

基于邻域粗糙集和支持向量机的企业财务 危机预警模型研究

祝建军，柳云鹏

(武汉纺织大学会计学院，湖北 武汉 430200)

摘 要：进行财务危机预警，科学选择预警指标是首先需要解决的问题。应用邻域粗糙集理论获得约简后的财务指标，并使用 T 检验和秩和检验验证约简指标合理性，结合基于支持向量机建立的企业财务危机预警模型对目标企业进行分类。研究结果显示，将邻域粗糙集理论和支持向量机结合，可以实现以较少的财务指标获得较高的分类精度。

关键词：邻域粗糙集；支持向量机；财务预警

Research on enterprise financial crisis warning model based on neighborhood rough set and support vector machine

Zhu Jianjun. Liu Yunpeng

(School of Accountancy, Wuhan University of Textile, Wuhan, China)

Abstract: Scientific selection of early warning indicators is the first problem of financial crisis early warning. This paper uses the theory of neighborhood rough set to figure out fewer financial indicators, and then uses the T test and rank sum test to verify the rationality of the reduction indicators. With the financial crisis warning model based on SVM, we classify whether the target enterprise is in a financial crisis. The results show that the combination of neighborhood rough set theory and SVM can achieve higher classification accuracy with fewer financial indicators.

Keywords: Neighborhood Rough Set; SVM; Financial Crisis Warning

当前，我国企业正处于内部市场经济体制改革进入深水区和外部全球经济一体化出现不确定性的关键时期，从宏观层面看，经济发展面临速度换挡节点、结构调整节点、动力转换节点；从微观层面看，经济发展过程中的诸多矛盾也体现各种组织形式的企业中，日益增加的财务风险是企业需要面对的问题之一。在激烈竞争且具有重大不确定性的环境中，如何及早发现和规避财务风险，建立在企业风险管理中发挥重要作用的财务危机预警系统，一直都是企业、学者和政府关注的重点问题之一。通过设置预警规则，并对市场变动和关键指标进行跟踪监测，财务危机预警系统能够对在一定内可能发生财务危机的企业进行识别，并提前发出警告。

任何企业自其诞生之日起就要与财务危机相伴，不论是富可敌国的跨国企业，还是开疆拓土的成长型公司，甚至一家公司在不同的发展阶段都可能要与其谋面。成立于 1908 年的

世界三大汽车生产商通用汽车 2009 年 6 月 1 日申请破产，这也是继克莱斯勒申请破产保护后，又一宗汽车业巨头破产案；有着 131 年辉煌历史的摄影胶卷业巨头柯达于 2012 年 1 月 19 日早间柯达提交了破产保护申请；在上世纪计算机发展浪潮中大放异彩的世界 500 强常客 IBM，也曾在 90 年代初遭遇连续三年巨额亏损，一度步入破产边缘；20 世纪 90 年代初，“巨人集团”以迅雷不及掩耳之势在中国商界取得了辉煌的成就，但在成立不到 6 年时就陷入了财务危机，并在之后申请破产；一直致力于打造互联网生态的上市科技型乐视科技于 2017 年 6 月遭遇严重的资金危机，公司发展前景面临重大不确定性。

在企业诸多危机中，财务危机可以直接影响企业的生存，因此在财务危机爆发前就对其预警，可以提醒企业提前进行应对。对企业管理者而言，其在财务危机爆发前就制定资金、材料采购、销售收款等方面的计划，把财务危机的影响降到最低；对企业投资者而言，提前发现财务危机可以帮助其作出正确的投资决策，并因投资失误造成的损失；对企业债权人而言，其可以根据危机预警提示做出提前收回贷款，要求债务人提供担保等决策，以保障债权安全。

1 文献综述

几十年来，理论界对如何准确预测企业财务危机进行了许多研究，重点主要集中在以下方面：

1) 企业财务危机的界定。因研究目的不同，不同学者对财务危机的界定和判断标准也出现了显著区别。Beaver (1966) 认为，当某企业出现破产、无法按期偿还债务、透支银行账户及未按时支付优先股股利中的任一事件时，可认为其出现了财务危机；Altman (1968) 依据美国破产法，认为“进入法定破产程序的企业”可判断其遭遇了财务危机；Deakin (1972) 将已经破产、无力偿还债务或正在进行清算的企业认定为发生财务危机的企业。由此可见，国外对财务危机的界定主要有：①已经破产；②进入破产程序；③无力还本付息；④资不抵债。但许多研究结果表明，我国资本市场为半强式有效市场，表现出了一定的信息不对称性，因此本文依据证监会对 ST 企业的认定标准判断企业是否发生财务危机。

2) 财务危机预测方法的研究。从传统的预测方法来看，统计判定分析是研究预测企业是否发生财务危机的重要方法。Beaver 应用单变量判定分析预测企业破产；Altman 应用多变量判定分析预测企业破产，简称“Z 分数模型”。

20 世纪末到 21 世纪初，随着社会经济和计算技术的急速发展和不断完善，财务危机预警理论与实务研究均进入发展的高峰期。以人工智能为代表新技术，因具有非线性、分布式运算能力等优点，被引入到企业财务预警领域的研究中，这些方法包括了时间序列回归预测、支持向量机、神经网络、案例推理、生存分析模型、随机游走估计等。

近年来支持向量机等技术的发展，给企业财务预测提供了新的工具和方法，应用这些工具和方法提高预测准确度成为该领域的重要发展方向。Ohlson (1980)，吴世农、卢贤义 (2001) 等运用了 Logistic 方法；Wilson & Sharda (1990)，Marco & Varetto (1994)，刘洪、何光军 (2004) 等采用了神经网络分析方法；Beyon & Peel (2001) 等采用粗糙集方法；Van Gestel (2006) 采用最小二乘支持向量机方法。

目前，许多学者利用改进的神经网络、集成方法建模、仿真技术以及各种动态预警技术对企业财务危机预警进行研究，或将多种技术结合起来构建模型，以达到更强的预测能力。

3) 建模指标的选择。Altman 等学者使用传统的财务指标作为判定变量；Beaver (1966) 选取现金流类、净收益类等 6 组财务指标，对企业是否发生财务危机进行一元判定研究；Gentry、Aziz、Emanuel 和 Lawson 以现金流量信息构建财务预警判定指标，并取得良好效果。周首华 (1996) 在 Z 值模型的基础上引入现金流量指标，建立了 F 分数模式 (Failure Score Model)。通过对前人研究进行总结可以发现，企业财务危机的建模指标可归纳为偿债能力类

指标、获利能力类指标、营运能力类指标、成长性指标、收益质量类指标等。从财务预警指标的选择看,选择过多的指标会造成信息的冗余,而指标过少又可能造成信息的缺失。部分研究对财务预警指标的选择主要基于指标间的相关性,对于筛选的标准也无固定的规则,因此包含一定主观因素。此外,许多研究采用的是统计方法对指标进行筛选,而统计数据通常需要服从某些分布规律,因此使得其应用范围受到限制。

为了应对这些问题,一些学者将粗糙集方法引入企业财务预警的研究中。粗糙集方法是由波兰学者 Z. Pawlak (1982) 提出的一种处理不确定性、模糊性的软计算工具,其基本思想是对数据集进行分类并归纳形成概念和规则,基于分类结果建立等价关系及对目标进行近似,从而实现知识发现、规则提取和属性约简。与其他处理不确定和不精确问题的理论相比,粗糙集理论最显著的区别是它不需要所处理数据集外任何先验信息或者附加信息,可以完全由已知数据来获得新的知识。粗糙集理论不包含处理不精确或不确定原始数据的机制,所以其与概率论、模糊数学等处理不确定或不精确问题的理论有较强互补性。因此,梁吉业 (2005) 认为粗糙集理论为企业财务危机研究指标的选取提供了一条与传统统计方法截然不同的新方法。

基于粗糙集易于处理、鲁棒性强、成本低廉等优点,许多学者也将其运用在企业财务预警的研究中。马若微 (2005, 2006) 利用粗糙集理论中的重要度原理和信息熵概念,对备选指标进行属性约简,并以此作为企业财务危机预警模型的指标;王宗军等 (2007),刘彦文、戴红军 (2007) 等运用经典粗糙集思想,以离散化的财务指标为基础,对财务危机预警问题进行研究。鲍新中 (2013) 运用层次聚类对目标企业进行分类,根据分类结果将企业财务状况分为健康、良好、中等、轻警、重警 5 类,运用神经网络构建财务预警模型,将分类结果作为输出层,将粗糙集约简后的指标作输入层。

粗糙集方法在财务危机预警邻域尤其是指标选择方面受到重视。但 Pawlak 粗糙集定义在经典的等价关系和等价类基础上,适于处理名义型变量。对于数值型属性,则需要离散化处理,把数值型属性转化为符号型属性才能使用粗糙集方法,而离散化处理容易造成信息损失,而且约简结果受到离散化效果的影响。

为了解决这一问题,部分学者提出了邻域粗糙集理论,并得到了广泛应用。Lin (1988) 提出邻域粗糙集模型的概念,Yao (1998) 和 Wu (2002) 论证了邻域粗糙集模型的性质,HuQ. et al. (2008) 以拓扑空间中邻域的概念为基础,研究了基于邻域粗糙集模型的数值数据特征选择算法。以空间中点的邻域为基础,邻域粗糙集模型可以将论域构造成单个的信息粒,因此邻域也可以被理解为论域中的基本信息粒子。与经典粗糙集方法相比,邻域粗糙集模型无须离散化就可直接处理数值型数据,因而其应用范围比经典粗糙集模型更广泛。而与企业财务预警相关的指标多以数值型存在,本文试图运用邻域粗糙集思想对财务危机预警备选指标进行属性约简筛选,并在此基础上构建企业财务危机预警模型。

2 基于邻域粗糙集模型的指标算法

对于信息系统 $IS = \langle U, C, V, f \rangle$, 其中 $U = \{x_1, x_2, \dots, x_n\}$ 表示非空有限集合, 为论域。 C 是 U 的属性集合, V 是值域, $f: U \times C \rightarrow V$ 是一个信息函数, 表示样本与其属性取值的映射关系。如果 $C = A \cup D$, 其中 A 是条件属性, D 是决策属性, 且 $A \cap D = \emptyset$, $D \neq \emptyset$ 。 $\Delta(x_i, x_j)$ 为距离函数, 表示元素和元素之前的距离。本文采用欧几里德距离

$\Delta(x_i, x_j) = (\sum_{k=1}^N (f(x_i, a_k) - f(x_j, a_k))^2)^{1/2}$ 作为距离计算函数。

在论域 U 中, $\forall x_i \in U$, $\delta \geq 0$, 称点集 $\delta(x_i) = \{x \mid x \in U, \Delta(x, x_i) \leq \delta\}$ 为 x 的 δ 邻域。将 U 和邻域关系 N 称为一个邻域近似空间, 记作 $NS = (U, N)$ 。 $\forall X \subseteq U$, 则子集 X 在邻域近似空间 NS 中的上近似和下近似分别为: $\overline{NX} = \{x_i \mid \delta(x_i) \cap X \neq \emptyset, x_i \in U\}$, $\underline{NX} = \{x_i \mid \delta(x_i) \subseteq X, x_i \in U\}$ 。

$\forall B \subseteq A$, 将条件属性子集 B 和决策属性 D 的邻域关系记为 N_B 。将 U 和 N_B 构成的邻域近似空间记为 $N_B S = (U, N_B)$ 。在邻域近似空间 $N_B S$ 中, 决策属性 D 将集合 U 划分为 N 个等价类 (X_1, X_2, \dots, X_N) 。 $\forall B \subseteq A$, 决策属性 D 关于条件属性子集 B 的上、下近似分别为: $\overline{N_B D} = \bigcup_{i=1}^N \overline{N_B X_i}$, $\underline{N_B D} = \bigcup_{i=1}^N \underline{N_B X_i}$ 。其中,

$$\overline{N_B X} = \{x_i \mid \delta_B(x_i) \cap X \neq \emptyset, x_i \in U\}, \quad \underline{N_B X} = \{x_i \mid \delta_B(x_i) \subseteq X, x_i \in U\}.$$

根据关于正域的结论可得, $N_B S$ 的下近似也是该近似空间的正域 ($Pos_B(D)$)。正域的大小反映了分类问题在给定属性空间中的可分离程度。正域越大, 表明各类的重叠区域即边界越少。定义决策属性 D 对条件属性子集 B 的依赖性为 $\gamma_B(D) = \frac{|Pos_B(D)|}{|U|}$, $0 \leq \gamma_B(D) \leq 1$ 。

该指标表明根据条件属性子集 B 对样本集合进行归类, 其中能够完全包含于决策属性 D 的样本所占全体样本的比率。

属性 α 属于条件属性子集 B , 如果 $\gamma_{B-\alpha}(D) < \gamma_B(D)$, 则 α 相对于 B 是必不可少的; 否则, $\gamma_{B-\alpha}(D) = \gamma_B(D)$, 则 α 相对于 B 是多余的。如果条件属性子集 B 满足, $\gamma_B(D) = \gamma_A(D)$, 则称 B 是 A 的一个约简。

$\forall \alpha \in A - B$, 属性 α 相对于条件属性子集 B 的重要度可定义为, $Sig(\alpha, B, D) = \gamma_{B \cup \alpha}(D) - \gamma_B(D)$ 。如果属性 α 加入条件属性子集 B 后, 决策属性 D 对 B 的依赖度没有发生变化, 可得属性 α 的重要度为 0, 说明其可从条件属性子集 B 中去掉。本文以胡清华 (2008) 等人提出的前向贪心约简算法为基础, 将重要度高于 10% 的指标纳入指标集合, 当新指标的加入不再影响分类精度时算法终止。

3 支持向量机的基本原理

支持向量机是一种通用的前馈网络类型，由 Vapnik 首先提出。支持向量机的主要思想是建立一个超平面作为决策曲面，使得正例和反例之间的隔离边缘最大化，其也是结构风险最小化方法的近似。其归纳原理如下，分类模型在测试数据上的误差率以训练误差率和一个依赖于 VC (Vapnik-Chervonenks Dimension) 维数的项的和为界。在线性可分模式下，支持向量机对于前一项的值为零，并且使第二项最小化，因此在分类问题上支持向量机能提供很好的泛化能力。在训练样本有限的情况下，支持向量机构造的分类模型具有很好的推广能力，而不需要所研究问题的先验知识。支持向量机可解决高维特征问题，从而避免“维数灾难”，且其总是寻找全局最优解，而避免仅提供局部最优解。

支持向量机的基本思想是根据定义的内积函数对输入空间进行非线性变换后，将其映射到一高维空间，并在高维空间中寻找一具有最大的分类间隔的线性超平面。在应用于分类问题时，支持向量机的原理如下：将输入的样本通过非线性函数映射到一特征空间中，在此空间中构造最优分类超平面，使两类样本（可推广到多类样本）具有最大分类间隔。该非线性函数仅与低维输入向量和特征空间的点积有关，其可用核函数 K 来代替。支持向量机的判

别函数为： $f(x) = \text{sng}(\sum_{i=1}^n a_i^* y_i K(x_i, x) - b^*)$ 。式中 y_i 是类别符号，且 $y_i \in \{0,1\}$ ； a_i^* 是

常数，且 $a_i^* \geq 0$ ； x_i 和 x 是向量； b^* 是分类域值。

4 研究步骤与结果分析

基于上述思路，本文提出一种基于邻域粗糙集—支持向量机的财务预警模型。通过运用邻域粗糙集理论对财务指标进行筛选，将筛选后的财务指标和目标企业是否被 ST 分别作为输入和输出，使用支持向量机得出企业是否发生财务危机的分类模型，并以此模型对测试样本的财务预警状况进行预测，随后与真实情况进行对比，判断分类模型的准确率。

4.1 确定样本，初选财务指标

与企业财务危机相关的财务指标众多，参考财务分析理论为基础并结合前人的研究成果，本文将其分类为偿债能力类指标、获利能力类指标、营运能力类指标、成长性指标、收益质量类指标共五类。随后，我们从每类指标中初选一定数量的指标组成初选指标体系，并以此为基础运用邻域粗糙集理论对其进行属性约简，将约简后的属性集作为最终的财务预警指标体系。我们初取了 16 个财务指标，并对其进行约简。具体指标如表 1 所示。

表 1 初选财务指标

偿债能力类指标	a1 流动比率	a3 有息负债/全部投入资本
	a2 资产负债率	a4 已获利息倍数
获利能力类指标	a5 权益净利率	a7 销售净利率
	a6 资产净利率	
营运能力类指标	a8 存货周转率	a10 总资产周转率
	a9 应收账款周转率	

成长性指标	a11 收入同比增长率	a13 总资产同比增长率
	a12 净利润同比增长率	
收益质量类指标	a14 扣除非经常损益后的净利润/净利润	a16 经营活动产生的现金流量净额/总负债
	a15 经营活动产生的现金流量净额/营业收入	

为避免运营模式、财务状况和市场规模等方面的不同对财务指标造成影响，我们拟选取沪深两市制造业上市企业为研究样本。本文以“因财务状况异常而被特别处理”（ST）作为企业是否发生财务危机的判断标准。通常情况下， $t-1$ 年度的财务报告将在 t 年公布，此外证监会对 ST 企业的判定标准包括：最近两个会计年度的审计结果显示的净利润均为负值；最近一份经审计的财务报告对上年度利润进行调整，导致连续两个会计年度亏损。因此本文选取沪深两市制造业企业 2009 年-2014 年的财务指标作为粗糙集的条件属性，并采用 2011 年-2016 年的企业财务状况作为粗糙集的决策属性。本文所采用的数据主要来自于锐思数据库以及上海证券交易所网站。本文使用软件 stata、matlab2014b 等软件进行数据处理。

为避免运营模式、财务状况和企业规模等因素的不同对财务指标造成影响，选取制造业的上市企业数据进行分析。样本选取标准为：（1）2011 年-2016 年被 ST 的制造业企业为样本，并根据同年度、同行业、相近规模的标准选取相应的未被 ST 的企业；（2）剔除了发行 B 股或 H 股的企业，其原因是发行 B 股或 H 股的企业除需满足 A 股发行条件外还需满足其他特殊要求，为保证样本的代表性和信息纯度，在样本选取过程中将其剔除。经过筛选后共选取了 496 家上市企业，其中 ST 企业 and 非 ST 企业各 248 家，（2011 年 142 家，2012 年 96 家，2013 年 52 家，2014 年 52 家，2015 年 74 家，2016 年 80 家）。

同时，为保证数据质量和模型计算需要，本文对数据进行了以下处理：（1）对所有指标进行归一化处理；（2）因少量 ST 企业的部分指标极端异常，为避免这些极端指标对数据造成影响，在 1%和 99%百分位上对所有指标进行了缩尾处理。由于篇幅所限，部分归一化处理后指标体系如表 2 所示。本文所用数据来源于 choice 和 ifind 数据库。

表 2 部分归一化处理后指标体系

股票代码	a1	a2	a3	a4	a5	a6	a7	...	a16
000048.SZ	0.0054	0.3826	0.3639	0.0012	0.9517	0.5582	0.8992	...	0.1298
000068.SZ	0.0045	0.2823	0.3575	0.0010	0.8740	0.0000	0.7761	...	0.1648
000403.SZ	0.0046	0.3367	0.3607	0.0012	0.9904	0.5150	0.8951	...	0.1172
000408.SZ	0.0135	0.2787	0.3571	0.0011	0.9200	0.3903	0.8595	...	0.1105
000509.SZ	0.0065	0.3884	0.3616	0.0012	0.9517	0.4962	0.8955	...	0.1081
...

4.2 预警指标的选择——备选指标属性约简

对初选财务指标数据归一化处理后，应用邻域粗糙集模型对初选财务指标进行约简。本文在 matlab2014b 中应用前向贪心算法对财务指标进行属性约简。算法思想如下：以空集为起点，每次计算全部剩余属性的属性重要度；以属性重要度为判断标准，将重要度最高的属性加入属性约简集合中；循环执行以上步骤，直到所有剩余属性的重要度都低于给定的重要度；程序结束后即可得到约简属性集合。

在财务危机预警模型中，论域由我们选取的样本企业构成，不同的财务指标和企业是否被 ST 分别构成条件属性和决策属性。由于该算法计算出的约简属性个数和属性重要度受到邻域 δ 大小的影响，我们将邻域 δ 赋值为各指标财务标准差乘以一系数，该系数以 0.1 为起

点，每次变化 0.1，其变化区间为 0.1 到 1。在基于前向贪心算法的粗糙集属性模型下，每一邻域 δ 都可以计算出一个约简后的财务指标集合。本文选取了重要度 10%以上的财务指标。表 3 列示了根据不同邻域 δ 计算出的约简财务指标体系，表 4 列示了根据不同邻域 δ 计算出的约简财务指标重要度。

表3 不同邻域 δ 计算出的约简财务指标体系

系数	约简财务指标体系												
0.10	2	6	8	10	11	12	15						
0.20	1	2	4	5	6	8	10	11	12	13	14		
0.30	1	2	4	5	6	7	8	10	11	12	13	14	
0.40	1	2	4	5	6	8	10	11	12	13	14	15	16
0.50	2	4	5	6	7	9	10	13	14				
0.60	1	2	4	5	6	9	10	11	12	13	14	15	
0.70	2	5	6	13	14	15							
0.80	2	5	6	13	14	15							
0.90	2	4	5	13	14	15							
1.00	2	4	5	13	14	15							

表4 不同邻域 δ 计算出的约简财务指标重要度

系数	约简财务指标重要度												
0.10	0.11	0.48	0.76	0.90	0.94	0.97	0.99						
0.20	0.85	0.27	0.40	0.52	0.63	0.66	0.69	0.72	0.74	0.76	0.77		
0.30	0.58	0.11	0.16	0.26	0.34	0.40	0.44	0.46	0.49	0.51	0.53	0.54	
0.40	0.54	0.99	0.14	0.20	0.25	0.29	0.31	0.35	0.36	0.39	0.43	0.42	0.43
0.50	0.46	0.97	0.13	0.16	0.19	0.28	0.23	0.25	0.26				
0.60	0.44	0.85	0.11	0.14	0.16	0.17	0.19	0.20	0.22	0.23	0.24	0.25	
0.70	0.36	0.73	0.99	0.12	0.15	0.16							
0.80	0.36	0.69	0.93	0.11	0.14	0.15							
0.90	0.32	0.58	0.77	0.97	0.13	0.11							
1.00	0.28	0.54	0.69	0.83	0.95	0.15							

为检验使用邻域粗糙集模型获得的约简财务指标是否合理，本文拟采用 T 检验判断 ST 和非 ST 企业的财务比率指标之间是否存在显著差异。同时，本文共选取了 496 家上市企业，为避免因样本分布规律不符合正态分布假设导致 T 检验失效，同时汇报了秩和检验的结果。表 5 列示了 ST 和非 ST 企业各财务指标的均值，以及 T 检验和秩和检验的结果。从检验结果中可以看出，除应收账款周转率外，ST 和非 ST 企业的其他各财务指标均在 5%水平上存在显著性差异。

表 5 样本描述性统计、T 检验及秩和检验结果

变量	均值		T 检验		秩和检验	
	非 ST	ST	p 值	t 值	Z 值	p 值
a1 流动比率	1.69	1.09	0.000	5.428	0.000	10.60
a2 资产负债率	50.49	73.15	0.000	-9.51	0.000	-8.81
a3 有息负债/全部投入资本	37.18	70.72	0.000	-5.31	0.000	-6.48

a4 已获利息倍数	24.25	-3.15	0.000	8.26	0.000	16.26
a5 权益净利率	25.32	-24.84	0.000	14.45	0.000	16.94
a6 资产净利率	15.42	-7.20	0.000	19.20	0.000	18.03
a7 销售净利率	12.16	-29.35	0.000	9.51	0.000	16.26
a8 存货周转率	5.38	4.56	0.031	2.16	0.003	3.02
a9 应收账款周转率	107.34	93.94	0.842	0.20	0.202	1.28
a10 总资产周转率	1.23	0.52	0.000	13.96	0.000	13.88
a11 收入同比增长率	15.74	-4.63	0.000	5.71	0.000	8.82
a12 净利润同比增长率	18.34	-3.11	0.000	11.19	0.000	12.07
a13 总资产同比增长率	38.96	-1290.25	0.000	7.37	0.000	10.66
a14 扣除非经常损益后的 净利润/净利润	91.62	5.67	0.000	4.54	0.000	-4.77
a15 经营活动产生的现金 流量净额/营业收入	10.87	-13.94	0.000	5.76	0.000	9.31
a16 经营活动产生的现金 流量净额/总负债	0.27	0.01	0.000	8.16	0.000	12.32

注：t 检验值为异方差双样本配对检验值；标准值是 p-值 ≤ 0.05 。原假设为：ST 公司的财务指标和非 ST 公司的财务指标在 0.05 的显著性水平上没有差异。

4.3 基于支持向量机的企业财务预警模型

基于以上数据预处理及约简获得的财务指标，本文拟使用支持向量机构建 ST 和非 ST 企业的分类模型。根据上述分析，构造了样本集 (x, y) ，其中 x 是约简后的财务指标集合， y 是决策属性，即企业是否发生财务危机的归类，根据前文选取的分类方法，非 ST 企业归类未发生财务危机的企业， $y = 0$ ；ST 企业归类发生财务危机的企业， $y = 1$ 。同时，支持向量机中不同的内积核函数将形成不同的算法，目前研究最多的核函数主要有多项式函数、径向基函数、Sigmoid 函数 3 种。本文中支持向量机选取的内积核函数是最常用的径向基函数，其形式为：

$$K(x, x_i) = \exp \left\{ -\frac{|x - x_i|^2}{\sigma^2} \right\}。$$

本文根据不同邻域 δ 大小获得的约简后财务指标体系和企业是否为 ST 企业的标签带入支持向量机中进行训练和测试。为提高预测结果的稳定性和准确性，本文分别采用了十折交

叉验证法和网格参数寻优法。十折交叉验证法是常用的精度测试方法，该方法将数据集分成十份，轮流将其中九份做为训练集，一份做测试，将十次结果的均值作为对算法精度的估计。网格参数寻优法式是获取支持向量机最优参数的方法，其先选定一组参数的范围，然后将它们的准确率用等高线连接起来，确定准确率最高的一段步长，再据此确定使模型最优的参数。表 2 汇报不同邻域 δ 下的约简后指标数量和支持向量机分类准确率。

表 2 约简后指标数量和支持向量机分类准确率

系数	0.1	0.2	0.3	0.4	0.5	0.6	0.7	0.8	0.9	1.0
约简后指标数量	7	11	12	13	9	12	6	6	6	6
支持向量机分类准确率	95.7%	96.2%	96.0%	95.7%	96.9%	95.7%	97.3%	97.3%	96.0%	96.0%

从表 2 我们可以看出，分别以系数 0.7 和 0.8 确定的邻域 δ ，约简得出的财务指标均为 a2 资产负债率、a5 权益净利率、a6 资产净利率、a13 总资产同比增长率、a14 扣除非经常损益后的净利润/净利润、a15 经营活动产生的现金流量净额/营业收入。从报表分析角度来看，约简后的六个财务指标分别反映了企业的偿债能力、获利能力、成长性和收益质量四个方面，可以认为综合反映了企业的经营情况和财务状况。随后，将网格寻优方法得出的参数 $c = 48.5029$ ， $g = 3.0314$ 代入基于支持向量机构建的企业财务预警模型中，可获得最高的分类准确率，97.3%。图 1 和图 2 均为 SVC 参数选择结果图。

图 1 SVC 参数选择结果图（等高线图）

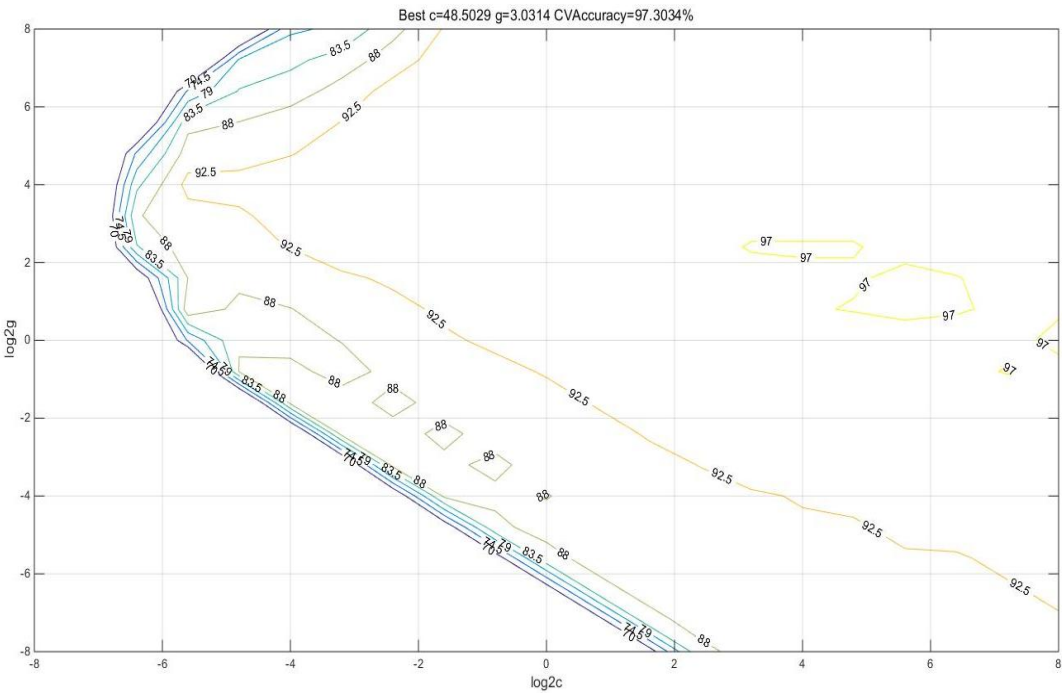
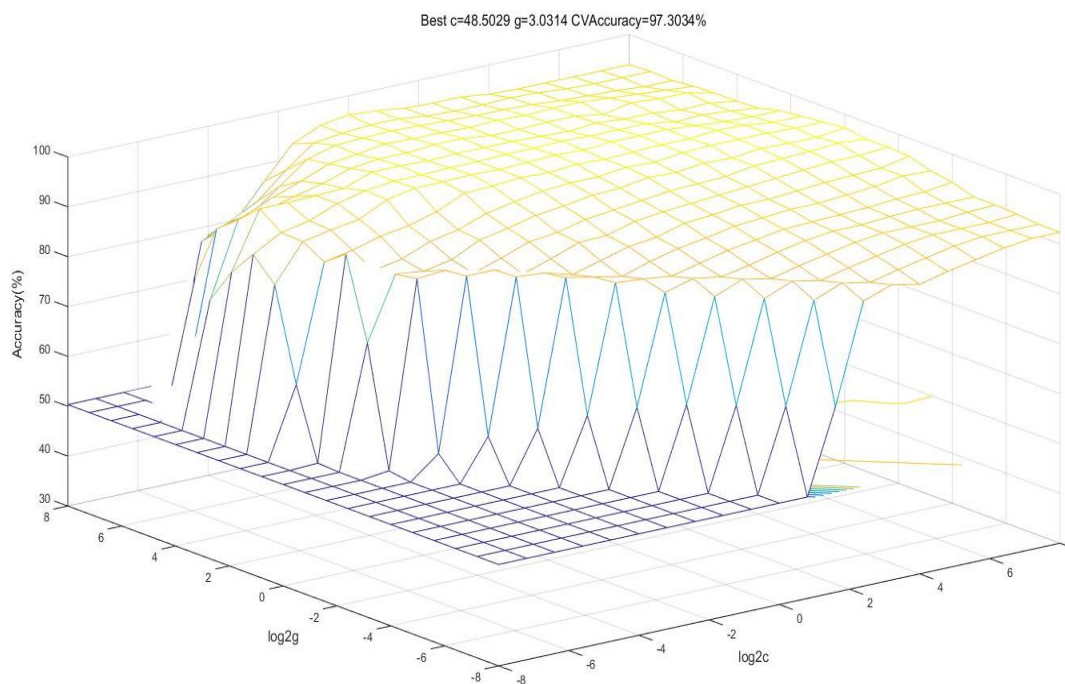


图 2 SVC 参数选择结果图（3D 视图）



4.4 结果分析

本文运用邻域粗糙集方法,将不同大小的邻域 δ 带入模型中得到十组约简后的财务指标集。将以 0.7 倍和 0.8 倍的各指标标准差作为邻域 δ 时,获得的六个约简财务指标带入基于支持向量机构建的企业财务预警模型中,可以获得最高的分类准确率。结果表明,在邻域粗糙集方法下,运用较少的指标反而达到了获得更高分类能力。在实际应用中,我们可以利用约简结果,仅重点观测这些更敏感反映企业财务状况的指标,这些数量较少、敏感性较高的指标可以降低监测成本,也可以改善了监测效率,且避免了监测资源分散和信息过度的情况。同时,支持向量机作为一种基于小样本学习理论的通用学习算法,具有严格的理论基础,本文运用基于支持向量机构建的企业财务危机预警模型取得了较高的分类准确率。

5 结论

随着市场竞争的日益激烈,准确的财务危机预警可以帮助企业为可能到来的危机提前准备,也能使股东、债权人、员工等利益相关者得利益获得更好的保障。以往在对企业进行财务危机预警时,需要面对多而杂的财务指标,而指标的选择是预警过程中需要考虑的重要问题。根据本文研究结果,应用邻域粗糙集理论得到约简后的财务指标,可实现仅利用较少的指标就对目标企业是否发生财务危机进行更准确的分类。同时,支持向量机不需要特定问题的经验知识,在训练样本有限的情况下,可很好地控制分类模型的推广能力,本文将邻域粗糙集理论和支持向量机结合获得较好的分类效果。因此,将邻域粗糙集理论和支持向量机结合应用在企业财务预警中具有重要的现实意义,提供一种切实可行、效果较好的解决方法。

参考文献:

- [1] Altman E I, Marco G, Varetto F. Corporate distress diagnosis: Comparisons using linear discriminant analysis and neural networks (the Italian experience) [J]. Journal of Banking & Finance, 1994, 18(3):505-529.
- [2] Altman E I. Financial Ratios, Discriminant Analysis and the Prediction of Corporate Bankruptcy[J]. Journal of Finance, 1968, 23(4):589-609.
- [3] Beaver W H. Financial Ratios As Predictors of Failure[J]. Journal of Accounting Research, 1966,

4(1):71-111.

[4] Beynon M J, Peel M J. Variable precision rough set theory and data discretisation: an application to corporate failure prediction[J]. Omega, 2001, 29(6):561-576.

[5] Deakin E B. A Discriminant Analysis of Predictors of Business Failure[J]. Journal of Accounting Research, 1972, 10(1):167-179.

[6] Fitzpatrick P J. A Comparison of the Ratios of Successful Industrial Enterprises with those of Failed Companies[J]. Analyse Molecular Do Gene Wwox, 1932:598-605.

[7] Hu Q, Yu D, Xie Z. Neighborhood classifiers[M]. Pergamon Press, Inc. 2008.

[8] Jiye Liang, K.S. Chin, Chuangyin Dang, et al. A new method for measuring uncertainty and fuzziness in rough set theory[J]. International Journal of General Systems, 2002, 31(4):331-342.

[9] Lin T Y. Granular computing on binary relations I: data mining and neighborhood systems[J]. Rough Sets in Knowledge Discovery, 1998(2):165-166.

[10] Pawlak Z, Skowron A. Rough sets and Boolean reasoning[J]. Information Sciences, 2007, 177(1):41-73.

[11] Pawlak Z. Rough Sets: Theoretical Aspects of Reasoning about Data[M]. Kluwer Academic Publishers, 1992.

[12] Vapnik V N. The Nature of Statistical Learning Theory[M]// The nature of statistical learning theory. Springer, 1995:988 - 999.

[13] Vapnik V, Levin E, Cun Y L. Measuring the VC-dimension of a learning machine[M]. MIT Press, 1994.

[14] Vapnik V. Estimation of Dependencies Based on Empirical Data (in Russian)[J]. 2015.

[15] 鲍新中, 杨宜. 基于聚类-粗糙集-神经网络的企业财务危机预警[J]. 系统管理学报, 2013, 22(3):358-365.

[16] 陈静. 上市公司财务恶化预测的实证分析[J]. 会计研究, 1999(4):31-38.

[17] 胡清华, 于达仁, 谢宗霞, 等. 基于邻域粒化和粗糙逼近的数值属性约简[J]. 软件学报, 2008, 19(3):640-649.

[18] 胡清华, 赵辉, 于达仁. 基于邻域粗糙集的符号与数值属性快速约简算法[J]. 模式识别与人工智能, 2008, 21(6):730-738.

[19] 梁吉业. 信息系统中的不确定性与知识获取[M]. 科学出版社, 2005.

[20] 刘碧森, 姚宇, 王玲. 基于支持向量机的企业财务预警模型[J]. 电子科技大学学报, 2007(s1):445-447.

[21] 刘洪, 何光军. 基于人工神经网络方法的上市公司经营失败预警研究[J]. 会计研究, 2004(2):42-46.

[22] 刘彦文, 戴红军. 基于粗糙集-神经网络的财务危机预警模型实证研究[J]. 科研管理, 2007, 28(6):138-142.

[23] 卢增祥, 李衍达. 交互支持向量机学习算法及其应用[J]. 清华大学学报(自然科学版), 1999, 39(7):93-97.

[24] 马超群, 吴丽华, MaChao-qun, 等. 基于邻域粗糙集和神经网络的财务预警研究[J]. 软科学, 2009, 23(11):123-126.

[25] 马若微, MaRuowei. 基于RS与ANN的上市公司财务困境预测模型的实证研究[J]. 南开管理评论, 2006, 9(3):85-91.

[26] 马若微. 基于粗糙集与信息熵的上市公司财务困境预警指标的确立[J]. 当代经济科学, 2005, 27(2):45-50.

[27] 宋鹏, 梁吉业, 曹付元. 基于邻域粗糙集的企业财务危机预警指标选择[J]. 经济管理, 2009(8):130-135.

- [28] 王宗军, 李红侠, 邓晓岚. 基于粗糙集的企业财务困境预警[J]. 统计与决策, 2007(4):137-140.
- [29] 吴世农, 卢贤义. 我国上市公司财务困境的预测模型研究[J]. 经济研究, 2001(6):46-55.
- [30] 杨淑娥, 黄礼. 基于 BP 神经网络的上市公司财务预警模型[J]. 系统工程理论与实践, 2005, 25(1):12-18.
- [31] 张根明, 向晓骥. 基于支持向量机的上市公司财务预警模型研究[J]. 科技管理研究, 2007, 27(4):234-235.
- [32] 周首华, 杨济华. 论财务危机的预警分析——F 分数模式[J]. 会计研究, 1996(8):8-11.