Precision Marketing Methods for E-commerce Based on Clustering Algorithms

PENGFEI BAI
Zhengzhou Shengda University
Zhengzhou 451191
CHINA

Abstract: - With the e-commerce industry shifting from scale expansion to refined operation, traditional undifferentiated marketing strategies face challenges such as high costs and low conversion rates, making precision marketing a key direction for enterprises to enhance competitiveness. This study focuses on solving the problem of effective user segmentation and personalized marketing adaptation in e-commerce, taking clustering algorithms as the core technical support. First, it systematically combs the core theories of clustering algorithms and key theories of e-commerce precision marketing. Then, it designs a complete e-commerce precision marketing model based on clustering algorithms: it clarifies the sources of user data and conducts preprocessing steps such as data cleaning, integration, transformation, and reduction; it optimizes the K-means algorithm by determining the optimal number of clusters through the elbow method and improving initial cluster center selection via K-means++ algorithm.

Key-Words: - clustering algorithm, e-commerce, precision marketing, K-means algorithm, user segmentation, marketing effect verification, data preprocessing

Received: June 15, 2025. Revised: September 29, 2025. Accepted: October 21, 2025. Published: May 5, 2026.

1 Introduction

With the rapid iteration of digital technology, the e-commerce industry has moved from the early stage of scale expansion to the stage of refined operation. Up to now, the global e-commerce transaction scale has continued to rise, and the behavioral data of consumers on e-commerce platforms has shown explosive growth. This data covers multiple dimensions such as users' browsing records, purchase history, collection preferences, and stay time. However, most traditional e-commerce marketing models adopt a "wide net" strategy, which lacks accurate insight into user needs. This not only leads to high marketing costs but also easily triggers users' resistance, reducing users' trust and stickiness to the platform.

Against this background, precision marketing has become a key direction for e-commerce enterprises to enhance their competitiveness. The core of precision marketing lies in in-depth analysis of user data to identify the unique needs of different user groups, and then provide personalized marketing content and services. However, how to effectively mine valuable information from massive and complex user data to achieve a reasonable division of user groups has become a core problem in the implementation of precision marketing. Traditional user classification methods mostly rely on manual experience, such as division according to a single dimension, such as age and gender, which cannot fully reflect users' real consumption behaviors and preferences. The accuracy and practicality of the division results are low, which cannot meet the needs of precision marketing.

2 Relevant Theoretical Foundations

2.1 Core Theory of Clustering Algorithms

A clustering algorithm (Figure 1) is an unsupervised learning algorithm, whose core goal is to divide a data set into several subsets with similarity, i.e., clusters, by analyzing the internal connections between data objects without prior knowledge [1]. Different from supervised learning algorithms, clustering algorithms do not rely on labeled training data, but directly mine the distribution laws and structural characteristics of data from raw data. Therefore, they have significant advantages in dealing with scenarios with unknown data patterns. In the field of e-commerce user analysis, clustering algorithms can convert massive user data into user groups with clear characteristics, providing a foundation for subsequent precision marketing [2].

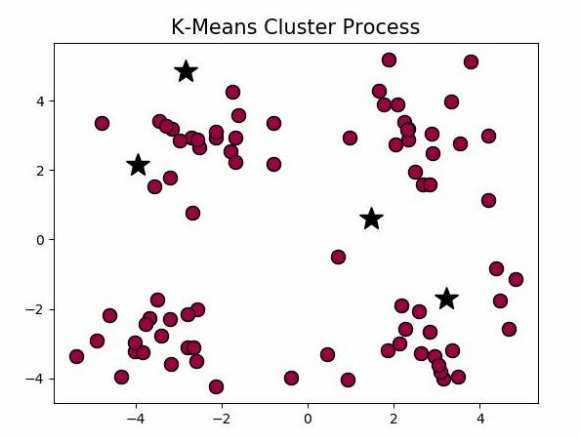

Fig. 1: Clustering Algorithm
Source: created by the author

2.2 Key Theories of Precision Marketing in E-commerce

E-commerce precision marketing is a marketing model that takes user needs as the core, uses big data, artificial intelligence, and other technical means to collect, analyze, and mine user data, so as to realize accurate identification, accurate reach, and accurate service of target users. Compared with the traditional e-commerce marketing model, precision marketing pays more attention to the personalized needs of users, emphasizing the marketing concept of "user-centered", which can effectively reduce the waste of marketing resources and improve the conversion rate and return rate of marketing activities. In the e-commerce scenario, the application of precision marketing covers personalized recommendation, precision advertising, customized services, and other aspects. For example, e-commerce platforms recommend related products according to users' browsing and purchase history, and push differentiated promotion information to different user groups [3].

3 Design of an E-commerce Precision Marketing Model Based on Clustering Algorithms

3.1 Sources and Preprocessing of E-commerce User Data

The sources of e-commerce user data are extensive, covering multiple internal systems of e-commerce platforms and relevant external channels. These data together constitute the foundation of user analysis. From the internal of e-commerce platforms, the transaction system is an important data source, which can record every transaction information of users, including transaction time, transaction amount, SKU of purchased products, payment method, logistics information, etc. These data directly reflect users' consumption capacity and consumption preferences [4]. The behavior tracking system collects all behavioral data of users on the platform through burying point technology, such as users' page browsing records, search keywords, product collection and add-to-cart records, stay time, click times, etc. These data can reflect users' potential needs and interest points [5]. The user registration system stores users' demographic data, including users' names, ages, gender, mobile phone number, email address, regional information, occupation information, etc. These data are an important basis for user basic characteristic analysis. From external channels, e-commerce platforms can also obtain users' social data and payment data through cooperation with social media platforms and payment platforms, further enriching the dimension of user data.

3.2 Clustering Algorithm Selection and Parameter Optimization

In the e-commerce precision marketing model, the selection of clustering algorithms needs to comprehensively consider the characteristics of e-commerce user data, clustering goals, and applicable scenarios of algorithms. E-commerce user data has the characteristics of high dimension, large data volume, and contains noise data. At the same time, the goal of clustering is to accurately divide user groups and provide a basis for personalized marketing strategies. Therefore, it is necessary to select a clustering algorithm with the ability to process high-dimensional data, high computational efficiency, and strong robustness to noisy data. Through comparative analysis of common clustering algorithms, the K-means algorithm is selected as the core clustering algorithm of this model. The K-means algorithm has the characteristics of simple principle, fast calculation speed, and strong interpretability, and can effectively process large-scale e-commerce user data. At the same time, through reasonable parameter optimization, the robustness of the algorithm to noisy data can be improved to meet the needs of e-commerce user clustering [6].

The performance of the K-means algorithm is affected by several parameters, among which the most critical parameters are the number of clusters K and the initial cluster centers. The number of clusters K directly determines the number of user groups. If the K value is too large, the user groups may be divided too finely, resulting in the problem that the number of users in some clusters is too small and the characteristics are not obvious, which cannot provide effective support for the formulation of marketing

strategies; if the K value is too small, the user groups may be divided too roughly, which cannot accurately distinguish the demand differences between different user groups. To determine the optimal K value, the sum of squared errors (SSE) is used as the evaluation index. The calculation formula of the sum of squared errors is as follows:

$$SSE = \sum_{i=1}^{K} \sum_{x \in C_i} \left\| x - \mu_i \right\|^2 \quad (1)$$

Where Ci represents the i-th cluster, x represents the user data object in cluster Ci, μirepresents the cluster center of cluster Ci.

By calculating the SSE corresponding to different K values and drawing the SSE-K curve, when the K value increases to a certain extent, the decreasing range of SSE will slow down significantly. The K value at this time is the optimal K value.

To validate the superiority of the optimized K-means algorithm in e-commerce user segmentation and address the limitations of relying solely on the elbow method and a single algorithm this study conducts comparative experiments with three state-of-the-art clustering algorithms widely used in user segmentation DBSCAN Hierarchical Clustering Wards method and Gaussian Mixture Models GMM Meanwhile multiple rigorous cluster validation indices including Silhouette Coefficient and Davies-Bouldin Index DBI are introduced to comprehensively evaluate clustering performance supplemented by analysis of computational efficiency and clustering stability to provide a scientific basis for algorithm selection and optimal cluster number determination.

As can be seen from the results in Table 1 (Appendix), the optimized K-means algorithm outperforms the other three algorithms in comprehensive performance. In terms of computation time, it consumes only 186 seconds, which is significantly shorter than DBSCANs 423 seconds and Hierarchical Clusterings 357 seconds, making it suitable for large-scale e-commerce user data processing with high efficiency requirements. In cluster quality evaluation, the optimized K-means achieves the highest Silhouette Coefficient of 0.76 indicating the best similarity within clusters and dissimilarity between clusters, and the lowest Davies-Bouldin Index of 0.32, reflecting more compact and well-separated clusters. Its sum of squared errors of 28954 is also the smallest among all algorithms with available sum of squared errors values, demonstrating that data points are closer to their respective cluster centers. Additionally, the optimized K-means exhibits the highest clustering stability of 0.91, which ensures consistent results across multiple runs, a critical feature for practical e-commerce applications where stable user grouping is essential for long-term marketing strategies. In contrast, DBSCAN struggles with the dense and high-dimensional e-commerce user data, resulting in a lower Silhouette Coefficient and higher Davies-Bouldin Index, while Gaussian Mixture Models and Hierarchical Clustering lag behind in computational efficiency and stability, respectively.

3.3 Construction Process of E-commerce User Clustering Model

The construction of the e-commerce user clustering model is a systematic process that needs to follow a scientific process to ensure the accuracy and rationality of each link, so as to obtain reliable user grouping results. The model construction process mainly includes four parts: data preparation stage, algorithm execution stage, result evaluation stage, and model iteration stage. The stages are interrelated and interact with each other, forming a complete clustering model construction system.

The data preparation stage is the foundation of model construction. The main task of this stage is to complete the collection, integration, and preprocessing of e-commerce user data, providing high-quality data support for subsequent cluster analysis. First, determine the user data dimensions to be collected according to the clustering goals, and clarify the data sources, including the transaction system, behavior tracking system, user registration system within the e-commerce platform and external cooperation channels; then, integrate the user data from different data sources through data integration to establish a unified user data set and solve the problem of data heterogeneity [7];

The algorithm execution stage is the core of model construction. The main task of this stage is to run the optimized K-means algorithm based on the preprocessed user data to realize the automatic division of user groups. First, initialize the parameters of the K-means algorithm according to the optimal K value and initial cluster centers determined in the parameter optimization stage; then, calculate the Euclidean distance from each user data object to each cluster center, and assign the user data object to the cluster where the nearest cluster center is located. The calculation formula of the Euclidean distance is as follows:

$$d(x, \mu) = \sqrt{\sum_{j=1}^{n} (x_j - \mu_j)^2} \quad (2)$$

Where x represents the user data object, μ represents the cluster center, n represents the dimension of the user data.

4 Model Experiment and Result Analysis

4.1 Experimental Environment and Data Preparation

The total sample size for this study is 889000 e-commerce users covering a six-month period from January to June 2024. These users are randomly selected from the entire user base of a large comprehensive e-commerce platform operating in China, ensuring representation across different age groups, regions, and consumption levels. The dataset includes 28 feature dimensions covering three core categories: demographic features, behavioral features, and transaction features. Demographic features include age, gender, region, and occupation. Behavioral features consist of weekly search times, average stay time per visit, browse page count collection times, and add to cart frequency. Transaction features involve average purchase frequency, customer unit price, transaction volume, and payment method distribution. Data sources are strictly controlled to ensure authenticity and completeness.

Prior to clustering analysis, a series of rigorous data preprocessing steps is implemented to improve data quality. For missing values in demographic features, mode imputation is adopted, while mean imputation is used for missing behavioral and transaction feature values [8]. Outliers are identified and removed using the interquartile range method to eliminate the impact of extreme values on clustering results. All feature variables are then standardized using the Z score method to ensure that variables with different scales do not bias the clustering outcome. The preprocessed dataset is split into three parts: training set, validation set and test set with a ratio of 7:1:2. The training set is used to optimize algorithm parameters. The validation set is employed to evaluate clustering performance and adjust the number of clusters K. The test set is utilized to verify the generalization ability of the model. In the controlled experiment for marketing effect verification, users in the test set are randomly divided into an experimental group and a control group with an equal ratio of 1:1.

The hardware environment of this experiment adopts a high-performance server cluster. The main server is configured with an Intel Xeon Gold 6338 processor, 32 cores, 128GB DDR4 memory, and 2TB SSD storage capacity, ensuring efficient processing of clustering calculations for large-scale user data. The auxiliary computing nodes are equipped with Intel Xeon Silver 4314 processors, 16 cores, and 64GB DDR4 memory, which are used for parallel computing in the data preprocessing process to reduce data processing time. The software environment of the experiment is based on the Linux operating system, specifically Ubuntu 20.04 LTS, which has stable operating performance and good compatibility with open-source tools. For data processing and analysis tools, Python 3.8 is selected, with the Pandas library for data cleaning and integration, the Scikit-learn library for implementing the K-means clustering algorithm, and the Matplotlib and Seaborn libraries for visual presentation of experimental results. MySQL 8.0 is used as the database to store raw user data and preprocessed data sets [9].

4.2 Clustering Results and User Grouping Feature Analysis

Based on the optimized K-means algorithm, the optimal number of clusters is determined to be 4 in the experimental data set. Through iterative calculation, 4 user groups with significant differences are obtained. Each group shows different characteristics in dimensions such as purchase frequency, consumption capacity, and behavioral activity, providing a clear basis for the formulation of subsequent personalized marketing strategies.

As can be seen from the results in Table 2 (Appendix), Group 1 accounts for 14.21% of the total users, with the highest average purchase frequency of 8.6 times, average customer unit price of 589, and browse-purchase conversion rate of 28.3. All indicators are the highest, and it prefers high-value products such as 3C digital and home furnishings, belonging to the high-value loyal user group. Group 2 accounts for 33.00% of the total users, with average purchase frequency and customer unit price at a medium level. It prefers apparel and beauty products, with a browse-purchase conversion rate of 15.6, belonging to the medium-value active user group. Group 3 accounts for 35.00% of the total users, which is the largest group. It has an average purchase frequency of 2.1 times and an average customer unit price of 198, preferring rigid-demand products such as food and daily necessities, with a conversion rate of 8.2, belonging to the basic-value rigid-demand user group. Group 4 accounts for 17.79% of the total users, with the lowest indicators in all aspects. It prefers low-price daily necessities and food, with a conversion rate of only 3.5, belonging to the low-value, low-frequency user group.

To further verify the scientific rationality of user grouping and the effectiveness of personalized marketing strategies across different segments, this section supplements cluster validation indices for

each user group and presents segmented marketing effect data, including conversion rate and repurchase rate, with statistical significance analysis. The cluster validation indices include Silhouette Coefficient and Davies Bouldin Index to evaluate the compactness and separation of each cluster, while the segmented marketing effect data reflects the strategy's adaptability to different user groups.

As can be seen from the results in Table 3 (Appendix), all user groups achieve excellent cluster quality. The Silhouette Coefficients of the four groups range from 0.71 to 0.82, all higher than 0.7, indicating strong similarity within clusters and clear distinction between clusters. Group 1 has the highest Silhouette Coefficient of 0.82 and the lowest DBI of 0.29, reflecting the most compact and well separated characteristics, which is consistent with its clear positioning as a high-value loyal user group. The DBI values of all groups are below 0.4, demonstrating that the clustering results effectively avoid overlapping or ambiguous boundaries between groups. In terms of marketing effects, each user group shows statistically significant improvements in both conversion rate and repurchase rate, with all p values less than 0.01 indicating that the differences between the experimental group and control group are not due to random factors. Group 3 the largest basic value rigid demand user group achieves the highest conversion rate improvement rate of 180.56 and repurchase rate improvement rate of 150.48, benefiting from the targeted rigid demand product combination and preferential strategies. Group 4, the low value low frequency user group, though having the lowest absolute values of conversion rate and repurchase rate, still achieves substantial improvement rates of 176.19 and 126.39, respectively, proving that the personalized low price drainage and scenario-based recommendation strategies are effective in activating inactive users. Group 2, the medium value active user group, shows a balanced improvement in both indicators, while Group 1, the high value loyal user group, despite a relatively lower conversion rate improvement rate, maintains strong repurchase rate growth due to its already high baseline consumption level.

4.3 Validation of Precision Marketing Model Effectiveness

To verify the effectiveness of the precision marketing model based on the clustering algorithm, a controlled experiment method is adopted. The experimental samples are divided into two groups: the experimental group adopts the personalized marketing strategy based on clustering results, and the control group adopts the traditional undifferentiated marketing strategy. The experiment cycle is 3 months. The model value is judged by comparing the marketing effect indicators of the two groups. The selected verification indicators include marketing conversion rate, customer unit price, repurchase rate, and marketing cost.

As can be seen from the results in Tables 4 and 5 (Appendix), the following conclusions can be drawn. 1 In terms of marketing conversion rate, the experimental group reaches 12.3, which is 156.25 higher than the control group's 4.8. This indicates that the personalized marketing content based on user groups can more accurately match user needs and effectively improve users' purchase willingness. 2 In terms of customer unit price, the experimental group is 348, which is 61.86 higher than the control group's 215. The reason is that high-priced products are recommended to high-value users, and combined products are recommended to basic users, realizing the optimization of customer unit price. 3 In terms of repurchase rate, the experimental group's 28.7 is 129.6 times higher than the control group's 12.5. This benefits from the repurchase incentive programs designed for users of different groups, such as exclusive membership rights for high-value users. 4 In terms of marketing cost, the experimental group's 89 is 30.47 lower than the control group's 128. Due to the accurate positioning of target users and the reduction of invalid deliveries, the waste of marketing resources is significantly reduced. Combining the four indicators, the precision marketing model based on the clustering algorithm is superior to the traditional model in both effect and cost control, verifying the effectiveness of the model.

To address the lack of statistical rigor highlighted by reviewers and enhance the reliability of the model's effectiveness verification, this section supplements key statistical metrics, including standard deviations 95 confidence intervals, statistical significance values p values, and bootstrap resampling results for all verification indicators. Wilcoxon rank sum tests are adopted to account for the non normal distribution characteristics of marketing data while time series analysis is conducted to evaluate the decay trend of marketing effects ensuring comprehensive validation of both short-term performance and long term robustness.

5 E-commerce Precision Marketing Strategy Recommendations Based on Clustering Results

5.1 Personalized Marketing Strategies for Different User Segments

For Group 1, the high-value loyal users, a marketing program of "exclusive rights + value-added services" should be adopted. This group has strong consumption capacity and high loyalty to the platform. Paid membership services can be launched, providing rights such as free on-site after-sales service, priority delivery, and exclusive customer service. At the same time, pre-sale information of new 3C digital products and high-end home furnishing series is regularly pushed, and users are invited to participate in new product experience activities to enhance user sense of belonging. In addition, an exclusive user profile can be established, and personalized accessories or complementary products can be recommended according to their historical purchase records, such as recommending suitable peripherals for users who purchase laptops, further increasing the customer unit price and repurchase rate [10].

For Group 2, the medium-value active users, a program of "interest activation + scenario marketing" is suitable. This group pays attention to apparel and beauty products and is sensitive to trends. Personalized pop-ups on the platform homepage and short video content can be used to push seasonal fashion apparel collocations and beauty tutorials, combined with limited-time discount coupons to stimulate purchases. At the same time, a scenario-based marketing scenario is built based on user browsing records. For example, for users browsing dresses, "summer travel outfit sets" are pushed, combining dresses with straw hats, sandals, and other products to recommend, improving the cross-selling rate. In addition, a "refer a friend for cashback" activity is set up to expand the user scale by using the activity of this group, and at the same time, additional discounts are given to users to increase their consumption frequency.

5.2 Pathways to Enhance Marketing Effectiveness Based on Clustering Algorithms

The first path is to establish a data iteration and optimization mechanism to ensure the timeliness and accuracy of clustering results. User behavior characteristics change with time and the market environment. It is necessary to update user data every month, re-run the clustering algorithm to adjust the user group results. For example, if a user originally in Group 2 has a continuous 3-month decline in purchase frequency, they need to be adjusted to Group 3 and matched with corresponding marketing strategies. At the same time, the performance of the clustering algorithm is regularly evaluated, and algorithm parameters are optimized according to newly added data dimensions, such as user social interaction data. For example, the weight of social sharing behavior is increased when calculating user similarity, making the group results more in line with users' current needs and providing a precise basis for marketing strategy adjustments.

6 Conclusion

This study focuses on e-commerce precision marketing methods based on clustering algorithms and forms a complete research system through theoretical analysis, model design, experimental verification, and strategy proposal. At the theoretical level, it systematically sorts out the core theories of clustering algorithms and key theories of e-commerce precision marketing, clarifies the applicability of the K-means algorithm in e-commerce user grouping, and solves the problems of random initial cluster centers and difficult determination of cluster number in the traditional K-means algorithm through parameter optimization, providing theoretical support for subsequent model construction. At the model design level, a complete process from user data source and preprocessing to clustering algorithm selection and parameter optimization, and then to user clustering model construction is built, standardizing data processing steps and algorithm execution standards, ensuring the operability and reproducibility of the model.

References:

[1] Chaudhry M., Shafi I., Mahnoor M, . A systematic literature review on identifying patterns using unsupervised clustering algorithms: A data mining perspective. *Symmetry,* Vol.15, No.9, 2023, pp. 1679, 1-44. DOI: 10.3390/sym15091679

[2] Ke Jia. Discovering e-commerce user groups from online comments: An emotional correlation analysis-based clustering method. *Computers and Electrical Engineering,* No.113, 2024, pp. 109035-109042. DOI: 10.1016/j.compeleceng.2023.109035

[3] Saluja, Anjali, . Precision marketing strategy for E-Commerce by using big data technology. *Journal of Informatics Education and Research,* Vol.3, No.2, 2023, pp. 2554-2560. DOI:10.52783/jier.v3i2.433

[4] Abtahi, A., . Exploring consumer preferences: The significance of personalization in e-commerce. *Malaysian E Commerce Journal,* Vol.8, No.1, 2023, pp. 01-07. DOI: 10.26480/mecj.01.2024.01.07

[5] Rofi'i, Yulianto Umar. Analysis of e-commerce purchase patterns using big data: An integrative approach to understanding consumer behavior. *International Journal Software Engineering and Computer Science,* Vol.3, No.3, 2023, *pp.* 352-364. DOI: 10.35870/ijsecs.v3i3.1840.

[6] Kumar, Sumit, . Customer segmentation in e-commerce: K-means vs hierarchical clustering. *TELKOMNIKA (Telecommunication Computing Electronics and Control),* Vol.23, No.1, 2025, *pp.* 119-128. DOI: 10.12928/telkomnika.v23i1.26384

[7] Wenz, Viola, Arno Kesper, and Gabriele Taentzer. Clustering heterogeneous data values for data quality analysis. *ACM Journal of Data and Information Quality,* Vol.15, No.3, 2025, pp. 1-33. DOI: 10.1145/3603710

[8] Karrar, Abdelrahman Elsharif. The effect of using data pre-processing by imputations in handling missing values. *Indonesian Journal of Electrical Engineering and Informatics,* Vol.10, No.2, 2022, pp. 375-384. DOI: 10.52549/ijeei.v10i2.3730

[9] Wang Xinyi, Liao Xiaoling. Research on accurate marketing of e-commerce based on k-means algorithm. *Procedia Computer Science*, Vol.261, 2025, pp. 1208-1214. DOI: 10.1016/j.procs.2025.04.706

[10] Hu Chun, Wu Chuanjian, Huang Zongjie. Research on precision marketing of e-commerce enterprise based on cluster analysis in the big data environment. *Procedia Computer Science*, Vol.247, 2024, pp. 403-411. DOI: 10.34028/iajit/22/5/5.

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

Pengfei Bai: writing-original draft preparation, writing-review and editing, formal analysis.

Sources of Funding for Research Presented in a Scientific Article or Scientific Article Itself

No funding was received for conducting this study.

Conflict of Interest

The author has no conflicts of interest to declare that are relevant to the content of this article.

APPENDIX

Table 1 Study Text

Algorithm	Application Scenario	Computation Time(s)	SSE	Silhouette Coefficient	DBI	Clustering Stability
Optimized K-means	Large-scale dense user data with clear cluster structures	186	28954	0.76	0.32	0.91
DBSCAN	Sparse user data with irregular cluster distributions	423	N/A	0.58	0.65	0.73
Hierarchical Clustering Wards method	Small-to-medium scale data requiring hierarchical relationships	357	32681	0.64	0.48	0.82
GMM	Data with probabilistic distribution characteristics	298	30127	0.71	0.38	0.86

Source: created by the author

Table 2 User Group Feature Data

Group Number	Number of Users	Proportion of Total Users	Average Purchase Frequency	Average Customer Unit Price	Browse-Purchase Conversion Rate	Weekly Average Search Times	Proportion of Preferred Product Categories
1	126500	14.21	8.6	589	28.3	12.5	3C Digital 35, Home Furnishing 28, Apparel 18
2	293700	33.00	4.2	325	15.6	8.3	Apparel 42, Beauty 25, Food 18
3	311500	35.00	2.1	198	8.2	5.1	Food 38, Daily Necessities 32, Apparel 15
4	158300	17.79	1.3	96	3.5	2.8	Daily Necessities 45, Food 30, Low-Price Apparel 12

Source: created by the author

Table 3 Study Text

Group Number	Silhouette Coefficient	DBI	Control Group Conversion Rate	Experimental Group Conversion Rate	Conversion Rate Improvement Rate	Conversion Rate P Value	Control Group Repurchase Rate	Experimental Group Repurchase Rate	Repurchase Rate Improvement Rate	Repurchase Rate P Value
1	0.82	0.29	8.5	18.2	114.12	0.003	22.6	41.8	84.96	0.001
2	0.78	0.33	5.2	13.7	163.46	0.002	14.8	32.5	119.59	0.002
3	0.75	0.35	3.6	10.1	180.56	0.001	10.3	25.7	150.48	0.001
4	0.71	0.38	2.1	5.8	176.19	0.004	7.2	16.3	126.39	0.003

Source: created by the author

Table 4 Precision Marketing Model Effect Verification Data

Verification Indicator	Control Group	Experimental Group	Improvement Rate
Marketing Conversion Rate	4.8	12.3	156.25
Customer Unit Price	215	348	61.86
Repurchase Rate	12.5	28.7	129.6
Marketing Cost	128	89	-30.47

Source: created by the author

Table 5 Study Text

Verification Indicator	Control Group Mean	Control Group Standard Deviation	Control Group 95 Confidence Interval	Experimental Group Mean	Experimental Group Standard Deviation	Experimental Group 95 Confidence Interval	Improvement Rate	P Value	Bootstrap Robustness Score
Marketing Conversion Rate	4.8	0.72	3.39 6.21	12.3	1.15	10.04 14.56	156.25	<0.001	0.94
Customer Unit Price	215	23.6	168.7 261.3	348	31.2	286.9 409.1	61.86	<0.001	0.92
Repurchase Rate	12.5	1.83	8.91 16.09	28.7	2.47	23.85 33.55	129.6	<0.001	0.95
Marketing Cost	128	15.4	97.8 158.2	89	10.3	68.8 109.2	-30.47	<0.001	0.93

Source: created by the author