level before building the evaluation model, so that subsequent analysis can be carried out more reasonably.

Since the partitioning around medoids (PAM) algorithm in the k-medoid clustering algorithm is insensitive to outliers and noise, it has high accuracy and can handle different types of data points. In this paper, combined with the random forest algorithm, the PAM algorithm is used for customer segmentation. On the basis of customer segmentation, multi-level fuzzy comprehensive evaluation is used to evaluate the loyalty of Internet catering customers, so as to dig out high-value and high-loyalty customers for the enterprise, which is conducive to saving enterprise resources, maintaining old customers and attracting new customers to provide suggestions and strategies.

## 2. Model Construction and Data Sources

### 2.1. Index System Construction

Customer segmentation needs to be based on certain standards and methods. After clarifying its purpose, it is necessary to study the methods of customer segmentation from different angles. Generally speaking, customer segmentation can be divided into four different angles: feature segmentation, value segmentation, benefit segmentation, and behavior segmentation. According to the needs of the research, this paper constructs the customer segmentation index system as shown in Table 1.

ISSN: 2688-9323

**Table 1.** Construction of internet catering customer segmentation index system

| Primary indicators | Secondary indicators (variables) |
|---|---|
| Feature segmentation | Gender (x1) |
| | Age (x2) |
| | Marital status (x3) |
| | Education level (x4) |
| | Occupation (x5) |
| | Birthplace (x6) |
| | Residence in the past year (x7) |
| Value segmentation | Average monthly income (x8) |
| | Monthly food and beverage consumption (x9) |
| Benefit segmentation | Average number of takeaway orders per week (x10) |
| | Takeaway cost per meal (x11) |
| Behavior segmentation | Acceptable delivery time (x12) |
| | Food delivery provider (x13) |

**Table 2.** Construction of the evaluation index system of Internet catering customer loyalty

| Target layer (A) | Criterion layer (B) | Index layer (C) |
|---|---|---|
| Internet catering customer loyalty (A) | Customer expectation (B1) | Expectation degree of platform experience (C1) |
| | | Merchant dish quality expectation degree (C2) |
| | | The overall expectation of food delivery is (C3) |
| | Platform information quality (B2) | Satisfaction level of platform dishes (C4) |
| | | Satisfaction level of dish photos on the platform (C5) |
| | | Platform business evaluation satisfaction (C6) |
| | | Platform release information satisfaction (C7) |
| | Platform service quality (B3) | Delivery staff service attitude satisfaction (C8) |
| | | Satisfaction with delivery time of couriers (C9) |
| | | Platform delivery fee satisfaction (C10) |
| | | Platform complaint handling satisfaction (C11) |
| | | Overall service satisfaction of the platform (C12) |
| | Platform system quality (B4) | Platform client satisfaction (C13) |
| | | Platform response speed satisfaction (C14) |
| | | Platform page design satisfaction (C15) |
| | | Platform system quality satisfaction (C16) |
| | Merchant product quality (B5) | Satisfaction with the types of merchant dishes (C17) |
| | | Merchant dish taste satisfaction (C18) |
| | | Satisfaction with the portion size of merchant dishes (C19) |
| | | Merchant dish hygiene satisfaction (C20) |
| | | Merchant dish quality satisfaction (C21) |
| | Business service quality (B6) | Merchant dish price satisfaction (C22) |
| | | Merchant dish packaging satisfaction (C23) |
| | | Merchant dish packaging fee satisfaction (C24) |
| | | Merchant dish freshness satisfaction (C25) |
| | | Merchant overall service quality satisfaction (C26) |
| | Customer satisfaction (B7) | Overall platform satisfaction (C27) |
| | | Merchant overall satisfaction (C28) |
| | | Food order demand (C29) |
| | Industry recommendation (B8) | Willingness to support the development of food delivery industry (C30) |
| | | Willingness to use food delivery platforms to order food (C31) |
| | | Willingness to recommend takeaway products (C32) |

The variables used to study catering customer loyalty are related customer satisfaction, including the customer's expectations of food delivery, platform information quality satisfaction, platform service quality satisfaction, platform system quality satisfaction, merchant product quality satisfaction, merchant service quality satisfaction, customer satisfaction and industry recommendation. Analyze the relationship between various factors and build a hierarchical customer loss index system, which can be summarized as: target layer (A), criterion layer (B), and index layer (C), as shown in Table 2.

### 2.2. Data Collection and Preprocessing

This paper studies the customer loyalty evaluation under the subdivision of Internet catering customers. In order to make the survey results more descriptive, this survey adopts online survey and collects data through questionnaire star survey. A total of 62 questionnaires were distributed in the pre-investigation, 61 questionnaires were recovered, and 58 valid questionnaires were finally obtained, with an effective rate of 95%. The Cronbach $\alpha$ coefficient of the pre-survey questionnaire item is 0.806, the KMO value is 0.819, and the Bartlett sphericity test is $p = 0.000 < 0.05$, and it is statistically significant.

A total of 931 questionnaires were distributed in the formal survey, 925 questionnaires were recovered, and 861 valid questionnaires were finally obtained, with an effective rate of 93.1%. The missing data is interpolated by mode. For abnormal values, the percentile distribution and box plot are mainly used for detection. For abnormal points, according to the statistical results of variable index characteristics, abnormal values beyond the extreme value range are converted into specified extreme values. For data of different dimensions, it mainly includes normalization of numerical variables and one-hot encoding of categorical data, as well as standardization of data of different dimensions.

The processed questionnaire data is tested for reliability and validity, the Cronbach $\alpha$ coefficient is 0.821, the KMO value in the structural validity test is 0.839, and the Bartlett sphericity test is $p = 0.000 < 0.05$. It is statistically significant, indicating that the reliability and validity of the questionnaire are high, and the quality of the questionnaire is high.

## 3. Internet Catering Customer Segmentation based on PAM Clustering

With the rapid development of the Internet economy era, the competition among Internet catering companies is mainly reflected in the attraction and retention of customers. In order to better manage customer relationships, this paper will use PAM clustering to segment customers according to the value of Internet catering customer base.

### 3.1. Customer Segmentation Variable Selection

The Internet catering customer data in this survey has a large latitude and many data types. Before using PAM clustering to segment Internet catering customers, variable selection must be carried out first. Commonly used variable selection methods are linear weighting and Pearson coefficient methods, but these two methods are not suitable for high-latitude multi-type data. In this regard, this paper uses the Random Forest algorithm in data mining to select the main variables of Internet catering customer characteristics. The Random Forest is a classifier that uses multiple trees to train and predict samples. The forest contains many decision trees with high prediction accuracy and forms a combined prediction model. Among all current classification algorithms, Random Forest has excellent classification accuracy, can be effectively used on large data sets, and can handle input variables with high-dimensional features.

Since the variable "weekly average takeout times" represents the consumption frequency of internet catering customers, it is the most important indicator for customer value segmentation evaluation, so the weekly average number of takeout orders is selected as the dependent variable in the random forest, and all remaining variables are used as independent variables. For the Random Forest, in the process of building the decision tree $i$, a few input variables are randomly selected to form the candidate variable subset $\Theta_i$, and the randomly selected input variables obtained by building a random forest are: $P=10$, $K=\sqrt{P}$. The prediction error rate of the classification regression tree $i$ based on out-of-bag observation OOB is $e_i(i=1,2,\cdots,500)$, the prediction error of the classification and regression tree $i$ is $e_i^j(j=1,2,\cdots,10)$, and the change in the prediction error of the classification and regression tree $j$ caused by adding noise to the input variable is $c_i^j=e_i-e_i^j$.

Finally, the change of $M=500$ prediction errors is obtained, $\bar{c}^j=\frac{1}{M}\sum_{i=1}^{M}c_i^j$ is the average change of the random forest's overall prediction error caused by adding noise to the input variable $j$, and measures the importance of the input variable $j$. A total of 500 decision trees were established using RStudio software. Since the average number of takeaway orders per week is divided into 4 levels, the number of candidate input variables for each node is 4. The prediction misjudgment rate based on out-of-bag observation OOB was 31.01%, see Table 3.

**Table 3.** Random forest prediction results

| Confusion matrix: | | | | |
|---|---|---|---|---|
| | 1 | 2 | 3 | 4 |
| 1 | 397 | 78 | 2 | 0 |
| 2 | 109 | 193 | 2 | 0 |
| 3 | 19 | 47 | 4 | 0 |
| 4 | 2 | 7 | 1 | 0 |

Figure 1 is a direct observation of the characteristics of the OOB misjudgment rate with the number of decision trees in the random forest:

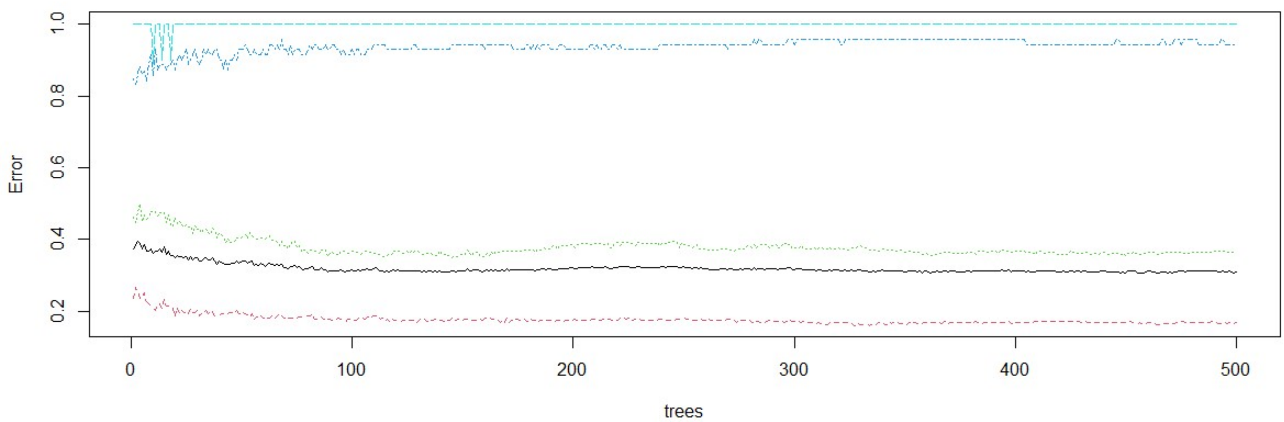


**Figure 1.** OOB false positive rate of random forest and decision tree

In Figure 1, the black line is the overall misjudgment rate, the red line is the misjudgment rate of "1" category prediction, the green line is the misjudgment rate of "2" category prediction, the dark blue line is the misjudgment rate of "3" category prediction, and the sky blue line is the misjudgment rate of "4" category prediction. It can be seen that the prediction effect of the model on the "1" class is better than that of the whole and other classes. After the number of decision trees reached 320, the misjudgment rates of various types remained basically stable.

When all observations are predicted, the prediction error is as small as 0.001, and the Importance component is used to evaluate the importance of the input variables, and the histogram of the average impact on the prediction accuracy after adding random noise to each input variable is drawn, see Figure 2:

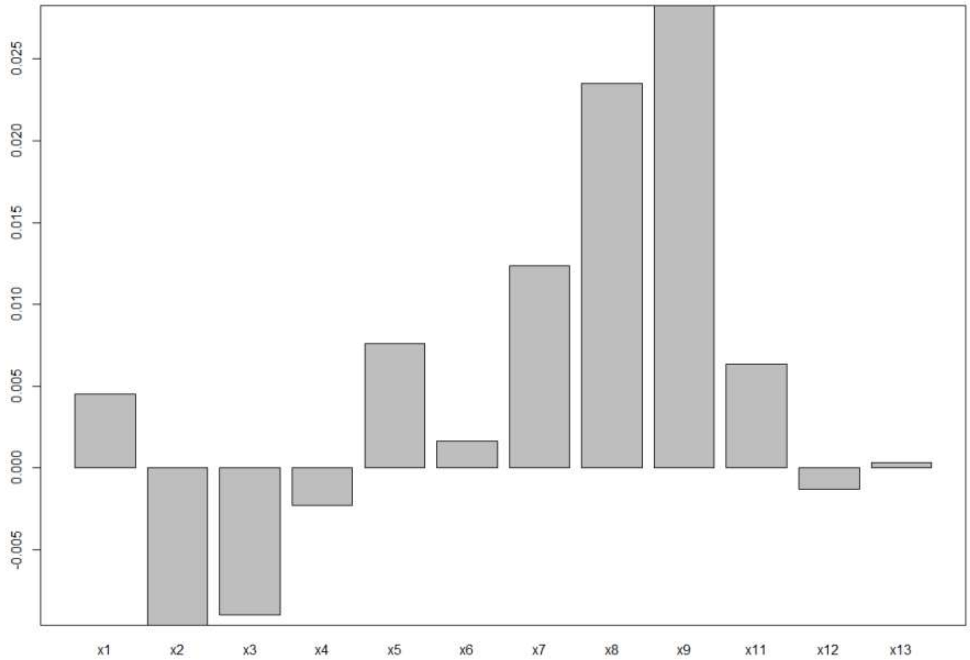


**Figure 2.** Histogram of input variable importance measures

Figure 2 shows that the variables x1, x2, x3, x5, x7, x8, x9, and x11 are input variables that are more important to the average number of takeaway orders per week. According to the height of the column chart, groups of people with different places of permanent residence, different monthly average income, and different monthly catering consumption in the past year have relatively different responses to Internet catering consumption.

## 3.2. PAM Clustering for Customer Segmentation

According to customer characteristics and important variables screened by random forest, use PAM clustering to cluster the sample data into 5 categories. The centroids of the five categories are observations 184, 43, 757, 264, and 699, respectively, including observations 162, 127, 167, 288, and 117 observation points. Then, these five types of customer samples can be divided into important customer retention, important development customer, important retention customer, general customer, and low-value customer according to customer value.

In order to prove the applicability and effectiveness of PAM clustering, this paper selects Calinski-Harabaz Index (CH) and Davies-bouldin Index (DBI) as cluster analysis evaluation indicators: CH is a sample-based measure of intra-class distance and inter-class variance matrix, and its judgment function is: $CH(k)=\frac{BGSS}{K-1}\Big/\frac{WGSS}{n-K}$. Where $n$ is the number of samples and $K$ is the number of categories. $WGSS=\frac{1}{2}[(n_1-1)\bar{d}_1^2+\cdots+(n_K-1)\bar{d}_k^2]$, $BGSS=\frac{1}{2}[(K-1)\bar{d}^2+(n-K)A_k]$, ($\bar{d}_j^2$ is the average distance between samples in the class $j$, $j=1,2,\cdots,K$; $\bar{d}^2$ is the average distance between all samples), $A_k=\frac{1}{n-K}\sum_{i=1}^{K}(n_i-1)(\bar{d}^2-\bar{d}_i^2)$. The larger the value of the CH index, the better the performance of the clustering result.

The DBI is the sum of the average distance from the samples in any two categories to the center of the class divided by the distance between the two center points, and the maximum value is taken, expressed as: $DBI = \frac{1}{K}\sum_{i=1}^{K}\max_{j \neq i}(\frac{\overline{C_i} + \overline{C_j}}{\|w_i - w_j\|_2})$. The smaller the DBI value, the smaller the intra-class distance, the larger the inter-class distance, and the better the clustering effect.

According to the above formula, the CH and DBI of PAM clustering can be calculated as: CH=85.63842, DBI=7.452282. It can be seen that the effect of PAM clustering is better, and the classification of customers is relatively balanced, so it is feasible to use PAM clustering as the method for segmenting Internet catering customers.

Based on the customer segmentation results, it is necessary to further calculate the value of each type of customer to identify the customer segmentation category. This paper uses principal components to calculate the average score of each type of customer groups on the principal components representing value, and on this basis, identifies the customer value levels of five types of customer representatives. Since the cumulative variance contribution rate of the first 12 principal components reached 86.34%, the first 12 principal components were selected.

This paper identifies the customer segmentation category itself based on customer value, so the first principal component can be understood as the value difference factor, and the average score of the value difference factor for each type of customer group is calculated, so as to quantitatively judge the value of customers, as shown in Table 4.

**Table 4.** Average score of value difference factors for different customer groups

| Different customer groups | Quantity | Proportion | Average score | Results |
|---|---|---|---|---|
| Category 1 | 162 | 18.82% | 12.8915 | Important to keep customers (KC) |
| Category 2 | 127 | 14.75% | 6.2295 | Important to retain customers (RC) |
| Category 3 | 167 | 19.39% | 10.8237 | Important development customers (DC) |
| Category 4 | 288 | 33.45% | 4.0234 | General customers (GC) |
| Category 5 | 117 | 13.59% | -0.3596 | Low-value customers (LC) |

As shown in Table 4, the average score of customer group value difference factor corresponding to category 1 is 12.8915, accounting for 18.82%, which is KC; the average score of customer group value difference factor corresponding to category 2 is 6.2295, accounting for 14.75%, which is RC; the average score of customer group value difference factor corresponding to category 3 is 10.8237, accounting for 19.39%, which is DC; The average score of customer group value difference factor corresponding to category 4 is 4.0234, accounting for 33.45%, which is GC; the average score of customer group value difference factor corresponding to category 5 is -0.3596, accounting for 13.59%, which is LC.

## 4. Customer Loyalty Evaluation based on Multi-level Fuzzy Comprehensive Evaluation

After the consumption is completed, the customer will evaluate the consumption behavior, and then generate customer satisfaction results. High satisfaction will prompt the customer to make repeated consumption, thus developing into a loyal customer of the merchant. Customer value and customer satisfaction are two important factors for customers to be loyal to the enterprise. Customer loyalty is the result of direct improvement on the basis of satisfaction, which will have direct benefits for businesses. In this regard, this paper uses customer satisfaction data to

ISSN: 2688-9323

evaluate the loyalty of Internet catering customers by constructing a multi-level fuzzy comprehensive evaluation model.

## 4.1. Construct Judgment Matrix

Since the customer loyalty evaluation index system has been constructed in Table 2, this paper constructs the judgment matrix according to the consistent matrix method proposed by Saaty. The scaling method of the judgment matrix $a_{ij}$ is shown in Table 5:

**Table 5.** Scaling method of each element in the judgment matrix

| $a_{ij}$ | Meaning |
|---|---|
| 1 | Index *i* and index *j* are equally important |
| 3 | Indicator *i* is slightly more important than indicator *j* |
| 5 | Indicator *i* is significantly more important than indicator *j* |
| 7 | Indicator *i* is strongly more important than indicator *j* |
| 9 | Indicator *i* is extremely more important than indicator *j* |
| 2, 4, 6, 8 | The median value of the above two adjacent judgments |
| Reciprocal | The two indicators are compared in reverse |

From this, the judgment matrixs can be obtained as Table 6 - Table 14:

**Table 6.** Judgment matrix of A

| | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 |
|---|---|---|---|---|---|---|---|---|
| B1 | 1 | 1/2 | 1/3 | 1/2 | 1/3 | 1/3 | 1/5 | 1/2 |
| B2 | 2 | 1 | 1/2 | 1 | 2/5 | 2/3 | 2/5 | 1 |
| B3 | 3 | 2 | 1 | 1/2 | 3/5 | 1 | 1/3 | 1/2 |
| B4 | 2 | 1 | 2 | 1 | 1/2 | 1/2 | 1/3 | 1 |
| B5 | 3 | 5/2 | 5/3 | 2 | 1 | 2 | 1/2 | 3 |
| B6 | 3 | 3/2 | 1 | 2 | 1/2 | 1 | 1/2 | 2 |
| B7 | 5 | 5/2 | 3 | 3 | 2 | 2 | 1 | 3 |
| B8 | 2 | 1 | 2 | 1 | 1/3 | 1/2 | 1/3 | 1 |

**Table 7.** Judgment matrix of B1

| | C1 | C2 | C3 |
|---|---|---|---|
| C1 | 1 | 1/3 | 1/5 |
| C2 | 3 | 1 | 1/2 |
| C3 | 5 | 2 | 1 |

**Table 8.** Judgment matrix of B2

| | C4 | C5 | C6 | C7 |
|---|---|---|---|---|
| C4 | 1 | 2 | 1/2 | 1/2 |
| C5 | 1/2 | 1 | 1/2 | 1/2 |
| C6 | 2 | 2 | 1 | 1/2 |
| C7 | 2 | 2 | 2 | 1 |

ISSN: 2688-9323

**Table 9.** Judgment matrix of B3

| | C8 | C9 | C10 | C11 | C12 |
|---|---|---|---|---|---|
| C8 | 1 | 1 | 1/2 | 2 | 1/3 |
| C9 | 1 | 1 | 1/2 | 2 | 1/3 |
| C10 | 2 | 2 | 1 | 3 | 1/2 |
| C11 | 1/2 | 1/2 | 1/3 | 1 | 1/5 |
| C12 | 3 | 3 | 2 | 5 | 1 |

**Table 10.** Judgment matrix of B4

| | C13 | C14 | C15 | C16 |
|---|---|---|---|---|
| C13 | 1 | 1/2 | 1 | 1/2 |
| C14 | 2 | 1 | 2 | 2 |
| C15 | 1 | 1/2 | 1 | 1 |
| C16 | 2 | 1/2 | 1 | 1 |

**Table 11.** Judgment matrix of B5

| | C17 | C18 | C19 | C20 | C21 |
|---|---|---|---|---|---|
| C17 | 1 | 1/2 | 1/2 | 1/2 | 1/3 |
| C18 | 2 | 1 | 1 | 1/2 | 1/2 |
| C19 | 2 | 1 | 1 | 1/2 | 1/2 |
| C20 | 2 | 2 | 2 | 1 | 1/2 |
| C21 | 3 | 2 | 2 | 2 | 1 |

**Table 12.** Judgment matrix of B6

| | C22 | C23 | C24 | C25 | C26 |
|---|---|---|---|---|---|
| C22 | 1 | 2 | 1 | 1/3 | 1/5 |
| C23 | 1/2 | 1 | 1/2 | 1/3 | 1/5 |
| C24 | 1 | 2 | 1 | 1/2 | 1/4 |
| C25 | 3 | 3 | 2 | 1 | 1/3 |
| C26 | 5 | 5 | 4 | 3 | 1 |

**Table 13.** Judgment matrix of B7

| | C27 | C28 | C29 |
|---|---|---|---|
| C27 | 1 | 1/2 | 1 |
| C28 | 2 | 1 | 2 |
| C29 | 1 | 1/2 | 1 |

**Table 14.** Judgment matrix of B8

| | C30 | C31 | C32 |
|---|---|---|---|
| C30 | 1 | 1 | 1/2 |
| C31 | 1 | 1 | 1/2 |
| C32 | 2 | 2 | 1 |

## 4.2. Judgment Matrix Consistency Test

Before using the judgment matrix to calculate the weight, it must be checked for consistency. First calculate the consistency index $CI$ : $CI = \frac{\lambda_{\max} - n}{n-1}$ , Among them, $\lambda_{\max}$ is the maximum eigenvalue of the judgment matrix, and $n$ is the index number of the judgment matrix. To measure the size of $CI$ , the average random consistency index $RI$ , is introduced, and its corresponding values are shown in Table 15:

**Table 15.** Average stochastic consistency index $RI$

| $n$ | 1 | 2 | 3 | 4 | 5 | 6 | 7 | 8 | 9 | 10 | 11 |
|---|---|---|---|---|---|---|---|---|---|---|---|
| $RI$ | 0 | 0 | 0.58 | 0.9 | 1.12 | 1.24 | 1.32 | 1.41 | 1.45 | 1.49 | 1.51 |

Finally calculate the random consistency index $CR$ : $CR = \frac{CI}{RI}$. It is generally believed that when $CR < 0.1$, the judgment matrix meets the consistency test, otherwise the judgment matrix should be readjusted. Through calculation, the consistency test results of the above judgment matrix can be obtained, as shown in Table 16:

**Table 16.** Judgment matrix consistency test results

| Judgment matrix | $\lambda$ | $n$ | $CI$ | $RI$ | $CR$ |
|---|---|---|---|---|---|
| A | 8.320 | 8 | 0.046 | 1.41 | 0.0326 |
| B1 | 3.004 | 3 | 0.002 | 0.58 | 0.0034 |
| B2 | 4.121 | 4 | 0.040 | 0.9 | 0.0444 |
| B3 | 5.017 | 5 | 0.004 | 1.12 | 0.0036 |
| B4 | 4.061 | 4 | 0.020 | 0.9 | 0.0222 |
| B5 | 5.088 | 5 | 0.022 | 1.12 | 0.0196 |
| B6 | 5.094 | 5 | 0.024 | 1.12 | 0.0214 |
| B7 | 3.000 | 3 | 0.000 | 0.58 | 0.0000 |
| B8 | 3.000 | 3 | 0.000 | 0.58 | 0.0000 |

From the test results in Table 16, it can be seen that all $CR$ values are less than 0.1, that is, all the above-mentioned judgment matrix consistency checks have passed, and the weight distribution of the matrix is reasonable. Then, the weight of the customer loyalty evaluation index can be obtained by using the eigenvalue method, as shown in Table 17.

From Table 17, it can be concluded that the weight of C is:

$$W_1 = (0.1094, 0.3090, 0.5816). \tag{1}$$

$$W_2 = (0.1953, 0.1381, 0.2761, 0.3905). \tag{2}$$

$$W_3 = (0.1350, 0.1350, 0.2414, 0.0743, 0.4143). \tag{3}$$

$$W_4 = (0.1682, 0.3952, 0.1976, 0.2390). \tag{4}$$

$$W_5 = (0.0965, 0.1577, 0.1577, 0.2437, 0.3444). \tag{5}$$

$$W_6 = (0.1040, 0.0689, 0.1161, 0.2231, 0.4879). \tag{6}$$

$$W_7 = (0.2500, 0.5000, 0.2500). \tag{7}$$

$$W_8 = (0.2500, 0.2500, 0.5000)\,. \tag{8}$$

**Table 17.** Weight value of Internet catering customer loyalty evaluation index

| B | Weights | C | Weights |
|---|---|---|---|
| B1 | 0.0449 | C1 | 0.1094 |
| | | C2 | 0.3090 |
| | | C3 | 0.5816 |
| B2 | 0.0817 | C4 | 0.1953 |
| | | C5 | 0.1381 |
| | | C6 | 0.2761 |
| | | C7 | 0.3905 |
| B3 | 0.0991 | C8 | 0.1350 |
| | | C9 | 0.1350 |
| | | C10 | 0.2414 |
| | | C11 | 0.0743 |
| | | C12 | 0.4143 |
| B4 | 0.0971 | C13 | 0.1682 |
| | | C14 | 0.3952 |
| | | C15 | 0.1976 |
| | | C16 | 0.2390 |
| B5 | 0.1877 | C17 | 0.0965 |
| | | C18 | 0.1577 |
| | | C19 | 0.1577 |
| | | C20 | 0.2437 |
| | | C21 | 0.3444 |
| B6 | 0.1316 | C22 | 0.1040 |
| | | C23 | 0.0689 |
| | | C24 | 0.1161 |
| | | C25 | 0.2231 |
| | | C26 | 0.4879 |
| B7 | 0.2645 | C27 | 0.2500 |
| | | C28 | 0.5000 |
| | | C29 | 0.2500 |
| B8 | 0.0934 | C30 | 0.2500 |
| | | C31 | 0.2500 |
| | | C32 | 0.5000 |

## 4.3. Calculation of Customer Loyalty Score by Fuzzy Comprehensive Evaluation

First, this paper creates an evaluation set=(very satisfied, satisfied, average, dissatisfied, very dissatisfied); Internet catering customer loyalty evaluation index set $U = U_i$ , $(i = 1, 2, \cdots, 8)$ , respectively represent B1, B2, B3, B4, B5, B6, B7 and B8. Then, the judgment matrix can be obtained by summarizing the satisfaction data of the questionnaire $R_i$, $(i = 1, 2, \cdots, 8)$. The matrix row vector is the factor of C, and the column vector is the percentage of customer satisfaction from high to low, specifically:

$$R_1 = \begin{cases} 0.381 & 0.343 & 0.142 & 0.076 & 0.058 \\ 0.353 & 0.445 & 0.075 & 0.061 & 0.066 \\ 0.386 & 0.395 & 0.111 & 0.055 & 0.053 \end{cases}. \tag{9}$$

$$R_2 = \begin{cases} 0.341 & 0.334 & 0.154 & 0.089 & 0.082 \\ 0.348 & 0.423 & 0.069 & 0.059 & 0.101 \\ 0.379 & 0.373 & 0.108 & 0.065 & 0.075 \\ 0.367 & 0.417 & 0.056 & 0.109 & 0.051 \end{cases}. \tag{10}$$

$$R_3 = \begin{cases} 0.376 & 0.358 & 0.107 & 0.074 & 0.085 \\ 0.325 & 0.453 & 0.064 & 0.066 & 0.092 \\ 0.404 & 0.368 & 0.086 & 0.067 & 0.075 \\ 0.372 & 0.391 & 0.092 & 0.107 & 0.038 \\ 0.353 & 0.395 & 0.098 & 0.083 & 0.071 \end{cases}. \tag{11}$$

$$R_4 = \begin{cases} 0.388 & 0.360 & 0.114 & 0.071 & 0.067 \\ 0.359 & 0.452 & 0.051 & 0.053 & 0.085 \\ 0.406 & 0.380 & 0.099 & 0.051 & 0.064 \\ 0.387 & 0.405 & 0.076 & 0.098 & 0.034 \end{cases}. \tag{12}$$

$$R_5 = \begin{cases} 0.392 & 0.336 & 0.128 & 0.086 & 0.058 \\ 0.338 & 0.466 & 0.068 & 0.056 & 0.072 \\ 0.376 & 0.392 & 0.117 & 0.065 & 0.050 \\ 0.362 & 0.429 & 0.082 & 0.099 & 0.028 \\ 0.347 & 0.415 & 0.102 & 0.065 & 0.071 \end{cases}. \tag{13}$$

$$R_6 = \begin{cases} 0.382 & 0.350 & 0.136 & 0.074 & 0.058 \\ 0.362 & 0.458 & 0.055 & 0.052 & 0.073 \\ 0.411 & 0.365 & 0.107 & 0.052 & 0.065 \\ 0.401 & 0.404 & 0.069 & 0.087 & 0.039 \\ 0.367 & 0.398 & 0.111 & 0.062 & 0.062 \end{cases}. \tag{14}$$

$$R_7 = \begin{cases} 0.343 & 0.352 & 0.144 & 0.079 & 0.082 \\ 0.355 & 0.437 & 0.071 & 0.068 & 0.069 \\ 0.375 & 0.377 & 0.128 & 0.062 & 0.058 \end{cases}. \tag{15}$$

$$R_8 = \begin{cases} 0.373 & 0.348 & 0.129 & 0.075 & 0.075 \\ 0.347 & 0.458 & 0.050 & 0.063 & 0.082 \\ 0.364 & 0.386 & 0.124 & 0.053 & 0.073 \end{cases}. \tag{16}$$

Then according to the weight value of each index, calculate the fuzzy comprehensive evaluation matrix $B$ of C, and $B = \{B_i\} = \{W_i \times R_i\}$, $i = 1, 2, \cdots, 8$. Can be calculated:

$$B=\begin{bmatrix}0.375 & 0.405 & 0.103 & 0.059 & 0.058\\ 0.363 & 0.389 & 0.091 & 0.086 & 0.071\\ 0.366 & 0.391 & 0.091 & 0.077 & 0.074\\ 0.380 & 0.411 & 0.077 & 0.066 & 0.066\\ 0.358 & 0.415 & 0.097 & 0.074 & 0.056\\ 0.381 & 0.395 & 0.100 & 0.067 & 0.058\\ 0.357 & 0.401 & 0.104 & 0.069 & 0.070\\ 0.271 & 0.298 & 0.076 & 0.048 & 0.058\end{bmatrix}. \quad (17)$$

Then perform defuzzification calculation on C to obtain the evaluation scores of B1, B2, B3, B4, B5, B6, B7, and B8: $F_i = 5B_{i1} + 4B_{i2} + 3B_{i3} + 2B_{i4} + B_{i5}$, where $i = 1, 2, \cdots, 8$:

$$\begin{aligned} F_B &= (F_1, F_2, F_3, F_4, F_5, F_6, F_7, F_8) \\ &= (3.980, 3.887, 3.895, 3.973, 3.945, 3.977, 3.909, 2.929) \end{aligned}. \quad (18)$$

Combined with B's weight index value $W$ :

$$W = (0.0449, 0.0817, 0.0991, 0.0971, 0.1877, 0.1316, 0.2645, 0.0934). \quad (19)$$

and fuzzy comprehensive evaluation matrix $B$ , and obtain the comprehensive evaluation results of customer loyalty $Q$:

$$Q = (0.357, 0.392, 0.095, 0.069, 0.064). \quad (20)$$

Its comprehensive evaluation score is: $F_Q = 3.84$. The comprehensive score can be identified as the loyalty of Internet catering customers, and finally the evaluation results of the overall loyalty of Internet catering customers are obtained:

**Table 18.** Evaluation results of internet catering customer loyalty

| Item | B1 | B2 | B3 | B4 | B5 | B6 | B7 | B8 | A |
|---|---|---|---|---|---|---|---|---|---|
| Score | 3.98 | 3.887 | 3.895 | 3.973 | 3.945 | 3.977 | 3.909 | 2.929 | 3.84 |
| Conclusion | satisfied | average | average | satisfied | satisfied | satisfied | satisfied | dissatisfied | higher |

It can be concluded from Table 18 that the overall satisfaction score of Internet catering customers is 3.84, which is greater than 3, between average and satisfactory, and the overall customer loyalty is relatively high. However, the lowest score of B8 is 2.929, indicating that catering customers are less willing to support the development of the food delivery industry, use food delivery platforms to order food, and recommend food delivery products. Catering merchants and platforms should take corresponding measures in time to increase the industry recommendation of customers.

## 5. Evaluation of Customer Loyalty based on Customer Segmentation

In the previous article, PAM clustering was used to subdivide Internet catering customers into LC, GC, RC, DC and KC, and the loyalty evaluation of overall catering customers was obtained through multi-level fuzzy comprehensive evaluation. This article will then establish the fuzzy comprehensive evaluation judgment matrix of the five types of customer segments, and obtain the loyalty evaluation results of the five types of customer segments (see Table 19).

**Table 19.** Segmented Customer Loyalty Evaluation

| B | $F_B$ | | | | |
|---|---|---|---|---|---|
| | LC | GC | RC | DC | KC |
| B1 | 4.058 | 4.005 | 4.120 | 3.840 | 3.950 |
| B2 | 3.821 | 3.910 | 3.830 | 3.790 | 3.990 |
| B3 | 4.026 | 3.650 | 4.420 | 3.700 | 4.220 |
| B4 | 4.305 | 4.280 | 4.430 | 4.380 | 2.513 |
| B5 | 2.020 | 4.300 | 4.300 | 4.360 | 4.020 |
| B6 | 3.788 | 3.580 | 4.400 | 4.280 | 4.215 |
| B7 | 3.933 | 3.790 | 3.990 | 3.983 | 3.985 |
| B8 | 2.976 | 2.920 | 2.790 | 2.950 | 2.885 |
| $F_Q$ | 3.506 | 4.449 | 4.068 | 3.987 | 4.220 |

From the multi-level fuzzy comprehensive evaluation, the evaluation indicators and loyalty scores of five types of subdivided customers can be obtained. It can be found that among LC, the platform system quality score is the highest, and the merchant product quality score is the lowest. Among GC, merchants have the highest product quality score and the lowest industry recommendation score. Among RC, the platform system quality score is the highest, and the industry recommendation score is the lowest. Among DC, the platform system quality score is the highest, and the industry recommendation score is the lowest. Among KC, the platform service quality score is the highest and the platform system quality score is the lowest. Among the five types of segmented customers, general customers have the highest loyalty score, and low-value customers have the lowest loyalty score.

## 6. Conclusion and Suggestion

Customer loyalty is the direct and sustainable source of corporate interests, and customer value and customer satisfaction are the two most important factors affecting customer loyalty. Based on user data in Lanzhou, this paper uses PAM clustering to divide Internet catering customers into LC, GC, RC, DC and KC, based on the value of five types of segmented customers, evaluate the loyalty of catering customers. The following conclusions are drawn: LC have the lowest loyalty score among the five types of customer segments, and GC have the highest loyalty score. The industry recommendation degree of each type of customer is generally not high, which shows that Internet catering in Lanzhou City has not really attracted the audience, let alone the purpose of attracting new customers, and there is still a lot of room for improvement.

Combined with the results of customer segmentation and customer loyalty evaluation, in order to maintain old customers and attract new customers, the following suggestions are put forward for Internet catering businesses and platforms: First, identify customer types and achieve precise marketing. The overall loyalty evaluation of Internet catering customers is relatively high, but when it comes to specific customer groups, different types of customers may have different loyalty. It is necessary to pay attention to identifying customer types and formulate targeted sales strategies based on group characteristics and consumption behavior. Second, improve product quality in an all-round way to ensure that production meets standards. The product quality of merchants is an important factor affecting customer loyalty. Internet catering merchants must ensure the quality of each product, and strictly control food safety and taste effects on the premise of ensuring visual effects. Third, improve overall customer satisfaction, so that customers have a strong willingness to recommend, thereby increasing the number of users. According to the analytic hierarchy process, customer satisfaction is the first major influencing factor of customer loyalty, and customer satisfaction is a direct influencing factor of customer recommendation. Internet catering merchants need to focus on ensuring that the quality of meals meets standards, grasp the freshness and health of meals, carefully package them, and price them reasonably; Internet catering platforms need to strengthen the construction of platform systems. Comprehensively improve catering customer satisfaction, and promote customers' long-term purchase and recommendation willingness.

## Acknowledgments

This work was supported by National Social Science Fund (21BTJ042), Gansu Higher Education Innovation Fund Project (2020A-068), Research Project of Lanzhou University of Finance and Economics (Lzufe2019B-012), and Innovation Fund of Anhui University of Finance and Economics (ACYC2021395).

## References

[1] Z.H. Sun, T.Y. Zuo, D. Liang, et al. GPHC: A heuristic clustering method to customer segmentation, Applied Soft Computing, Vol.111 (2021), p.107677.

[2] K. Tabianan, S. Velu, and V. Ravi, K-means clustering approach for intelligent customer segmentation using customer purchase behavior data, Sustainability, Vol.14 (2022) No.12, p.7243.

[3] Y. Deng, and Q. Gao: A study on e-commerce customer segmentation management based on improved K-means algorithm, Information Systems and e-Business Management, Vol.18 (2020), p.497-510.

[4] Y. Li, J. Qi, X. Chu, et al. Customer segmentation using K-means clustering and the hybrid particle swarm optimization algorithm, The Computer Journal, Vol.66 (2023) No.4, p.941-962.

[5] A.S. Harish, and C. Malathy: Customer segment prediction on retail transactional data using k-means and markov model, Intelligent Automation & Soft Computing, Vol.36 (2023) No.1, p.589-600.

[6] P. Anitha, and M.M. Patil: RFM model for customer purchase behavior using K-Means algorithm, Journal of King Saud University-Computer and Information Sciences, Vol.34 (2022) No.5, p.1785-1792.

[7] S. Hwang, M. Lee, E. Park, et al. Determinants of customer brand loyalty in the retail industry: A comparison between national and private brands in South Korea, Journal of Retailing and Consumer Services, Vol.63 (2021), p.102684.

[8] I. Sasono, A.D. Jubaedi, D. Novitasari, et al. The impact of e-service quality and satisfaction on customer loyalty: Empirical evidence from internet banking users in Indonesia, The Journal of Asian Finance, Economics and Business, Vol.8 (2021) No.4, p.465-473.

[9] M.Z. Hoq, and M. Amin: The role of customer satisfaction to enhance customer loyalty, African Journal of Business Management, Vol.4 (2010) No.12, p.2385-2392.

[10] F. Safari, N. Safari, and G.A. Montazer, Customer lifetime value determination based on RFM model, Marketing Intelligence & Planning, Vol.34 (2016) No.4, p.446-461.

[11] C.C.H. Chan: Intelligent value-based customer segmentation method for campaign management: A case study of automobile retailer, Expert Systems with Applications, Vol.34 (2008) No.4, p.2754-2762.

[12] C.C.H. Chan, Y.R. Hwang, and H.C. Wu: Marketing segmentation using the particle swarm optimization algorithm: a case study, Journal of Ambient Intelligence and Humanized Computing, Vol.7 (2016), p.855-863.