International Journal of Computer Information Systems and Industrial Management Applications
ISSN 2150-7988 Volume 18 (2026) pp. 1-14
© Cerebration Science Publishing

https://doi.org/10.70917/ijcisim-2026-0115
Article

A Cloud Computing-Based Framework for Enterprise Marketing Data Analysis and Customer Behavior Prediction in the Digital Economy Era

Ruimei Wang *

Gannan University of Science and Technology, Ganzhou 341000, Jiangxi, China; ruimeiwang1981@163.com

Abstract: This paper first establishes a data collection and quality control mechanism, combining statistical analysis tools and data mining techniques to enhance data value. It defines the boundaries of customer behavior, introduces survival analysis theory, and characterizes the temporal distribution patterns of customer churn. Based on the Markov chain model, it designs a transition probability matrix and prediction framework for customer behavior to capture the dynamic characteristics of state transitions. Using three years of micro-store order data from a certain snack chain brand as a sample, the consumption characteristics of 1,454 high-frequency customers are extracted. Survival analysis is combined to validate the negative correlation between customer system association and survival time, and Sharpley's additional explanatory value is used to analyze the key influencing factors of customer ordering and order cancellation behavior. In the survival experiment, the higher the association degree among units within the customer system, the worse the customer's survival status. Customers with significant fluctuations in consumption amounts exhibit stronger survival resilience, with the longest survival time reaching 23,731 hours. Feature A9 has the most significant positive impact on customer ordering intent, while feature B8 has the most significant negative impact on customer order cancellation intent, making them key variables for precise intervention in corporate marketing decisions.

Keywords: marketing decisions; survival analysis; Markov chain model; sharpley incremental value; behavior prediction

1. Introduction

The digital economy has not only reshaped the global economic structure and triggered significant changes in both domestic and overseas markets, but has also disrupted traditional business models [1-2]. Faced with significant changes in both internal and external marketing environments, if businesses remain stagnant and cling to traditional marketing strategies, they will experience a continuous decline in market competitiveness and may even suffer severe losses [3-5]. Under such circumstances, large, medium, and small enterprises at different stages of development and of varying scales have come to a consensus that leveraging digitalization to drive production and development has become a necessity, whether during crises or in normal times [6-7]. The deep integration of technology and marketing strategies not only enhances a company's sensitivity to changes in the external marketing environment but also improves its ability to connect with customers, its marketing planning and execution capabilities, and its ability to apply marketing mix strategies [8-10]. Therefore, against the backdrop of the digital economy, systematically exploring new models of corporate marketing strategy will drive enterprises to gradually adapt their marketing business processes and organizational structures, thereby facilitating the transition of corporate marketing strategies from “Marketing 4.0” to “Marketing 5.0” [11-14].

The rapid development of internet big data and cloud computing has established a vast platform for online shopping and social interactions, while innovations in marketing software are transforming

Copyright: © 2026 by the authors

people's shopping habits and lifestyles [15-16]. By leveraging big data and cloud computing technologies to process and store vast amounts of marketing information data in corporate information repositories, businesses can gain technical support for optimizing marketing strategies and selecting the most optimal marketing models [17-19]. Companies can utilize these technologies to understand market trends, further refine marketing approaches, expand sales channels, identify valuable information content, and enhance their core competitive strengths [20-22]. Therefore, companies should actively respond to this significant change, formulate online marketing strategies, and comprehensively innovate marketing models that organically integrate online and offline marketing to create greater commercial value for the company [23-24].

The big data era is an era of innovation, with various marketing strategies emerging continuously. The advent of cloud computing technology enables companies to broadly learn from various marketing strategies while predicting customer consumption behavior, leveraging strengths, addressing weaknesses, and optimizing marketing strategies based on industry characteristics and their own needs. Literature [25] establishes a comprehensive cloud computing data processing workflow tailored for enterprise marketing and operational development, significantly enhancing the enterprise's discovery efficiency and employee productivity, holding significant implications for achieving long-term enterprise development. Literature [26] analyzes the impact of digital technologies, including cloud computing, on enterprise marketing systems, finding that the introduction of digital technologies effectively enhances the potential for enterprise efficiency, scalability, and competitiveness, providing reference for future enterprise marketing integration strategies. Literature [27] indicates that cloud computing technology provides essential support for enterprises to create personalized marketing strategies and deepen connections with customers. It enables scalable solutions to improve operational efficiency and facilitate real-time data analysis, offering wise decision-making for optimizing enterprise marketing efforts. Literature [28] elucidates the connection between corporate marketing and customer behavior and proposes an IoT cloud computing service platform based on intelligent agents. By predicting customer demand for marketing products, this platform effectively improves corporate supply chain management and marketing strategies. Literature [29] explores the mechanisms through which cloud computing systems influence enterprise data management and marketing practices, finding that cloud-based social daily marketing significantly impacts customer consumption behavior. Therefore, leveraging cloud computing systems to advance the implementation of personalized marketing strategies can maximize the effectiveness of enterprise marketing efforts.

This paper explains the concept and importance of data-driven decision-making and proposes methods for data collection and analysis. It defines the boundaries of customer behavior and introduces the theory and methods of survival analysis. Combining the Markov chain model, it designs a predictive framework for customer behavior state transitions. Using data from the WeChat store of a certain leisure snack chain brand as a sample, it explores the impact of the degree of association between various units within the customer system on customer survival status. It sets two scenarios—customer orders and order cancellations—to identify the key factors influencing customer behavior.

2. Data-Driven Enterprise Marketing Data Analysis and Customer Behavior Prediction Methods

In the digital economy era, the corporate marketing environment is characterized by massive amounts of data, dynamic user behavior, and real-time decision-making. Traditional experience-driven marketing models are no longer able to meet market demands. How to efficiently extract the value of marketing data and accurately predict customer behavior has become a core issue for businesses seeking to enhance their competitiveness and operational efficiency. While existing research has focused on marketing data analysis, it still has shortcomings in areas such as data quality control, dynamic modeling of customer behavior, and long-term trend prediction: on one hand, standardized methods for comprehensive data collection and quality control mechanisms have yet to be established; on the other hand, customer churn prediction often relies on static models, which have limited ability to dynamically capture behavioral transition patterns. To address these issues, this paper proposes a cloud-based enterprise marketing data analysis and customer behavior prediction framework that integrates data engineering, survival analysis, and Markov chain theory to systematically explore the entire process from data collection to behavior prediction.

2.1. Data Collection and Analysis Methods

2.1.1. Data Collection Methods

Data collection is a critical foundation for marketing decisions. To obtain comprehensive and accurate data, companies need to adopt different data collection methods. First, clarify the sources and collection methods of data. Data sources include internal and external data. Internal data includes a company's sales data, customer data, inventory data, etc., which can be collected through internal systems. External data includes market research, competitor analysis, industry trend analysis, etc., which can be collected through online surveys, purchasing third-party databases, and other methods.

During the data collection process, companies must focus on data quality control. Data quality directly impacts the accuracy and effectiveness of marketing decisions. Therefore, companies must take measures to ensure data quality. On one hand, companies can establish data cleaning and filtering mechanisms to exclude outliers, missing values, duplicate values, and other poor-quality data. On the other hand, companies can adopt standardized and normalized data processing methods to improve data accuracy and readability. In addition to basic sources, collection methods, and quality control, there are other methods and techniques for data collection. For example, companies can obtain more comprehensive market trends and consumer demand through long-term, continuous data monitoring and collection, and expand the scope and efficiency of data collection through digital channels such as social media and online surveys.

2.1.2. Data Analysis Methods

Data analysis is one of the key steps in marketing decision-making. Through data analysis, one can gain a deep understanding of market trends, consumer needs, and industry development trends, providing scientific and precise basis for marketing decisions. First, statistical analysis tools and techniques form the foundation of data analysis. These tools include Excel, SPSS, SAS, R, etc., which can be used to organize, clean, analyze, and visualize collected data. Through statistical analysis tools, data can be categorized, summarized, analyzed for correlations, and subjected to regression analysis to extract valuable information. Secondly, data mining and machine learning methods are also important tools in data analysis. Data mining is a process of automatically discovering useful information hidden within large datasets, including techniques such as clustering analysis, association rules, and time series analysis. Machine learning is an artificial intelligence method that improves the accuracy and efficiency of data analysis by training models. These methods can help businesses uncover patterns and trends within data, providing more precise recommendations for marketing decisions. Different data analysis methods and tools have varying scopes of application and usage conditions. Businesses should select the most appropriate methods and tools based on actual circumstances such as data sources, data quality, and data processing requirements. Additionally, data analysis must be combined with real-world business scenarios and market demands for in-depth exploration and analysis, providing more comprehensive and accurate data support for marketing decisions.

2.2. Survival Analysis of Customer Churn

2.2.1. Definition of Customer Behavior

Considering the research objectives and practical application needs of this paper, based on the above understanding of relevant concepts of customer behavior, the customer behavior studied in this paper in the context of mobile commerce is defined as transaction behavior and customer loyalty based on the buying and selling relationship between customers and enterprises. The specific customer behavior definition process is shown in Figure 1.

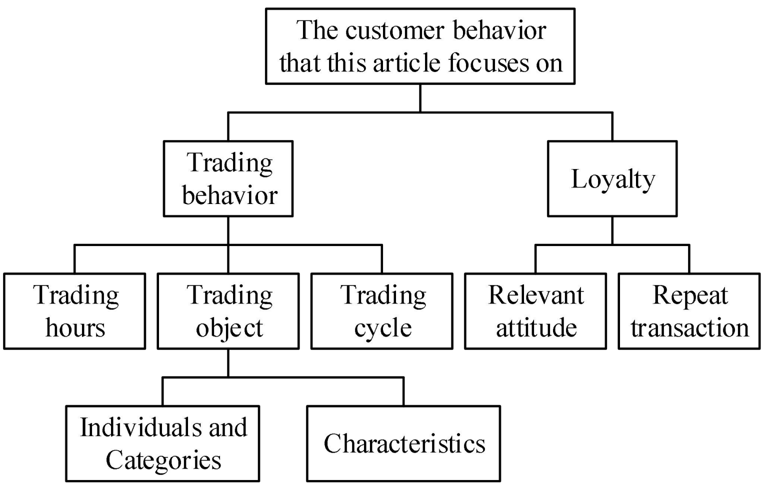

Figure 1. Definition of the customer behavior focused on in this article.

Customer loyalty encompasses two aspects: repeat transaction behavior and customer-related attitudes. Customer-related attitudes can be simply understood as customer satisfaction and brand loyalty.

Customer transaction behavior includes three aspects: transaction time, transaction counterpart, and transaction cycle.

Generally, companies can analyze customer transaction time from different dimensions (such as monthly, weekly, daily, or hourly) based on different business needs. For example, analyzing the days of the week or specific times of the day when individual customers or customer groups most frequently engage in transaction behavior. Such analysis results can provide a basis for the reasonable allocation of business resources.

In comparison, the transaction cycle dimension is relatively more complex, encompassing the cycle in which customer transaction behavior occurs, the transaction cycle for similar or identical products or product categories, and the patterns of change in transaction cycles. For example, when parents purchase school supplies for their children, some parents may choose to buy them whenever they are needed, while others may opt to concentrate their purchases on weekends or days off due to work commitments. Thus, the transaction cycles of these two groups of parents will differ. Additionally, influenced by school holiday schedules, the purchasing cycle for school supplies typically increases before and after holidays compared to the periods around the start of the school term. Such changes in transaction cycles also require attention, and analyzing transaction cycles can help businesses promptly track customer transaction trends.

Transaction targets can be further categorized into preferences for individual transaction targets and categories, as well as preferences for the characteristics of transaction targets. Individual and category preferences refer to customers' preferences for specific products and product categories during the transaction process. The characteristics of transaction targets refer to the special status of products, such as regular-priced items, promotional items, new products, etc. Analyzing transaction objects can help businesses understand customers' product preferences, sensitivity to promotions, and acceptance of products, among other factors.

2.2.2. Related Functions of Survival Analysis

When conducting research using survival analysis methods, three functions are generally used to characterize the distribution of survival time: the survival function $S(t)$, the probability density function $f(t)$, and the hazard function $h(t)$. In survival analysis, the three functions are estimated through the analysis of relevant survival data from the observed sample, and the survival model of the sample population is then derived.

(1) Survival function

The survival function, also known as the cumulative survival function, is denoted by $S(t)$ and represents the probability that the survival time of the study subjects exceeds time t. It can be expressed by the following formula:

$$\begin{aligned} S(t) &= P(\text{The individual survival time is greater than or equal to } t) \\ &= P(T \geq t) = 1 - F(t) \end{aligned} \tag{1}$$

Among them, $F(t)$ refers to the distribution function of the survival time T of the research object, and the survival probability of the research object is represented by $S(t)$, which refers to the probability that the research object has not experienced a specific event after t time. When only complete data exists in the observation sample, the survival probability can be expressed as:

$$\begin{aligned} S(t) &= P(T \geq t) \\ &= \frac{\text{The number of samples still alive after time } t}{\text{The total number of samples at the beginning of the observation}} \end{aligned} \tag{2}$$

In this case, t represents the current survival time of the research subject. However, when the observation sample contains deleted data, the survival probability needs to be expressed using the following formula:

$$S(t_k) = P(T \geq t_k) = p_1 \times p_2 \times \cdots \times p_k \tag{3}$$

In this context, $p_1, p_2, \ldots, p_k$ represent the survival probabilities of the study subjects at different time intervals. As can be seen from the above formula, the survival probability of the study subjects is the cumulative sum of the survival probabilities at multiple time intervals. Therefore, the survival probability obtained in this scenario is referred to as the cumulative survival probability.

Correspondingly, the graph of the survival function $S(t)$ is called the survival curve. The trend of the survival curve reflects the values of survival time and survival probability. A smooth survival curve indicates longer survival time or higher survival probability; conversely, a steep curve indicates shorter survival time or lower survival probability. Typically, survival functions are used to compare and analyze multiple survival patterns.

(2) Probability density function

The probability density function, also known as the density function, is denoted by $f(t)$. Its corresponding graph is the density curve. The proportion and maximum probability of a specific event occurring within any observation time interval can be determined from this density curve graph. Its corresponding function expression is:

$$\begin{aligned} f(t) &= \lim_{\Delta t \to 0} \frac{P(\text{An individual has A death event in the interval } (t, t+\Delta t))}{\Delta t} \\ &= \lim_{\Delta t \to 0} \frac{P(t < T \le t + \Delta t)}{\Delta t} \end{aligned} \tag{4}$$

The survival function can be regarded as the integral of the probability density function, expressed as

$$S(t_k) = P(T \ge t_k) = \int_t^{\infty} f(t)dt \tag{5}$$

Furthermore, the probability density function can also be expressed using the following formula:

$$f(t) = -\frac{dS(t)}{dt} \tag{6}$$

(3) Hazard function

The hazard function, also known as the risk function or hazard probability, is denoted by $h(t)$. It is the most basic function in survival analysis and is typically used to represent the probability of a specific event occurring in a unit of time after a specific point in time, given that the observed object is in a state of survival at that point in time. It is a type of conditional probability and can be expressed by the following formula:

$$\begin{aligned} h(t) = \lim_{\Delta \to 0} \frac{1}{\Delta t} P(&\text{An individual whose existence time is } t \\ &\text{experiences life and death events within } (t, t+\Delta t)) \end{aligned} \tag{7}$$

Therefore, the conditional probability of a specific event occurring within the observation period $(t, t+\Delta t)$ is represented by $h(t)\Delta t = P(t < T < t + \Delta t \mid T \ge t)$. This function can be used to study whether the subject will experience a specific event at a specific point in time when the survival time reaches a certain value.

(4) The relationship between the three functions

Mathematically, the above three functions can be converted into one another. Typically, knowing one of the functions allows the other two to be calculated using their mathematical relationships. The formulas above describe the mathematical relationship between the survival function $S(t)$ and the probability density function $f(t)$. The relationship between the hazard function $h(t)$ and the other two functions can be expressed by the following formula:

$$h(t) = \frac{f(t)}{S(t)} = \frac{f(t)}{1 - F(t)} = -\frac{d \ln S(t)}{dt} \tag{8}$$

It can also be written as:

$$S(t) = \exp\left[-\int_0^1 h(t)dt\right] \tag{9}$$

2.3. Markov Chain-Based Customer Behavior Prediction Model

2.3.1. Markov Chains, Transition Probabilities

Definition 1: Let $\{X(t), t \in T\}$ be a stochastic process, where the parameter set is $T = \{t_1, t_2, \cdots, t_n\}$. For any n parameters $t_1 < t_2 < \cdots < t_n$ in T, and such that:

$$P\{x(t_1) = i_1, x(t_2) = i_2, \cdots, x(t_{n-1}) = i_{n-1}\} > 0 \tag{10}$$

Established in I, either state $i_1, i_2, \cdots, i_{n-1}$ or i_n is present.

$$\begin{aligned} &P\{x(t_n) = i_n \mid x(t_1) = i_1, x(t_2) = i_2, \cdots, x(t_{n-1}) = i_{n-1}\} \\ &= P\{x(t_n) = i_n \mid x(t_{n-1}) = i_{n-1}\} \end{aligned} \tag{11}$$

Then, $\{X(t), t \in T\}$ is called a Markov chain.

Markov chains can be divided into discrete-parameter Markov chains and continuous-parameter Markov chains. Considering the characteristics of customer browsing behavior, this paper only considers discrete-parameter Markov chains.

Definition 2: Let I be the state space of the discrete-parameter Markov chain $\{X(t), t \geq 0\}$. The conditional probability is defined as the k-step transition probability of $\{X(t), t \geq 0\}$.

$$p_{ij}(m,h) = P\{X(m+k) = j \mid X(m) = i\}, i, j \in I, m \geq 0, k \geq 0 \tag{12}$$

The transition probability represents the probability that a process known to be in state i at time m will be in state j after k time units.

The transition probability has the following properties:

(1) Each element p_{ij} in the matrix is non-negative, i.e., $p_{ij} \geq 0$

(2) The sum of the elements in each row of the matrix is 1, i.e., $\sum_{j=1}^{n} p_{ij} = 1$

Definition 3: $\{X(t), t \geq 0\}$ is a Markov chain if, for any $i, j \in I$, the following always holds:

$$P\{X(n+1) = j \mid X(n) = i\} = P\{X(m+1) = j \mid X(m) = i\} \tag{13}$$

Then, $\{X(t), t \geq 0\}$ is called a homogeneous Markov chain, whose transition probabilities do not depend on the initial time. At this point, the k-step transition probability can be denoted as $P_i(k)$. If the state space I is a finite set $I = \{0, 1, 2, \cdots, j\}$, the corresponding k-step transition matrix is denoted as

$$\begin{aligned} P(k) &= \begin{bmatrix} p_{00}(k) & p_{01}(k) & \cdots & p_{0j}(k) \\ p_{10}(k) & p_{11}(k) & \cdots & p_{1j}(k) \\ \cdots & \cdots & \cdots & \cdots \\ p_{j0}(k) & p_{j1}(k) & \cdots & p_{ji}(k) \end{bmatrix} \\ &= \left[\sum_{r=0}^{\infty} p_r(k-1) p_{rj} \right] = P(k-1) \times P = p^{(k-1)} p = p^k \end{aligned} \tag{14}$$

When $k = 1$, it is called a one-step transition probability matrix, denoted as P.

2.3.2. Analysis and Design of Customer Behavior Prediction Models

The Markov chain prediction method assumes that the next visited state is only related to the current state and is unrelated to any previous states. This assumption has certain limitations in practical applications. However, the focus of this paper is to establish a relatively simple model to find patterns in customer visit behavior. Therefore, although the actual customer visit process is not a Markov chain without aftereffects, such a model can still provide an effective and simple approximate framework for analyzing this state transition process.

Suppose we analyze n representative products, each corresponding to states $1,2,\cdots,n$. Additionally, we introduce an end state E. Here, the end state E is a virtual state, and cookie file information is used to divide customer visit records into independent sequences, with state E serving as the dividing point between sequences. If a customer remains on a webpage for 20 minutes without any clicks, the visit is considered ended, thereby dividing the customer's browsing behavior into several independent visit sequences.

First, establish the transition probability matrix for customer visit interests, where D represents the customer's visit log data over a period of time, and the customer's visit behavior is a discrete-parameter stochastic process. This paper assumes that customer visits approximately satisfy Markovian and homogeneous properties. Here, we select n items to analyze customer visit behavior, where $p_{ij}(i,j=1,2,\cdots,n)$ represents the probability of a customer transitioning from item i to item j. The specific calculation process for p_{ij} is as follows: analyze all customer behaviors during the selected period where customers transition from state i to other states, denote the number of times such transitions occur as n_{ij}, and derive:

$$p_{ij}=\frac{n_{ij}}{\sum_{j=1}^{n}n_{ij}} \tag{15}$$

From this, we can establish a probability matrix P for the shift in customer interest during that period.

$$P=\begin{bmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \cdots & \cdots & \cdots & \cdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{bmatrix} \tag{16}$$

Then, determine the initial state. Since customer access behavior has n mutually incompatible states, its initial distribution is

$$I^{(0)}=(i_1^{(0)},i_2^{(0)},\cdots,i_n^{(0)}) \tag{17}$$

In the equation, $i_t^{(0)}(t=1,2,\cdots,n)$ represents the probability of visiting product t at time 0.

Finally, establish a Markov prediction model, as shown below.

$$\begin{aligned} &I^{(1)}=I^{(0)}\cdot P \\ &I^{(2)}=I^{(1)}\cdot P=I^{(0)}\cdot P^2 \\ &\cdots\cdots \\ &I^{(k+1)}=I^{(k)}\cdot P=I^{(0)}\cdot P^{(k+1)} \end{aligned} \tag{18}$$

3. Empirical Research on Enterprise Marketing Data Analysis and Customer Behavior Prediction

3.1. Data Collection and Preprocessing

This article exports all completed orders from January 1, 2022, to January 1, 2025, from the backend system of a certain leisure snack chain brand's WeChat store, totaling 12,957 orders involving 7,521 customers. Each order record includes the customer's order time, payment time, delivery address, recipient's name, contact phone number, product price, order notes, product refund amount, product shipping time, and transaction completion time.

During the preliminary preprocessing phase, the following data cleaning and transformation steps were performed:

(1) Changes in customer phone numbers during the observation period were not considered. Assuming no customer changed their phone number during this period, the contact phone number was used as the primary key to uniquely identify the customer ID. If a customer continued to make purchases after changing their phone number, they were treated as a new customer.

(2) To exclude suspected fraudulent flash sales, orders with a unit price below 10 yuan are considered invalid and deleted, such as orders with a shipping fee of 0.01 yuan. To exclude suspected order padding, orders for which the order number cannot be queried and are marked as “no logistics required” in the system are considered invalid and deleted.

(3) Closed orders are not considered. Closed orders refer to orders where the customer initiates a full refund before the seller ships the goods. Although an order number is generated, the transaction is not successful, and the customer has not purchased any goods, so it does not reflect the customer's purchasing behavior. Therefore, such orders are deleted. In contrast, if the customer receives the goods but is dissatisfied and decides to return them in full, such orders are considered valid completed orders and should be retained, with the refund status recorded.

(4) Data analysis revealed that there were virtually no instances of the same customer changing their delivery province within the 36-month period. Therefore, it can be assumed that each customer's delivery province is unique. Since the collected Weidian data primarily consists of customers within Province A as the largest group, to balance customer distribution, this study categorizes customer locations based on delivery addresses into “within the province” and “outside the province.”

(5) Since the study requires determining customer churn based on purchase intervals, and two purchases form a time interval, this paper focuses on customers who have made at least two purchases, excluding those who have only made one purchase to date. After data preprocessing, this study extracted 1,454 customers who had made at least two purchases, including their consumption characteristics from the first purchase to the end of the observation period, such as the number of purchases, the first purchase time, and the most recent purchase time.

The results of extracting some customer consumption characteristics are shown in Table 1. The variables in the table are used for analyzing the factors influencing customer churn risk. The customer ID is the anonymized form of the customer's contact phone number, and the definitions and explanations of the remaining variables are as follows:

(1) Total consumption amount. This refers to the total order amount spent by the customer on Weidian up to the end of the observation period. It is measured in yuan.

(2) Average amount spent per purchase. This is the ratio of the total consumption amount to the number of purchases. It is measured in yuan.

(3) Standard deviation of consumption amount. This is the square root of the arithmetic mean of the squared deviations between each customer's consumption amount and the average amount spent per purchase, reflecting the variability in each customer's consumption amount.

(4) Average purchase interval. Each interval is formed by two consecutive purchases. For n purchases, there are n-1 intervals. The average of these n-1 intervals is the average purchase interval. If a customer has only made two purchases, the average purchase interval is equal to the single purchase interval itself. Measured in days.

(5) Average shipping speed. The shipping speed for each purchase is calculated as the shipping time minus the order payment time. The average of the n shipping speeds experienced by the customer over n purchases is the average shipping speed, which measures the logistics experience provided by the merchant to the customer. Measured in hours.

(6) Slowest shipping speed. This refers to the maximum value of the n shipping speeds experienced by the customer over n purchases, measuring the worst logistics experience provided by the merchant to the customer. Measured in hours.

Table 1. Extraction results of some customer consumption characteristics.

Customer id	Total consumption amount	The average amount spent each time	Standard deviation of consumption amount	Average purchase interval	Average delivery speed	The slowest delivery speed
95	302.42	100.81	25.38	12.09	10.83	21.57
18	419.58	139.86	8.02	61.38	9.05	16.38
46	297.12	148.56	17.51	25.45	12.26	18.05
72	964.28	192.86	9.05	30.06	6.04	16.77
83	309.55	103.18	22.48	19.42	8.71	19.38
29	426.71	142.24	11.37	21.55	9.48	22.04
66	833.26	119.04	19.02	17.73	9.27	18.36
05	607.58	101.26	5.86	29.44	9.15	25.27

3.2. Customer Survival Analysis

If there is a large-scale collapse of customer units within the customer system, it can be considered that the entire customer system has collapsed (failed). The number of times unsatisfactory values are added between two collapses can be used as the time, and this time can be regarded as the survival time of the customer system. If the degree of connectivity between units within the customer system is analogized as a "treatment plan," survival analysis can be used to explore the survival time of the customer system under different degrees of connectivity between units within the system. This further analyzes the relationship between the evolution of the customer system and the degree of connectivity between units within the system.

To investigate the self-organizing evolutionary behavior of the customer system under different conditions, six experimental schemes were designed to obtain relevant data. The difference between the six experimental schemes lies in the varying degrees of interconnection among the units within the customer system, ranging from low to high. The interconnection degrees in the first five schemes are 1 to 5. In the sixth scheme, the interconnection degree is set as a random integer between 1 and 5. For each experimental scheme, 20,000 unsatisfactory values were randomly added. In this study, a customer system was considered to have collapsed (died) when more than 50% of its internal units failed. The survival time of the system was defined as the number of times unsatisfactory values were added from the start until the system collapsed. The data on the number of system collapses and survival times for each experimental scheme are shown in Table 2. Direct observation of the data reveals significant differences in the survival times of the customer systems across the experimental schemes. For example, in Scheme 1, where the degree of interconnection is relatively low, only 5 system failures occurred, with an average survival time of 26,240 hours. The average survival times for Schemes 2, 3, 4, 5, and 6 were 12,248 hours, 385 hours, 146 hours, 68 hours, and 124 hours, respectively. The survival time of the system decreases rapidly as the degree of interconnection between units within the customer system increases. When the degree of interconnection among units within the client system is a random number between 1 and 5, its survival time is greater than that of Scheme 5 but less than that of Scheme 4.

Table 2. Customer System Lifetime.

Plan	Number of crashes	The maximum survival time of the system/h	The shortest survival time of the system/h	Average survival time of the system/h
1	5	26240	9048	16836
2	12	13752	6155	12248
3	196	903	2	385
4	1074	561	2	146
5	1364	255	2	68
6	602	2797	2	124

Using the data obtained from the experiments, the number of additions from the last system crash to the 20,000th addition of unsatisfactory values was set as the truncated data. First, the Kaplan-Meier process in survival analysis was used for analysis, yielding the average survival time and median survival time for each experimental scheme, as shown in Table 3. For Scheme 1, the average survival time and median survival time were 16,367.37 hours and 13,028.00 hours, respectively. As the degree of interconnection among units within the customer system increased, the average survival time and median survival time for Schemes 2–6 decreased rapidly, with a decline rate approximating an exponential pattern.

Table 3. Mean survival time and median survival time(h).

Plan	Average survival time				Median survival time			
	Estimated value	Standard deviation	95% confidence interval		Estimated value	Standard deviation	95% confidence interval	
			Lower bound	Upper bound			Lower bound	Upper bound
1	16367.37	6018.36	5022.97	26018.38	13028.00	2754.37	7001.55	18274.87
2	9039.57	1375.47	7003.38	13963.82	9013.00	1534.85	6538.27	13028.22
3	648.38	16.36	569.48	695.48	536.00	20.58	553.28	635.48
4	201.42	9.48	185.37	233.14	167.00	10.11	131.48	201.07
5	35.27	0.93	32.11	39.52	29.00	0.93	25.47	30.18
6	98.66	11.49	67.37	141.45	16.00	1.36	16.84	23.22

The customer system survival curve is shown in Figure 2. From the survival curve diagram, it can be seen that the six schemes can be clearly divided into two groups in terms of survival rate, namely, schemes 1 and 2 form one group, and schemes 3, 4, 5, and 6 form another group. Schemes 1 and 2 have better survival rates, while schemes 3, 4, 5, and 6 have poorer survival rates.

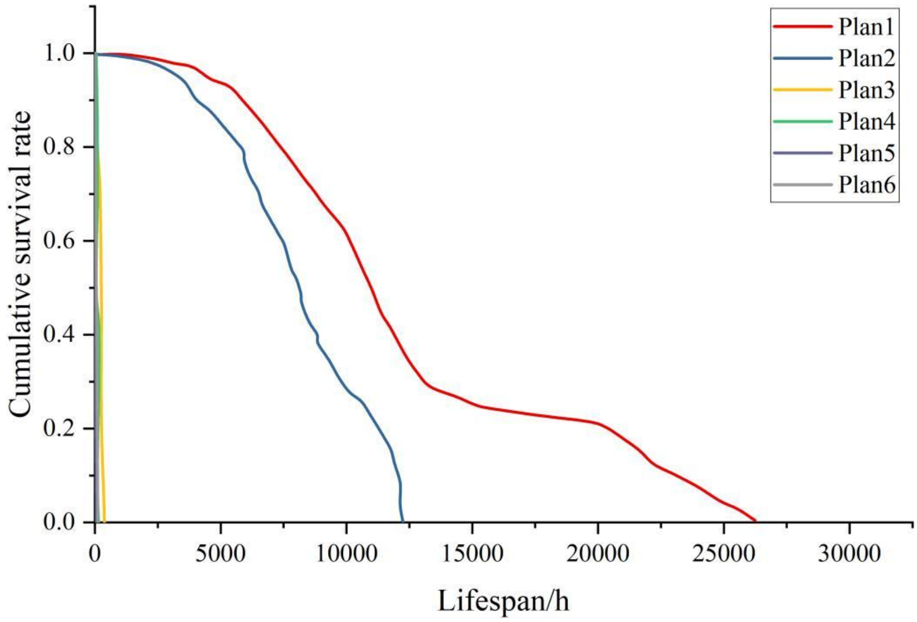

Figure 2. Survival curve of the customer system.

Similarly, this paper plots survival curves for fluctuations in consumption amounts. The comparison results of survival curves for different levels of consumption amount fluctuations are shown in Figure 3. It can be observed that there are significant differences in the survival curves of customers with varying degrees of consumption amount fluctuations. Customers with larger consumption amount fluctuations have higher survival rates, with the longest survival time reaching 23,731 hours. Upon reviewing the raw data, it was found that customers with larger consumption amount fluctuations had almost all participated in store promotional activities. Their larger consumption amount fluctuations were likely due to participating in a promotional activity during a purchase, and typically, customers' purchasing power during promotional activities is often greater than during non-promotional periods.

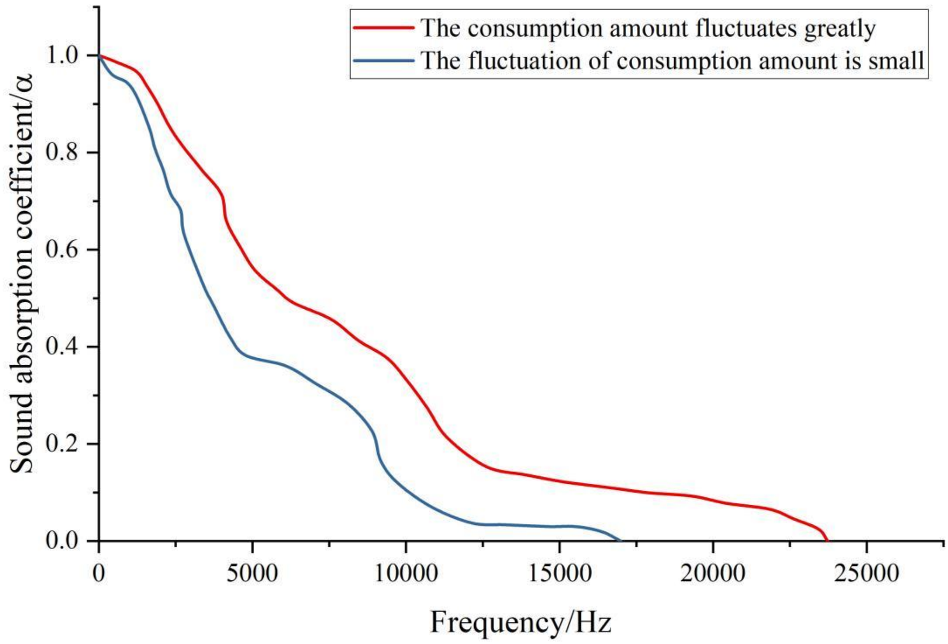

Figure 3. Comparison results of survival curves.

3.3. Predicting Customer Behavior

3.3.1. Customer Order Forecasting

In customer order forecasting, the explanatory power of the prediction model's results is equally significant for business decision-makers. Therefore, this section employs Sharpley's additional explanatory value to conduct an explanatory analysis of customer order forecasting results, investigating the influence of various features on whether customers engage in purchasing behavior. In the customer order scenario, feature values are numbered A1 to A15, and the Sharpley additional explanatory values for sample features are shown in Figure 4. It can be observed that feature A9 has the most significant impact on customer order intent and is positively correlated with the strength of customer order intent, while feature A12 has the weakest impact on customer order intent.

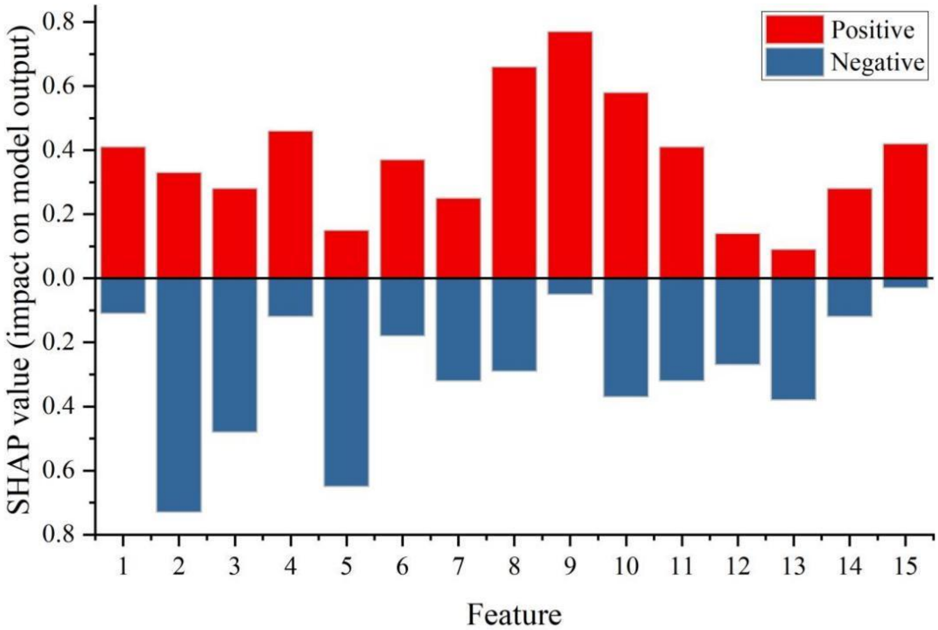

Figure 4. Additional Sharpley interpretation values of sample features.

3.3.2. Order Cancellation Prediction

In the scenario of customer order cancellation, the feature values are numbered B1 to B15, and the Sharpley additional explanatory values of the sample features are shown in Figure 5. It can be seen that feature B5 has the most significant positive impact on customer order cancellation intention, while feature B8 has the most significant negative impact on customer order cancellation intention.

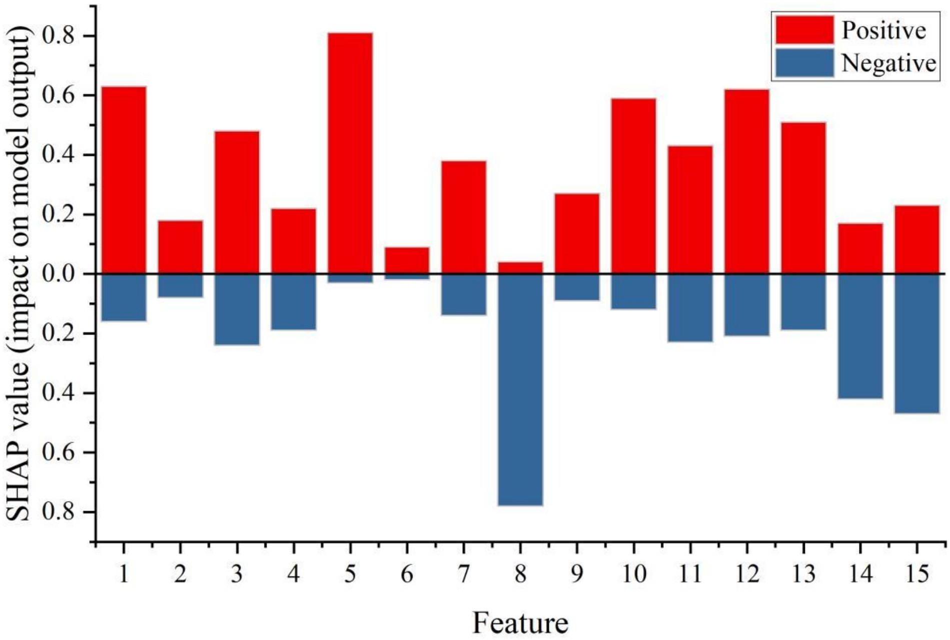

Figure 5. Additional Sharpley interpretation values of sample features.

4. Conclusion

This paper constructs a comprehensive framework for marketing decision-making that includes data collection, quality control, survival analysis, and Markov prediction, and verifies its effectiveness through empirical research.

The survival experiment results show that there are significant differences in the survival times of customer systems across different experimental schemes. In Scheme 1, which has a lower degree of closeness, only 5 system crashes occurred, with an average survival time of 26,240 hours. The average survival times for Schemes 2, 3, 4, 5, and 6 were 12,248 hours, 385 hours, 146 hours, 68 hours, and 124 hours, respectively. Customers with higher fluctuations in consumption amounts demonstrated stronger survival resilience, with the longest survival time reaching 23,731 hours. Sharpley's additional explanatory value analysis indicated that feature A9 had the most significant impact on customers' willingness to place orders and was positively correlated with the strength of their willingness to order. Feature B5 had the most significant positive impact on customers' willingness to cancel orders, while feature B8 had the most significant negative impact on customers' willingness to cancel orders.

Funding

This paper is supported by the Western Project of National Social Science Foundation of China, "Research on the Characteristics of Strategic Critical Metal Trade Network and the Improvement of Industrial International Competitiveness" (23XGL013).

References

1. Semenov, I. V. (2021). Marketing Strategies of the Company Development within the Transition to a Digital Economy. Socio-economic Systems: Paradigms for the Future, 1373-1383.
2. Pasholikov, M. (2023, January). Digital economy: a marketing approach. In 2nd International Conference on Computer Applications for Management and Sustainable Development of Production and Industry (CMSD-II-2022) (Vol. 12564, pp. 280-286). SPIE.
3. Miao, L. (2021). Influence of big data technology on enterprise marketing strategy. Ind Eng Innov Manag, 4(2), 55-60.
4. Tkachenko, V. (2017). Formation of marketing management system of the enterprise. Baltic Journal of Economic Studies, 3(5), 208-213.
5. Peng, L., & Duan, X. (2020, April). Marketing Strategy of Modern Enterprise under the Background of Big Data. In 2020 International Conference on Computer Information and Big Data Applications (CIBDA) (pp. 1-3). IEEE.

6. Menz, M., Kunisch, S., Birkinshaw, J., Collis, D. J., Foss, N. J., Hoskisson, R. E., & Prescott, J. E. (2021). Corporate strategy and the theory of the firm in the digital age. Journal of Management Studies, 58(7), 1695-1720.
7. Kobets, K., Terentieva, N., Shkvyria, N., Lysytsia, N., & Siemak, I. (2024). Digitalization and its impact on the development of contemporary marketing strategies. Economic Affairs, 69(2), 1021-1040.
8. BABUKH, I., AVANESYAN, N., POVNA, S., BILANENKO, O., SAKUN, V., & ABAKUMOV, D. (2025). STRATEGIC OBJECTIVES FOR INNOVATIVE DEVELOPMENT OF MARKETING AND E-COMMERCE AS PART OF DIGITALIZATION OF CORPORATE BUSINESS. Journal of Theoretical and Applied Information Technology, 103(7).
9. Li, L., & Zhang, J. (2021). Research and analysis of an enterprise E-commerce marketing system under the big data environment. Journal of Organizational and End User Computing (JOEUC), 33(6), 1-19.
10. Han, S., Li, C., Li, P., Zhao, M., & Liu, N. (2022, March). Enterprise Marketing Strategy Based on Data Mining Technology. In The International Conference on Cyber Security Intelligence and Analytics (pp. 353-358). Cham: Springer International Publishing.
11. Budrin, A. G., & Soldatova, A. V. (2018). Economic analysis of marketing activities of an enterprise. π-Economy, 11(3), 173-184.
12. Zhang, Z., & Xu, D. (2025). Enterprise marketing data mining method based on decision tree algorithm. Procedia Computer Science, 261, 1172-1178.
13. Li, N. (2025). Big data-driven enterprise management and market decision-making framework. Service Oriented Computing and Applications, 1-12.
14. Pavlenchyk, N., Horbonos, F., Pavlenchyk, A., Skrynkovskyy, R., & Pawlowski, G. (2021). Increasing the competitiveness of enterprises based on the use of marketing management tools. Agricultural and Resource Economics: International Scientific E-Journal, 7(3), 77-89.
15. Dong, X. (2025). The impact of cloud computing technology on enterprise network marketing and coping strategies. Journal of Computational Methods in Sciences and Engineering, 14727978251321329.
16. Yi, Y. (2020, October). Research on E-commerce Marketing and Enterprise Performance Based on Cloud Service Platform. In 2020 13th International Conference on Intelligent Computation Technology and Automation (ICICTA) (pp. 614-618). IEEE.
17. Ali, C. S. M., & Zeebaree, S. (2025). Cloud-Based Web Applications for Enterprise Systems: A Review of AI and Marketing Innovations. Asian Journal of Research in Computer Science, 18(4), 427-451.
18. Gao, Z., Zhao, Y., & Kong, C. (2023). E-Commerce Marketing and Enterprise Performance Based on Cloud Service Platform. Revista Ibérica de Sistemas e Tecnologias de Informação, (E62), 440-449.
19. Esteves, M., Piedade, M. B., & Pereira, A. (2018). A Cloud Computing Solution for Digital Marketing: A Case Study for Small and Medium-Sized Enterprises. In Interactive Mobile Communication Technologies and Learning: Proceedings of the 11th IMCL Conference (pp. 843-849). Springer International Publishing.
20. Zhang, G. (2024). Exploring the enterprise marketing strategy innovation in the era of big data. Academic Journal of Business & Management, 6(11), 152-157.
21. Avram, M. G. (2014). Advantages and challenges of adopting cloud computing from an enterprise perspective. Procedia Technology, 12, 529-534.
22. Abdullah, C. M., & Zeebaree, S. (2025). The Role of Machine Learning in Enhancing Marketing Strategies within Cloud-based Enterprise Systems. Asian Journal of Research in Computer Science, 18(5), 32-54.
23. Vasiljeva, T., Shaikhulina, S., & Kreslins, K. (2017). Cloud computing: Business perspectives, benefits and challenges for small and medium enterprises (case of Latvia). Procedia Engineering, 178, 443-451.
24. Attaran, M., & Woods, J. (2019). Cloud computing technology: improving small business performance using the Internet. Journal of Small Business & Entrepreneurship, 31(6), 495-519.
25. Jia, D., & Wu, Z. (2022). Development model of enterprise green marketing based on cloud computing. Wireless Communications and Mobile Computing, 2022(1), 4931374.
26. Saleh, R. A., & Zeebaree, S. R. (2025). Transforming enterprise systems with cloud, AI, and digital marketing. International Journal of Mathematics, Statistics, and Computer Science, 3, 324-337.
27. Khurshed, A. K., & Zeebaree, S. (2025). Web Technology and Cloud Computing in Enterprise System: The Role of AI for Digital Marketing. Asian Journal of Research in Computer Science, 18(5), 554-580.
28. Yang, W., Chen, Y., Chen, Y. C., & Yeh, K. C. (2021). Intelligent agent-based predict system with cloud computing for enterprise service platform in IoT environment. Ieee Access, 9, 11843-11871.
29. Subramaniyan, P., & Muthuraman, K. (2023). Effectiveness of Cloud Technologies in Social Media Marketing and its Impact on Customer Buying Behaviour. Virtual Economics, 6(4), 60-72.