

会计师事务所可持续发展战略研究

——基于聚类与决策树方法

邱吉福

(集美大学 工商管理学院, 福建 厦门 361021)

[摘要] 聚类和决策树方法有效的结合, 不仅可以实现对会计师事务所多个因素的多角度考察, 同时也可以避免人为假设因素的影响及敏感特征因素的丢失。研究发现, 会计师事务所的可持续发展和其自身的条件具有很强的相关性, 优秀的会计师事务所应当从内涵发展和外延拓展方面实现做大做强。

[关键词] 审计市场; 数据挖掘; 聚类; 决策树; 会计师事务所; 可持续发展战略

[中图分类号] F239.4 **[文献标识码]** A **[文章编号]** 1672-8750(2012)02-0077-07

一、引言

许多专家学者经过长期的研究和探索, 发现在审计市场中存在规模经济。我国一些专家学者在早期对审计市场的研究中, 基本上是从市场集中度角度对审计市场进行研究, 随着时代的变迁, 现在也逐渐开始将规模经济和市场结构优化理论应用于审计市场的研究, 从这方面来看, 我国学界对审计市场已有了一个新的认识。

许汉友等运用灰色系统分析方法, 假设业务结构、分所数、CPA(注册会计师)人数、CPA的年龄结构和学历结构对会计师事务所运营绩效可能产生影响, 通过实证分析得出结论: 影响会计师事务所运营绩效的主要因素是CPA的年龄结构和学历结构, 而业务结构的影响甚微^[1]。姜虹等指出, 审计市场呈现出规模经济不明显、集中度偏低等特点, 过度竞争导致审计资源浪费, 市场结构尚未形成寡头垄断格局^[2]。杨永森等在2007年指出审计市场具有规模经济性, 在2009年指出运营效率是影响和衡量会计师事务所业绩的重要指标, 它不仅影响会计师事务所的核心竞争力, 同时也影响会计师事务所后续的发展。他们首先对DEA(数据包络分析)中的CCR模型和“超效率”模型进行分析, 然后实证研究了我国会计师事务所运营效率偏低的普遍特征^[3-4]。王兵等对2003年至2009年的前百家会计事务所排名变迁进行初步分析, 认为由于市场经济发展和政府力量的推动, 我国注册会计师行业发展呈现良好态势, 行业总收入快速增长, 注册会计师行业学历水平不断提高, 年龄结构进一步优化。同时他还认为, 随着事务所强强联合, 我国会计师事务所的整体实力已经得到大幅提升, 但相比国际四大会计师事务所而言, 还存在一定的差距^[5]。贾宪威等对我国2008年的前百家会计师事务所综合排名信息运用两阶段数据包络分析法进行了实证分析, 认为CPA的素质是会计师事务所运营效率的重要因素, 而适当的事务所规模也会对会

[收稿日期] 2011-11-07

[基金项目] 福建省高校服务海西建设重点项目(Z80107)

[作者简介] 邱吉福(1965—), 男, 福建永春人, 集美大学工商管理学院副院长, 副教授, 硕士生导师, 主要研究方向为会计、审计实务。

计师事务运营效率产生一定影响^[6]。朱志红运用 AHP(层次分析法)从分析会计师事务所知识管理绩效评价要素出发,建立了会计师事务所知识管理绩效评价的多级模糊综合评价模型^[7]。胡晓莉将 BSC(平衡计分卡)的思想引入会计师事务所的建设中来,从财务、顾客、内部流程、学习与成长四个方面设计战略性绩效考核思路,找出影响自身发展的因素作为管理的重点,但作者同时也指出,设定了相应的指标之后还必须为各指标确定权重,否则仍无法运用 BSC^[8]。

以上专家学者对审计市场或会计师事务所的研究中,由于所应用的数据模型不同,或者由于在对数据指标选择上具有主观性,他们的研究结果或多或少存在一定的差异。从另一个角度来看,这些研究或者是先基于假设,然后验证假设的置信度,或者只是运用单一模型实证分析影响会计师事务所发展的某些指标,未能全面论证影响会计师事务所做大做强的重要因素。

中国注册会计师协会(以下简称为“中注协”)自 2003 年至今已经连续多年发布“全国前百家会计师事务所信息”,2003 年公布的数据指标包括总收入、CPA 人数、CPA 年龄结构和学历结构,2004 增加了分所数及分所总收入指标,2005 年又增加了总收入中的审计收入和兼营资产评估收入等指标,2005 年首次将综合得分纳入评价指标体系,包含了培训完成率、处罚(惩戒)应减分值及领军人才数指标。从公布的数据来看,全国前百家会计师事务所做大做强特征持续显现,然而,我们不能简单地从 CAP 人数和总收入来判断会计师事务所的核心竞争力和可持续发展能力,而是应当从多个数据指标对之综合评价。

中注协规定的方法简便实用,但信息量较小,忽略了其他许多方面的影响因素,比如注册资本规模、盈利能力、分红情况、纳税情况、管理人员素质等,这使得它的评价不够全面^[9]。

本研究的目的是不是探讨中注协的评价指标和评价方法是否正确完整,而是针对中注协公布的 2010 年和 2011 年会计师事务所综合评价前百家的信息,通过聚类分析方法对各项指标进行分簇,在分簇的基础上运用 ID3 决策树方法,实证研究优秀会计师事务所竞争力和快速发展的影响因素,并指出我国会计师事务所可持续发展的有效途径。

二、数据挖掘技术

数据挖掘(Data Mining),习惯上又称为“数据库中的知识发现”,它是从大量存储的数据中找出不同数据的共同特征,寻求归并和分簇,发现数据集中的个体关联性和影响度,因此在数据挖掘中我们常常需要应用数据库技术建立数据仓库,从海量数据中寻找规则。数据挖掘吸收了各学科领域(如统计学、人工智能、数学建模等)的思想,并运用计算机强大的计算和并发处理能力,发现规则、法则、规律、方程或概念等。在数据挖掘技术中,常用的数据方法包括了分簇、聚类、关联分析以及决策树等^[10]。

决策树方法是利用逼近离散函数值来构建一棵树,从而发现数据中蕴涵的分簇规则。应用较广泛的决策树方法是 Ross Quinlan 提出的 ID3 算法。ID3 算法是通过构建学习系统来减少一棵树的深度,并采用最优化的原则,使决策树生成的叶子节点数、叶子节点的深度都达到最小值。

三、实证分析

本文首先运用聚类方法将前百家事务所进行分簇,然后通过 ID3 决策树方法寻找影响会计师事务所排名的决定因素,这样做具有以下优势:(1)采用多个重要因素从不同角度进行考察^[11];(2)避免单独采用任何一项因素衡量都存在的理由不足之嫌;(3)排除人为假设的主观因素的影响;(4)对噪声数据有很好的健壮性,避免一些敏感的因素数据被剪枝处理;(5)分簇精度高,模式简单。

具体的分析过程如图 1 所示。

(一) 聚类分析

聚类分析可以依据分簇的目标将数据分成许多簇,而被分在同一簇的数据会有某些特征是相似的,分在不同簇的数据会有某些特征是不相同的。在实际应用聚类分析方法时,一般分为以下几个步骤。

1. 确定主题

本文研究的目的是我国前百家会计师事务所的影响因素,主要的评价指标包括会计师事务所的总收入、CAP 人数、培训完成率、领军人才数、处罚(惩戒)应减分值,辅助指标包括分所数、从业人数、人均收入、股东人数、境外分部数量、CPA 年龄结构及学历结构。这些都是聚类分析时需要提取的与事务所排名相关评价指标有关的数据。

2. 数据来源

本研究的数据是对中注协公布的“2011 年会计师事务所综合评价前百家信息”和“2010 年会计师事务所综合评价前百家信息”的数据进行整理而得,具有权威性、真实性和可靠性。

3. 数据准备

本研究建立数据库来存储数据。数据库的一部分数据来自于原始数据,可以直接从中注协公布的“前百家会计师事务所信息”中提取,如收入、从业人数、分所数等,而另一部分统计数据是利用原始数据运算而得,如年龄结构比例、学历结构比例、人均收入等。

4. 数据预处理

从中注协公布的前百家会计师事务所的数据看,笔者认为,“总收入”、“从业人员数量”等评价指标不能完全说明一个事务所是否优秀,总收入和事务所的规模相关,从业人员包含了各学历及各年龄段所有人员,因此在运用数据模型分析前要对数据进行预处理,选取能够说明会计师事务所可持续发展相关的数据指标,并对缺失值和异常值进行处理。对于诸如“总收入”、“从业人员数量”等指标,笔者认为它们和排名不具有直接相关性,因此取而代之的是 CPA 人均收入、CPA 各年龄段和各学历层次的百分比数据,获得能够体现会计师事务所核心竞争力的指标。同时,因为公司结构变化对公司绩效的影响具有滞后性,所以笔者选取“2011 年会计师事务所综合评价前百家信息”中的“总收入”和“CPA 人数”,计算 CPA 人均收入,选取“2010 年会计师事务所综合评价前百家信息”的“处罚(惩戒)应减分值”及“综合评价”等指标,获得连续两年均在前百家排名中的 82 家会计师事务所综合信息,构成新的数据集,如表 1 所示。

表 1 82 家会计师事务所综合信息

会计师事务所	综合排名	股东人数	领军人才数量	CPA 人均收入	处罚(惩戒)应减分值	分所数量	境外分部数量	CPA 年龄结构(%)			CPA 学历结构(%)		
								40 岁及以下	41 至 60 岁	61 岁及以上	本科及以下	硕士及双学历	博士
普华永道中天	1	2	6	350.37	—	9	—	96.02	3.98	0.00	73.89	25.81	0.29
德勤华永	2	2	3	324.60	—	6	—	93.71	6.29	0.00	67.83	31.33	0.84
安永华明	3	2	10	226.15	—	7	—	97.46	2.54	0.00	68.01	31.64	0.35
毕马威华振	4	2	1	242.45	—	3	—	99.23	0.77	0.00	64.81	35.19	0.00
中瑞岳华	5	42	14	79.27	4	21	2	62.54	32.33	5.13	89.98	9.53	0.49
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
上海迈伊兹	82	5	0	118.01	—	0	0	88.89	11.11	0.00	96.30	3.70	0.00

数据来源:中注协 2010 年 4 月 30 日和 2011 年 7 月 25 日讯。

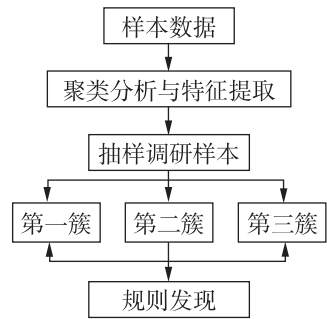


图 1 会计师事务所分簇流程

5. 模型构建

K-means 是聚类算法中最常用的分簇方法。在使用 K-means 方法之前,必须先决定分簇结果的数量,也就是定义 K 的值;当 K-means 初始化时,会先任意选择 K 个数据点作为聚类中心,接着依据所有数据点与每个聚类中心的距离,将所有数据点分配到最近的簇中,然后再重新产生每一个聚类中心,找出新的聚类中心,并依据新的聚类中心来重新分配数据点。这些步骤会反复执行到每一个聚类中心都不再改变为止。在这个过程中我们发现:(1)随机选择的初始聚类中心点不适合时,会造成分簇效率不高,从而降低分簇的可靠性。(2)每一次重复都要重新计算每一个数据到各聚类中心的距离,如果数据量较大时,其迭代的次数是相当惊人的。

由于采用传统的 K-means 聚类分析方法可能造成聚类结果的不稳定性,降低分簇的可靠性,因此我们采用改进的 K-means 聚类分析方法,首先确定 K 个初始聚类中心,然后借用三角形不等边定律,逐步减少每次迭代的计算量,简化计算比较过程。

6. 事务所细分结果

聚类算法的一个特点是在每次迭代中都要考察每个样本的分簇是否正确。如果分簇不正确,就要调整,在全部样本调整完后,再修改聚类中心,进入下一次迭代;如果在一次迭代算法中,所有的样本都被正确分簇,则不会有调整,聚类中心不会再有变化^[12]。因此,我们首先采用 K 为 10 对前 82 家事务所进行细分,得到 10 簇会计师事务所的主要特征,然后再次采用 K 为 3 进行聚类,得到正确的前百家事务所三个簇,同时分析三簇事务所的主要特征^①。

(1) 选取 K 值为 10

对这些会计师事务所进行分析,利用改进的 K-means 算法细分成 10 个簇,如表 2 所示。

从这个结果中我们可以看到:第 4、第 6、第 9 三簇事务所的 CPA 人均收入比较高;第 7 簇事务所的分所数量较多;第 6 和第 9 簇事务所的人均收入、CPA 年轻及高学历层次人员都比较高。

表 2 82 家会计师事务所细分结果(K=10)

簇	事务所个数	领军人才数量	处罚(惩戒)应减分值	分所数量	CPA 人均收入	股东人数	境外分部数量	CPA 年龄结构(%)			CPA 学历结构(%)		
								40 岁及以下	41 至 60 岁	61 岁及以上	本科及以下	硕士及双学历	博士
6	1	0.0	—	0.0	498.2	2.0	0.0	88.9	11.1	0.0	77.8	22.2	0.0
9	4	5.0	—	6.3	285.9	2.0	0.0	96.6	3.4	0.0	68.6	31.0	0.4
4	4	0.8	—	3.3	130.7	21.5	0.3	49.6	46.7	3.7	93.8	6.2	0.0
1	7	1.3	0.6	4.1	96.2	13.4	0.1	71.7	24.2	4.1	89.9	9.7	0.4
10	12	2.0	1.3	4.8	76.8	24.8	0.3	64.9	28.3	6.8	93.6	5.9	0.4
7	11	6.9	5.0	13.9	67.8	27.2	1.1	71.6	23.8	4.6	90.5	8.8	0.8
2	8	0.9	2.4	5.8	53.8	13.6	0.1	67.6	27.0	5.4	90.1	9.6	0.3
3	10	0.3	4.1	3.5	50.9	13.0	0.0	60.8	31.0	8.3	94.9	4.6	0.5
8	9	0.6	2.3	4.9	46.2	21.2	0.0	44.1	37.8	18.1	95.0	4.9	0.1
5	16	0.8	0.9	5.7	39.1	15.4	0.1	62.3	29.8	7.9	93.5	6.3	0.2

注:按人均收入排序。

(2) 选取 K 值为 3

假设 clust0、clust1、clust2 分别代表三簇事务所,我们把 82 个事务所聚成 3 簇,每簇的事务所数量、数量比等指标如表 3 所示。

从这个结果我们可以看出:clust0 簇类事务所人均收入平均超过 328 万元,其处罚(惩戒)应减分

^①在选取不同 K 值的循环过程中,当 K 值为 10 时,人均业务收入最高的本土“四大”会计师事务所浮出水面,同时,其置信度也是较高的,因此,首先采用 K 为 10 对前百家事务所进行细分,得到 10 簇会计师事务所的主要特征。聚类分析是为后续探索前百家事务所重要影响因素的规则作数据预处理,通过循环测试,当 K 值为 3 时,不仅在应用决策树发现规则中得到高集中度的会计师事务所主要影响因素,同时其置信度也是较高的。

值均为0,领军人才数平均近4人,CPA年龄40岁以下平均值超过95%,高学历CPA平均占近30%,实证了国内大型优秀的会计师事务所不仅具有高学历、年轻化的CPA团队,而且具有健全有效的内部控制制度。

表3 82家会计师事务所细分结果(K=3)

簇	事务所个数	领军人才数量	处罚(惩戒)应减分值	分所数量	CPA人均收入	股东人数	境外分部数量	CPA年龄结构(%)			CPA学历结构(%)		
								40岁及以下	41至60岁	61岁及以上	本科及以下	硕士及双学历	博士
0	5	4.00	—	5.00	328	2.00	0.00	95.06	4.94	0.00	70.47	29.24	0.30
1	43	2.02	2.45	7.09	55	18.28	0.28	61.78	29.40	8.82	92.02	7.66	0.32
2	34	1.56	1.96	4.79	72	19.65	0.24	63.30	30.46	6.24	93.76	5.82	0.42

(二) 模型应用及战略策略

根据以上“选取K值为10”的结果,各会计师事务所可以根据自身情况制定可持续的发展战略。

(1) 2、3、4、5、6、8这几簇事务所没有或极少顶尖人才,因此应该加强领军人才的引进或培养。

(2) 2、3、7、8这几簇事务所内控制度建设薄弱,因此应健全内部控制制度。

(3) 第6、第9簇事务所基础较好,但在境外没有设立分支机构,因此应加速国际化网络布局。

(4) 4、7、8、10这几簇事务所股东人数较多,因此应更加注重提升股东的素质和能力,鼓励他们探索“走出去”的路径、方式和步骤,同时协调好“走出去”与“请进来”的关系,不仅要做大,而且还要做强。

(5) 除第6、第9簇外,所有事务所CPA学历普遍偏低,这说明事务所应当重视CPA队伍的后续教育(包括学历教育和业务培训),鼓励他们继续学习和深造,注重人才培养和储备。

运用聚类方法对会计师事务所的细分,不仅可以使我们了解每家会计师事务所在本行业中的地位,同时也能够让我们发现会计师行业本身与其他同行相比有哪些不足之处。对于任何企业来说,在激烈的竞争环境中实现可持续发展都显得至关重要。

四、发现规则

前面我们已经运用聚类分析方法对82家会计师事务所进行分簇,而决策树是一种利用逼近离散函数挖掘数据中蕴涵的分簇规则的方法,因此,笔者对聚类分析结果运用ID3决策树方法,进一步验证前期分簇的有效性,同时寻求影响82家会计师事务所排名的重要因素。

(一) 应用ID3进一步细分

ID3算法通过循环逐步求精,直到获得最优的决策树,而应用其他的决策树方法,几乎都加入剪枝优化处理,在实际应用中将丢失一些重要的影响因素。以“簇”为构造目标,以CPA人均收入、领军人才数、CPA年龄结构(包括40岁以下、41岁至60岁、61岁及以上)、分所数、学历结构(包括本科及以下、硕士及双学历、博士)、股东人数为数据指标,然后根据ID3决策树模型建立决策树,如图2所示。

(二) 规则发现

通过运用ID3方法构建的82家事务所决策树和规则表,我们不仅发现了4条规则(如表4),同时也验证了采用K=3对82家事务所进行聚类的正确性:对clust0的置信度达到100%;对clust1的置信度达到

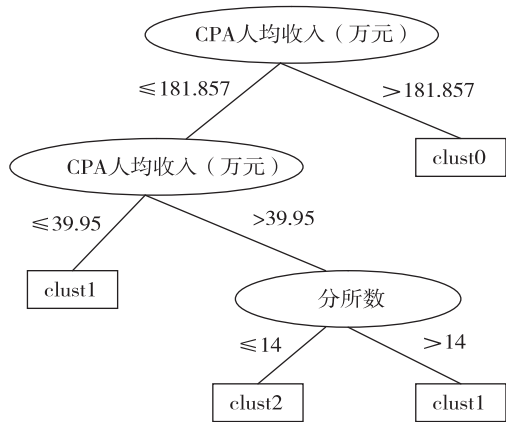


图2 会计师事务所综合评价指标决策树

100%的有4家,达到置信度93%以上的有16家;对于clust2的置信度亦超过50%。根据规则表我们发现,CAP人均收入和分所数是决定会计师事务所综合排名的主要因素,会计师事务所在发展过程中,不但要加大人才培养,而且要扩展规模。

表4 决策树发现规则表

规则序号	规则	类名	置信度	支持记录序号
1	人均业务收入(万元) ≤ 39.95	clust1	93.75	16.0
2	人均业务收入(万元) ≤ 181.857 AND 人均业务收入(万元) > 39.95 AND 分所数量 ≤ 14	clust2	57.895	57.0
3	人均业务收入(万元) ≤ 181.857 AND 人均业务收入(万元) > 39.95 AND 分所数量 > 14	clust1	100.0	4.0
4	人均业务收入(万元) > 181.857	clust0	100.0	5.0

(三) 发展战略制定

根据上面分析的结果,笔者认为随着审计市场的国际化,会计师事务所要实现可持续发展,必须制定出合理的发展战略。

1. 加强人才培养,制定人才发展战略规划。对于注册会计师行业而言,人才就是资源,人才就是收入。首先,要加大注册会计师领军人才的引进力度和内在培养力度,积极物色高层次人才,鼓励和支持年轻注册会计师攻读博士学位或参加注册会计师领军人才的全国统一考试。其次,要重视人才的培养和储备,根据注册会计师所学的专业和特长制定培养目标及发展规划,鼓励和支持会计师事务所的各层次人才参加技术职称和执业资格全国统一考试,保持一定规模的注册会计师数量。最后,要重视人才的后续教育,加大人才培训的投入,鼓励和支持注册会计师及其他从业人员参加行业协会举办的培训,更新知识,开阔视野,增强服务意识,不断提高从业人员的职业能力和服务水平,提升注册会计师团队核心竞争力。

2. 推动跨越式发展,实现规模经济效益。首先,扩大审计业务范围。进一步落实国办56号文件精神和各省级人民政府实施意见,将非盈利组织的财务报表纳入注册会计师审计范围。其次,大力拓展新业务范围。2009年我国注册会计师行业非鉴证业务收入占业务收入的比重不足17%,而国际上注册会计师行业非传统业务收入占业务收入的比重近60%^[13]。我国大型会计师事务所与国际“四大”会计师事务所相比,差距悬殊:国际“四大”会计师事务所较早介入税务、管理咨询等新业务领域,服务独到,经验丰富。我国大型会计师事务所可以借鉴国际“四大”会计师事务所的经验,加速拓展税务、管理咨询、内部控制评价、并购重组、投资决策、业绩评价等新业务领域。最后,创新经营管理模式。中注协应改革现行管理体制,而会计师事务所应创新经营管理模式,转移核心竞争力,实施联合兼并,新设分支机构,深化与境内外事务所的合作,联手打拼国际市场,加速事务所国际化网络布局,实现规模经济效益。

3. 完善内控制度,推进文化建设。会计师事务所要进行以事务所文化为基础的人本管理,合理制定质量控制制度、风险管理制度、收入分配制度、财务管理制度和激励约束机制,调动事务所从业人员的积极性、创造性,增强凝聚力^[14];要优化事务所发展环境,创造成就事业的良好氛围,坚持“以事业留人,以感情留人,以待遇留人”,注重改善从业人员工作条件及生活待遇,拓展空间;抓好文化建设,促进企业文化逐渐形成,充分调动和发挥从业人员的积极性、创造性,吸引更多的高水平注册会计师不断加入进来,共同发展,这样不仅能做大,而且能够做强。

五、结论

通过聚类和决策树相结合的分析,本文认为,会计师事务所的可持续发展不仅与其自身的条件有

很强的相关性,而且与其规模大小密切相关。我国的会计师事务所经营规模不够大,CPA 队伍年龄偏大,学历偏低,领军人才数量偏少。会计师事务所要实现可持续发展,必须做到:一方面,中注协应改革现行管理体制,会计师事务所应创新经营管理模式,加大资金投入,大力引进和培养领军人才,优化 CPA 团队结构,另一方面,事务所本身要加大人才培养力度,制定人才发展战略规划,要实行以会计师事务所文化为基础的人本管理,结合自身实际情况,准确定位,开拓创新,沿着国际化方向,走出创新高效的可持续发展之路。本文局限之处在于仅对 2010 年和 2011 年 82 家大型会计师事务所相关数据进行实证分析,这对研究结果可能存在一定的影响。

参考文献:

- [1]许汉友,丁长青.中国大型会计师事务所运营绩效关键影响因素之灰色分析[J].审计与经济研究,2007(2):26-32.
- [2]姜虹,李美群,聂铁良.规模化和集中度视角下的审计市场结构研究[J].中国注册会计师,2010(8):57.
- [3]杨永森,杨美丽,张彬.我国会计师事务所规模经济效应的实证研究[J].财会月刊,2007(5):39.
- [4]杨永森.基于超效率模型的中国会计师事务所效率研究[J].中国注册会计师,2009(5):41.
- [5]王兵,陈惠珍.百家会计师事务所排名变迁的初步分析[J].中国注册会计师,2010(3):73.
- [6]贾宪威,刘文,侯路,等.我国会计师事务所审计运营效率的实证研究[J].财会月刊,2010(7):66.
- [7]朱志红.会计师事务所知识管理绩效定量评价模型[J].会计之友,2010(1):37.
- [8]胡晓莉.平衡计分卡及其在会计师事务所中的应用[J].中国注册会计师,2010(5):61.
- [9]李俊,夏斌.层次分析法、模糊综合评价法联用的会计师事务所综合评价[J].财会月刊,2010(11 下旬):81.
- [10]Xu Rui, Wunsch D II. Survey of clustering algorithms[J]. IEEE Transactions on Neural Networks,2005,16:645-678.
- [11]耿建新,房巧玲.我国会计师事务所规模研究——基于审计市场经验数据的聚类分析[J].会计研究,2005(3):22.
- [12]崔文迪,蔡佳佳.基于 K-means 算法和 FCM 算法的聚类研究[J].现代计算机:专业版,2007(10):7.
- [13]韩洁,徐蕊.中国注册会计师行业迈出跨越式发展关键一步[EB/OL]. [2011-12-12]. http://news.xinhuanet.com/fortune/2010-08/13/c_13444099.htm.
- [14]邱吉福.大型会计师事务所发展战略实证研究[J].集美大学学报:哲学社会科学版,2011(4):46.

[责任编辑:黄燕]

The Strategy Research of Sustainable Development of Accounting Firms: Based on Clustering and Decision Tree

QIU Ji-fu

Abstract: The effective combination of clustering and decision tree method can not only study the accounting firms from multi-perspectives, but also avoid the influence of man-made hypothesis and the loss of the sensitive characters and important factors. The study found that the accounting firm's sustainable development and its own condition have a strong correlation, and an excellent accounting firm should enrich its connotation and expand its extension to make it larger and stronger.

Key Words: audit market; data mining; clustering; decision tree; accounting firm; the strategy of sustainable development