



基于变精度粗糙集神经网络的企业失败风险预测研究

尹鹏¹, 王宗军¹, 肖德云²

1 华中科技大学 管理学院, 武汉 430074

2 武汉理工大学 经济学院, 武汉 430070

摘要:描述粗糙集、变精度粗糙集和神经网络的相关概念和应用机理,在此基础上构建基于变精度粗糙集的强耦合粗糙集神经网络,即变精度粗糙集神经网络模型,并以此作为预测中国上市企业失败风险的研究方法。借助信息熵和T检验等手段对财务指标和非财务指标进行筛选,建立企业失败风险预测评价指标体系。在时间跨度上,选择 $(t-3)$ 年作为研究起点,避免人为夸大预测精度。将被特别处理(ST)的上市企业作为界定企业失败的标准,以160家ST企业和160家配对非ST企业作为测试样本,利用FUSINTER技术离散化相关数据,运用粗糙集、变精度粗糙集、粗糙集神经网络和变精度粗糙集神经网络等模型进行实证研究。分析结果表明,在由78家ST和78家非ST企业组成的检验样本中,变精度粗糙集神经网络的预测准确率要高于其他模型,且所需训练时间更短,生成规则更少,训练的整体效果最为理想,由此可以认为该模型比较适用于对中国上市企业进行失败风险的预测研究。

关键词:中国上市企业; 风险预测; 变精度粗糙集神经网络模型; FUSINTER 数据离散化法

中图分类号:F272.1

文献标识码:A

文章编号:1672-0334(2010)04-0015-12

1 引言

随着中国股票市场的不断完善和发展,对企业经营失败的预测成为指导投资者和企业债权人等利益相关者规避风险、做出有效投资决策的重要手段,但由于信息不对称等现象的存在,许多投资者或企业债权人难以获知企业的真实经营状况,上市企业尽管有披露自身财务状况的义务,但何时披露取决于企业自身,其利益相关者无法提前获取关键信息。因此,如果能有效预测企业的失败风险,对于保护广大投资者切身利益以及促进中国上市企业的健康发展都具有极其重要的意义。

对上市企业失败风险的预测主要是运用相关模型进行定量化的研究,从总体上看包含两个方面,一是预测指标体系的构建,二是评价模型的选择或建立。指标体系的构建不仅是指标的选取(主要是财务指标和非财务指标),还涉及指标的选取依据和筛选方法,评价模型是选择已有模型或创建新模型。目前国内外许多学者已在这些方面取得了大量突破

性进展^[1],但由于指标体系和评价模型的选择具有一定的时代特性和对使用环境的严格要求,因此当以中国上市企业作为研究对象时,不能简单照搬国外或早期已有的预测方法,必须结合当前中国企业特征以及所处国内外金融环境等因素加以分析,这种预测的复杂性和不确定性给研究工作造成一定困难,如何合理选取指标和模型(或建立新模型)进而提高对企业失败风险预测的准确性将是本研究的重点。

2 相关研究评述

对企业失败风险的预测一直是管理学的重要研究领域之一,相关的预测方法不断推陈出新且被日益完善,总体而言,可将其分为统计方法和智能技术两大类^[2]。多元线性判别式分析(multiple discriminant analysis, MDA)是国外早期统计方法的代表,由于该方法从整体层面研究影响公司未来发展的因素,因此在公司出现财务危机前两年具有较高的预

收稿日期: 2010-03-18 修返日期: 2010-05-25

基金项目: 国家自然科学基金(70872033)

作者简介: 尹鹏(1980-),男,江西永新人,华中科技大学管理学院博士研究生,研究方向:企业经营困境评价和金融风险管理等。E-mail: davidyin07@sina.com

测精度,但MDA对评价对象有严格的假设要求(如正态分布、等协方差、先验概率等),因此在实际运用中其有效性受到一些学者的质疑。

随着计算机的普及和相关技术的发展,从20世纪90年代开始,一批先进的智能技术被陆续运用到预测研究中。如Swicegood、Ryu和Boyacioglu等采用神经网络(neural networks, NN)预测企业的失败风险^[3-5],根据Swicegood等的研究结果,反向式传播神经网络(back propagation trained neural network, BPNN)模型在预测区域银行和社区银行失败风险时的正确率分别为81.400%和78.250%^[3]。

尽管NN模型具有较高的预测准确度,但它最大的不足在于当其输入维数过高时,隐含规则将以几何级数增加,这会导致网络结构迅速扩大;此外,神经网络学习速率固定,会存在局部最小点问题^[6]。这些缺陷都将导致NN模型收敛速度变慢,训练时间延长,严重时甚至可能引发网络瘫痪^[7]。

1982年Pawlak提出一种基于数集理论的非参数方法,即粗糙集理论(rough set theory, RST)^[8]。RST的一大优点是它不需要数据的先验知识和额外信息,且可以在不损失任何知识信息的情况下去除多余属性信息系统中的冗余属性^[9],因此一些学者将RST属性约简的优势和NN模型相结合,把RST中的线性信息系统导入到神经网络中进行训练,构造了粗糙集神经网络模型(rough neural networks, RNN)^[10]。

Jelonek等通过实证研究认为,RST方法中的属性约简可有效地减少NN模型中的输入变量,因此可将基于代表性数据约简的RST与BPNN模型较好地结合、互补,从而达到提高混合模型预测精度的目的^[10]。经检验,数据约简后的神经网络学习时间比约简前提高4.720倍,数据的误分类小于11.060%。

Ahn等最先将RNN模型应用到企业失败风险的预测中,他们将2D约简数据预处理算法运用到模型的计算中(所谓2D是指信息系统的约简包含横向和纵向两方面,其中横向约简指属性约简,纵向约简指去除相互矛盾或重复的评价对象)^[11],该算法可以有效防止神经网络中的过拟合问题,同时还能减少神经网络的训练时间^[12]。

中国学者在进行相关研究时虽然起步较晚,但也取得了一定成绩^[7,13]。刘彦文等在对中国上市公司财务危机预警研究中运用RST模型对神经网络的输入进行预处理,通过知识约简的方法消除冗余信息,同时利用动量添加法和参数自适应算法修正传统BP算法,提高网络训练速度和预测效率,经计算,模型综合准确率为82%^[7]。

尽管国内外学者对企业失败风险的预测研究(特别是运用RNN模型)已取得一定成绩,但仍存在以下几点不足。①在RNN模型中,RST的分类必须是完全正确或肯定的,没有某种程度的“包含”或“属于”,因此对误分类非常敏感,且缺乏对噪声数据的适应能力^[14];②RST模型只能处理离散化数据,而传统离散化方法大多采用卡方统计来衡量变量的

拟合优度,进而对数据进行离散化处理,且一次只能离散化决策表子区域中的一组条件属性,因此这类方法对数据的处理过程较为繁琐和费时;③在选取一些预测指标时缺乏相关理论依据做支撑,存在一定的主观性。

本研究借助一种改进的RNN模型,即基于变精度粗糙集的强耦合粗集神经网络(various precision rough set neural networks, VPRNN),预测企业的失败风险,该模型利用变精度粗糙集(various precision rough set, VPRS)降低噪声数据的干扰,在确保不影响关键信息的情况下,对原始数据进行有效的规则提取和精简,同时通过采用一种更为先进的离散化技术(即FUSINTER方法)对数据进行离散化处理,并在此基础上指导神经网络的设计和运行。VPRNN模型最大的优势是精简网络结构,提升网络训练速度,可以将小规模数据中获取的评价规则运用到更大规模的数据预测中,具有一定的抗干扰性和推广泛化能力。

3 VPRNN模型的构建

3.1 理论基础

VPRNN模型是在NN模型和VPRS模型的基础上建立的,因此先对这两类模型做一简要介绍。由于目前已有大量文献对NN模型进行深入的研究和探讨,故本研究将不再赘述,仅介绍VPRS模型。

VPRS模型与RST理论最大的区别在于模型中允许错误分类的存在,这使VPRS模型可以更准确、有效地处理数据信息^[15]。为衡量错误分类的大小,Ziarko在VPRS中引入概率值 β ,并将其值域定义为 $[0.000, 0.500)$ ^[15];An等则将 β 值定义为正确分类率,其值域也相应地改为 $(0.500, 1.000]$ ^[16]。在本研究中将采用An等的定义。Ziarko运用分类质量来衡量VPRS中出现错误分类的概率,即^[15]

$$\gamma^\beta(P, D) = \frac{\text{card}\left(\bigcup_{\Pr(Z|X_\lambda) \geq \beta} \{X_\lambda \in E(P)\}\right)}{\text{card}(U)} \quad (1)$$

其中, $\gamma^\beta(P, D)$ 为论域 U 能以概率 β 被归入一定决策类的元素比率, P 为条件属性集合 C 的子集, D 为决策属性集合, $\text{card}(\cdot)$ 为集合的基数, X_λ 为按条件属性划分的集合, λ 为划分后对应的集合数, $E(P)$ 为 P 中的任意集合, $\Pr(Z|X_\lambda)$ 为条件集与决策集交集的数目与条件集数目的比值, Z 为论域 U 中的子集。

此外,VPRS模型中对于任意一个给定的概率值 β ,将其正域定义为

$$\text{POS}_P^\beta(Z) = \bigcup_{\Pr(Z|X_\lambda) \geq \beta} \{X_\lambda \in E(P)\} \quad (2)$$

其中, POS 为正域; $Z \subseteq U$ 且 $P \subseteq C$ (以下相同)。

相应的, β 负域和 β 边界域分别定义为

$$\text{NEG}_P^\beta(Z) = \bigcup_{\Pr(Z|X_\lambda) \leq 1-\beta} \{X_\lambda \in E(P)\} \quad (3)$$

$$\text{BND}_P^\beta(Z) = \bigcup_{1-\beta < \Pr(Z|X_\lambda) < \beta} \{X_\lambda \in E(P)\} \quad (4)$$

其中, NEG 为负域, BND 为边界域。

特别地,当 $\beta = 1$ 时, $\text{POS}_P^\beta(Z)$ 和 $\text{NEG}_P^\beta(Z)$ 等同于RST中的下近似集和上近似集^[17]。

在 VPRS 中,对于任意一个 β 值进行近似约简 $RED^\beta(C,D)$,必须满足以下两个条件^[18],即

$$\gamma^\beta(C,D) = \gamma^\beta(RED^\beta(C,D),D) \quad (5)$$

且 $RED^\beta(C,D)$ 中不存在任意一个子集,使其在相同 β 值条件下能够产生相同的分类质量。上述条件中, $RED^\beta(C,D)$ 为条件属性集合 C 关于决策属性集合 D 的一个 β 近似约简。

3.2 模型构建

VPRNN 模型的结构如图 1 所示。与 NN 模型相似,VPRNN 模型的第一层和最后一层分别为输入层和输出层,但与之不同的是,该模型的隐含层被分为两层,其中一层为离散化层,另外一层为规则层。

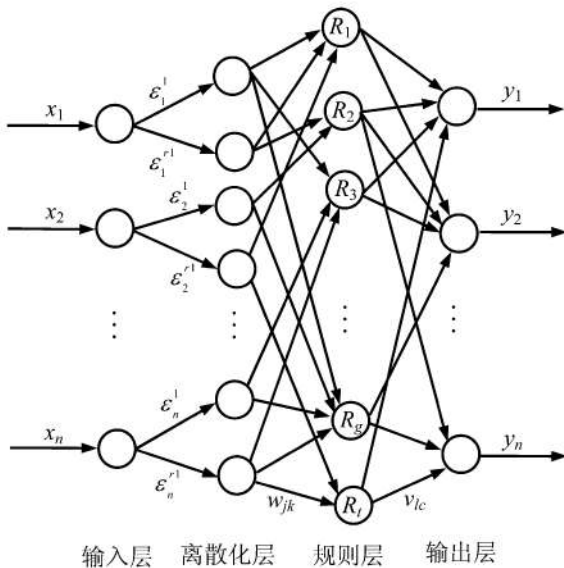


图1 VPRNN 模型
Figure 1 Model of VPRNN

在 VPRNN 模型中,第一层的输入向量为 $x, x = (x_1, x_2, \dots, x_n)^T, x_\eta (\eta = 1, 2, \dots, n)$ 为输入特征变量。

第二层离散化层中,采用 FUSINTER 算法对神经元数据进行离散化处理,分别将输入变量离散化为 r_i 个值,将高斯函数作为激励函数,即

$$\mu_i^j = \exp\left[-\frac{x_\eta - c_{ij}}{\delta_{ij}}\right] \quad (6)$$

其中, μ_i^j 为激励函数(本研究将其定义为高斯函数); c_{ij} 为 x_η 离散化到第 j 档的平均值, δ_{ij} 为 x_η 离散化到第 j 档的方差, j 为条件属性数, $j = 1, 2, \dots, r_i, r_i$ 为输入变量离散化后的不同值。

第三层规则层是离散化层和输出层的连接层,该层的主要作用是根据 VPRS 模型获取相应的评价规则。将第 k 个神经元的输出变量 R_k 定义为

$$R_k = f' \left(\sum_{j=1}^q w_{jk} \varepsilon_i^j + \beta_k \right) \quad (7)$$

$k = 1, 2, \dots, s$

其中, f' 为规则层的激励函数; w_{jk} 为输入层至规则层的连接权系数; ε_i^j 为离散化层变量; β_k 为规则层单元的阈值; s 为规则数; q 为基于规则 k 下的条件属性数。本研究将 f' 定义为正切 Sigmoid 函数,即 tansig 。由图 1 可知,输入层与规则层可能有多种映射关系,故连接权系数的个数与规则层神经元个数并不相等,因此 k 的取值并不等同于规则层神经元个数;此外,(7) 式中的 j 与(6) 式中的 j 含义相同,都指条件属性数,每个条件属性表示上一层中的一个神经元与下一层中的一个神经元相对应的关系,所不同的是,(6) 式中的 j 指第一层与第二层间的条件属性数,而(7) 式中的 j 则指第二层与第三层间的条件属性数,因此两者取值范围不同。

第四层输出层由所有输出神经元组成, y_ϕ 为输出特征变量, $\phi = 1, 2, \dots, n$, 每一个神经元都代表输出层中相应的决策结果 R_p , 其计算公式为

$$R_p = f'' \left(\sum_{l=1}^p v_{lc} R_l + \theta_c \right) \quad (8)$$

$c = 1, 2, \dots, t$

其中, f'' 为输出层的激励函数; v_{lc} 为输出层的连接权系数; θ_c 为输出层单元的阈值; l 为第三层所有神经元与第四层某一神经元相关的规则数; t 为决策数目; p 为基于规则 c 下的决策属性数。本研究将 f'' 定义为对数 Sigmoid 函数,即 logsig ^[14]。权值参数 w_{jk} 和 v_{lc} 以及神经阈值 β_k 和 θ_c 均可通过 BP 算法训练后求出。

从 VPRNN 模型的构造可以看出,相对于 NN 或 RNN 模型而言,VPRNN 所产生的规则将更简洁,也更准确,经过鲁棒性实验检验的结果也表明,VPRNN 可以在不同层级的噪音环境中取得良好的表现。可以认为,VPRNN 模型进一步完善了 RNN 中的近似概念,使其可以从不相关的预测变量值中挖掘出数据的相关性,因此运用该模型预测中国上市企业的失败风险具有一定的优势^[19]。

4 中国上市企业失败风险预测实证研究

4.1 评价指标的初选

选择评价指标是建立预测模型的关键和前提,但由于构建企业失败风险指标体系并不存在统一的标准,因此许多学者在进行相关研究时往往根据自身需要,通过不断论证和反复比较建立一套指标体系。如 Dimitras 等在查阅 1932 年 ~ 1994 年间的相关文献后发现,不同国家的学者对评价指标的使用偏好不尽相同,他们找出几种使用频率较高的指标,包括 $\frac{\text{营运资本}}{\text{总资产}}$ (16 次)、 $\frac{\text{流动资产}}{\text{总资产}}$ (在 15 个关于企业经营困境预测的模型中被使用到)、 $\frac{\text{流动资产}}{\text{流动负债}}$ (12 次)、 $\frac{\text{息税前利润}}{\text{总资产}}$ (12 次)、 $\frac{\text{净收入}}{\text{总资产}}$ (11 次)^[20]。Ravi Kumar 等在整理 1968 年 ~ 2006 年间与企业失败风险预测相关的文献后列出 126 类评价指标,这些指标几乎囊括了管理学的各个方面,如管理水平、研发水平、市场

发展动向、市场地位、国际市场份额以及套现方式等^[2]。

目前中国学者在进行企业风险预测时采用的指标种类也比较多,根据章之旺对相关文献的统计,可将其大致分为会计比率类、现金流量类、市场收益类和市场收益方差类^[21]。但遗憾的是,许多学者并未将选择评价指标的标准做出详细的说明,这使得构建指标体系缺乏一定的科学性和严谨性。

本研究构建指标体系的思路如下。由于VPRNN模型采用一种非线性的平行结构模式,在处理数据时对变量间的相关性敏感度不高^[22],因此在指标的初选过程中尽可能多地选择各类指标,使其涵盖的信息量更为全面,然后运用信息熵方法对所有指标进行初步筛选,并在此基础上构建预测指标体系。本研究选取的指标包括关键财务指标、调整后财务指标、非财务指标等三大类共计51项,具体指标名

称如表1所示。

4.2 评价指标的筛选

指标的初步筛选采用信息熵理论^[23],该方法能够运用确切的数值度量准确描述知识的相关特性,本研究正是利用信息熵的这种精确度量性衡量评价体系中各指标的重要程度,进而剔除重要度较轻的指标。

将信息熵 $H(A)$ 定义为

$$H(A) = - \sum_{h=1}^m A(X_h) \cdot \ln A(X_h) \quad (9)$$

其中, $A(X_h) = \frac{|X_h|}{|U|}$, $|X_h|$ 为集合 X_h 的基数, $|U|$ 为集合 U 的基数。

此外,设 $M = (U, A, V, f)$ 为一个决策信息系统, U 为对象的集合; A 为属性集合; V 为属性值的集合, $V = \cup_{a \in A} V_a$, V_a 为属性 $a \in A$ 的范围; f 为信息函数,它

表1 初选指标及其相应代码

Table 1 Primary Selected Indicators and Their Relevant Codes

指标代码	指标名称	指标代码	指标名称	指标代码	指标名称
C ₁	营业收入	C ₁₈	速动比率	C ₃₅	$\frac{\text{负债}}{\text{有形净资产}}$
C ₂	净利润	C ₁₉	债务资本比率	C ₃₆	股本收益率
C ₃	长期债务结构适合率	C ₂₀	有形净值负债率	C ₃₇	流动资产周转率
C ₄	$\frac{\text{营运资本}}{\text{总资产}}$	C ₂₁	成本费用利润率	C ₃₈	固定资产周转率
C ₅	债务权益比	C ₂₂	现金流量收入比	C ₃₉	资产周转率
C ₆	$\frac{\text{主营利润}}{\text{总资产}}$	C ₂₃	资本充足率	C ₄₀	主营业务增长率
C ₇	$\frac{\text{经营性现金流}}{\text{负债}}$	C ₂₄	无形资产增长率	C ₄₁	营运投资周转率
C ₈	每股收益	C ₂₅	现金负债比率	C ₄₂	应收账款周转率
C ₉	每股净资产	C ₂₆	利息保障倍数	C ₄₃	营业外收支净额
C ₁₀	$\frac{\text{非主营利润}}{\text{总资产}}$	C ₂₇	债务资产比率	C ₄₄	资产负债率
C ₁₁	净资产收益率	C ₂₈	存货依赖比率	C ₄₅	总资产增长率
C ₁₂	市盈率	C ₂₉	存货周转率	C ₄₆	现金流动负债比率
C ₁₃	总资产报酬率	C ₃₀	现金比率	C ₄₇	存货流动负债比率
C ₁₄	净利润增长率	C ₃₁	现金流量收入比	C ₄₈	营业利润增长率
C ₁₅	流动比率	C ₃₂	资本扩张率	C ₄₉	董事会规模
C ₁₆	主营业务鲜明率	C ₃₃	$\frac{\text{总债务}}{\text{EBITDA}}$	C ₅₀	国家股比率
C ₁₇	累计保留盈余率	C ₃₄	行业债务结构特征值	C ₅₁	高管持股比率

注:EBITDA为利息折旧及摊销前利润。主营业务鲜明率又称为主营业务利润比重或主营业务利润贡献率,其计算公式为 $\frac{\text{企业当年主营业务利润}}{\text{企业当年利润总额}}$,该指标反映企业主营业务的鲜明程度,该值越大说明公司主营业务越突出、越鲜明。

表2 运用信息熵方法筛选后的指标及其重要性

Table 2 Indicators Filtering by Information Entropy and Their Relevant Importance

指标代码	指标名称	指标重要性	指标代码	指标名称	指标重要性
C ₁	营业收入	0.094	C ₂₇	债务资产比率	0.058
C ₂	净利润	0.086	C ₂₉	存货周转率	0.067
C ₈	每股收益	0.070	C ₃₂	资本扩张率	0.014
C ₉	每股净资产	0.089	C ₃₆	股本收益率	0.026
C ₁₁	净资产收益率	0.076	C ₃₉	资产周转率	0.083
C ₁₂	市盈率	0.084	C ₄₀	主营业务增长率	0.060
C ₁₃	总资产报酬率	0.013	C ₄₁	营运投资周转率	0.031
C ₁₅	流动比率	0.064	C ₄₂	应收账款周转率	0.070
C ₁₈	速动比率	0.063	C ₄₃	营业外收支净额	0.055
C ₂₁	成本费用利润率	0.018	C ₄₅	总资产增长率	0.053
C ₂₂	现金流量收入比	0.022	C ₄₆	现金流动负债比率	0.026
C ₂₃	资本充足率	0.062	C ₄₇	存货流动负债比率	0.023
C ₂₆	利息保障倍数	0.043	C ₄₈	营业利润增长率	0.078

为 U 中的每个对象的每个属性指定一个信息值,即 $f: U \times A \rightarrow V$ 。属性在 $a \in A$ 中的重要性定义为

$$S_A(a) = |H(A) - H(A - \{a\})| \quad (10)$$

其中, $S_A(a)$ 为属性 a 的重要性, $H(A)$ 为属性 A 的信息熵值, $H(A - \{a\})$ 为属性 $(A - a)$ 的信息熵值。如 $S_A(a) = 0$, 则称 $a \in A$ 在 A 中是冗余的; 若 $S_A(a) > 0$, 则称 $a \in A$ 在 A 中是必须的。

利用信息熵方法筛选指标的具体步骤如下。①根据(9)式计算出属性集 A 的信息熵 $H(A)$ 和 $H(A - \{a\})$, $A = \{a_1, a_2, \dots, a_n\}$; ②用(10)式计算所有属性的重要性, 根据计算结果删去重要性为 0 对应的所有指标, 并将剩余指标组成新的集合 B , 记为 $B = \{b_1, b_2, \dots, b_s\} (s < n)$; ③按照已生成的信息表计算集合 B 中各指标的相关性, 如指标相关性大于某一设定阈值(本研究将其设置为 0.750), 则去除重要性较小的指标。

本研究在 CCER 数据库和巨潮资讯网站 (www.cninfo.com.cn) 2006 年 ~ 2007 年年报中选取 855 家企业的相关数据, 通过信息熵方法并结合 MATLAB 编程计算, 筛选后的结果如表 2 所示。

在对指标集进行第一次筛选时, 剔除了 3 个重要性为 0 的指标, 得到指标集 B , 通过进一步计算各指标间的相关性(限于篇幅, 相关性未一一列出), 找出阈值大于 0.750 的指标, 删去其中重要性较小的指标。例如, 经计算总资产报酬率与净资产收益率的相关系数为 0.806, 大于设定的阈值 0.750, 通过比较两项指标的重要性发现, 净资产收益率的重要性为 0.076, 而总资产报酬率的重要性为 0.013, 故剔除重要性较小的总资产报酬率指标。至此, 可以在基于信息熵的筛选方法上初步建立一套上市企业失败

风险预测指标体系, 表 3 列出该指标体系和相应的指标说明。

从表 3 可以发现, 虽然股本收益率和 *EBITDA* 等重要财务指标在国外相关研究中使用较多, 但经过信息熵方法的筛选, 这些指标均不适用于本研究的预测研究。由此可见, 企业失败风险预测指标体系的建立不能完全照搬国外已有的研究成果, 应当根据实际情况进行适当的调整。

4.3 样本选取和数据分类

国外相关研究在界定企业经营是否失败时通常将通过法律程序申请破产保护作为主要的判断依据^[24], 但这种判断标准对于中国上市企业并不适用。主要是因为中国的上市企业资格作为一种稀有的“壳”资源受到广泛青睐, 即使陷入破产危机, 上市企业也会借助自身或其他企业的力量通过各种大规模的资产重组摆脱财务困境; 另外, 各级地方政府十分重视上市企业的发展, 会运用各种手段保护上市企业, 使其免遭退市威胁。因此, 中国自 1986 年实施《企业破产法》以来, 目前还没有发现一家破产的上市企业。

针对这种现象, 中国学者在进行相关研究时通常将因财务状况异常而被特别处理(ST)的上市企业视为经营失败企业^[25,26]。陈晓等认为将 ST 作为界定企业经营失败标志的原因有以下两点, 从摆脱 ST 的公司来看, 大部分都是经过资产重组才摘掉 ST 的帽子, 这表明 ST 可以在一定程度上反映公司已陷入失败风险; ST 指标作为一个客观事实可以有效地度量^[25]。因此, 本研究也将采用这一标准。

在时间跨度的选择上, 本研究选用企业出现经营失败前 3 年的数据, 这是因为按照中国证券监督

表3 上市企业失败风险预测指标体系和相应指标说明
Table 3 Indicator System of Corporate Failure Risk and Definitions of Variables

一级指标	二级指标	指标说明
赢利状况	营业收入	公司每年总收入
	净利润	公司总收入扣除业务成本、折旧、利息、税款及其他开支后未做分配的利润
	每股收益	$\frac{\text{营业利润}}{\text{期末总股本}}$
	每股净资产	$\frac{\text{所有者权益合计}}{\text{期末总股本}}$
	净资产收益率(营业利润)	$\frac{\text{营业利润}}{\text{所有者权益合计}}$
	市盈率	$\frac{\text{价格}}{\text{每股收益}}$
	营业外收支净额	营业外收入 - 营业外支出
偿债能力	流动比率	$\frac{\text{流动资产}}{\text{流动负债}}$
	速动比率	$\frac{\text{流动资产} - \text{存货}}{\text{流动负债}}$
	资本充足率	$\frac{\text{所有者权益合计}}{\text{资产总计}}$
	债务资产比率	$\frac{\text{负债合计}}{\text{资产合计}}$
资产运营状况	存货周转率	$\frac{\text{营业成本}}{\frac{\text{当年存货净额} + \text{前一年存货净额}}{2}}$
	资产周转率	$\frac{\text{主营业务收入净额}}{\frac{\text{当年资产合计} + \text{前一年资产合计}}{2}}$
	应收账款周转率	$\frac{\text{主营业务收入净额}}{\frac{\text{当年应收账款} + \text{前一年应收账款}}{2}}$
发展能力	主营业务增长率	$\frac{\text{本期主营业务收入} - \text{上期主营业务收入}}{\text{上期主营业务收入}} \times 100\%$
	总资产增长率	$\frac{\text{本期总资产} - \text{前一期总资产}}{\text{前一期总资产}} - 1$
	营业利润增长率	$\frac{\text{本期营业利润} - \text{上期营业利润}}{\text{上期营业利润}} - 1$

管理委员会颁布的相关管理条例,如果上市企业在($t-1$)年和($t-2$)年连续两年出现财务困境现象,该企业肯定将被视为ST公司,在这种状况下对企业进行预测没有任何实际意义;此外,上市企业年报披露的最后期限是每年的4月30日,如果仅根据前1年的数据进行分析会因为时间间隔较近而夸大预测效果。

本研究从CCER数据库和巨潮资讯网2003年~2007年的数据里找出245家ST公司,其中7家企业因为数据缺失而被剔除,同时选取238家非ST企业作为配对样本,这些配对企业无论从资产规模还是

企业类型或选择的时间都与配对的ST企业接近或类似(各指标间的差距不超过10%)。此外,本研究将ST企业的决策属性定义为1,将非ST企业的决策属性定义为0。所有研究样本情况如表4所示。

为验证ST企业与非ST企业各财务指标间是否存在差异,通常会对所有变量采用配对样本T检验,但该检验的前提条件是各变量应服从正态分布,为此先对所有配对变量进行K-S正态分布检验,如果变量不服从正态分布,则运用Wilcoxon配对符秩检验检查各变量间是否存在差异。表5列出所有变量的检验结果。

表4 研究样本描述
Table 4 Description of Test Samples

年份	ST 公司数	数据缺失企业数	有效样本数	配对样本数	所有样本合计
2003 年	38	2	36	36	72
2004 年	45	0	45	45	90
2005 年	36	5	31	31	62
2006 年	61	0	61	61	122
2007 年	65	0	65	65	130
总计	245	7	238	238	476

表5 所有变量的正态性检验和配对检验结果
Table 5 Paired-Samples T Test and One-Sample Kolmogorov-Smirnov Test of All Variables

变量	样本名称	数量	均值	标准差	K-S 正态分布检验	Wilcoxon 配对符秩检验	配对样本 T 检验
营业收入	ST 公司	238	848 272 822.800	3 286 514 400.000	0.000***	0.000***	/
	非 ST 公司	238	628 306 087.100	6 948 766 367.000	0.000***		
净利润	ST 公司	238	43 258 808.200	211 895 207.700	0.000***	0.000***	/
	非 ST 公司	238	276 010 068.200	517 591 345.500	0.000***		
每股收益	ST 公司	238	-0.116	0.627	0.000***	0.000***	/
	非 ST 公司	238	0.492	0.743	0.001***		
每股净资产	ST 公司	238	-0.271	1.835	0.000***	0.000***	/
	非 ST 公司	238	3.439	2.373	0.003***		
总资产增长率	ST 公司	238	0.087	0.928	0.000***	0.000***	/
	非 ST 公司	238	0.269	0.524	0.000***		
流动比率	ST 公司	238	0.825	0.866	0.000***	0.000***	/
	非 ST 公司	238	1.513	1.402	0.000***		
速动比率	ST 公司	238	0.597	0.766	0.000***	0.000***	/
	非 ST 公司	238	1.036	1.280	0.000***		
资本充足率	ST 公司	238	-1.465	10.933	0.000***	0.000***	/
	非 ST 公司	238	0.501	0.201	0.645		
存货周转率	ST 公司	238	-50.727	781.293	0.000***	0.047	/
	非 ST 公司	238	91.005	610.971	0.000***		
债务资产比率	ST 公司	238	2.441	10.932	0.000***	0.000***	/
	非 ST 公司	238	0.500	0.188	0.507		
资产周转率	ST 公司	238	0.433	0.262	0.111	/	0.000***
	非 ST 公司	238	0.569	0.281	0.361		
营业外收支净额	ST 公司	238	90 550 749.000	229 098 106.300	0.000***	0.251	/
	非 ST 公司	238	15 806 863.200	79 152 424.500	0.000***		
净资产收益率 (净利润)	ST 公司	238	0.894	13.701	0.000***	0.680	/
	非 ST 公司	238	0.082	0.216	0.000***		
主营业务增长率	ST 公司	238	-0.137	0.586	0.248	0.000***	/
	非 ST 公司	238	0.226	0.602	0.000***		
市盈率	ST 公司	238	132.857	306.201	0.000***	0.797	/
	非 ST 公司	238	136.543	333.308	0.000***		
营业利润增长率	ST 公司	238	0.819	11.965	0.000***	0.000***	/
	非 ST 公司	238	-0.357	26.494	0.000***		
应收账款周转率	ST 公司	238	27.736	55.927	0.000***	0.526	/
	非 ST 公司	238	24.690	63.971	0.000***		

注:表中3种检验方式均采用双尾检验;***为1%显著性水平,**为5%显著性水平,*为10%显著性水平。

表6 经 FUSINTER 离散化后条件属性分类区间
Table 6 Condition Attributes Intervals for FUSINTER Discretisation

属性	区间0	区间1	区间2	区间3
营业收入	[94 493.000, 63 702 829.000)	[63 702 829.000, 248 333 101.110)	[248 333 101.110, 39 067 561 000.390)	
净利润	[-1 051 841 398.000, 3 113 274.230)	[3 113 274.230, 20 212 454.081)	[20 212 454.081, 1 068 707 005.712)	
每股收益	[-4.461, -0.176)	[-0.176, 0.038)	[0.038, 1.335)	[1.335, 3.814)
每股净资产	[-8.139, -0.275)	[-0.275, 0.028)	[0.028, 1.168)	[1.168, 4.623)
流动比率	[0.006, 0.373)	[0.373, 0.847)	[0.847, 7.102)	
速动比率	[0.004, 0.267)	[0.267, 0.535)	[0.535, 7.125)	
资本充足率	[-123.022, -0.023)	[-0.023, 0.048)	[0.048, 0.371)	[0.371, 1.000)
债务资产比率	[0.047, 0.433)	[0.433, 0.562)	[0.562, 0.888)	[0.888, 124.022)
存货周转率	[0.004, 4.391)	[4.391, 11.557)	[11.557, 178.939)	
资产周转率	[0.014, 0.222)	[0.222, 0.466)	[0.466, 1.336)	
主营业务增长率	[-1.046, -0.026)	[-0.026, -0.161)	[-0.161, 136.652)	
总资产增长率	[-0.993, -0.275)	[-0.275, -0.004)	[-0.004, 0.135)	(0.135, 9.089)
营业利润增长率	[-0.193, 2.446)	[2.446, 5.431)	[5.431, 10.504)	(10.504, 276.552)

通过以上检验,除存货周转率因配对符秩检验的值为0.047(小于0.050)而被保留外,剔除显著性大于10%的4个指标,即营业外收支净额、净资产收益率(净利润)、市盈率和应收账款周转率,形成最终的评价指标体系。

4.4 数据离散化处理

由于VPRNN模型只能处理离散型数据,因此对本研究中出现的连续型数据必须先进行离散化处理。传统的数据离散化方法包括最小熵法、Chi-Merge和Chi2等方法^[27],这些方法大多采用卡方统计来衡量变量的拟合优度,一次仅能离散化决策表子区域中的一组条件属性,因此对数据的处理过程较为繁琐。目前使用较多的是一种连续变量离散化方法(continuous variable discretisation, CVD),与传统方法相比,CVD的最大优势是可以一次性离散化决策表中的所有条件属性,因而可以进行大规模的运算。

本研究采用一种更为先进的CVD技术,即FUSINTER方法。该方法是一种自下而上的算法,其原理是根据某些特殊的优选评价方法(如二次方程熵法)将所有的条件属性划分到普通组。FUSINTER技术最大的优点是不需要专家参与就可以运用客观的运算将原始数据划分到某些预定的区间中^[28],避免因专家评判所带来的主观影响。本研究中的数据处理结合二次方程熵法,同时将缺省值定义为 $\alpha=0.975$ 且 $\lambda=1$, α 为离散化处理结果的准确性, λ 为离散化

样本的覆盖率。离散化后的分类区间如表6所示。

4.5 结果分析和讨论

本研究采用ROSETTA软件对决策表进行约简并产生对应的规则,RST模型和VPRS模型产生的规则数分别为256条和22条,运用VPRS筛选规则时选取精度大于0.750、覆盖度大于0.050的规则集,并令 β 值域在(0.750,1]的范围内。在生成规则的基础上采用RNN模型和VPRNN模型计算训练样本的误判率、训练时间、规则数等重要指标,表7给出了应用RNN模型和VPRNN模型在计算过程中产生的各类指标的具体数值。

表7 RNN模型和VPRNN模型分析比较
Table 7 Comparison and Analysis between RNN Model and VPRNN Model

	规则数	需调节权值数	训练时间(秒)	平均均方差	
				训练集样本	校验集样本
RNN	256	856.000	15.032	0.003	0.005
VPRNN	22	328.000	6.221	0.004	0.006

表7数据显示,RNN模型训练时间为15.032秒,VPRNN模型的训练时间仅为6.221秒,比RNN模型大大缩短;VPRNN模型所需调节权值数更少,仅为

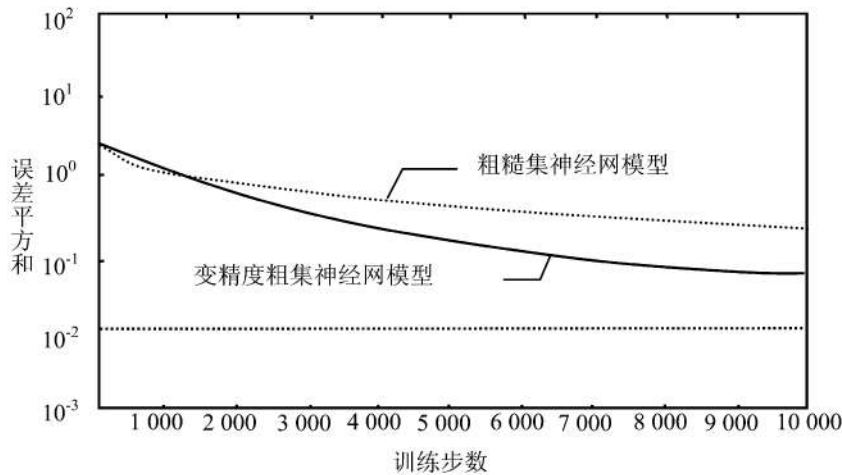


图2 不同模型训练过程比较

Figure 2 Comparison of Training Course of Different Models

328.000,但生成的规则数却更为精简。此外,RNN的校验误差要明显高于VPRNN,这意味着VPRNN模型的分類能力(特别是针对未知样本的分類能力)要优于RNN模型。因此,就整体预测效果而言,VPRNN模型比RNN模型更好。图2显示了两类模型训练过程中的差异化比较,下降的实线为变精度粗集神经网络模型的训练过程;下降的虚线为粗糙集神经网络模型的训练过程,数值 10^{-2} 代表的是事先设定的目标值或期望值,与该值越接近表示训练效果越好。由图2可知,变精度粗集神经网络模型下降速度更快,表示其学习的效率较高;其训练结果与期望值更接近,意味着其学习准确性更好,因而该模型要优于粗糙集神经网络模型。

表8是运用VPRNN模型对训练样本中的ST企业和非ST企业进行正确分类的样本数统计。测试样本为160家ST企业和160家非ST企业(本研究将三分之二的数作为测试样本,三分之一的数作为检验样本),检验的标准是将预测结果和实际分类情况对比,0为非ST企业,1为ST企业。经计算,160家非ST企业中有4家被误判为ST企业,即出现第I类错误的比例为2.500%;160家ST企业中有两家被误判为非ST企业,即出现第II类错误的比例为1.250%,总的判断准确率为98.125%。

表8 t-3年预测结果(测试样本)
Table 8 Prediction Results of Year t-3
(Test Sample)

		预测结果		预测准确率 (%)
		0	1	
实际类别	0	156	4	97.500
	1	2	158	98.750
实际分类比率 (%)		98.734	97.531	98.125

在检验样本中,78家非ST企业中有8家被误判为ST企业,78家ST企业中有5家被误判为非ST企业,出现第I类错误、第II类错误的比例分别为10.256%和6.410%,总的预测准确性为91.667%,如表9所示。由此可见,VPRNN模型具有较高的预测准确度。

表9 t-3年预测结果(检验样本)
Table 9 Prediction Results of Year t-3
(Examine Sample)

		预测结果		预测准确率 (%)
		0	1	
实际类别	0	70	8	89.744
	1	5	73	93.590
实际分类比率 (%)		93.333	90.123	91.667

4.6 与其他模型预测准确度的对比

为了对VPRNN模型的预测准确度进行更客观的评价,将其与其他模型的预测结果做横向比较。本研究采用RST、VPRS和RNN作为对比模型,其数据均来源于相同的测试样本。表10列出所有模型的预测结果。

由于VPRS具有一定容错性,相对RST而言在进行正确分类时会有一部分原本符合规则的评估对象纳入分类样本中,因而其正确分类率要高于RST。RNN结合RST和NN的优势,使其预测正确率略高于单独使用这两类模型。而VPRNN模型中的VPRS允许错误分类的出现,确保VPRNN模型能够对统计意义上的数据模式进行有效描述,同时将运算过程

表10 相同测试样本下各模型预测精度
Table 10 Test Sample Prediction Accuracy of Various Models

	粗糙集规则分类				粗糙集神经网络分类	
	RST 模型		VPRS 模型		RNN 模型	VPRNN 模型
	正确分类样本	不能决策样本	正确分类样本	不能决策样本		
ST 公司	141	0	148	0	150	156
非 ST 公司	135	0	143	0	152	158
整体预测	276	0	291	0	302	314
预测精度(%)	86.250	/	90.938	/	94.375	98.125

表11 相同检验样本下各模型预测精度
Table 11 Examine Sample Prediction Accuracy of Various Models

	粗糙集规则分类				粗糙集神经网络分类	
	RST 模型		VPRS 模型		RNN 模型	VPRNN 模型
	正确分类样本	不能决策样本	正确分类样本	不能决策样本		
ST 公司	58	2	68	1	69	70
非 ST 公司	67	1	66	0	67	73
整体预测	125	3	134	1	136	143
预测精度(%)	80.128	/	85.897	/	87.179	91.667

中某些不确定问题合理分类,使一些原本在 RNN 模型中难以正确分类的评价对象可以通过 VPRNN 的近似决策功能有效地处理,故其预测准确性也相应得到提高。此外,在粗糙集规则分类中,检验规则均从测试样本中提取,因此不存在“不能决策”样本。

表11 列出对检验样本中78家 ST 企业和78家非 ST 企业进行预测的结果。由于检验样本中的评价对象较少,且相应规则由测试样本生成,因此当这些规则用于评价全新的对象时,各模型的预测准确度出现不同程度的下降,同时会产生一部分不能决策的样本数。通过以上横向比较可以发现,VPRNN 模型的整体预测效果最好。

5 结论

本研究运用一种基于 VPRS 的强耦合粗糙神经网络(VPRNN)模型作为预测中国上市企业失败风险的方法,将被 ST 的上市企业作为经营失败的标志,通过信息熵等方法筛选出13个财务指标作为最终的评价指标体系,运用 FUSINTER 技术对数据进行离散化处理,选取476家上市企业作为研究样本,通过实证研究得到以下结论。①在企业出现失败风险前两年,ST 企业与对应的非 ST 企业在营业收入、净利润、每股收益等13个指标上存在显著性差异,因而在其他模型中也可运用这些指标进行预测研究;② FUSINTER 算法在离散化数据时能够一次性处理决策表中的所有条件属性,且不需要专家参与,与传统

离散化方法相比更为客观、有效;③ VPRNN 模型类似于在局部逼近神经网络,这使其具有较快的收敛速度,且 VPRNN 模型的结构较之 RNN 模型更为精简,所需调整的权值数更少,因此训练时间大幅缩减;④由于 VPRNN 模型是一种具有近似决策能力的神经网络,因而有较强的抗噪声数据干扰能力和处理小样本数据时的推广泛化能力;⑤通过与 RST、VPRS 和 RNN 模型的横向比较发现,VPRNN 模型的预测准确度最高。

VPRNN 模型虽然具有较高的预测精度,但操作过程相对复杂,后续研究可开发出一种方便实用的软件以简化实际运算中的繁琐过程;在进行离散化过程中,应用不同方法会对预测结果产生一定影响,本研究并未将不同的离散化结果进行对比、分析,这也是未来研究的工作之一。

参考文献:

- [1] 郑育书,刘沂佩. 企业失败预测模型研究——来自台湾的证据[J]. 管理科学,2008,21(4):82-95.
Cheng Y S, Liu Y P. Study on the Prediction of Business Failure: Evidence from Taiwan[J]. Journal of Management Science, 2008, 21(4): 82-95. (in Chinese)
- [2] Ravi Kumar P, Ravi V. Bankruptcy Prediction in Banks and Firms via Statistical and Intelligent Tech-

- niques——A Review [J]. *European Journal of Operational Research*, 2007, 180(1): 1-28.
- [3] Swicegood P, Clark J A. Off-site Monitoring Systems for Predicting Bank under Performance: A Comparison of Neural Networks, Discriminant Analysis and Professional Human Judgment [J]. *Intelligent Systems in Accounting, Finance & Management*, 2001, 10(3): 169-186.
- [4] Ryu Y U, Yue W T. Firm Bankruptcy Prediction: Experimental Comparison of Isotonic Separation and Other Classification Approaches [J]. *IEEE Transactions on Systems, Management and Cybernetics-Part A: Systems and Humans*, 2005, 35(5): 727-737.
- [5] Boyacioglu M A, Kara Y, Baykan Ö K. Predicting Bank Financial Failures Using Neural Networks, Support Vector Machines and Multivariate Statistical Methods: A Comparative Analysis in the Sample of Savings Deposit Insurance Fund (SDIF) Transferred Banks in Turkey [J]. *Expert Systems with Applications*, 2009, 36(2): 3355-3366.
- [6] 张华伦, 孙毅. 企业财务危机预警 Rough-Fuzzy-ANN 模型的建立及应用 [J]. *运筹与管理*, 2006, 15(2): 103-107.
Zhang H L, Sun Y. Establishment of Rough-Fuzzy-ANN Model for Forecasting Enterprise Financial Distress and Its Application [J]. *Operations Research and Management Science*, 2006, 15(2): 103-107. (in Chinese)
- [7] 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.
Liu Y W, Dai H J. Empirical Study of the Financial Risk Prediction Based on the Rough Set and ANN Mode [J]. *Science Research Management*, 2007, 28(6): 138-142. (in Chinese)
- [8] Pawlak Z. Rough Sets [J]. *International Journal of Information and Computer Sciences*, 1982, 11(5): 341-356.
- [9] Fan T F, Liu D R, Tzeng G H. Rough Set-based Logics for Multicriteria Decision Analysis [J]. *European Journal of Operational Research*, 2007, 182(1): 340-355.
- [10] Jelonek J, Krawiec K, Slowinski R. Rough Set Reduction of Attributes and Their Domains for Neural Networks [J]. *Computational Intelligence*, 1995, 11(2): 339-347.
- [11] Ahn B S, Cho S S, Kim C Y. The Integrated Methodology of Rough Set Theory and Artificial Neural Network for Business Failure Prediction [J]. *Expert Systems with Applications*, 2000, 18(2): 65-74.
- [12] Hashemi R R, Le Blanc L A, Rucks C T, Rajaratnam A. A Hybrid Intelligent System for Predicting Bank Holding Structures [J]. *European Journal of Operational Research*, 1998, 109(2): 390-402.
- [13] 马若微. 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.
Ma R W. The Prediction Model of Financial Distress of Listed Manufacturing Based on RS and ANN [J]. *Nankai Business Review*, 2006, 9(3): 85-91. (in Chinese)
- [14] 张东波, 王耀南, 黄辉先. 基于变精度粗糙集的粗糙神经网络 [J]. *电子与信息学报*, 2008, 30(8): 1913-1917.
Zhang D B, Wang Y N, Huang H X. Rough Neural Network Based on Variable Precision Rough Set [J]. *Journal of Electronics & Information Technology*, 2008, 30(8): 1913-1917. (in Chinese)
- [15] Ziarko W. Variable Precision Rough Set Model [J]. *Journal of Computer and System Sciences*, 1993, 46(1): 39-59.
- [16] An A, Shan N, Chan C, Cercone N, Ziarko W. Discovering Rules for Water Demand Prediction: An Enhanced Rough-set Approach [J]. *Engineering Applications of Artificial Intelligence*, 1996, 9(6): 645-653.
- [17] Pawlak Z, Skowron A. Rough Sets: Some Extensions [J]. *Information Sciences*, 2007, 177(1): 28-40.
- [18] 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.
- [19] Xue F, Ke K L. Five-Category Evaluation of Commercial Bank's Loan by the Integration of Rough Sets and Neural Network [J]. *System Engineering-Theory & Practice*, 2008, 28(1): 40-45.
- [20] Dimitras A I, Zanakis S H, Zopounidis C. A Survey of Business Failures with an Emphasis on Prediction Methods and Industrial Applications [J]. *European Journal of Operational Research*, 1996, 90(3): 487-513.
- [21] 章之旺. 现金流量在财务困境预测中的信息含量实证研究——来自 2003-2004 年度 ST 公司的新证据 [J]. *中国管理科学*, 2004, 12(6): 23-28.
Zhang Z W. An Empirical Study on Cash Flows' Information Content in Financial Distress Prediction——Evidence from ST Firms during the Period of 2003-2004 [J]. *Chinese Journal of Management Science*, 2004, 12(6): 23-28. (in Chinese)
- [22] Su C T, Hsu J H. Precision Parameter in the Variable Precision Rough Sets Model: An Application [J]. *Omega*, 2006, 34(2): 149-157.
- [23] Lai S C, Mayer-Kress G, Sosnoff J J, Newell K M. Information Entropy Analysis of Discrete Aiming

- Movements [J]. *Acta Psychologica*, 2005, 119 (3) : 283-304.
- [24] Gillan S, Martin J D. Corporate Governance Post-Enron: Effective Reforms, or Closing the Stable Door? [J]. *Journal of Corporate Finance*, 2007, 13 (5) : 929-958.
- [25] 陈晓, 陈治鸿. 中国上市公司的财务困境预测 [J]. *中国会计与财务研究*, 2000, 3 (2) : 55-72.
Chen X, Chen Z H. Prediction of Chinese Listed Companies' Financial Distress [J]. *China Accounting and Finance Review*, 2000, 3 (2) : 55-72. (in Chinese)
- [26] 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.
Wu S N, Lu X Y. A Study of Models for Predicting Financial Distress in China's Listed Companies [J]. *Economic Research Journal*, 2001 (6) : 46-55. (in Chinese)
- [27] Wang Z J, Li H X. Financial Distress Prediction of Chinese Listed Companies: A Rough Set Methodology [J]. *Chinese Management Studies*, 2007, 1 (2) : 93-110.
- [28] Griffiths B, Beynon M J. Expositing Stages of VPRS Analysis in an Expert System: Application with Bank Credit Ratings [J]. *Expert Systems with Applications*, 2005, 29 (4) : 879-888.

Prediction of Companies' Failure Risk Based on Various Precision Rough Neural Networks

YIN Peng¹, WANG Zong-jun¹, XIAO De-yun²

1 School of Management, Huazhong University of Science & Technology, Wuhan 430074, China

2 School of Economics, Wuhan University of Science & Technology, Wuhan 430070, China

Abstract: Described the relevant notation and mechanism of the Rough Sets Theory (RST), Various Precision Rough Sets (VPRS) and Neural Networks (NN), introduced an advanced model-Variety Precision Rough Neural Networks (VPRNN) based on VPRS and NN model, and utilized it as a research method for predicting the Chinese listed corporates' failure risk. Applying the information entropy and T-test to filter various financial and non-financial indicators, consequently, we constructed the enterprise failure risk prediction indicator system. In order to avoid exaggerating prediction result, we chose the data of year ($t-3$) as the starting point. Meanwhile, considering the Special Treatment (ST) company as the failure enterprise, we selected 160 ST listed companies and 160 non-ST matched companies as test samples and discretized all the data by FUSINTER. Applying the RST, VPRS, RNN and VPRNN models to predict 78 examine samples, we can conclude that the VPRNN model's accuracy is higher than other models, moreover, the training time needed as well as the rules generated is shorter and less than other models. Therefore, it can be inferred that the VPRNN model could be more useful and suitable for the prediction of the failure risk for Chinese listed companies.

Keywords: Chinese listed companies; risk prediction; various precision rough neural networks model; FUSINTER data discretisation

Received Date: March 18th, 2010 **Accepted Date:** May 25th, 2010

Funded Project: Supported by the National Natural Science Foundation of China (70872033)

Biography: YIN Peng, a Jiangxi Yongxin native (1980 -), is a Ph. D. candidate in the School of Management at Huazhong University of Science and Technology. His research interests include evaluation of enterprise financial distress, financial risk management, etc.

E-mail: davidyin07@sina.com

□