

## 客户细分的三支决策方法\*

黄顺亮<sup>1,2+</sup>, 李建林<sup>3</sup>, 王琦<sup>1</sup>

1. 山东理工大学 商学院, 山东 淄博 255000
2. 里贾纳大学 计算机系, 加拿大 里贾纳 S4S0A2
3. 南京信息职业技术学院 计算机与软件学院, 南京 210023

### Three-Way Decision Approach to Customer Segmentation\*

HUANG Shunliang<sup>1,2+</sup>, LI Jianlin<sup>3</sup>, WANG Qi<sup>1</sup>

1. Business School, Shandong University of Technology, Zibo, Shandong 255000, China
  2. Department of Computer Science, University of Regina, Regina S4S0A2, Canada
  3. School of Computer and Software, Nanjing College of Information Technology, Nanjing 210023, China
- + Corresponding author: E-mail: [huangshunliang@163.com](mailto:huangshunliang@163.com)

HUANG Shunliang, LI Jianlin, WANG Qi. Three-way decision approach to customer segmentation. Journal of Frontiers of Computer Science and Technology, 2014, 8(6): 743-750.

**Abstract:** To reduce the risk cost of customer segmentation, this paper proposes a new approach based on three-way decisions theory. This paper also constructs the customer segmentation model based on three-way decisions theory which includes constructing evaluation function, computing threshold and the procedure of application. The model balances the cost and interest of customer segmentation and is different from the traditional models. By introducing the concept of boundary, it provides the third choice in customer relation management. Finally, this paper gives an example to illustrate the procedure of application and the superiority of the new approach. The introduction of three-way decisions theory provides a new view for customer segmentation which can minimize risk cost. Under certain conditions, three-way decisions are superior to two-way decisions.

**Key words:** customer segmentation; three-way decisions; two-way decisions; evaluation index

---

\* The Humanities and Social Sciences Youth Foundation of the Ministry of Education of China under Grant No. 11YJCZH070 (教育部人文社科青年基金); the Young Scientists Award Foundation (Doctor Foundation) of Shandong Province under Grant No. BS2010DX027 (山东省优秀中青年科学家科研奖励基金(博士基金)); the Higher School Science and Technology Plan of Shandong Province under Grant No. J11LG54 (山东省高等学校科技计划); the Higher School Humanities and Social Sciences Plan of Shandong Province under Grant No. J12WF86 (山东省高等学校人文社科计划).

Received 2013-06, Accepted 2013-08.

CNKI网络优先出版 2013-10-11, <http://www.cnki.net/kcms/detail/11.5602.TP.20131011.0954.002.html>

**摘要:**为了降低客户细分中的决策风险代价,提出了一种客户细分的三支决策方法。在三支决策理论框架下,建立了客户细分模型,给出了客户价值的评价函数和三支决策阈值。该模型综合考虑了细分的成本和利润,改变了传统的二分模式,采用三支决策的边界域概念,使得在制定客户关系维系策略时有第三种选择存在。最后,通过实例分析说明了三支决策方法的应用过程和优势。三支决策理论的引入,为客户细分提供了新的思路和方法,可以最小化决策风险代价,具有很大的优越性。

**关键词:**客户细分;三支决策;二支决策;评价指标

**文献标志码:**A **中图分类号:**TP391

## 1 引言

三支决策理论是姚一豫教授在研究决策粗糙集的过程中,为合理解释粗糙集的3个域而提出的新理论<sup>[1-2]</sup>。2012年,姚一豫教授又给出了一个系统化的三支决策理论框架<sup>[3]</sup>。它是对二支决策的拓展,并具有很多二支决策没有的优势<sup>[4]</sup>。当人们掌握的信息不足以做出完全拒绝或接受决策时,增加延迟决策这一选项将有利于降低决策风险代价。三支决策理论提出后,很快引起了有关学者的注意,并进行了多方面的研究,取得了很多成果<sup>[5-7]</sup>。

贾修一等人编辑的第一本关于三支决策的专著<sup>[5]</sup>为希望介入三支决策理论研究的学者提供了很好的资源。在该书中,汇集了众多学者的研究成果。苗夺谦和高灿给出了一个基于三支决策理论的半监督学习方法。基于三支决策思想,将无标记数据划分为正域、负域和无标记样本,通过对正域样本的协同训练逐渐标识所有无标记数据。李华雄、周献中和赵佳宝提出了基于三支决策粗糙集的代价敏感分类方法,有效地将属性约简与代价分类方法相结合。刘盾、李天瑞和梁德翠将决策粗糙集用于政策决策,给出了一个三支政策决策方案,并通过实例演示了该方案的优越性。贾修一和商琳提出了基于三支决策的垃圾邮件过滤算法,它是三支决策在代价敏感学习中的另一个应用。赵姝和张燕平用三支决策解释构造性算法,将三支决策与神经网络有效地结合在一起,为三支决策理论提出了一个新的研究方向。姚静涛和张燕将三支决策与博弈论和粗糙集结合起来,给出了博弈粗糙集一个新的解释,并通过博弈论获得了三支决策所需要的阈值。邓晓飞、王洪凯和姚一豫通过三支决策构造了自适应粗糙集近似。同经典近似相比,自适应近似进一步给出了上

下近似中的粒结构。

三支决策的相关研究向人们展示了该理论的普适性和优越性,但其应用领域仍旧有待拓展。根据客户细分问题的特点,本文将三支决策理论引入到客户细分应用中。

客户细分的概念来源于20世纪50年代出现的市场细分概念<sup>[8]</sup>。它是按照一定的标准将企业的现有客户划分为不同的客户群。通过客户细分,企业可以更好地识别不同客户群体的需求及其对企业的价值,以此指导企业的客户关系管理,达到吸引客户、保持客户、建立客户忠诚的目的<sup>[9-10]</sup>。Donner认为,正确的客户细分能够有效地降低成本,同时获得更强、更有利可图的市场渗透<sup>[11]</sup>。

现有的客户细分从技术的角度来看,主要有经验描述法、传统统计法、非传统统计法等<sup>[9-15]</sup>。这些方法中,经验描述法的主观性过强,传统统计法分析过程复杂,耗时较多,并且细分后得到的客户群过多,造成分析代价过大,而新兴的非传统统计方法(主要是数据挖掘等基于人工智能技术的非数值计算方法)多使用确定性分类模型。由于客户行为、信息采集过程中的复杂性导致的不确定性,使得现有方法在解决客户细分问题时存在很多不足,尤其缺乏对细分过程中的风险收益分析。通常来讲,企业可以根据单一因素,也可根据多个因素对客户进行细分。选用的细分标准越多,相应的客户群体也就越多,每个群体的客户数量就越小。相反,选用的细分标准越少,相应的客户群体也就越少,每个群体的客户数量也就越大。从理论上来说,细分的程度越高,也就是每个群体的客户数量越小,越容易定位用户的需求,并针对需求制定相应的营销策略,实现所谓的精准营销。但随之而来的是细分成本问题,当实

现客户细分的成本接近或者超过了所获利润时,细分策略显然是不合算的。因此,从成本-效益分析的角度来看,客户细分同样存在一个度的问题,即如何保持细分增加的成本与带来的利润之间的均衡问题。

在这个问题上,三支决策理论提供了一个很好的理论框架。它在考虑分类时,不是单纯地采用传统的“非此即彼”二分模式,而是将对象分为3个域:正域、负域和边界域。边界域的存在使得分类代价可以被控制在一个更为合理的范围内。具体应用在客户细分中,就是在分类时,不是直接将客户分为高价值和低价值客户,而是存在一个边界域群体,从而使得在制定客户关系维系策略时有第三种选择存在,而这种选择是以利益最大化为基准的。

按照这个思路,本文给出了基于三支决策的客户价值评价模型以及分类阈值确定方法,并提出了三支决策在客户细分中的应用策略。该问题的研究,既为三支决策理论提供了应用实证,同时也为客户细分的研究提供了新的思路和方法,具有重要的理论价值和现实意义。

## 2 三支决策理论

三支决策可以如下形式化描述<sup>[3]</sup>:

设  $U$  是有限、非空实体集,  $A$  是有限条件集。基于条件集  $A$ , 三支决策的主要任务是将实体集  $U$  分为3个两两不相交的域, 分别记为 POS、NEG 和 BND, 分别称为正域、负域和边界域。

对应于3个域, 可以构造三支决策规则, 分别称为接受、拒绝和不承诺规则。在实际应用中, 针对3个域可以给出不同的实际解释(不限于接受或拒绝)。为了实现三支决策, 在应用中需要构造实体评价函数和指定阈值。一般来讲, 在评价函数的值域上可以建立某种序关系, 这样可以通过决策状态值比较实体满足条件的程度。本文引入文献[3]提出的全序评价函数作为构造实体评价函数的基础。

设  $(L, \leq)$  是一个全序关系集, 其中  $\leq$  是全序关系, 即  $\leq$  是自反、反对称和传递关系, 且集合  $L$  中任意两个元素是可比较的。基于全序关系及一对阈值, 一个三支决策模型定义如下。

定义1 设  $(L, \leq)$  是一个全序集, 其中  $\leq$  是全序

关系,  $\alpha, \beta$  是满足条件  $\beta < \alpha$  的两个元素, 即  $(\beta \leq \alpha) \wedge \neg(\alpha \leq \beta)$ 。集合  $L^+ = \{t \in L | \alpha \leq t\}$ , 表示指定接受值区; 集合  $L^- = \{b \in L | b \leq \beta\}$ , 表示指定拒绝值区。给定评价函数  $v: U \rightarrow L$ , 三支决策可以定义为:

$$POS_{(\alpha, \beta)}(v) = \{x \in U | \alpha \leq v(x)\}$$

$$NEG_{(\alpha, \beta)}(v) = \{x \in U | v(x) \leq \beta\}$$

$$BND_{(\alpha, \beta)}(v) = \{x \in U | \beta < v(x) < \alpha\}$$

三支决策的有效性取决于所选用的评价函数及阈值。根据对评价函数的不同语义解释, 可以得出不同的三支决策模型。而阈值的选择实际上是不同度量函数的取舍或折中。下文将会给出在客户细分问题中具体的三支决策模型表示。

## 3 客户细分指标体系

### 3.1 客户细分指标体系分析

客户价值的评价标准有很多, 但现在很难找到一个统一的能被所有人接受的指标体系。综合已有的研究<sup>[16-19]</sup>, 并结合调研的企业现状, 本文构建的指标体系总体上包含如下3类指标。

- (1) 当前价值指标: 单位毛利润、总购买量;
- (2) 潜在价值指标: 忠诚度、信用度、品牌传播价值;
- (3) 成本指标: 服务成本。

3类指标中, 前两类指标是效益型指标, 第三类是成本型指标。各项指标具体描述如下。

(1) 单位毛利润: 利润是衡量客户价值最重要的指标之一。这里以客户提供的单位产品毛利润作为评价依据, 为数值型指标。

(2) 购买量: 客户购买量与其带来的利润之间并不一定是正比关系, 但可为企业扩大规模, 提高市场占有率和知名度带来很大的益处。因此, 它也是客户当前价值的重要体现。这里以客户的购买历史数据、合同数据等为依据, 统计得到该指标。

(3) 服务成本: 是指与客户服务相关的所有成本。这一指标数据可以根据历史经验数据估计得到, 为数值型指标。

(4) 忠诚度: 是客户从情感上表现出的对公司或品牌的偏好程度, 以及行为上表现出来的对公司或品牌的持续性。按照客户对公司或品牌的不同依赖程度, 可以将其分为3类: 坚定的忠诚者、中度的忠诚

者、多变者。客户忠诚度可以通过问卷调查以及和公司的交易记录来确定。

(5)信用度:是客户在与公司进行交易过程中表现出来的信用情况,比如合同执行情况、付款情况。通常也将其分为3类:信用良好、信用一般和信用不良。该指标可以通过客户交易记录和相关调查得到。

(6)品牌传播价值:也就是客户的口碑价值。客户的口碑会影响企业当前客户后续的购买模式和潜在客户的未来购买模式。口碑的效应值主要根据客户的规模、地位、影响力等来确定。该指标可以根据市场调研资料得到。本文采用3级打分标准对其予以量化,分别对应好、一般、差。

### 3.2 客户细分指标分类及其重要性分析

根据市场调研和经验数据,将上述的价值指标体系中的指标按照对客户价值影响的重要程度分为两个层次。

(1)一类指标:单位毛利润、购买量、服务成本、忠诚度。

(2)二类指标:信用度、品牌传播价值。

对客户价值进行评价的时候,可以先采用一类指标,分离出高价值的客户群和低价值的客户群以及待分析客户群。然后可在下一个阶段分析时采用二类指标,并只对待分析的客户进行分析,从而减少了数据操作的复杂性。当然在实际应用中,评价指标如果很多的情况下,可以将其分为更多的层次。

## 4 基于三支决策理论的客户细分模型

### 4.1 三支决策理论框架下的客户细分模型表示

为研究的方便,首先给出客户价值细分的形式化描述。

定义2 一个客户价值信息表是一个四元组  $T=(U, AT, V=\{V_a|a \in AT\}, f)$ 。其中  $U$  是现有客户的全集,即问题讨论的论域; $AT$  是客户属性集,即客户价值评价的指标; $V$  是各属性取值的值域,即指标取值范围; $f$  是信息函数,它是  $U \rightarrow V$  的一个映射,定义了每个客户在各个指标上的取值。

这里  $AT=\{a_1, a_2, a_3, a_4, a_5, a_6\}$ , 分别代表评价指标:单位毛利润、购买量、服务成本、忠诚度、信用度、品牌传播价值。

定义3 设  $T=(U, AT, V=\{V_a|a \in AT\}, f)$  是一个信息表,  $R \subseteq U \times U$  是定义在论域  $U$  上的等价关系,如果  $R$  满足:

$$IND(R)=\{(x, y)|(x, y) \in U \times U, x, y \in U, \forall a \in R, \exists (f(x, a)=f(y, a))\} \quad (1)$$

$[x]_R=\{y|y \in U, xRy\}$  是包含元素  $x \in U$  的等价类,可简记为  $[x]$ 。

定义4 设  $T=(U, AT, V=\{V_a|a \in AT\}, f)$  是一个信息表,对于  $C \subseteq U, x \in U$ , 定义其  $(\alpha, \beta)$  下近似和上近似分别为:

$$\begin{cases} \underline{apr}_{(\alpha, \beta)}(C)=\{x \in U|P(C|[x])\} \\ \overline{apr}_{(\alpha, \beta)}(C)=\{x \in U|P(C|[x])>\beta\} \end{cases} \quad (2)$$

这里阈值对  $(\alpha, \beta)$  满足条件  $0 < \alpha < \beta < 1$ ,  $P(C|[x])$  表示实体  $y$  属于等价类  $[x]$  时,  $y$  属于集合  $X$  的条件概率。

$(\alpha, \beta)$  下近似和上近似可将  $C \subseteq U$  分为3个域,即  $(\alpha, \beta)$ -正域  $POS_{(\alpha, \beta)}(C)$ 、 $(\alpha, \beta)$ -负域  $NEG_{(\alpha, \beta)}(C)$  和  $(\alpha, \beta)$ -边界域  $BND_{(\alpha, \beta)}(C)$ :

$$\begin{cases} POS_{(\alpha, \beta)}(v)=\{x \in U|P(C|[x])\} \\ NEG_{(\alpha, \beta)}(v)=\{x \in U|P(C|[x])<\alpha\} \\ BND_{(\alpha, \beta)}(v)=\{x \in U|\beta < P(C|[x]) < \alpha\} \end{cases} \quad (3)$$

3个概率区间是两两不相交的,并且它们的并集是实体集  $U$ 。这里,正域中的客户是高价值客户群体;负域中的客户是低价值客户群体;而边界域中的客户则是在目前条件下难以明确判断,需进一步观察的客户群体。客户细分问题的目标就是在给定指标体系下,利用阈值对  $(\alpha, \beta)$  将客户分属上述3个域。其中的核心问题是(价值)评价函数和阈值的确定。

### 4.2 评价函数与阈值的确定

对于任一实体  $x \in U$ ,它在客观上存在两种状态,即满足与不满足给定条件,据此可以将实体集  $U$  划分为两部分,分别表示为  $C$  和  $C^c$ ,即  $U=C \cup C^c$ 。当所掌握的信息不确定或不完整时,可能获得一个实体的描述。由于无法确切地给出实体  $x$  的状态,需要通过实体评价函数对实体状态进行估计。这种情况下,评价函数可以定义为实体  $x$  满足或不满足条件的概率。在客户价值细分问题中,条件集就是关于客户的各项指标。对于客户来说,如果简单地说两



种实际状态是高价值和低价值,根据指标而采取的动作有3种:接受 $a$ (实际分入高价值类)、拒绝 $r$ (实际分入低价值类)、延迟 $n$ (实际分入边界区域),那么可以得到一张三支决策代价表,如表1所示。

Table 1 Three-way decisions cost table

表1 三支决策代价表

决策动作	实体的客观状态	实体的客观状态
	满足条件 $p$	不满足条件 $n$
接受实体 $a$	$\lambda_{ap}$	$\lambda_{an}$
拒绝实体 $r$	$\lambda_{rp}$	$\lambda_{rn}$
延迟决策 $n$	$\lambda_{np}$	$\lambda_{nn}$

如果一个实体属于 $C$ ,则 $\lambda_{ap}$ 、 $\lambda_{rp}$ 、 $\lambda_{np}$ 分别表示采取3种决策动作的代价;如果一个实体不属于 $C$ ,则 $\lambda_{an}$ 、 $\lambda_{rn}$ 、 $\lambda_{nn}$ 分别表示作出3种决策的代价。基于贝叶斯决策理论,阈值对 $(\alpha, \beta)$ 可以通过如下总的决策代价函数来确定:

$$R(\alpha, \beta) = R(a|x) + R(r|x) + R(n|x) \quad (4)$$

其中,

$$\begin{cases} R(a|x) = \lambda_{ap} P(C|x) + \lambda_{an} (1 - P(C|x)) \\ R(r|x) = \lambda_{rp} P(C|x) + \lambda_{rn} (1 - P(C|x)) \\ R(n|x) = \lambda_{np} P(C|x) + \lambda_{nn} (1 - P(C|x)) \end{cases} \quad (5)$$

$R(a|x)$ 、 $R(r|x)$ 、 $R(n|x)$ 分别表示3种决策动作的代价。假设决策代价表中的值满足以下条件:

$$\begin{cases} (c_0) \lambda_{ap} & \lambda_{np} & \lambda_{rp} & \lambda_{rn} & \lambda_{nn} & \lambda_{an} \\ (c_1) (\lambda_{rp} - \lambda_{np}) (\lambda_{an} - \lambda_{nn}) > (\lambda_{np} - \lambda_{ap}) (\lambda_{nn} - \lambda_{rn}) \end{cases} \quad (6)$$

那么,代价函数 $R(\alpha, \beta)$ 的最小值所对应的阈值对 $(\alpha, \beta)$ 可以表示为:

$$\begin{cases} \alpha = \frac{\lambda_{an} - \lambda_{nn}}{(\lambda_{an} - \lambda_{nn}) + (\lambda_{np} - \lambda_{ap})} \\ \beta = \frac{\lambda_{nn} - \lambda_{rn}}{(\lambda_{nn} - \lambda_{rn}) + (\lambda_{rp} - \lambda_{np})} \end{cases} \quad (7)$$

其中, $0 < \alpha, \beta < 1$ 。

另外,文献[20]专门研究了阈值对 $(\alpha, \beta)$ 的自适应求法,有兴趣的读者可参考。

### 4.3 客户细分的实施步骤

应用中,对客户实施细分时,采用分阶段动态分类

方式。根据第3章中对指标体系的分析,把细分过程分为两个阶段,每个阶段采用一组指标,具体步骤如下。

步骤1 采集客户信息,并将采集来的数据进行标准化处理,包括无量纲化和归一化处理。

步骤2 第一阶段采用指标组1中的指标对客户进行分类。

步骤2.1 确定决策代价表中的数据,即给出表1中的数据。

步骤2.2 计算阈值 $\alpha$ 、 $\beta$ 。

步骤2.3 分别计算每个客户的 $P(C|x)$ ,并根据阈值 $\alpha$ 、 $\beta$ 将其划入正域、负域或者边界域。

步骤3 对于边界域中的客户,继续分类。重复步骤2,但采用指标组2,重新计算决策代价和阈值 $\alpha$ 、 $\beta$ ,并将边界域中的客户划入正域、负域或者边界域。

步骤4 根据实际应用的需要,可确定是否需要继续对边界域中的元素进行划分,如果不需要继续划分则停止,否则变换指标,重复上面的过程即可。

## 5 实例分析

### 5.1 数据采集与处理

本文的实例分析数据来源于某养殖饲料公司,为保密起见,经过了技术处理,但不影响对分析方法的说明。

数据采集过程中,对于客户的忠诚度、信用度、品牌传播价值等指标值均采用调查和专家经验相结合的方法来确定。客户的单位毛利润、购买量、服务成本则根据销售记录结合专家经验来确定。数据预处理过程中,将单位毛利润、购买量、服务成本均进行离散化、无量纲化处理。为简单起见,离散区间取固定区间,不影响问题说明。各类决策代价之间的关系,根据专家经验,遵循以下原则进行设定:

(1) 获得一个新客户的成本是保持一个老客户成本的5倍。

(2) 挽留一个不忠诚客户的成本是保持一个老客户成本的10倍。

(3) 客户保持率提高5%,利润会提高25%。

(4) 保持客户关系的服务成本大约占该客户贡献利润的0.5%比较合理。

(5) 将一个高价值客户误判为低价值客户,从而导致客户流失所造成的损失是将低价值客户误判为高价值客户导致的服务成本增加的10倍。

(6) 将一个高价值客户误判为边界域客户, 从而导致从该客户得到的利润降低, 但同时服务成本也降低。

(7) 将一个低价值客户误判为边界域客户, 从而导致服务成本增加, 但利润并没改变。

下面给出经过离散化和标准化后的数据, 属性  $C$  是专家标记的分类, 筛选过后的原始数据共 300 条, 限于篇幅, 这里只给出部分数据如表 2 所示。

Table 2 Customer value evaluation table

表 2 客户价值评价表

	$a_1$	$a_2$	$a_3$	$a_4$	$a_5$	$a_6$	$C$
$x_1$	3	3	1	2	2	3	1
$x_2$	3	1	1	3	2	2	1
$x_3$	2	2	1	2	3	1	1
$x_4$	1	1	1	1	3	1	0
$x_5$	1	3	1	2	2	3	1
$x_6$	2	2	2	3	1	1	1
$x_7$	1	2	2	1	1	1	0
$x_8$	1	1	2	1	3	1	0
$x_9$	1	1	1	1	1	1	0
$x_{10}$	3	3	1	3	2	3	1

## 5.2 客户细分计算

将采集的数据 200 条用于学习, 100 条用于测试。其中关键的参数计算方法如下:

(1) 根据式 (7) 计算阈值  $\alpha$ 、 $\beta$ 。

(2) 根据文献 [21] 提供的方法, 用  $O(P(C|[x]))$  作为  $P(C|[x])$  的估算值。

根据贝叶斯公式,  $P(C|[x]) = \frac{p([x]|C)p(C)}{p([x])}$ , 而

$P(C^c|[x]) = \frac{p([x]|C^c)p(C^c)}{p([x])}$ , 可以计算  $O(P(C|[x]))$  来代替  $P(C|[x])$ :

$$O(P(C|[x])) = \frac{P(C|[x])}{P(C^c|[x])} = \frac{p([x]|C)p(C)}{p([x]|C^c)p(C^c)} \quad (8)$$

其中,  $C^c$  表示不属于类  $C$  的对象集合;  $p(C)$  是先验概率,  $p(C^c) = 1 - p(C)$ 。  $p([x]|C)$  和  $p([x]|C^c)$  按照下式计算:

$$\begin{cases} p([x]|C) = p(v_1, v_2, \dots, v_n|C) = \prod_{i=1}^n p(v_i|C) \\ p([x]|C^c) = p(v_1, v_2, \dots, v_n|C^c) = \prod_{i=1}^n p(v_i|C^c) \end{cases} \quad (9)$$

其中,  $p(v_i|C) = \frac{|m(a_i, v_i) \cap C|}{|C|}$  表示某个属性  $a_i = v_i$  在

分类  $C$  中的概率。式 (9) 假设各属性相互独立。

利用上述计算公式和 4.3 节给出的分类步骤, 从 200 条学习数据中学得分类代价并计算阈值, 然后用于 100 条测试数据的分类。学习得到的决策代价如表 3 所示。

Table 3 Customer segmentation decision cost table

表 3 客户细分决策代价表

决策动作	客户高价值 p	客户低价值 n
接受高价值 a	0.05	0.40
拒绝高价值 r	0.90	0.05
再分析 n	0.10	0.20

将测试结果与采用二支决策一次性分类进行对比, 二支决策分类采用模糊综合评价<sup>[22]</sup>给出客户综合评价价值, 然后根据评价进行分类。比较的主要参数是平均分类代价。实验结果如表 4 所示。

Table 4 Experimental results table

表 4 实验结果对比表

分类方法	正域	负域	中间域	平均分类代价
二支决策	73	27	0	0.084 5
三支决策	48	27	25	0.065 5

通过观察实验结果, 可以得出如下结论: 三支决策方法与二支决策相比, 分类代价明显减小。从实验数据看, 本例中采用二支决策情况下的分类代价将比采用三支决策时的分类代价增加 29%。三支决策方法的分类正确率并不肯定比二支决策更高, 这是因为在每个阶段都只是采用部分指标进行分类, 而有部分客户被划入延迟决策的边界域中。但本文追求的不是准确率更高, 而是总的分类代价本文最小(或者说收益最大)。从这个目标来看, 采用三支决策方法比二支决策有明显的优势。

## 5.3 三支决策客户细分方法的应用说明

通过上面的实例分析可以看出, 客户细分的三支决策方法, 其优势不在于分类准确率的提高, 而在于分类代价的减小。在实际应用过程中, 不一定一次性地将上述过程全部完成, 而是根据信息的采集

情况动态地分阶段进行。根据每个阶段分类的结果,可以采取不同的策略。这种策略的实施,一方面降低了数据操作的复杂性,减少了计算量;另一方面还可以减小决策风险代价,有实用性。

## 6 结束语

本文提出了一种基于三支决策理论的客户细分方法。三支决策理论的引入,为客户细分提供了新的思路和方法,尤其在不完全信息状态下,可以避免决策盲动风险,最大化决策收益。另外,本文对三支决策的应用也表明,对于3个域的解释,不一定只是拒绝、接受和延迟决策,而可能是对应3种不同的策略。同时,在很多情况下,三支决策可能只是作为二支决策的中间过程来使用,最终结果仍旧是二支的,但本文在应用时提出了可以保留三支结果,以采取3种不同的策略。这也为三支决策的理论和应用提供了新的解释。

致谢 感谢加拿大里贾纳大学的姚一豫教授在课题研究过程中给予的指导和帮助。

## References:

- [1] Yao Yiyu. Three-way decision: an interpretation of rules in rough set theory[C]//LNCS 5589: Proceedings of the 4th International Conference on Rough Sets and Knowledge Technology (RSKT '09), Gold Coast, Australia, Jul 14-16, 2009. Berlin, Heidelberg: Springer, 2009: 642-649.
- [2] Yao Yiyu. Three-way decisions with probabilistic rough sets[J]. Information Sciences, 2010, 180(3): 341-353.
- [3] Yao Yiyu. An outline of a theory of three-way decisions[C]//LNCS 7413: Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing (RSCTC '12), Chengdu, China, Aug 17-20, 2012. Berlin, Heidelberg: Springer-Verlag, 2012: 1-17.
- [4] Yao Yiyu. The superiority of three-way decisions in probabilistic rough set models[J]. Information Sciences, 2011, 181(6): 1080-1096.
- [5] Jia Xiuyi, Shang Lin, Zhou Xianzhong, et al. Three-way decisions theories and applications[M]. Nanjing: Nanjing University Press, 2012.
- [6] Jia Xiuyi, Zheng Kan, Li Weiwei, et al. Three-way decisions solution to filter spam email: an empirical study[C]//LNCS 7413: Proceedings of the 8th International Conference on Rough Sets and Current Trends in Computing (RSCTC '12), Chengdu, China, Aug 17-20, 2012. Berlin, Heidelberg: Springer-Verlag, 2012: 287-296.
- [7] Zhou Bing, Yao Yiyu, Luo Jigang. A three-way decision approach to email spam filtering[C]//LNCS 6085: Proceedings of the 23rd Canadian Conference on Advances in Artificial Intelligence (AI '10), Ottawa, Canada, May 31-Jun 2, 2010. Berlin, Heidelberg: Springer-Verlag, 2010: 28-39.
- [8] Kotler P. Marketing management—analysis, planning, implementation, and control[M]. 13th ed. New Jersey: Prentice Hall, 2008.
- [9] Liu Yingzi, Wu Hao. A summarization of customer segmentation methods[J]. Journal of Industrial Engineering and Engineering Management, 2006, 20(1): 53-57.
- [10] Ye Qiang, Lu Tao, Yan Xiangbin, et al. Research on dynamic customer segmentation in customer relationship management[J]. Journal of Management Sciences in China, 2006, 9(2): 44-52.
- [11] Donner S. What can customer segmentation accomplish? [J]. Bankers Magazine, 1992, 175(20): 72-81.
- [12] Yang Yongheng, Wang Yonggui, Zhong Xudong. Customer relation management: concept, drivers and dimension[J]. Nankai Business Review, 2002(2): 48-52.
- [13] Zhang Guofang, Jin Guodong. Research on customer segmentation theories and their application strategies[J]. Journal of Huazhong University of Science and Technology: Social Sciences Edition, 2003(3): 101-104.
- [14] Chen Mingliang, Li Huaizu. Study on value segmentation and retention strategies of customer[J]. Group Technology & Production Modernization, 2001, 18(4): 23-27.
- [15] Wang Hong. Customer value analysis based on rough sets data mining technology[D]. Harbin: Harbin Engineering University, 2006.
- [16] Qi Jiayin. Customer value evaluation, modeling and decision[M]. Beijing: Beijing University of Posts and Telecommunications Press, 2005.
- [17] Dong Dahai, Quan Xiaoyan, Qu Xiaofei. Customer value construct[J]. Journal of Dalian University of Technology: Social Sciences Edition, 1999(12): 47-52.
- [18] Xiong Benfeng. Review of customer value theory[J]. Journal of Chongqing Technology and Business University: Social Sciences Edition, 2003(6): 57.

- [19] Ling Lei. Study on cluster value based on rough sets theory[D]. Jinan: Shandong University, 2006.
- [20] Jia Xiuyi, Li Weiwei, Shang Lin, et al. An adaptive learning parameters algorithm in three-way decisions[J]. Acta Electronica Sinica, 2011, 39(11): 2520-2525.
- [21] Yao Yiyu, Zhou Bing. Naive Bayesian rough sets[C]//LNCS 6401: Proceedings of the 5th International Conference on Rough Sets and Knowledge Technology (RSKT '10), Beijing, China, Oct 15-17, 2010. Berlin, Heidelberg: Springer-Verlag, 2010: 719-726.
- [22] Yang Lunbiao, Gao Yingyi, Ling Weixin. Principle and applications of fuzzy mathematics[M]. 5th ed. Guangzhou: South China University of Technology, 2011.
- [12] 杨永恒, 王永贵, 钟旭东. 客户关系管理的内涵、驱动因素及成长维度[J]. 南开管理评论, 2002(2): 48-52.
- [13] 张国方, 金国栋. 客户细分理论及应用策略研究[J]. 华中科技大学学报: 社会科学版, 2003(3): 101-104.
- [14] 陈明亮, 李怀祖. 客户价值细分与保持策略研究[J]. 现代生产与管理技术, 2001, 18(4): 23-27.
- [15] 王宏. 基于粗糙集数据挖掘技术的客户价值分析[D]. 哈尔滨: 哈尔滨工程大学, 2006.
- [16] 齐佳音. 客户价值评价、建模及决策[M]. 北京: 北京邮电大学出版社, 2005.
- [17] 董大海, 权小妍, 曲晓飞. 客户价值及其构成[J]. 大连理工大学学报: 社会科学版, 1999(12): 47-52.
- [18] 熊本峰. 关于客户价值理论的述评与思考[J]. 重庆工商大学学报: 社会科学版, 2003(6): 57.
- [19] 凌磊. 基于粗糙集理论的客户价值研究[D]. 济南: 山东大学, 2006.
- [20] 贾修一, 李伟津, 商琳, 等. 一种自适应求三枝决策中决策阈值的算法[J]. 电子学报, 2011, 39(11): 2520-2525.
- [22] 杨纶标, 高英仪, 凌卫新. 模糊数学原理及应用[M]. 5版. 广州: 华南理工大学出版社, 2011.

### 附中文参考文献:

- [5] 贾修一, 商琳, 周献中, 等. 三支决策理论与应用[M]. 南京: 南京大学出版社, 2012.
- [9] 刘英姿, 吴昊. 客户细分方法研究综述[J]. 管理工程学报, 2006, 20(1): 53-57.
- [10] 叶强, 卢涛, 闫相斌, 等. 客户关系管理中的动态客户细分方法研究[J]. 管理科学学报, 2006, 9(2): 44-52.



HUANG Shunliang was born in 1977. He received the Ph.D. degree from Shandong University in 2009. Now he is an associate professor and master supervisor at Shandong University of Technology. His research interests include rough sets theory and data mining, etc.

黄顺亮(1977—),男,山东郯城人,2009年于山东大学获得博士学位,现为山东理工大学商学院副教授、硕士生导师,主要研究领域为粗糙集理论与应用、数据挖掘等。



LI Jianlin was born in 1974. He received the Ph.D. degree from Southeast University in 2006. Now he is an associate professor and senior engineer at Nanjing College of Information Technology, and the member of CCF and IEEE. His research interests include automatic control, artificial intelligence and data mining, etc.

李建林(1974—),男,湖南洪江人,2006年于东南大学获得博士学位,现为南京信息职业技术学院副教授、高级工程师,计算机与软件学院副院长,CCF会员,IEEE会员,主要研究领域为自动控制、人工智能、数据挖掘等。



WANG Qi was born in 1972. She received the M.S. degree from Tianjin University of Finance & Economics in 2004. Now she is an associate professor at Shandong University of Technology. Her research interests include marketing and enterprise public relation, etc.

王琦(1972—),女,山东淄博人,2004年于天津财经大学获得硕士学位,现为山东理工大学副教授,主要研究领域为市场营销、企业公共关系等。