

# 基于灰色关联度和 K-Means++ 的电子商务客户价值分类<sup>①</sup>



冀慧杰, 倪 枫, 刘 姜, 赵 焱

(上海理工大学 管理学院, 上海 200093)

通讯作者: 倪 枫, E-mail: [nifeng@usst.edu.cn](mailto:nifeng@usst.edu.cn)

**摘 要:** 现有的模型大多采用 RFM 模型和 K-means 对客户价值进行分类, 对指标权重的确定大多采用 AHP 法, 没有考虑到 RFM 模型指标相互之间的联系. 首先根据 RFM 模型选择平均购买时间间隔, 客户在一定时间内的购买频率, 平均每笔订单交易金额和客户的活跃时间构造 RFMT 模型来衡量客户价值. 其次使用灰色关联度确定各指标权重. 最后针对 K-means 的缺点, 运用改进 K-means (K-means++) 和肘部法则对 RFMT 模型进行聚类分析. 该模型能对客户群进行更加细致的划分, 既能帮助电子商务企业识别出需要重点关注的客户即已流失客户和新客户群体, 同时将该企业客户划分为价值由高到低的客户群, 对不同客户群提出具体的营销建议.

**关键词:** 客户价值; RFMT 模型; 灰色关联度; K-means++; 肘部法则

引用格式: 冀慧杰, 倪枫, 刘姜, 赵焱. 基于灰色关联度和 K-Means++ 的电子商务客户价值分类. 计算机系统应用, 2020, 29(9): 249-254. <http://www.c-s-a.org.cn/1003-3254/7562.html>

## Classification of E-Commerce Customer Value Based on Grey Correlation Degree and K-Means++

Ji Hui-Jie, Ni Feng, Liu Jiang, Zhao Yi

(Business School, University of Shanghai for Science & Technology, Shanghai 200093, China)

**Abstract:** The combine model of the RFM model and K-means is used to classify customer value and AHP method is mostly used to determine the weight of indicators, without considering the relationship between the indicators of RFM model. In this study, firstly, we select the average time interval, the customer purchase frequency in a period of time, average transaction money of each order, and customer active time to structure RFMT model in order to measure the customer value. Then, determine the index weight by using grey correlation degree. Finally, aiming at the shortcomings of K-means, K-means ++ and elbow law are used to carry out cluster analysis of RFMT model. This model can make a more detailed division of customer base. It can help e-commerce enterprises to identify the customers that need to be focused on. Meanwhile, the enterprise customers can be divided into customer groups with high value to low value, and put forward specific marketing suggestions for different customer groups.

**Key words:** customer value; RFM model; grey correlation degree; K-means++; elbow law

客户关系到企业的生存与发展, 赢得客户的青睐就能获得市场. 市场营销理念在发展过程中越来越重视客户的需求, 以顾客为核心, 把满足顾客需求作为首要任务. 不断提高顾客满意度, 培养忠诚客户对企业至

关重要. 根据中国电子商务报告, 2018 年我国电子商务交易规模继续扩大并保持高速增长态势, 全年实现电子商务交易额 31.63 万亿元, 同比增长 8.5%. 随着网购消费者数量逐渐增加, 对客户的管理更加有必要而且

① 基金项目: 国家自然科学基金 (11701370)

Foundation item: National Natural Science Foundation of China (11701370)

收稿时间: 2020-01-07; 修改时间: 2020-02-13; 采用时间: 2020-03-11; csa 在线出版时间: 2020-09-04

方便有效. 通过细分客户, 分析每一类客户群体的购买行为以获得不同客户群体的价值, 有利于企业实施差异化营销策略. 网络的发展真正实现了精准营销, 电子商务企业可以方便地获得客户的购买记录, 包括购买内容、购买时间以及购买金额等, 通过客户购买数据可以分析客户偏好, 准确衡量客户价值, 制定出适合相应细分客户群的营销策略.

RFM模型可用来衡量客户利润贡献度和忠诚度, 广泛应用于零售及其它行业. Chen等使用RFM模型对在线零售商的客户进行分析, 通过聚类算法和决策树识别每个客户群的主要特征<sup>[1]</sup>. 徐翔斌等提出改进的RFP模型, 用利润指标替换金额指标对电子商务网站客户进行细分, 但是对指标设定的权重是随机的, 缺乏一定的客观性<sup>[2]</sup>. 徐文瑞用层次分析法(AHP)确定RFM模型中的各指标权重<sup>[3]</sup>. 包志强等通过改进RFM模型对百度外卖客户进行价值分析, 用主成分分析确定各指标权重<sup>[4]</sup>. 熊兰等通过对零售企业的产品类别进行区分, 将客户的价值按产品类别区分, 创建多层级的客户价值模型<sup>[5]</sup>. Zhang细化了RFM指标, 利用K-means聚类算法为体育文具产业进行客户数据挖掘给予指导<sup>[6]</sup>. 综合各学者研究来看, RFM模型仍是根据客户消费行为衡量客户价值的有效方法, 本文借鉴以往学者对RFM模型各指标的改进并结合本文研究构建RFMT模型. 其次, 以往文献关于权重的确定大多采用层次分析法(AHP), 有些学者会采用客观法例如主成分分析和熵权法, 然而皆未考虑RFM模型各指标之间存在关联的问题. 本文基于灰色系统理论利用灰色关联度为改进的RFMT模型指标确定权重, 综合整个因素指标空间的影响, 分析各指标对衡量客户价值的影响程度大小. 最后, 由于K-means算法存在对初始聚类中心以及聚类数敏感等缺点, 本文运用K-means++和肘部法则对其进行改进, 进而对在线零售商的客户的购买行为进行分析, 并将客户分为价值由高到低的客户群.

## 1 模型建立

企业在进行客户关系管理时, 经常使用RFM模型来分析客户价值, 尤其是零售业的客户管理. RFM模型通过3个指标来衡量客户的价值, 即Recency(最近购买时间)、Frequency(购买频率)以及Monetary(购买金额). 其中最近购买时间 $R$ 指的是客户最后一次购

买产品的时间与对客户购买行为分析时间的间隔, 间隔越小表明客户对产品印象越深刻, 因此对广告、促销活动等反应越快, 更容易发生重复购买行为; 购买频率 $F$ 表明客户在一段时间内购买产品的次数, 频率越大意味着客户频繁购买产品, 对企业品牌或该类产品已形成偏好, 更容易发展为忠诚客户; 购买金额 $M$ 是客户直接为企业带来的价值. 高价值客户往往购买频率高、购买间隔时间短以及购买金额高.

### 1.1 改进的RFM模型

RFM模型的优点在于收集数据容易, 3个指标能够清楚地解释顾客的行为. RFM模型能够清楚表征客户当前价值, 然而在如今市场竞争激烈的情况下, 客户的潜在价值即未来能够带给企业的价值更加重要. 王客富认为客户的保持时间与忠诚度有密切联系<sup>[7]</sup>. 客户的忠诚度则是客户潜在价值的体现, 因此扩展RFM模型对客户忠诚度进行度量. 传统的RFM模型指标不够全面, 并且指标的权重确定大多以主观判断为主, 影响结果的客观性. 为了更准确地对在线零售商的客户进行当前和潜在价值分析, 提出改进的RFM模型, 选取客户的平均购买时间间隔( $R$ )、客户在活跃时间内的购买频率( $F$ )、平均每笔订单交易金额( $M$ )和客户在一年内的活跃时间( $T$ )4个指标构建RFMT模型. 为了抵消购买频率与购买金额之间的多重共线性关系, 用平均订单交易金额代替总购买金额; 客户的购买频率描绘客户在一年内的订单数量, 订单数量越多, 客户的购买金额越大, 为企业贡献的价值越大, RFM的3个指标衡量了客户的当前价值.  $F$ 和 $T$ 指标是对客户潜在价值的衡量, 客户的活跃时间能够反映客户对企业的忠诚度, 活跃时间越长, 购买频率越大, 说明客户的满意度越高, 越容易发生购买行为. 所构造的RFMT模型各指标的计算公式如下:

$$R = \frac{T_L - T_F}{BT} \quad (1)$$

$$M = \frac{M_T}{F} \quad (2)$$

$$T = T_L - T_F \quad (3)$$

其中,  $BT$ 表示客户在活跃时间内发生购买行为的天数;  $M_T$ 表示总购买金额;  $T_L$ 表示最远一次购买时间;  $T_F$ 表示最近一次购买时间.

### 1.2 基于灰色关联度确定指标权重

RFM模型中各指标对客户价值的影响程度是不

同的,通过对RFM模型中的指标赋予不同的权重能够帮助企业更准确地分析客户行为,采取科学的管理措施.人们所面临的实际决策问题的各因素指标之间的关系表面上看不明确,实际上却并不是独立的,常存在相互联系,这是灰色的关联性在起作用<sup>[8]</sup>.灰色关联分析的基本思想是通过确定参考数据列和比较数据列的几何形状相似程度来判断其联系是否密切,通常可以运用此方法来分析各个因素对于结果的影响程度.采用灰色关联法确定指标权重的优点在于其可以挖掘出数据本身具有的关联性,从而避免了如层次分析法等所包含的主观因素的影响.本文在衡量电子商务客户价值时所采用的实验数据是有限的,根据以往文献对客户价值的衡量在有限的信息下提炼出了RFMT的4个指标.RFMT这4个指标是根据客户购买时间、客户购买金额以及客户的订单量4个指标计算得到,由计算公式可得各个指标相互之间存在关联,RFMT模型及指标关联性如图1所示.各指标在时间和购买行为上存在一定的关联性,因此可以将客户价值看作是一个灰色系统,通过灰色关联度寻求RFMT各指标对客户价值的影响程度.通过灰色关联度<sup>[9,10]</sup>确定各指标权重,指标的权重值即代表了对客户价值的影响程度.

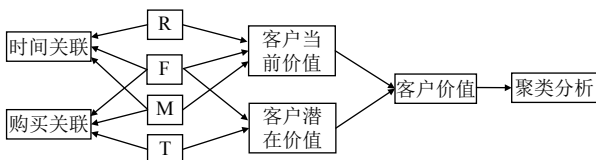


图1 RFMT客户价值分类模型

利用灰色关联度确定指标权重的计算方法如下:

步骤1. 根据数据处理结果选取购买时间间隔、购买频率、平均购买金额以及客户活跃时间4个指标变量形成矩阵:

$$(X'_1, X'_2, \dots, X'_n) = \begin{pmatrix} X'_1(1) & \cdots & X'_1(4) \\ \vdots & \ddots & \vdots \\ X'_n(1) & \cdots & X'_n(4) \end{pmatrix} \quad (4)$$

步骤2. 从4个指标中选出各自的最优值作为参考数据列,最优值代表最理想的客户类型,即客户价值最高.记为:

$$X'_0 = (X'_0(1), X'_0(2), X'_0(3), X'_0(4)) \quad (5)$$

步骤3. 指标数据值的范围相差较大,采用均值化

方法进行无量纲化处理.无量纲化后的数据形成新的矩阵:

$$x_i(k) = \frac{x'_i(k)}{\frac{1}{m} \sum_{k=1}^m x'_i(k)} \quad (6)$$

$$(X_0, X_1, \dots, X_n) = \begin{pmatrix} X_0(1) & \cdots & X_0(4) \\ \vdots & \ddots & \vdots \\ X_n(1) & \cdots & X_n(4) \end{pmatrix} \quad (7)$$

步骤4. 计算第*i*个数据的第*k*个指标与参考序列相比较的关联系数:

$$\xi_{ik} = \frac{\min_i \min_k |X_0(k) - X_i(k)| + \rho \cdot \max_i \max_k |X_0(k) - X_i(k)|}{|X_0(k) - X_i(k)| + \rho \cdot \max_i \max_k |X_0(k) - X_i(k)|} \quad (8)$$

其中,  $k = 1, 2, 3, 4$ ,  $\rho$ 为分辨系数,  $0 < \rho < 1$ , 一般取 $\rho = 0.5$ .

$$\text{因此得到关联系数矩阵: } \begin{pmatrix} \xi_{11} & \cdots & \xi_{14} \\ \vdots & \ddots & \vdots \\ \xi_{n1} & \cdots & \xi_{n4} \end{pmatrix} \quad (9)$$

步骤5. 对每一个指标的关联系数求平均值:

$$\bar{W}_k = \frac{1}{n} \sum_{i=1}^n \xi_{ik}, k = 1, 2, 3, 4 \quad (10)$$

步骤6. 做归一化处理即可得到各指标的权重:

$$W_k = \bar{W}_k / \sum_{k=1}^4 \bar{W}_k \quad (11)$$

## 2 改进 K-means 聚类

K-means 算法是数据挖掘最经典的聚类算法,也是客户细分经常使用的方法.算法的基本思想是根据样本之间的距离分类.先指定所需分类类别  $K$  类,随机生成  $K$  个聚类中心,将各个样本分到离聚类中心最近的一类,通过不断迭代更新聚类中心,达到最优分类效果.算法采用误差平方和准则函数 SSE 作为聚类准则函数.K-means 算法的缺点是对初始聚类中心的选取是随机的,聚类的结果和所需时间依赖初始质心的选择,并且需要预先设定聚类数.为了解决 K-means 对初始点敏感的问题,Arthur 等提出了 K-means++ 算法,对 K-means 随机选择初始聚类中心的方法进行了改进,即在选取初始聚类中心时使各聚类中心点之间的距离尽可能远,结果证明 K-means++ 能显著改善分类结果的最终误差<sup>[11]</sup>.为了改善聚类数设定的问题,本文利用 SSE 评价聚类结果,通过肘部法则确定最佳聚类数.算



法的具体步骤如下:

$$SSE = \sum_{i=1}^k \sum_{x \in C_i} \|x - \mu_i\|^2 \quad (12)$$

其中,  $x$  为每一簇内的样本点,  $\mu_i$  表示第  $i$  个簇的聚类中心点。

#### 算法 1. K-means++ 算法

- 1) 从数据集中随机选取一个样本作为初始聚类中心;
- 2) 首先计算每个样本与当前已有聚类中心之间的最短距离(即与最近的一个聚类中心的距离), 用  $D(x)$  表示; 接着计算每个样本被选为下一个聚类中心的概率  $\frac{D(x)^2}{\sum_{x \in X} D(x)^2}$ ;
- 3) 重复第 2) 步直到选择出共  $K$  个聚类中心.
- 4) 使用选出的聚类中心用 K-means 聚类.
- 5) 比较不同聚类数得到的  $SSE$  值, 利用肘部法则确定最终聚类数目.

## 3 实验分析

### 3.1 实验数据

数据集从 UCI Machine Learning Repository 数据库中获得, 这是一个跨国数据集. 该在线零售商总部位于英国, 成立于 1981 年, 主要销售各种场合的独特礼品. 该企业的主要客户来自英国和欧洲各地, 之前主要通过邮寄目录, 通过电话下订单的模式经营. 目前, 该企业已完全转向了电子商务, 积累了大量客户的数据, 因此可根据积累的数据集对客户进行管理. 该数据集其中包含 2010 年 12 月 1 日至 2011 年 12 月 9 日之间在英国注册的非商店在线零售的所有交易. 该公司的许多客户都是批发商, 因此客户与企业的每笔交易中可能包含多种商品, 每笔交易有唯一的发票编号, 默认为一次购买. 数据集包含了 4296 个客户在一年内的订单数据, 共有 406 829 条交易记录, 包含 7 个变量, 各变量含义如表 1 所示. Python 是目前最流行的用来做数据分析与数据挖掘的语言, 本文利用 Python 软件进行实验. 对该数据集处理, 构造 RFMT 模型以及进行聚类分析, 验证所提客户价值分类模型的可行性, 为电子商务企业的客户管理提供参考.

### 3.2 数据预处理

首先从原始数据集中提取出有用的信息经过分析得到的 RFMT 数据表如表 2 所示. 如编号为 12346 的客户, 平均购买时间间隔为 1 天, 购买频率为 4 次, 平均购买金额为 22.25, 活跃时间为 1 天, 说明该客户在一天内发生了 4 次购买交易.

表 1 数据集的变量描述

变量名	变量含义
发票编号	编号唯一, 每笔交易包含多个产品
产品编号	每个产品的编号唯一
数量	每笔交易中每个产品的数量
价格	每个产品的单位价格
客户ID	客户编号
发票日期	每笔交易生成的日期和时间
国家	收获地址, 即客户所在国家

表 2 RFMT 指标数据样本

客户ID	$R$	$F$	$M$	$T$
12346	1	4	22.25	1
12347	52.5	151	23.83	315

其次利用灰色关联度确定各指标权重. 从 4 个指标中选出最优值即从购买时间间隔数据中选取最小值, 从购买频率、平均购买金额以及客户活跃时间数据中分别选出最大值组成参考数据列, 参考数据列各指标构成了企业理想的客户即价值最高的客户. 经过灰色关联分析后得出的各指标权重值如表 3 所示.

表 3 灰色关联度确定指标权重值结果

指标	$R$	$F$	$M$	$T$
权重值	0.201	0.318	0.165	0.316

表 3 结果表明, 指标  $F$  和  $T$  的权重值相对较大, 平均购买金额  $M$  的权重最小. 分析可知该权重值结果以客户潜在价值最大是影响客户价值的最重要因素. 对电子商务企业来说, 客户的潜在价值即客户忠诚度才是最重要的. 客户平均购买金额与购买频率有关, 购买频率越大则总购买金额越大, 这在实际中反映了购买金额的大小不能准确体现出客户的价值, 即一次大金额的购买行为可能没有多次频繁小批量购买的客户带来的价值更高, 因此购买频率更能体现出客户的当前价值. 客户活跃时间越久则忠诚度越高, 流失倾向越低,  $T$  指标代表了客户价值中的潜在价值, 因此  $T$  指标的权重值较高. 平均购买时间间隔与购买频率息息相关, 购买频率越多则平均购买时间间隔越小,  $R$  指标越低, 意味着客户需求量大且对企业产品形成偏好. 分析权重结果及各指标的实际意义, 平均购买金额越大的购买频率和活跃时间对客户价值高低影响最大, 平均购买金额影响最小.

最后由于选取的平均购买时间间隔、购买频率、平均购买金额以及客户活跃时间 4 个指标的意义不

同、数据范围差异较大,为了消除对聚类结果的影响,对数据按下列公式进行无量纲化处理.

$$x'_i = \frac{x_i - x_{\min}}{x_{\max} - x_{\min}} \quad (13)$$

### 3.3 基于 K-means++的聚类结果

本文首先使用 K-means++算法进行聚类,设置分类数目  $K$  的范围为  $[1, 8]$ , 根据聚类代价函数即误差平方和  $SSE$  绘制分类效果图如图 2 所示. 然后根据肘部法则聚类数应为  $K=2$ , 但分为两个簇时误差较大,考虑到对客户分群的需要,确定聚类数为  $K=4$ , 这时的误差平方和较小且能对客户群进行细致的划分. 设定聚类数目为 4 时,分群结果及客户价值排名如表 4 所示.

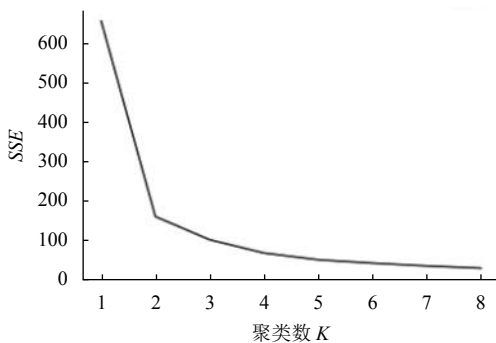


图 2 不同聚类数目下 SSE 值

表 4 聚类结果

类别	聚类个数	聚类中心点				价值排名
		R	F	M	T	
客户群1	2100	0.025338	0.004590	0.529550	0.031900	4
客户群2	831	0.189331	0.032259	0.529951	0.845631	1
客户群3	452	0.519005	0.008978	0.529595	0.714772	3
客户群4	913	0.237222	0.010138	0.529693	0.424273	2

根据聚类结果每一客户群的聚类中心点并结合上节各指标权重值对每一客户群体的消费行为进行分析,发现客户群在平均购买金额上没有体现出太大的差异,但结合频率来看总购买金额差异较大.这也正好说明了该企业需要关注的是客户的购买频率和购买时间间隔.表 4 结果表明在平均购买时间间隔、购买频率和客户活跃时间 3 个指标上各客户的行为差异显著.

客户群 1 在购买频率和活跃时间上是所有群体中表现最差的,但是平均购买时间间隔是最小的,结合购买时间间隔和活跃时间来看可以说明客户群 1 的客户处于两种状态:已流失和新客户.客户群 2 的活跃时间

是最长的、购买频率和购买金额最高,平均购买时间间隔也较短.客户群 3 的活跃时间较长,但是在平均购买时间间隔和购买频率上表现均较差.客户群 4 在平均购买时间间隔上和购买频率上表现较为显著,但是活跃时间较短.

本文根据该实验数据集进行聚类分析得出的结果,结合对每一客户群各个指标的分析,按照价值排名和客户群的行为特征将该企业客户分别定义为:重要保持客户、重要发展客户、重点关注客户和一般价值客户.

对重要保持客户,即客户群 2. 该客户群体人数为 831 人,他们价值排名第一,是企业所有客户中最为理想的客户类型. 该客户群的生命周期最长、购买最频繁、交易的金额也最大,说明该客户群是满意度最高的,已对该企业产品形成偏好,能持续为企业带来价值. 面向该类客户群,企业需要积极提供售后服务,定期回访与关怀,主要是向该客户咨询其关于产品和服务的看法,对该客户群采取服务差异化的策略,维持好与该客户群的关系,发展该客户群成为企业的忠诚客户群.

对重要发展客户,即客户群 3. 该客户群价值排名第二,有 913 人. 该客户群在平均购买时间间隔、购买频率和购买金额上仅次于重要保持客户群,但活跃时间较低. 为避免该客户群发生流失行为,企业应该投入更多的精力和资源去关注这类客户,面向该客户群体进行宣传以实现精准营销,并给予该客户群一些购买优惠,促使该客户发生持续购买行为,延长该类客户的生命周期,使该客户群发展为企业的高价值客户.

对重点关注客户,即客户群 1. 客户群 1 的人数是最多的,同时因为此类客户处于两个状态:已流失或者新客户,这说明企业中存在大量流失的客户或新增很多新客户. 对电子商务企业来说,这两类客户至关重要. 客户流失率较大的话,企业很难培养出庞大的忠诚客户群,同时该企业的口碑会受到影响. 新客户较多时,企业应该尽量保留住该客户群. 如果流失客户人数占比较多,企业需要关注导致客户流失的原因,分析流失客户群购买的产品,发现问题后及时调整其产品和服务策略,降低流失率;如果新客户数量居多,企业需要分析是由于特殊节日还是最近的营销活动所带来的新客户,同时在未来的一段时间里需要经常面向这些新客户宣传,在客户心中留下印象,关注该类客户的动态,促使新客户发生重复购买行为,使其成为长期客户.

对一般价值客户,即客户群4.虽然他们的活跃时间相对较长,但购买频率较少,平均购买时间间隔也较长,说明该类客户不经常发生购买行为,只在有需要的时候会想到该企业并发生购买行为.且该客户人数较少,为企业带来的价值较低,是低价值客户,因此企业只需做好广告营销,不用投入过多精力去关注该客户群.

#### 4 结语

电子商务企业能够收集到丰富的客户购买行为信息,通过对客户行为信息进行分析有助于企业更好地实施营销活动.针对现有模型普遍基于RFM和K-means对客户价值进行分类,对指标权重的确定大多采用AHP法,没有考虑到RFM模型指标相互之间的联系.本文对真实的电子商务客户庞大的交易信息进行分析,根据数据集所包含的客户行为数据的信息,提出改进的RFMT模型选取合适指标,采用灰色关联的权重确定方法,并结合数据挖掘算法对客户进行聚类分析.使用改进的RFMT模型及灰色关联度的权重确定方法,能更好地衡量客户价值;根据肘部法则确定聚类数目,运用K-mean++聚类分析的结果误差更小,基于RFMT模型对客户群的分类更加细致,能帮助企业识别出需要重点关怀的客户和重点关注的客户,提高企业对流失客户群的防范意识和对新客户群体的重视.本文对客户价值细分的方法也适用于其它电子商务企业,该模型能够对客户价值进行合理的辨别,通过识别出不同客户群的价值及行为特征,有助于企业针对不同客户群体采用不同的客户关系管理手段,合理安排资源投入,实现效率最大化.

#### 参考文献

- 1 Chen DQ, Sain SL, Guo K. Data mining for the online retail industry: A case study of RFM model-based customer segmentation using data mining. *Journal of Database Marketing & Customer Strategy Management*, 2012, 19(3): 197-208.
- 2 徐翔斌, 王佳强, 涂欢, 等. 基于改进RFM模型的电子商务客户细分. *计算机应用*, 2012, 32(5): 1439-1442.
- 3 徐文瑞. 基于RFM模型的顾客消费行为与顾客价值预测研究. *商业经济研究*, 2017, (19): 44-46. [doi: 10.3969/j.issn.1002-5863.2017.19.015]
- 4 包志强, 赵媛媛, 赵研, 等. 基于改进RFM模型的百度外卖客户价值分析. *西安邮电大学学报*, 2019, 24(1): 105-110.
- 5 熊兰, 高炳. 基于RFM多层级客户价值模型的客户细分研究. *商业经济研究*, 2017, (5): 55-57. [doi: 10.3969/j.issn.1002-5863.2017.05.020]
- 6 Zhang L. Design of a sports culture data fusion system based on a data mining algorithm. *Personal and Ubiquitous Computing*, 2020, 24(1): 75-86. [doi: 10.1007/s00779-019-01273-6]
- 7 王克富. 基于数据挖掘技术的AFH客户分类应用研究. *技术经济与管理研究*, 2012, (11): 24-28. [doi: 10.3969/j.issn.1004-292X.2012.11.005]
- 8 邓聚龙. 灰理论基础. 武汉: 华中科技大学出版社, 2002. 87-437.
- 9 周斌. 由灰色关联度确定权重的客观多目标决策法. *昆明理工大学学报(理工版)*, 2003, 28(5): 159-161.
- 10 沈安琪, 高琳琳, 刘义坤, 等. 灰色关联与聚类分析方法在油层层系优化重组中的应用. *数学的实践与认识*, 2019, 49(9): 130-137.
- 11 Arthur D, Vassilvitskii S. K-means++: The advantages of careful seeding. *Proceedings of the Eighteenth Annual ACM-SIAM Symposium on Discrete Algorithms*. New Orleans, LA, USA. 2007. 1027-1035.