Heliyon 9 (2023) e13384

Contents lists available at ScienceDirect

Heliyon

journal homepage: www.cell.com/heliyon

Research article

Research on customer lifetime value based on machine learning algorithms and customer relationship management analysis model

Yuechi Sun, Haiyan Liu *, Yu Gao

School of Economics and Management, China University of Geosciences (Beijing), No. 29 Xueyuan Road, Haidian District, Beijing 100083, China

ARTICLE INFO

Keywords:
Data mining
Machine learning
Customer lifetime value
Customer segmentation

ABSTRACT

Customer lifetime value is one of the most important tasks for enterprises to maintain customer relationships. However, due to the limitations of using a single data mining method, the measurement of customer lifetime value under the condition of noncontractual relationship has always been a research difficulty. This paper focuses on customer value measurement and customer segmentation based on customer lifecycle value theory, and carries out customer value measurement and customer segmentation research from the perspective of customer value, and constructs customer segmentation model. This paper first conducts feature engineering, such as data selection, data preprocessing, data transformation, and knowledge discovery, and then conducts customer value segmentation based on machine learning algorithms and customer relationship management analysis models and builds a customer value segmentation identification model under the condition of noncontractual relationship. Finally, empirical analysis is carried out with the real customer transaction data of the actual online shopping platform, which verifies the validity and applicability of the customer segmentation method and value calculation method proposed in this paper.

1. Introduction

Since the 1980s, customer relationship management (CRM) has increasingly become the focus of academia and industry. Facts have proven that maintaining good relationships with specific customers can increase corporate profits and increase the significant advantages of enterprises in market competition [1]. Enterprises can classify customer categories based on customer consumption behavior, customer lifetime value and other information, identify customer value, and apply it to CRM [2]. Scientific and effective research on customer lifetime value is the primary task of customer relationship management, which is of great significance to enterprise marketing plans and strategies. Scholars have confirmed that effective customer lifetime value segmentation research can bring more profits to enterprises [3]. However, due to the inadequacy of previous research on customer data processing and the limitation of single data mining methods, it is difficult to fully mine customer behavior information. The accurate measurement of customer life cycle value under non contractual relationships has always been a research difficulty. Through a literature review, the current calculation of customer lifetime value under noncontractual relationships mainly includes Pareto/NBD and other probability models [4,5], VAR, and other time series models, as well as machine learning algorithms [6]. Especially in recent years, with the wide application of Big Data in business intelligence and analytics, companies can collect a large amount of personal transaction data from customers at low cost. These can be analyzed using an array of methods including: neural network models, decision tree models,

* Corresponding author.
E-mail addresses: rexsyc@163.com (Y. Sun), liuhy@cugb.edu.cn (H. Liu), gaoyucugb@163.com (Y. Gao).

https://doi.org/10.1016/j.heliyon.2023.e13384
Received 13 May 2022; Received in revised form 29 January 2023; Accepted 30 January 2023
Available online 13 February 2023

random forests, generalized additive models (GAMs), multiple adaptive regression splines (MARS), classification and regression trees (CART), support vector machines (SVM) and other machine learning algorithms. These methods are useful for the application of customer lifetime value calculation.

In view of the large number of customer lifetime value models, each model has its own strengths and limitations. The classical probability model uses probability distribution to describe customers' future purchase behavior, while the econometric model explains customers' different purchase behavior through a series of covariates. Most of these deductive models are based on certain management theories and mathematical relationship assumptions and have strong explanatory ability. Therefore, they have always been the mainstream research method in the field of marketing and consumer behavior research. Different from traditional deduction-based research methods, the induction-based machine learning algorithm allows researchers to directly identify hidden customer behavior patterns through data. Such algorithms do not rely on theoretical assumptions, are especially suitable for complex relationship modeling brought by consumer big data, and usually have universality and excellent prediction ability.

In view of the fact that the research on customer lifetime value under the condition of noncontractual relationship has always been a hot issue in academic circles, this paper introduces data feature engineering to preprocess the data, then carries out the data mining process, subdivides the customer value based on machine learning algorithms and CRM analysis model, and constructs the customer value segmentation identification model under the condition of noncontractual relationship. It also provides marketing strategies for different customer groups for the shopping platform, retains more high-quality customers, and maintains the quantity and quality of customer groups.

The structure of this paper is as follows: Section 2 reviews the previous related research and determines the problems of this study; Section 3 includes the methodology and data mining process, mainly introducing the methods used in this paper, data processing, and data mining process; Section 4 provides results and discussion, discussing the research content and experimental results; and Section 5 provides conclusions and prospects, research conclusions, and prospects for future work.

2. Relevant research review

2.1. Research on customer life cycle value

After CRM became a key research issue, enterprises began to compete at the customer level [7,8]. Researchers first put forward the concept of customer perceived value [9]; that is, customers weigh the perceived value when obtaining products or services with the cost paid for obtaining products or services and finally obtain the overall evaluation of the effectiveness of products and services. At present, there are two perspectives to define customer value. One is to define customer value based on the customer perspective [10, 11]. Customer value is the difference between total customer value and total customer cost; the other is customer value based on the enterprise perspective [11]. Customer value is the customer lifetime value (CLV) within the customer life cycle. Customer lifetime value refers to the present value of all profits created by customers for the enterprise in the whole process of maintaining a relationship with the enterprise [12]. In 1985, Jackson [13] first proposed the definition of customer lifetime value based on the traditional net present value method, i.e., the net present value of the benefits that customers bring to enterprises in their future consumption period (i.e., customer life cycle) and defined the measurement model of customer lifetime value, which laid the foundation for subsequent research on customer lifetime value. At present, the mainstream view is to study customer value from the perspective of enterprises. This paper identifies the lifecycle value of each customer by analyzing the customer's purchasing behavior information and adopts effective data analysis and mining methods to classify customers according to the difference in customer value to identify customers who can bring greater profits to the business. Through accurate operation strategies, customer satisfaction and loyalty can be comprehensively improved, and businesses can obtain more profits [14].

When the enterprise and the customer have a contractual relationship, the customer only signs a long-term business contract with one enterprise and smoothly carries out trading activities. When both parties do not renew the contract or stop trading activities after performing the contract, it can be considered that the customer has been permanently lost. When the customer trades with the enterprise again, it will be treated as a new customer by the enterprise. For example, insurance business and loan business [15] are typical contractual relationships. In the case of a contract, because the enterprise maintains a contractual relationship with its customers, the loss of customers is observable to the enterprise. The lifetime value of customers can be calculated from the perspective of the whole customer group. The cash flow brought by customers to the enterprise in the future and the maintenance rate of segment customer groups can be calculated according to the past transactions of the enterprise [16]. In noncontractual relationships, customers may have transactions with multiple enterprises at the same time, and the cost of customers switching between several enterprises is very low. When customers suspend transactions with an enterprise, it does not mean that customers have lost. After a period of time, customers may buy again, and enterprises still treat them as old customers. For example, the retail industry is a typical noncontractual relationship [17]. The customer retention rate can be specifically expressed as the proportion of customers who continue to purchase enterprise products and services in the next year. This kind of model usually does not consider the heterogeneity between individual customers and can be calculated directly according to the formula. The change between different formulas mainly comes from the different assumptions about the customer purchase cycle, whether the customer profit changes in the weekly period, etc.

2.2. Customer life cycle value model

In a noncontractual relationship, the loss of customers is unobservable for the enterprise. At this time, the calculation of customers' lifetime value focuses on estimating the probability or frequency of customers' repeated purchases in the future based on the recency

frequency and monetary (RFM) model and other data. Scholar Arthur Hughes proposed the RFM model in the 1990s, which laid the theoretical foundation for the calculation of the customer lifetime value [18]. Among them, the latest consumption time of recency (R) refers to the time interval from the last consumption of the customer to the current one; frequency (F) consumption frequency refers to the total consumption frequency of customers in a period of time t; and monetary (M) consumption amount refers to the total consumption amount of the customer in a period of time t. Researchers believe that RFM can well reflect customers' historical consumption behavior, and a customer lifetime value prediction model can be established based on three variables R, F, and M [19]. Academic circles have put forward more calculation strategies of customer lifetime value under noncontractual relationships, such as probability models [4,5], econometric models [20], and random forest models [21,22]. Among them, the probability model is a classical method to estimate the lifetime value of individual noncontract customers, and the machine learning model is a new method that has attracted more attention with the development of consumer big data in recent years.

As the concept of customer lifetime value is increasingly accepted, to achieve the best customer value, more effective marketing plans and strategies are adopted [23]. It is very meaningful to study a more effective model to predict the customer lifetime value. Considering the distribution of customer value in the time of trading activities, it is an important reference index of CRM strategy to introduce the concept of customer lifetime value and evaluate the current and future value of different types of customers [24,25]. Çavdar and Ferhatosmanoğlu [26] used a model to estimate the lifetime value of customers in the aviation industry and estimate their CLV. This basic model is enhanced through customers' social network information to include customers' indirect contributions. This model is used to provide an example for potential customer analysis for airlines' CRM applications.

Customer churn prediction models play an important role in organizations that retain customers in saturated and rapidly changing markets (such as telecommunications and banking). Óskarsdóttir et al. [27] considered the life cycle value of a single customer and divided the characteristics of the customer group by calculating the expected maximum profit measure. Monalisa et al. [28] divided customers into superstar customers, typical customers, and dormant customers, analyzed the customer lifetime value, and implemented the strategy of managing customers based on the type of portfolio. Research on customer value is pushed to a higher level by the customer life cycle, which promotes the transition from the customer value concept to the customer relationship value concept.

However, scholars have found the shortcomings of the customer lifetime value model in the following aspects: (1) All customer lifetime value models are too conceptual and idealized, but they cannot be applied in practice. In particular, some models are more affected by other factors and are difficult to calculate in practice [29]. (2) The customer lifetime value only considers the profit brought by the customer to the enterprise in the calculation, and the customer value is equal to the customer's profit, but the customer's profit is only a part of the customer's value and cannot fully represent the customer's value. (3) In the information age, the information channels for customers to obtain enterprise products are richer, and customers will have more choice opportunities when choosing goods, resulting in intensified competition among enterprises for customers, and many customers do not form a complete life cycle; therefore, there is a serious lack of data when calculating customer value.

2.3. The machine learning model

The big data revolution has produced data with the characteristics of volume, variety, and speed [30–32]. It provides abundant research data for machine learning algorithms to play a role in the research field of customer lifetime value. Scholars use machine learning methods to mine high-value customer behavior information hidden behind the data. For a long time, the study of marketing and consumer behavior has been based on interpretation law, using theoretically-based theory and easy interpretation of parameter models, and is no exception to customers' life value. With the continuous progress of technology and methods, data mining technology will bring about the transformation of research methods in the field of marketing and consumer behavior from deduction to induction [33]. In solving the research of consumer behavior mining, more researchers began to study customer lifetime value and its components based on machine learning algorithms, such as establishing customer churn prediction models through machine learning methods, such as classification tree [34], GAM [35,36], SVM [37], and random forest [21]. Chen and Fan [38] used support vector regression to predict dynamic customer lifetime value and used customer behavior variables and controlled variables about multiple promotions to select the best promotion.

The study found that the algorithm based on machine learning has excellent predictive capabilities and can better predict model variables than traditional interpretative models. To study the application of machine learning in customer churn prediction, Ma and Xia [39] constructed a machine learning-based churn prediction model, compared it with prediction models, such as logistic regression and decision tree, after feature selection on the telecom customer dataset, and achieved a better prediction effect. Vafeiadis et al. [40] compared the performance of different machine learning methods when solving the problem of customer churn prediction in the telecommunications industry. The experimental results show that SVM is superior to artificial neural networks, decision trees, and Bayesian classifiers in terms of prediction hit rate, coverage, and promotion coefficient. Martínez et al. [41] used three machine learning methods, logistic lasso regression, limit learning machine, and gradient improvement trees, to predict customers' future purchases, and the results show that the effect of gradient improvement trees is improved. Jain et al. [42] showed that feature extraction is a very important part of machine learning algorithms.

Through a literature review, it can be seen that the calculation of individual noncontract customer lifetime value has always been the focus of research. Especially with the development of data mining calculations and computer computing abilities in recent years, the customer lifetime value calculation model based on machine learning algorithms is becoming increasingly mature. Academic circles have carried out different explorations around the performance, accuracy, and scope of application of different models. For example, Chen et al. [43] compared and analyzed the traditional parameter model represented by Pareto/NBD with the deep neural network algorithm represented by a convolution neural network in the prediction of customer lifetime value in the game industry.

According to different task types and data types, the performance of machine learning algorithms is not consistent. Finding the universal optimal solution still needs to be explored. As shown in Table 1, the advantages and disadvantages of traditional algorithms and machine learning algorithms in the field of customer segmentation research.

Based on the above research, this paper believes that, first, the traditional methods still have their merits, and their prediction effect cannot be ignored. For example, Coussement et al. [44] believe that on the basis of appropriate data preprocessing, the effect of simple logical regression prediction is no worse than that of random forest, support vector machine and other algorithms; Second, for different data sets, the performance of various machine learning algorithms is different, and the prediction of a single method will have some deviation. Therefore, the comprehensive calculation based on multiple methods is an ideal modeling strategy with strong universality.

In view of this, in terms of customer type segmentation, this paper combines RFM model and machine learning algorithm, introduces the Knowledge Discovery in Database method for data mining, and constructs a customer value segmentation recognition model under the condition of noncontractual relationship. In terms of the method of calculating future customer value, this paper draws lessons from the negative two distribution BG/NBD model and parameter heterogeneous distribution gamma-gamma model, avoids the errors caused by the uncertainty of purchase possibility, exit possibility, and consumption amount caused by the heterogeneity of customers with different customer values, improves the accuracy of customer value, contributes to the correlation analysis at the user level, and strengthens the explanation of the intermediary effect.

3. Methodology and data mining process

3.1. Methodology

3.1.1. RFM model

The RFM model was proposed by Arthur Hughes in 1994, which laid a theoretical foundation for the calculation of customer lifetime value [18].

R represents the time when the customer last purchased the product. If the R value is smaller, it indicates that customers have consumed the products or services of the enterprise in the shorter time before the statistical time point, which means that the customer keeps paying attention to the enterprise's products in a short time, and the customer is more likely to be touched by the enterprise's marketing information. Therefore, such customers are more likely to make repeated purchases in a short time, and the enterprise can obtain greater income by investing less cost in such customers. When the R value is large, it indicates that the customer has not exhibited purchase behavior for a long time, and the competitiveness of the enterprise's products or services decreases.

F refers to the frequency of customers purchasing enterprise products or services in the statistical period. Generally, the larger the F value of customers, the higher the frequency of customers purchasing enterprise products or services, the higher the customer loyalty and the greater the value of customers. The customer's purchase frequency is often considered in combination with the customer's purchase proximity. When the customer's purchase frequency is high, the closer the purchase time is from now, the more likely the customer will continue to consume the enterprise's products. When the purchase time is far from now, such customers used to be valuable customers for some time, but they are now highly likely to be lost.

M refers to the total amount of customers' consumption of enterprise products or services within the statistical time. Generally, under the constraint of disposable resources, the higher the amount of customers' purchases of enterprise products or services, the lower the amount of consumption of alternative products or services, the more loyal customers and the higher value. However, due to the problem of multiple collinearity between monetary and recency or frequency, the average consumption amount of customers is used to replace the total amount to reduce the problem of collinearity between M and R or F.

Table 1
Advantages and disadvantages of previous traditional algorithms and machine learning algorithms in the field of customer segmentation research.

Serial No	Research methods	Research object	Research limitations
Literature 1 [45]	Probability models: Pareto/NBD and BG/NBD Machine learning algorithms: generalized additive model and support vector machine	Retailer data during large online promotions	The classical probability model has weak ability to capture and track extreme changes in data; The overall prediction result of machine learning algorithm is lower than the actual mean value.
Literature 2 [46]	K-means clustering method, K-medoids method and Fuzzy RFM model	Online sales transaction data	Only customer segmentation research has been conducted, and there are only three categories of customer segmentation:very potential (loyal) customers, potential customers and non potential customers.
Literature 3 [47]	Fuzzy c-means cluster and the RFM model	Customer consumption data of e-commerce platform	According to the characteristics of customers, the fuzzy c-means cluster and the RFM model is used to segment customers, but the customer life cycle value is not calculated.
Literature 4 [48]	Tree clustering, RFM model	Customer transaction data	This research combines Tree clustering with RFM model, and only conducts customer segmentation research.
Literature 5 [49]	RFM model, K-Means and Fuzzy C-Means algorithms	Online retail data	This study only calculates the number of customer clusters according to customer characteristics.

3.1.2. BG/NBD predicted purchase expectation

The BG/NBD model [50] is used to describe repeat purchase behavior in the context of noncontractual customer relationships. That is, users can buy products at any time without time constraints. The model can use historical user transaction data to predict the transaction times and turnover rate of each user in the future. BG/NBD is an improved version of the classic model, which has the following assumptions:

H1: When the customer is active, the transaction volume of each customer obeys the Poisson distribution of the transaction rate λ; that is, the interval of transaction time obeys the Poisson distribution of the transaction rate λ. The formula is as Eq. (1):

$$f\left(t_j \middle| t_{j-1}; \lambda\right) = \lambda e^{-\lambda\left(t_j - t_{j-1}\right)}, t_j > t_{j-1} \geq 0 \tag{1}$$

H2: The probability density function of nonuniformity of customer transaction rate λ obeys *gamma* distribution. The formula is as Eq. (2):

$$f(\lambda|r,\alpha) = \frac{\alpha^r \lambda^{r-1} e^{-\lambda\alpha}}{\Gamma(r)}, \lambda > 0 \tag{2}$$

H3: After each transaction, the probability that the customer becomes silent is p, and the customer churn point follows the binomial distribution. The formula is as Eq. (3):

$$P = p(1-p)^{j-1}, j = 1, 2, 3, \cdots \tag{3}$$

H4: The nonuniformity probability density function of probability p follows a *beta* distribution. The formula is as Eq. (4):

$$f(p|a,b) = \frac{p^{a-1}(1-p)^{b-1}}{B(a,b)}, 0 \leq p \leq 1 \tag{4}$$

where $B(a,b)$ is a *beta* function, which can be expressed by the *gamma* function. The formula is as Eq. (5):

$$B(a,b) = \Gamma(a)\Gamma(b) \,/\, \Gamma(a+b) \tag{5}$$

Based on the above assumptions, there are four key expressions for predicting customer purchase behavior:

(1) The probability $P(X(t))$ of one purchase in a period of time with length x, and the formula is as Eq. (6):

$$\begin{aligned} P(X(t) = x|r,\alpha,a,b) = {} & \frac{B(a,b+x)}{B(a,b)} \frac{\Gamma(r+x)}{\Gamma(r)x!} \left(\frac{\alpha}{\alpha+t}\right)^r \left(\frac{t}{\alpha+t}\right)^x \\ + \delta_{x>0} & \frac{B(a+1,b+x-1)}{B(a,b)} \left[1 - \left(\frac{\alpha}{\alpha+t}\right)^r \left\{\sum_{j=0}^{x-1} \frac{\Gamma(r+j)}{\Gamma(r)j!} \left(\frac{t}{\alpha+t}\right)^j\right\}\right] \end{aligned} \tag{6}$$

(2) The expected number of transactions $E(X(t))$ in a period of t, and the formula is as Eq. (7):

$$E(X(t)|r,\alpha,a,b) = \frac{a+b-1}{a-1}\left[1 - \left(\frac{\alpha}{\alpha+t}\right)^r {}_2F_1{}^2\left(r, b; a+b-1; \frac{t}{\alpha+t}\right)\right] \tag{7}$$

(3) The expected number of transactions $E(Y(t))$ of each customer in a period of t, and the formula is as Eq. (8):

$$\begin{aligned} & E(X) = E(Y(t)|s,\beta,a,b,x,t_x,T) \\ = {} & \frac{(a+b+x-1)}{1+\sigma(\beta+T/\beta+t_x)^{S+x}} \cdot \left[1 - \left(\frac{\beta+T}{\beta+t^x+T}\right)^{S+x} {}_2F_1{}^2\left(s+x, b+x, a+b+x+1, \frac{t}{\beta+T+t}\right)\right] \end{aligned} \tag{8}$$

(4) The maximum likelihood function of the four parameters for calculating the number of transactions, and the formula is as Eq. (9):

$$\begin{aligned} & L(s,\beta,a,b|x,tx,T) \\ = {} & \frac{B(a,b+x)\Gamma(s+x)d}{B(a,b)\Gamma(s)(\beta+T)^{S+x}} + \sigma\frac{B(a+1,b+x+1)\Gamma(s+x)d}{B(a,b)\Gamma(s)(\beta+T)^{S+x}} = A_1 \cdot A_2 \cdot (A_3 + \delta_{x>0}A_4) \end{aligned} \tag{9}$$

where, and the specific calculation formulas are shown in Eqs. (10)–(13):

$$A_1 = \frac{\Gamma(r+x)\alpha^r}{\Gamma(r)} \tag{10}$$

$$A_2 = \frac{\Gamma(a+b)\Gamma(b+x)}{\Gamma(b)\Gamma(a+b+x)} \tag{11}$$

$$A_3 = \left(\frac{1}{\alpha+T}\right)^{r+x} \tag{12}$$

$$A_4 = \left(\frac{a}{b+x-1}\right)\left(\frac{1}{\alpha+t_x}\right)^{r+x} \tag{13}$$

After taking the logarithm, the four formulas of A_1, A_2, A_3, A_4 are solved simultaneously, the logarithm and the minimum parameter s, β, a, b value are calculated, which are brought into the expected transaction times $E(Y(t))$ of each customer in a period of time, and the expected value is calculated in Formula (8). The model can predict customers' expectation E(X) of purchase times in the next 40 weeks based on customer behavior data in the past two years.

3.1.3. The gamma-gamma model predicts the consumption amount

The gamma-gamma model [51] of the average consumption amount of customers is subject to the following assumptions:

(1) The average purchase amount of individual customers follows the *gamma* distribution. The formula is Eq. (14):

$$f(m_x|p,v,x) = \frac{(vx)^{px} m_x^{px-1} e^{-vxm_x}}{\Gamma(px)} \tag{14}$$

where x is the number of purchases in customer history, p is the scale parameter of the *Gamma* distribution, and v is the shape parameter of the *Gamma* model.

(2) Due to the individual to heterogeneity between customers, the shape parameter v follows the *Gamma* distribution with scale parameter q and shape parameter γ. The formula is as Eq. (15):

$$f(m_x|p,q,\gamma,x) = \frac{\Gamma(px-q)}{\Gamma(px)\Gamma(q)} \frac{\gamma^q m_x^{px-1} x^{px}}{(\gamma+m_x x)^{px+q}} \tag{15}$$

(3) The distribution function of the expected transaction amount m_x is obtained from the Bayesian posterior probability. The formula is as Eq. (16):

$$g(v|p,q,\gamma,m_x,x) = \frac{(\gamma+m_x)^{px+q} v^{px+q-1} e^{-v(\gamma+m_x x)}}{\Gamma(px+q)} \tag{16}$$

In summary, the expected transaction amount of customers in transaction behavior. The formula is as Eq. (17):

$$E(v|p,q,\gamma,m_x,x) = \frac{(\gamma+m_x x)p}{px-q-1} = \left(\frac{q-1}{px+q-1}\right)\frac{\gamma p}{q-1} + \left(\frac{px}{px+q-1}\right)m_x \tag{17}$$

Finally, the expected purchase times of customers in a certain period of time in the future and the average transaction amount $E(M)$ of customers are obtained.

Combined with Jackson's idea of defining customer lifetime value with the discounted value of cash flow, the predicted value of CLV is obtained by predicting the expected purchase times and expected transaction amount [52,53]. The formula is as Eq. (18):

$$CLV = \frac{X(\overline{X})E(\overline{M})}{(1+p)^n} \tag{18}$$

3.2. Data mining process

Data mining, also known as knowledge discovery in databases (KDD), refers to the process of automatically searching for information with special relationships hidden in a large amount of data. It is a technology to find its law from a large amount of data. Data mining technology is often used to enhance the ability of information retrieval systems. As shown in Fig. 1, the knowledge discovery process generally includes data source, data selection, data preprocessing, data transformation, and knowledge discovery.

The method used in this study is KDD, which uses a data-driven method to obtain knowledge from online retail datasets and classify customers according to their purchase patterns. By identifying the customer's recency, frequency, and monetary value, the cluster model is used to analyze the customer's purchase pattern. This study is divided into the following stages.

3.2.1. Data sources

The longitudinal data used in this study come from an internet insurance sales platform. The link address of the dataset used in this article is https://archive.ics.uci.edu/ml/datasets/Online+Retail+II. This online retail dataset is generated from nonstore online retail transactions registered in the UK. The company sells unique occasional gifts, and many of the company's customers are wholesalers. The dataset-related information introduction is shown in Table 2. The dataset contains 9 attributes: Dataset characteristics, Number of

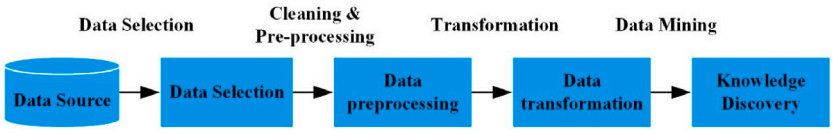


Fig. 1. KDD process flow.

instances, Region, Attribute characteristics, Number of attributes, Donation date, Related tasks, Lack of value, and Network hits.

3.2.2. Data selection

This paper uses Python software to view the attribute summary information of the data in the data list through the pandas. dataframe.info () function. Data summary information includes a list of all columns with their data types and quantities. Table 3 shows the attribute information of the original data. The attributes of the original data include "InvoiceNo", "StockCode", "Description", "Quantity", "InvoiceDate", "UnitPrice", "CustomerID", and "Country". In addition to the attribute information of the data, it also includes the meaning of each attribute information.

3.2.3. Data preprocessing

Data are easily disturbed by noise during acquisition and storage, resulting in problems with the accuracy, integrity, and consistency of the collected data. These problems in the data will lead to the inability to perform data mining directly or the data mining effect is not good. Therefore, to ensure the accuracy and credibility of the mining results, it is necessary to preprocess the data before data mining.

In this paper, data preprocessing mainly addresses repeated data processing, missing value processing, and outlier processing. First, this paper uses the *Pandas. Dataframe. The duplicated()* function of Python 3.8.12 software was used to check that there were duplicate values in the dataset and delete the duplicate items in the dataset through the *Pandas. Dataframe. Drop_duplicates()* function. In addition, this article uses the *Pandas. Dataframe. The Dropna()* function was used to eliminate missing values after removing duplicate data and missing values from the dataset. This article examines the data attribute values after data selection processing. If it is found that the invoice number code starts with the letter "C", it means that the order has been canceled, and the transaction has not been completed, so this part of the outlier data also needs to be deleted.

This paper $_{\text{uses}}$ the *heatmap()* function in the *missingno* library of Python 3.8.12 software to detect missing values. If the attribute value is missing, stripes will be displayed in the figure, so you can visually see the missing value and observe its position. Colors in the drawing represent features that lack relevance. As shown in Fig. 2, no missing value is found, and there is no null value in any column in the data. After the target data preprocessing is completed, the data conversion stage can be entered.

3.2.4. Data transformation

R is calculated by calculating the difference between the deadline and the latest invoice date, F is calculated by accumulating the customer's order quantity, and M is calculated by accumulating the total amount of all customer orders. Therefore, the converted dataset contains the customer ID, recency, frequency, and monetary value. Through the above calculation process, the R, F, and M values of each customer can be calculated, and the extraction of RFM original eigenvalues is completed. Next, the data conversion of RFM original eigenvalues is carried out.

The R, F, and M values of each customer are standardized and marked as R-score, F-score, and M-score, respectively. The indicator whose value of each dimension is lower than the average of the dimension is marked as 0, and the indicator whose value of each dimension is higher than the average of the dimension is marked as 1; that is, when index > avg(), record as 1, while, when index < avg (), record as 0. Through the standardization of the three indicators, the construction of the RFM matrix is completed, and, finally, the RFM combination value is obtained.

3.2.5. Knowledge discovery

It is calculated according to a single R-score, F-score, M-score label; for example: R-score > avg (R-score), F-score > avg (F-score), M-score > avg (M-score), which means that the customer has recently purchased, the purchase frequency is higher than the average purchase frequency of all customers, and the purchase amount is higher than the average purchase amount of all customers. Therefore, the label of "important customer retention" is attached. There are two cases for each dimension indicator, marked as 1 (high) or 0 (low). According to the knowledge of arrangement and combination, eight combinations can be obtained. According to the method of combination, customers are classified into groups. The customer classification and customer characteristics are shown in Table 4. In this paper, customers are divided into important value customers, important development customers, important protection customers, important retained customers, general value customers, general development clients, general retention clients, and lost customers.

4. Results and discussion

4.1. Customer classification based on the RFM model

As a quantitative analysis model, the RFM model describes the importance of customers through three-dimensional attribute values and classifies them. The three dimensions are the latest purchase time (R), the number of purchases in a certain period (F), and the total

Table 2
Dataset-related information introduction.

Dataset characteristics	Multivariate, Order, Time Series, Text	Number of instances	541909	Region	Business
Attribute characteristics	Integer, Real	Number of attributes	8	Donation date	2019-09-21
Related tasks	Classification, Regression, Clustering	Lack of value	Yes	Network hits	106790

Table 3
Attribute information of raw data.

Attribute	Information	Meaning
InvoiceNo	Invoice number	A 6-digit integral number uniquely assigned to each transaction. If this code starts with the letter 'c', it indicates a cancelation.
StockCode	Product (item) code	A 5-digit integral number uniquely assigned to each distinct product.
Description	Product (item) name	Description of product name
Quantity	Numeric	The quantities of each product (item) per transaction.
InvoiceDate	Invoice date and time	The day and time when a transaction was generated.
UnitPrice	Unit price	Product price per unit in sterling (Â£).
CustomerID	Customer number	A 5-digit integral number uniquely assigned to each customer.
Country	Country name	The name of the country where a customer resides.

In this paper, the "CustomerID", "InvoiceDate", "UnitPrice", and "Quantity" data columns of the selected dataset constitute the target data of this study. In this paper, the "InvoiceDate", "UnitPrice", and "Quantity" data columns are selected mainly to construct the feature data of the RFM model.

Fig. 2. Missingness map.

Table 4
Customer classification and customer characteristics.

Customer type	R value	F value	M value
Important value customer	High	High	High
Important development customer	High	Low	High
Important protection customer	Low	high	High
Important retention customer	Low	Low	High
General value customer	High	High	Low
General development customer	High	Low	Low
General retention customer	Low	High	Low
Lost customer	Low	Low	Low

amount of purchases in a certain period (M).

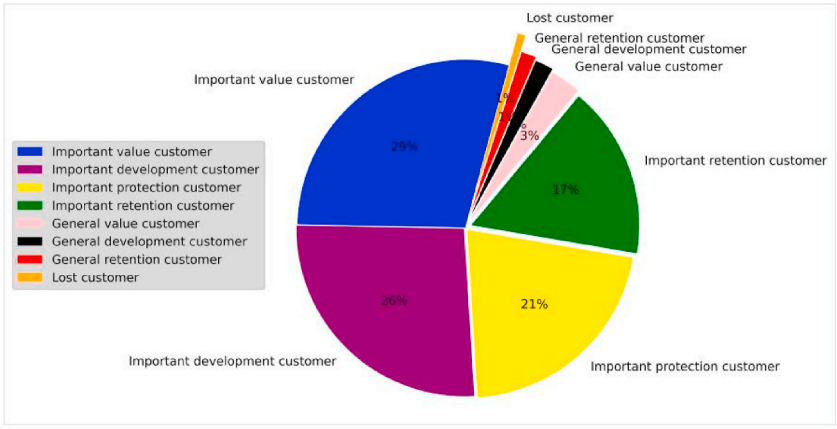
According to the customer classification results in Table 4, the relevant data of each customer group are counted. Fig. 3 shows statistics on the proportion of people and the consumption amount of each customer group.

Any data analysis should be based on business, and it is meaningful to combine the data of business activities. Specifically, why the user value of RFM is distributed like this requires the platform to have sufficient understanding so that the reasons behind the loss of users can be found. Improvement comes from three aspects: improving activity, increasing retention rate, and increasing payment rate.

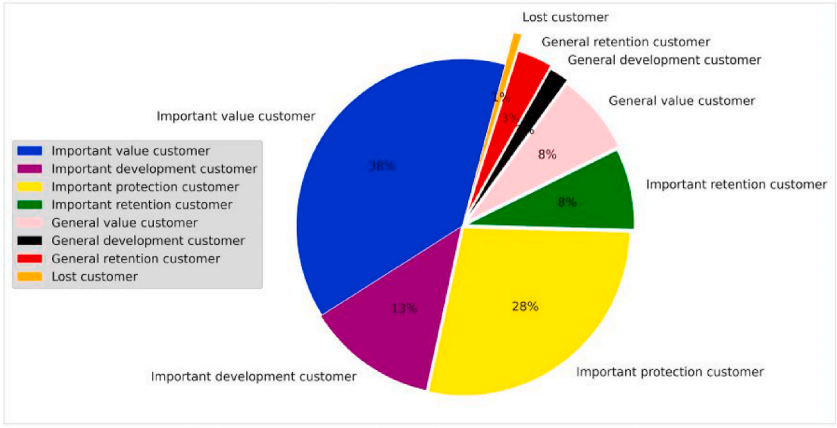
To increase the activity, increase the activity of general customers and low value customers. Then, turn it into high-quality customers; next, improve the retention rate by interacting with important retained customers to improve the retention rate of these users; improve the payment rate means maintain and develop the loyalty of important customers and maintain the good income of the enterprise. For each customer group, this paper provides operators with reference operation strategies. Table 5 shows the operational strategies adopted by different customer groups.

4.2. Response rate analysis based on machine learning

In this paper, the standardization process was completed with the three indicators of the RFM model. The R-score, F-score, and M-score values in the constructed RFM matrix are used as independent variables to predict the response of the target variable. The customer has a response of 1; otherwise, it is 0. This paper directly uses the logistic regression, decision tree, SVM, K-nearest neighbors, naive Bayes, random forest, AdaBoost, gradient boosting decision tree, and multilayer perceptron algorithms from the sklearn library



(a) Proportion of people in different customer categories



(b) Consumption proportion of different customer categories

Fig. 3. Proportion of people and consumption amount of different customer groups. (a) The proportion of the number of different customer groups in the total number, and (b) the proportion of the consumption amount of different customer groups in the total consumption amount.

Table 5
Operational strategies for different customer groups.

Customer type	Behavior characteristics	Operation strategy
Important value customer	Recently, this customer group has purchased, with high purchase frequency and high consumption, and they are the main consumers.	Upgrade to the very important person (VIP) customers, provide personalized services, and tilt more resources.
Important development customer	Recently, this customer group has purchased, with low purchase frequency and high customer unit price. They may be a new wholesaler or enterprise purchaser.	Provide member points service and provide a certain degree of discount to improve the retention rate of customers.
Important protection customer	Recently, this customer group has not bought, but the purchase frequency is high and the consumption is high.	Introduce the latest products/functions/upgraded services through SMS and email to promote customer consumption.
Important retention customer	Recently, this customer group has not bought, and the purchase frequency is low, but the customer unit price is high.	Introduce the latest products/functions/upgrade services, promotional discounts, etc., through SMS, email, phone, etc., to avoid the loss of customers.
General value customer	Recently, this customer group has purchased, with high purchase frequency, but low consumption.	Introduce the latest products/functions/upgraded services to promote customers' consumption.
General development customer	Recently, this customer group has purchased, with low purchase frequency and low consumption. They may be new customers.	Provide community services, introduce new products/functions, and promote customers' consumption.
General retention customer	Recently, this customer group has not bought, with high purchase frequency and low consumption.	Introduce new products/functions to arouse this part of customers.
Lost customer	Recently, this customer group has not bought, with low purchase frequency and low consumption, which has been lost.	This part of customers can be aroused by promotion and discount. When the resource allocation is insufficient, this part of users can be temporarily abandoned.

in the Python language for validation.

During the experiment, this paper uses four cross-validation methods: leave-one-out cross-validation (LOOCV), standard cross-validation, stratified k-fold cross-validation and shuffle-split cross-validation. By comparing the results of the models, finally this paper uses the LOOCV method in the cross validation method. The LOOCV method is suitable for small sample datasets and can make full use of all data, so the deviation of model results will be reduced [54]. The specific process is as follows: This paper uses Python version 3.8.12 software and Scikit-Learn version 1.0.1 in the machine learning framework, selects the train_test_split model in the sklearn. model_selection library, and sets the test_size parameter of the train_test_split model to 0.3; that is, 70% and 30% of the datasets are divided into training datasets and test datasets, respectively. Among them, the training dataset is used to train the model, and the test dataset is used to evaluate the model. In addition, the area under the curve (AUC) and accuracy index were used to evaluate the performance of the algorithm.

The proportion of response in marketing activities is generally relatively low. The target variable category data in this sample data are a typical unbalanced dataset, so this paper mainly uses the AUC value and the accuracy value to evaluate the model. From the evaluation results of the algorithms in Table 6, it can be seen that the AUC values and accuracy values of boosting fusion algorithms, such as AdaBoost, and gradient boosting decision trees are relatively best. Shortcomings of machine learning algorithms in this paper: First of all, the sample data in this paper is a typical unbalanced data set, and the performance of different algorithms in handling unbalanced data sets is very different, especially the information extraction ability of algorithms in feature data needs to be improved. Secondly, the accuracy of machine learning algorithms involved in this paper needs to be improved, especially the accuracy of some algorithms is relatively low. In addition, the parameters optimization of the algorithm need to be carried out continuously, and the performance of the algorithms can be improved by finding the optimal parameters of the algorithms.

In this paper, the optimal algorithm gradient boosting decision tree is selected for retraining, and the effectiveness of the model is evaluated by the yield curve. As seen from the income curve in Fig. 4, after the model calculation, among the 20% customers with the highest response rate, 53% of the customer groups can respond. From the perspective of CRM strategy, the enterprises will choose the customer group with the highest response rate as the target group of marketing.

4.3. Calculating customer lifetime values

To calculate CLV, the BG/NBD model and the Gama-Gama model are used. To use this model, we changed the "R" and "T" variables to weeks. Through the model, the expected purchase amount of each customer in the dataset can be calculated. In addition, all these purchases are summed to obtain the expected revenue of the company.

4.3.1. BG/NBD model training

Using the **BG/NBD** model and the Gama-Gama model, the values of α and r of the gamma distribution and a and b of the β distribution are calculated, a: 0.06, α: 9.30, b: 1.13, r: 1.78. The required parameters are fitted by the model: α and r of the gamma distribution and a and b of the β distribution, and the confidence interval is given, as shown in Table 7.

4.3.2. BG/NBD model visualization

The BG/NBD model predicts the expected number of future transactions within a predefined period and the possibility that the customer is still alive. The gamma-gamma model will be used in combination with the BG/NBD model to predict the expected monetary value of the purchase based on the current dollar value discounted at the predetermined discount rate. Plotting the estimated gamma distribution λ. Fig. 5 shows the customer's purchase tendency curve.

Draw the estimated beta distribution of P, where p is the probability of the customer quitting immediately after the transaction, and Fig. 6 is the probability change curve of the customer quitting immediately after the transaction.

Use the frequency/recency matrix of the expected transaction to visualize the possibility of the customer's existence. The matrix calculates the expected number of transactions in the next time period of the customer's frequency/recency. Fig. 7 visualizes the frequency/recency matrix.

Fig. 8 shows the thermodynamic diagram of customer retention probability.

As shown in Fig. 9, the chart groups all customers in the calibration period by the number of repeat purchases (x-axis) and then averages their repeat purchases in the retention period (y-axis). The orange line and blue line represent the actual results of the model

Table 6
The AUC value and accuracy value benchmarking table of nine algorithms.

Classifier	Accuracy	AUC
K-nearest neighbors	90.25%	74.65%
Logistic regression	91.36%	86.14%
Support vector machine	91.14%	67.25%
Decision tree	87.14%	68.67%
Random forest	89.98%	73.94%
AdaBoost	94.60%	93.82%
Gradient boosting decision tree	93.80%	95.12%
Naive Bayes	88.18%	85.71%
Multilayer perceptron	91.14%	87.17%

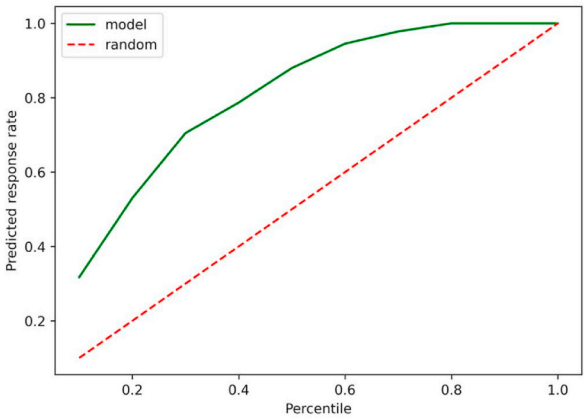


Fig. 4. Yield curve.

Table 7
Value and confidence interval distribution of model parameters.

	Coef	Se (coef)	Lower 95% bound	Upper 95% bound
α	9.30268544	0.337551659	8.641084188	9.964286691
r	1.784611036	0.052432477	1.681843382	1.887378691
a	0.059187585	0.006403364	0.046636992	0.071738178
b	1.132869909	0.081850595	0.972442743	1.293297074

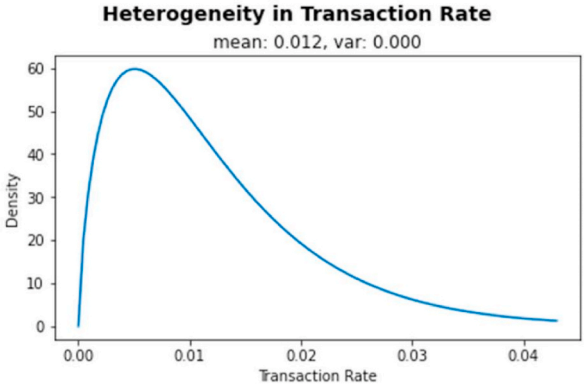


Fig. 5. Customer's purchase tendency curve.

prediction and y-axis, respectively. The model can predict the behavior of the customer group in the sample very accurately.

According to the customer transaction history, the historical probability of survival can be calculated according to the trained model. Customers' purchase frequency is low, and the possibility of loss is high. As shown in Fig. 10, when a customer has a transaction, the probability of retention immediately increases significantly but then decreases gradually.

4.3.3. Gamma-gamma model fitting

The gamma-gamma model is used to fit the dataset and estimate the average transaction value of each customer. After applying the gamma-gamma model, the average transaction value of each customer in a lifetime can be estimated. Assuming that the monthly discount rate is 0.01%, and the monthly discount rate is 12.7%, the life cycle value of each customer in the next 12 months is finally calculated. Table 8 shows the average future transaction value and lifetime value (LTV) of the top 20 customers in the next 12 months.

4.4. Discussion

Currently, e-commerce is booming, and paying attention to customer value on e-commerce platforms is one of the important links to win the competition. By analyzing customers' purchasing behavior information and adopting effective data analysis and mining methods, businesses can identify customers who will bring more profits. Through precise marketing means, customer satisfaction and

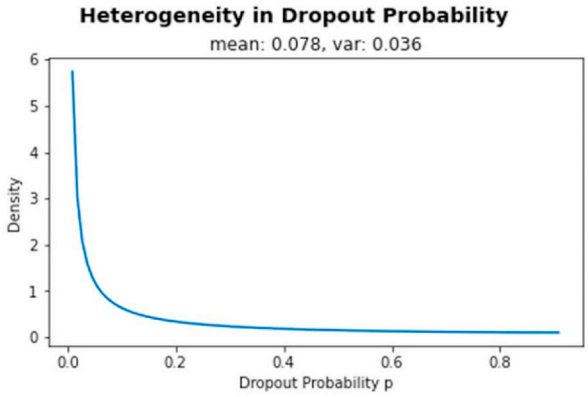


Fig. 6. Probability curve of customer exiting immediately after transaction.

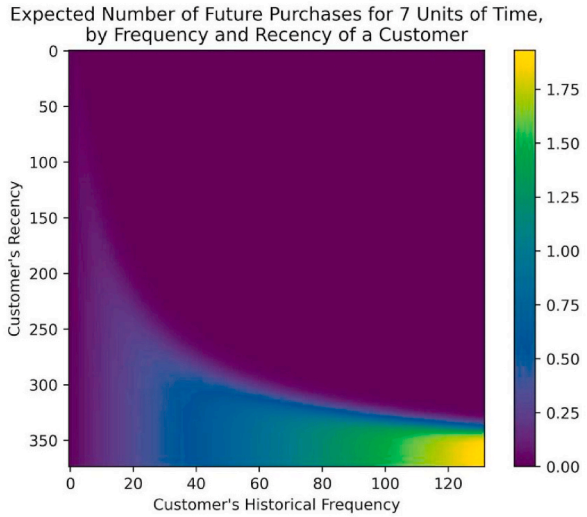


Fig. 7. Visualize frequency/recency matrix.

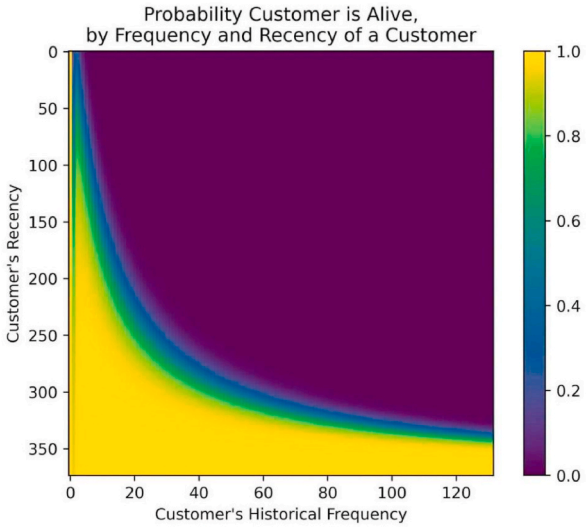


Fig. 8. Thermodynamic diagram of customer retention probability.

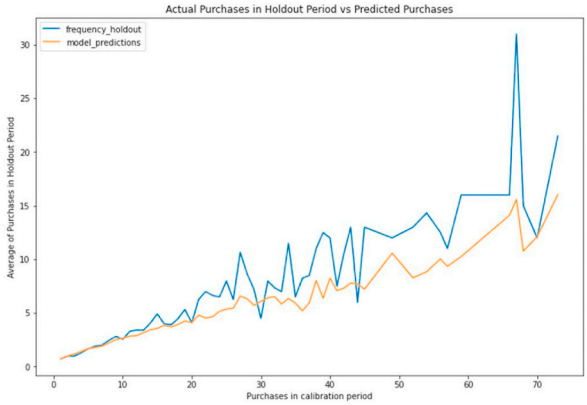


Fig. 9. Comparison between actual and estimated purchases during the extension period.

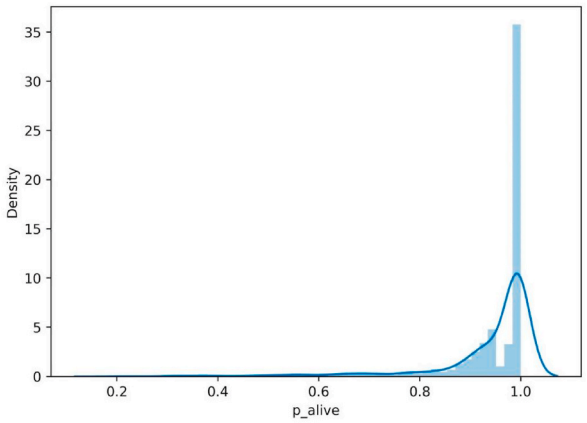


Fig. 10. Customer survival curve.

Table 8
Average future transaction value and LTV of the top 20 customers in the next 12 months.

CustomerID	Monetary_value	Predicted_purchases	P_alive	Predicted_Sales	LTV
16446.00	168469.60	0.27	0.95	78051.29	231839.79
14646.00	6366.71	2.98	1.00	6213.28	207563.20
18102.00	9349.48	1.70	1.00	8951.86	170553.51
17450.00	7404.69	1.76	1.00	7105.09	140365.47
14096.00	4071.43	2.56	1.00	3823.47	109633.22
14911.00	1093.66	8.27	1.00	1087.70	101031.68
12415.00	7860.21	1.14	1.00	7323.92	93502.77
14156.00	2787.08	2.74	1.00	2723.13	83911.25
17511.00	3305.06	1.80	1.00	3185.63	64389.23
16029.00	2034.81	2.47	0.99	1986.70	55086.32
16684.00	4394.29	1.09	1.00	4107.72	50134.14
13694.00	1888.16	2.24	1.00	1839.43	46175.80
15311.00	677.73	5.66	1.00	674.36	42847.99
13089.00	893.71	4.20	1.00	885.32	41748.31
17949.00	2042.73	1.87	1.00	1977.81	41531.17
15769.00	2674.33	1.43	1.00	2555.71	41106.12
14298.00	2023.09	1.66	1.00	1948.76	36273.71
14088.00	3904.26	0.90	1.00	3570.38	36039.42
15061.00	1611.67	1.87	1.00	1563.96	32771.31
12931.00	3807.18	0.83	1.00	3482.59	32506.03

loyalty can be comprehensively improved so that enterprises can obtain more profits.

In terms of customer type segmentation, this paper combines RFM model with machine learning algorithm, introduces the knowledge discovery method in data mining, and constructs a customer value segmentation recognition model under the condition of non contractual relationship. The sample data in this paper is a typical unbalanced dataset, which improves the performance of the algorithm on the unbalanced dataset. In addition, by comparing four mainstream cross-validation methods, this paper finally selects the leave-one-out cross-validation method that is more conducive to the performance of the model.

Based on the theory of customer lifetime value, this paper subdivides customers from the perspective of customer value. Based on a machine learning algorithm and CRM analysis model, this paper constructs a customer value segmentation identification model under the condition of a noncontractual relationship and conducts an empirical analysis using the real customer transaction data of an actual online shopping platform, which verifies the effectiveness and applicability of the customer segmentation method and value calculation method proposed in this paper. In the era of online shopping becoming increasingly popular, this paper formulates the corresponding CRM strategy according to the customer group. Using an accurate CRM strategy can greatly tap the consumption potential of consumers in the promotion period and promote the growth of commodity consumption; it can enable enterprises to optimize sales and pricing, maximize benefits, contribute to the competitive advantage of enterprises in the industry, and promote the sustainable development of enterprises.

5. Conclusion and outlook

This paper studies the connotation of customer lifetime value under the noncontractual relationship from two aspects of customer segmentation and calculation methods, and constructs the customer segmentation mode under the noncontractual relationship and the customer value segmentation recognition model based on machine learning algorithm. This paper uses feature engineering to process the data, and comprehensively applies RFM model, machine learning algorithm and other methods to ensure the rationality and feasibility of the research results. The main innovation points of this paper are as follows: (1) Define the measurement method of customer value, build a customer segmentation model based on customer value under the noncontractual relationship, mine the characteristics of segmented customer groups, and enrich the existing research on customer value segmentation; (2) In the preprocessing stage of customer consumption behavior information, feature engineering is used to map the original data space to the new feature vector space through the data preprocessing and transformation process, so that in the new feature space, the machine learning model can better mine the feature vector information and improve the classification results of the model; (3) The traditional RFM model and machine learning algorithms are combined to expand the research on customer segmentation and customer lifetime value under the condition of noncontractual relationship.

Customer value segmentation is of great significance to improve the management ability of enterprises. However, due to the limitation of time and data attribute information, the research on some problems in this paper has not been further deepened: First, the division of lost customers and new customers has not been deeply analyzed; Secondly, the characteristic information of customer heterogeneity, such as customers' different occupations, income, personalized needs, consumption ability and so on, is not considered. Therefore, future research can focus on the following aspects: First, divide new customers and lost customers, observe the growth path of new customers, find out the reasons for losing customers, and help product decision-making; Secondly, consider mining market segments from multiple dimensions, such as customer occupation, actual consumption ability, etc; Thirdly, we should build customer portraits according to customer consumption information and conduct accurate research on product recommendations.

Author contribution statement

Yuechi Sun: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Haiyan Liu: Conceived and designed the experiments; Contributed reagents, materials, analysis tools or data; Wrote the paper.

Yu Gao: Conceived and designed the experiments; Performed the experiments; Analyzed and interpreted the data.

Funding statement

This work was supported by the National Natural Science Foundation of China [71673256].

Data availability statement

Data associated with this study has been deposited at https://archive.ics.uci.edu/ml/datasets/Online+Retail+II.

Declaration of interest's statement

The authors declare no competing interests.

References

[1] M. Mohan, M.W. Nyadzayo, R. Casidy, Customer identification: the missing link between relationship quality and supplier performance, Ind. Market. Manag. 97 (2021) 220–232.
[2] D. Simões, J. Nogueira, Learning about the customer for improving customer retention proposal of an analytical framework, J. Mark. Anal. 10 (2022) 50–63.
[3] F. Safari, N. Safari, G.A. Montazer, Customer lifetime value determination based on RFM model, Market. Intell. Plann. 34 (2016) 446–461.
[4] Y. Li, Y. Zhang, F. Luo, W. Zou, Y. Zhang, K. Zhou, Customer tiered purchase forecast by mobile edge computing based on Pareto/NBD and SVR, China Commun. 18 (2021) 1–10.
[5] Z. Li, X. Zhang, G. Wang, D. Shi, W. Shi, Research on user value prediction model based on Pareto/NBD, J. Beijing Univ. Posts Telecommun. 22 (2020) 7–14.
[6] Y. Sun, D. Cheng, S. Bandyopadhyay, W. Xue, Profitable retail customer identification based on a combined prediction strategy of customer lifetime value, Midwest Soc. Sci. J. 24 (2021) 104–127.
[7] H.T. Tsou, Y.W. Huang, Empirical study of the affecting statistical education on customer relationship management and customer value in hi-tech industry, Eurasia J. Math. Sci. Technol. Educ. 14 (2018) 1287–1294.
[8] K.K. Hari Kunasekaran, Y. Zheng, W. Wang, Research on customer relationship management based on data mining, Asia-Pacific, J. Converg. Res. Interchang. 6 (2020) 61–77.
[9] J.Y. Young, The effects of customer's perceived value of the Korean restaurants on customer satisfaction and behavioral intention, J. Hosp. Tour. Stud. 19 (2017) 101–122.
[10] A.M. Estrella-Ramón, M. Sánchez-Pérez, G. Swinnen, K. VanHoof, A marketing view of the customer value: customer lifetime value and customer equity, South Afr. J. Bus. Manag. 44 (2013) 47–64.
[11] H. Zhang, X. Liang, S. Wang, Customer value anticipation, product innovativeness, and customer lifetime value: the moderating role of advertising strategy, J. Bus. Res. 69 (2016) 3725–3730.
[12] R.T. Rust, V.A. Zeithaml, K.N. Lemon, Driving customer equity: how customer lifetime value is reshaping corporate strategy/r.t. rust, v.a. zeithaml, k.n. lemon, J. Market. 68 (2000) 109–127.
[13] B.B. Jackson, Build customer relationships that last, Harv, Bus. Rev. 63 (1985) 120–130.
[14] J. Xiao, X. Liu, L. Xie, D. Liu, J. Huang, A cost-sensitive semi-supervised ensemble model for customer targeting, Chinese J. Manag. Sci. 26 (2018) 186–196.
[15] J. Xiao, D. Liu, X. Gu, S. Wang, Dynamic classifier ensenmble selection model for bank customer's credit scoring, J. Manag. Sci. China. 18 (2015) 114–126.
[16] V. Kumar, W. Reinartz, Creating enduring customer value, J. Market. 80 (2016) 36–68.
[17] M. Clemente-Císcar, S. San Matías, V. Giner-Bosch, A methodology based on profitability criteria for defining the partial defection of customers in non-contractual settings, Eur. J. Oper. Res. 239 (2014) 276–285.
[18] H. Abbasimehr, M. Shabani, A new methodology for customer behavior analysis using time series clustering, Kybernetes 50 (2021) 221–242.
[19] L. Paul, T.R. Ramanan, An RFM and CLV analysis for customer retention and customer relationship management of a logistics firm, Int. J. Appl. Manag. Sci. 11 (2019) 333–351.
[20] J.S. Thomas, R.C. Blattberg, E.J. Fox, Recapturing lost customers, J. Market. Res. 41 (2004) 31–45.
[21] L. Yang, Z. Bai, Y. Kou, Random forest algorithm based on RFM model for civil aviation customer churn analysis, Comput. Mod. (2021) 100–104.
[22] T. Gattermann-Itschert, U.W. Thonemann, Proactive customer retention management in a non-contractual B2B setting based on churn prediction with random forests, Ind. Market. Manag. 107 (2021) 134–147.
[23] H. Koosha, A. Albadvi, Customer lifetime valuation using real options analysis, J. Mark. Anal. 3 (2015) 122–134.
[24] S. Gupta, D. Hanssens, B. Hardie, W. Kahn, V. Kumar, N. Lin, N. Ravishanker, S. Sriram, Modeling customer lifetime value, J. Serv. Res. 9 (2006) 139–155.
[25] H. Castéran, L. Meyer-Waarden, W. Reinartz, Modeling customer lifetime value, retention, and churn, in: Handb. Mark. Res., Springer International Publishing, Cham, 2017, pp. 1–33.
[26] A.B. Çavdar, N. Ferhatosmanoğlu, Airline customer lifetime value estimation using data analytics supported by social network information, J. Air Transport. Manag. 67 (2018) 19–33.
[27] M. Óskarsdóttir, B. Baesens, J. Vanthienen, Profit-based model selection for customer retention using individual customer lifetime values, Big Data 6 (2018) 53–65.
[28] S. Monalisa, P. Nadya, R. Novita, Analysis for customer lifetime value categorization with RFM model, Procedia Comput. Sci. 161 (2019) 834–840.
[29] C. Fang, H. Liu, Research and application of improved clustering algorithm in retail customer classification, Symmetry 13 (2021) 1789.
[30] N. Lozada, J. Arias-Pérez, G. Perdomo-Charry, Big data analytics capability and co-innovation: an empirical study, Heliyon 5 (2019), e02541.
[31] A.A. Morán-Reyes, Towards an ethical framework about big data era: metaethical, normative ethical and hermeneutical approaches, Heliyon 8 (2022), e08926.
[32] D. Wiltshire, S. Alvanides, Ensuring the ethical use of big data: lessons from secure data access, Heliyon 8 (2022), e08981.
[33] S. Erevelles, N. Fukawa, L. Swayne, Big data consumer analytics and the transformation of marketing, J. Bus. Res. 69 (2016) 897–904.
[34] J. Vijaya, E. Sivasankar, Computing efficient features using rough set theory combined with ensemble classification techniques to improve the customer churn prediction in telecommunication sector, Computing 100 (2018) 839–860.
[35] K. Coussement, D.F. Benoit, D. Van den Poel, Preventing customers from running away! Exploring generalized additive models for customer churn prediction, in: Sustain. Glob. Marketpl., , Springer International Publishing, Cham, 2015, 238–238.
[36] J. Zhang, M. Zhi, Y. Zhang, Combined generalized additive model and random forest to evaluate the influence of environmental factors on phytoplankton biomass in a large eutrophic lake, Ecol. Indicat. 130 (2021), 108082.
[37] D. Boughaci, A.A. Alkhawaldeh, Three local search-based methods for feature selection in credit scoring, Vietnam J. Comput. Sci. 5 (2018) 107–121.
[38] Z.Y. Chen, Z.P. Fan, Dynamic customer lifetime value prediction using longitudinal data: an improved multiple kernel SVR approach, Knowl. Base Syst. 43 (2013) 123–134.
[39] W. Ma, G. Xia, Prediction model of customer churn based on deep neural network, Comput. Technol. Dev. 29 (2019) 76–80.
[40] T. Vafeiadis, K.I. Diamantaras, G. Sarigiannidis, K.C. Chatzisavvas, A comparison of machine learning techniques for customer churn prediction, Simulat. Model. Pract. Theor. 55 (2015) 1–9.
[41] A. Martínez, C. Schmuck, S. Pereverzyev, C. Pirker, M. Haltmeier, A machine learning framework for customer purchase prediction in the non-contractual setting, Eur. J. Oper. Res. 281 (2020) 588–596.
[42] H. Jain, A. Khunteta, S. Srivastava, Telecom churn prediction and used techniques, datasets and performance measures: a review, Telecommun. Syst. 76 (2021) 613–630.
[43] P.P. Chen, A. Guitart, A.F. del Rio, A. Perianez, Customer lifetime value in video games using deep learning and parametric models, in: 2018 IEEE Int. Conf. Big Data (Big Data), IEEE, 2018, pp. 2134–2140.
[44] K. Coussement, S. Lessmann, G. Verstraeten, A comparative analysis of data preparation algorithms for customer churn prediction: a case study in the telecommunication industry, Decis. Support Syst. 95 (2017) 27–36.
[45] D. Cheng, Y. Sun, W. Xue, Robustness measurement of non contractual customers' lifetime value: a study on the comprehensive calculation of classical methods and machine learning algorithms, Manag. Rev. 31 (2019) 83–98.
[46] E.M. –AMIK BSI Yogyakarta, Komparasi metode clustering k-means dan k-medoids dengan model fuzzy RFM untuk pengelompokan pelanggan, Evolusi J. Sains Dan Manaj. 6 (2018) 106–113.
[47] S.S. Prasetyo, M. Mustafid, A.R. Hakim, Penerapan fuzzy c-means kluster untuk segmentasi pelanggan e-commerce dengan metode recency frequency monetary (RFM), J. Gaussian 9 (2020) 421–433.
[48] Y. Ming, W. Zhang, Z. Huang, X. Chen, Customer segmentation based on RFM purchase tree, J. Shenzhen Univ. Sci. Eng. 34 (2017) 306.
[49] A.J. Christy, A. Umamakeswari, L. Priyatharsini, A. Neyaa, RFM ranking – an effective approach to customer segmentation, J. King Saud Univ. – Comput. Inf. Sci. 33 (2021) 1251–1257.

[50] B. Wu, L. Yang, Y. Chen, An empirical study of purchase rate and dropout rate between mobile and PC customers, J. Syst. Manag. 29 (2020) 924–933.
[51] C. Wang, F. Lin, Forecast and analysis of customer economic value in retail industry based on commodity category, J. Commer. Econ. (2018) 55–58.
[52] P. Jasek, L. Vrana, L. Sperkova, Z. Smutny, M. Kobulsky, Comparative analysis of selected probabilistic customer lifetime value models in online shopping, J. Bus. Econ. Manag. 20 (2019) 398–423.
[53] D. Krstevski, G. Mancheski, Managerial accounting modeling customer lifetime value: an application in the telecommunication industry, Eur. J. Bus. Soc. Sci. 5 (2016) 64–77.
[54] T.T. Wong, Performance evaluation of classification algorithms by k-fold and leave-one-out cross validation, Pattern Recogn. 48 (2015) 2839–2846.