

**The Study of Financial Distress Prediction
of Listed Companies Based on the Combination
Forecasting Techniques**

A Dissertation Submitted for the Degree of MaSTer

Candidate: Zhongwu Li

Supervisor: Prof. Guanhua Zhao

School of Accountancy

Shandong University of Finance and Economics

中图分类号:

密级: 公开

学科分类号:

论文编号: KJ2010120201026

硕 士 学 位 论 文

基于组合预测技术的 上市公司财务困境预测问题研究

作者姓名: 李忠武

申请学位级别: 管理学硕士

指导教师姓名: 赵冠华

职 称: 教 授

学 科 专 业: 会计学

研 究 方 向: 财务决策支持系统

学 习 时 间: 自 2010 年 9 月 1 日 起 至 2013 年 6 月 30 日 止

学位授予单位: 山东财经大学

学位授予日期: 2013 年 6 月

摘 要

我国的资本市场已走过了二十多年的历程，企业所面临的竞争日趋激烈，企业经营稍有不慎就可能陷入财务困境，上市公司因财务状况异常而陷入困境的情况却屡见不鲜。因此，利用上市公司的财务数据，建立上市公司财务困境预测模型以揭示风险，已成为上市公司管理者、投资者和债权人等相关利益方共同关注的问题。

本文共选取来自于多个行业的 146 家上市公司的财务数据，包括 73 家被特别处理（用 ST 表示）的上市公司的财务数据和与之配对的 73 家运行良好的公司的财务数据，把它们分为 49 对训练样本和 24 对测试样本。根据国内外学者对企业财务困境预测的研究成果，选择了 20 个财务指标构建了财务预警指标体系。利用 T 检验和非参数检验对原始预测变量进行显著性检验，使得预测的指标体系更具有合理性，并运用因子分析法消除指标间的共线性，得到了 7 个因子。分别采用多元判别分析、Logit 回归分析和最小二乘支持向量机对训练样本数据进行了分类判断，构建了三个单项财务困境预测模型；并在此基础上，构建了一个线性组合预测模型和一个基于神经网络的组合预测模型，最后利用测试样本对五个模型进行了有效的验证，并对五个检验结果进行了比较。研究结果表明：组合预测模型对上市公司财务困境的预测效果，跟单一方法比较，预测精度没有显著提高，但组合预测模型更具有稳定性；基于神经网络方法的组合效果好于线性组合模型。

关键词：财务预测 逻辑回归 支持向量机 组合预测

Abstract

As of this year, China's capital market has gone through the course of the last twenty-three years, enterprises are facing increasingly fierce competition, the slightest mistake business may fall into a financial crisis, the crisis of listed companies because of abnormal financial position is not uncommon. Therefore, the financial data of listed companies, listed companies' financial crisis early warning model to reveal the risks, has become the common concerns of managers of listed companies, investors and creditors and other stakeholders.

This paper selected financial data of 146 listed companies from multiple industries, including 73 well-listed company's financial data, and paired 73 special treatment (ST said) the company's financial data, divided into 49 pairs of training samples and 24 of the test sample. Research scholars warning of enterprise financial crisis, asked the 20 financial indicators to build a financial early warning indicator system. T-test and non-parametric test will be more original predictor filter indicator system, making predictions is more reasonable, and the use of factor analysis to eliminate collinearity between indicators, the seven factors. Were used multiple discriminant analysis, logistic regression analysis and support vector machine classification to determine the training sample data to construct three single financial crisis early warning model; constructed on this basis, a linear combination of the prediction model and a neural network combined forecasting model, and finally the use of the test sample of five model effective verification on five test results were compared. The research results show that the combination forecasting model based on neural network methods predict the effect of corporate financial distress, compared with a single method, the prediction accuracy is not significantly increased, but the model is more stability.

Keywords: Financial Distress Prediction Logit Regression
Support Vector Machine Combining Forecast

目录

第 1 章 绪论	1
1.1 研究背景及意义	1
1.1.1 研究背景	1
1.1.2 研究意义	1
1.2 文献综述	3
1.2.1 财务困境的界定	3
1.2.2 国外研究综述	4
1.2.3 国内研究综述	8
1.3 研究内容、方法、创新点及框架	10
1.3.1 研究内容	10
1.3.2 研究方法	10
1.3.3 创新点	11
1.3.4 财务困境预测框架	11
第 2 章 财务困境预测模型方法介绍	12
2.1 基于单模型的财务困境预测方法	12
2.1.1 多元判别分析 (MDA)	12
2.1.2 逻辑回归模型 (Logit)	13
2.1.3 最小二乘支持向量机 (LS-SVM)	15
2.2 基于组合预测技术的财务困境预测方法	24
2.2.1 线性组合预测模型	24
2.2.2 神经网络组合预测模型	25
第 3 章 样本的选取和预测指标的选取及预处理	29
3.1 样本选取	29
3.2 预测指标体系	29
3.2.1 预测指标选取原则	29
3.2.2 预测指标的选取	30
3.3 预测指标的预处理	32
3.3.1 预警指标的显著性检验	32
3.3.2 预警指标主成分的提取	36
第 4 章 上市公司财务困境预测模型的构建和检验	41
4.1 单模型预测结果分析	41
4.1.1 判别分析模型	41

4.1.2 逻辑回归模型	43
4.1.3 支持向量机模型.....	45
4.2 组合模型预测结果分析	46
4.2.1 线性组合模型	46
4.2.2 基于基于 ANN 的组合预测模型	47
4.3 财务预警模型判别结果比较分析.....	48
第五章 研究结论及不足	50
5.1 研究结论	50
5.2 研究不足及未来研究方向	50
参考文献	52
附 录	55
攻读硕士学位期间发表的学术成果	57
致 谢	58

第 1 章 绪论

1.1 研究背景及意义

1.1.1 研究背景

1929 年的世界经济困境成为国际上企业财务困境预测研究的触发根源。随着企业经济环境、竞争环境以及经营情况的多样化和复杂化，国际上的著名专家学者都进行了企业困境、企业风险管理等方面的研究，目前企业财务困境预测已成为研究的一个热点问题。

中国证券市场历经二十多年的发展，资本市场不断的成熟壮大，截至 2012 年底，我国共有上市公司 2342 家，8000 多万股民，总市值高达 21.48 万亿元，名列全球第 3 位。但是我国证券市场存在的时间短发展尚不成熟，上市公司退市机制不完善，使得公司管理者与投资者不关心公司实际的经营业绩与财务状况，只注重于各自短期利益，这样导致企业出现破产及陷入其他财务困境现象的现象也越来越严重，严重地破坏了证券市场的秩序，更损害了广大投资者的利益。

为了保护投资者和债权人的利益，中国证监会和沪、深证交所制定了一系列制度。其中在 1998 年，中国证券监督管理委员会颁布《关于上市公司状况异常期间的股票特别处理方式的通知》的文件，“ST”制度在我国正式启动。对于监管者和广大的投资人来说，通过构造合理的上市公司财务困境预测体系，正确地预测企业财务困境，促进证券市场的健康有序发展，有着极为重要的现实意义。

目前我国对上市公司的财务信息披露有较严格的规定，尽管如此还是不能解决证券市场存在着很多的问题，比如上市公司提供的信息若有所失真，那么这会严重的误导投资者的决策，造成投资损失。对于上市公司本身，我国上市公司的总体经营业绩仍然不尽人意，部分上市公司出现了严重的财务困境，如*ST 炎黄和*ST 创智已被宣告将退市，给投资者造成巨大损失，所以建立一套财务预警体系显得尤为重要。再加上证券市场制度难以有效的保障广大投资者的利益，因而对上市公司破产预测或财务困境预测的需求，越来越多的专家学者开始探索使用统计方法建立上市公司财务困境的预测模型。

1.1.2 研究意义

财务困境已经被认为是许多企业所面临的最重要的威胁之一。大多数企业的财务

困境都是由财务状况正常到逐步恶化，是一个渐进的过程。因而，企业财务困境的发生是有先兆的，是可以进行预测的。当企业陷入财务困境后，将投资者、贷款人、管理层和员工将受到严重影响，因此，利用上市公司的财务数据，通过构造模型来对财务困境给予及时和有效的预测，已经成为学术界与实务界研究的一个热门课题。实践中正确地预测企业财务困境，对于经营者防范财务困境，对于保护投资者和债权人的利益，对于政府管理部门监控上市公司质量和证券市场风险，企业管理层及早拟定企业再生策略，最终使企业走出困境都具有重要的现实意义^[1]。

(1)有利于投资者的投资决策。投资者进行投资的目的是为了获取高额投资收益，只有把资金投向发展潜力好的企业，这一目的才能实现。但是我国证券市场是弱式有效市场，投资者不能利用公开的信息获取超额收益。因而，对于投资者来说，通过建立较准确的财务困境预测模型，检查被投资单位是否有陷入财务困境的危险，及时调整自己的投资方向，降低投资风险，防止潜在损失，提高投资收益有助于投资者做出正确的投资决策，防止被投资单位资不抵债时，投资者的投资血本无归。

(2)有利于证券监管机构加强对上市公司的监管。证券监管部门对上市公司的管理，如ST制度，主要以事后监管为主对上市公司进行约束，因而具有一定的滞后性。通过上市公司财务困境预测研究，利用上市企业前几年披露的财务数据提前进行监测，加强事前监管；证券监管部门可以及早的发现上市公司财务状况的异常，判断上市公司未来陷入财务困境的概率，加强对上市公司的事前监管，及时采取措施，更好地维持证券市场的秩序，保证证券市场健康有序的发展。此外，财务困境预测也有利于政府进行宏观上的管理。国家经济管理部门需要掌握企业经济效益等各个方面的情况才能制定宏观经济政策，调控经济运行。财务困境预测有助于政府部门及时了解企业陷入财务困境的信息，减少破产成本和因员工失业而造成的社会动荡，改善资源的宏观配置计划，实现资源的优化配置，提高决策的科学化。

(3)有利于上市公司的健康发展和公司管理水平的提高。对于管理者来说，财务困境预测模型可以起到警示的作用。通过财务困境预测模型的预测结果，管理者可以及早发现企业财务状况的变化，找出影响企业财务状况的核心指标，有针对性的调整企业的经营策略，并采取有效的措施解决问题，制定合理的财务计划，防止财务状况的进一步恶化，避免公司陷入财务困境，使上市公司能够长期稳定的健康发展，减小陷入财务困境的概率。通过财务困境的预测研究，可以发现公司是否有破产的威胁，采取相应的措施改善经营管理以避免破产的发生以此改善公司的财务状况，防患于未

然。

(4)对银行和债权人的正确决策有重要意义。对于金融机构等债权人而言,发觉上市公司发生财务困境的可能性,以保证其贷款的安全。而财务困境预测模型的预测结果可以帮助银行等金融机构识别企业的财务状况,监控企业的信用风险,避免不良贷款的发生,降低贷款风险。债权人通过财务困境预测可以判断企业的偿债能力,从而提前采取相应的措施加大清收贷款的力度。

1.2 文献综述

1.2.1 财务困境的界定

财务困境的英文可以被翻译为 Financial Distress, 也被称为财务失败(Financial Failure)。目前,对于财务困境的定义有多种解释。

Beaver (1966) 年的研究中,由 79 家“财务困境公司”构成的样本包括:59 家破产公司, 16 家拖欠优先股股利公司和 3 家拖欠债务的公司,即在该研究中, Beaver 把破产、拖欠优先股股利、拖欠债务界定为财务困境^[2]。

Altman (1968) 认为,公司在破产前要经历三个阶段:失败阶段、无力偿还债务阶段和破产阶段。而在这三个阶段中,失败阶段是公司出现财务困境的开端,而无力偿还债务阶段则表明公司已经进入财务困境状态了^[3]。

Carmichael (1972) 认为财务困境是企业履行义务时受阻具体表现为:流动性不足,权益不足,债务拖欠和资金不足四种形式^[4]。

Morris (1997) 列出了企业陷入财务困境的标志,其严重程度逐次递减:(1)债权人申请破产清算,企业自愿申请破产清算,或者被指定接收者完全接收;(2)公司股票在交易所被停止交易;(3)针对持续经营,会计师出具了保留意见;(4)与债权人发生债务重组;(5)债权人寻求资产保全^[5]。

Ross(1999) 将企业财务困境定义为以下四方面:第一,企业失败,即企业清算后仍然不能偿还债权人的债务;第二,法律上的破产,即企业或者是债权人向法院申请破产;第三,会计意义上的破产,即企业资产不足以抵偿债务,账面净资产为负第四,技术层面的破产,即企业到期无力偿还本息^[6]。

谷棋和刘淑莲(1999)将财务困境定义为“企业无力支付到期债务或费用而出现资不抵债的一种经济现象。由于资金管理技术性失败而引发的支付能力不足,通常是暂时的和比较次要的困难,一般可以采取一定的措施加以补救,如通过协商,求得债权

人让步，延长偿债期限，或通过资产抵押等借新债还旧债^[7]。

从以上大多数财务困境研究来看，财务困境的主要内涵包括企业无力支付到期债务、资金流动性较差破产这几方面。而申请破产和破产清算又总是被看成是典型的财务困境状况，所以财务困境研究常被称为破产研究，而财务困境预测也常被称为破产预测。

沪深证券交易所股票上市规则自 1998 年 1 月实施到目前为止，先后经过六次修订，其中对于特别处理的规定也进行了较大修改，日趋完善、合理。

在《关于对存在股票终止上市风险的公司加强风险警示等有关问题的通知》规定上市公司出现财务状况或者其他状况异常，导致投资者难以判断公司前景，权益可能受到损害的，对公司股票交易实行特别处理，在其股票简称前冠以“ST”字样。

2003 年 5 月 8 日实施退市风险警示制度之后，在 2004 年沪深证券交易所修订的《上市规则》中规定上市公司出现财务状况异常或者其他异常情况，投资权益可能受到损害的，对该公司股票交易实行特别处理。特别处理分为警示存在终止上市风险的特别处理和其他特别处理，分别在公司股票简称前冠以“*ST”和“ST”字样，并且股票报价的日涨跌幅限制为 5%。

我国学者在进行相关研究时，大都以被 ST 的企业作为陷入财务困境的企业。ST(Special Treatment, 简称 ST)是指：若公司出现财务状况异常(连续两年亏损或每股净资产低于股票价格)或其他状况异常，就要予以特别处理。在本文的研究中也选择被 ST 的企业作为陷入财务困境的企业。因此本文采用沪深两市 A 股上市公司股票简称前冠以“*ST”和“ST”字样的上市公司作为研究对象。为了表示简便起见，以下被特别处理的公司无论是*ST 公司还是 ST 公司统称为 ST 公司。

1.2.2 国外研究综述

自 20 世纪 30 年代开始，西方学者陆续对企业的财务困境进行研究。到目前为止，各国学者已经提出了很多成熟的预测模型包括传统统计预测模型和人工智能预测模型。传统统计预测模型主要有单变量预测模型、多变量预测模型，人工智能预测模型包括支持向量机模型和人工神经网络模型，并且运用这些预测模型结合实际数据进行了大量的实证研究。

Fitz Patrick (1931) 利用单个的财务指标做破产的预测模型。他在文章“A Comparison of ratios of Successful Industrial Enterprises with those of Failed Firms”中以

19 对公司为样本，将样本分为破产和非破产两组，利用单个的财务指标做破产的预测模型，比较不同时期的财务指标在两组间的差异，以判断公司是否陷入财务困境。他发现净资产收益率和股东权益负债比这两个财务指标的预测能力最强^[8]。

Beaver (1966) 提出了较为成熟的单变量模型，在文章《财务比率与财务失败》中，建立了一个一元线性判定模型，选取了 1954~1964 年间 79 家发生财务困境的企业和相同数量、相同行业、相近资产规模的成功企业，对两类企业的 30 个财务比率在企业破产前 1~5 年的预测能力进行研究。结果发现，债务保障率(现金流量/负债总额)、资产收益率(净收益/资产总额)、资产负债率(负债总额/资产总额)对于预测财务失败是最有效的，而且其中债务保障率预测的准确率最高。并且离财务失败日越近，误判率越低，预测能力越强^[9]。

尽管 Beaver 的一元判别分析方法计算简单，运用方便，具有较高的预测能力，但运用一元判别法建立财务困境预测模型比较简单，但是使用一个或几个财务比率来概括公司的财务特征显然是片面的，并且选用不同财务比率对同一公司进行预测的结果之间可能存在差别，会影响最终的判断。因此，随着学者对统计方法的不断深入研究，多元判别法逐渐取代一元判别法。

Altman (1968) 首次将多元线性判别分析法运用到财务困境预测领域，其样本由 1946-1965 年间提出破产申请的 33 家企业及其相对应的 33 家同行业健康公司配对组成。他从 22 个备选财务指标中选取了营运资产/资产总额、留存收益/资产总额、息税前利润/总资产、权益市价/总负债和本期销售收入/总资产 5 个判别能力最强的比率构造 Z-score 模型，所得到的判别函数为：

$$Z=0.12*X1+0.14*X2+0.33*X3+0.06*X4+1.0*X5$$

即通过产生一个总的判别分，成为 Z 值，并根据 Z 值来判断企业是否破产。在破产前一年的预测准确率较 Beaver 的模型有较大提高，但破产前五年的预测准确性却不如 Beaver 的模型。这就是被后来的学者称之为 Z-Score Model 的 Z 记分模型^[10]。

随着经济环境出现变化，原来的 Z 模型已经无法更准确地预测企业的财务困境。1977 年，Altman, haldeman 和 Narayanan 在 Z-Score Model 的基础上，对 Z-Score 模型进行了改进和发展，提出了一种能够更准确预测企业财务失败的新模型 Zeta 模型^[11]。

多元判别法避免了一元判别法的片面性和不同变量预测结果的矛盾，通过多个变量来反映公司的财务状况，提高了预测精度。但是，该方法需要严格的前提假设，如

变量总体要服从联合正态分布，两组样本的协方差矩阵要相同。而实际上大部分变量都是不服从正态分布的，假设条件很难满足，大多数是在近似情况下运用，从而在一定程度上影响了模型的预测准确率。

Meyer & Pifer (1970) 首次将线性概率模型 (LPM) 用于银行破产预测研究。他们以 1948-1965 年间的 39 家失败银行与相同时间、相同地区、同业时间相似的正常经营银行为配对样本，采用 Altman 研究的变量，将各个财务比率对时间作回归，找出时间趋势，再到线性概率模型中建立判别模型。研究结果表明在失败前 1~2 年，预测准确性达到 80%。但是线性概率模型一个很大的缺点是它的拟合概率值可能大于 1 或者小于 0，而概率值必须要在 0 和 1 之间^[12]。为避免这一缺陷，Logit 模型应运而生。

为了克服多元线性判别分析存在的假设上的局限性，Ohlson (1980) 运用 Logit 逻辑回归方法进行财务困境预测研究。Ohlson (1980) 首次使用 Logit 模型分析财务困境预测问题，对 1970~1976 年间的 105 家破产公司和 2058 家非破产公司组成的配对样本进行分析，使用 9 个变量建立了预测一年内失败、两年内失败和一年或两年内失败的 3 个 Logistic 模型。他发现至少有 4 类显著影响公司破产概率的变量：公司规模、资本结构、经营业绩和当前资产的变现能力^[13]。研究表明，预测准确率分别为 96.12%。

Logit 模型的缺点是对于变量之间的相关性和判别函数变量之间可能存在的多重共线性问题，Logit 的缺点在于模型适用于样本数量多于 200 个，否则就存在参数估计的有偏性；当样本点存在完全分离时，模型参数的最大似然估计可能不存在，故 Logit 模型的有效性存在问题；Logit 模型对中间区域的判别敏感性较强，导致判别结果不稳定。

进入 80 年代后，人工神经网络以非线性、非参数、自适应学习等特征，成为财务困境预测模型的一个强有力工具。Odom 和 Sharda (1990) 最早应用人工神经网络 (ANN) 进行财务困境预测，以 1975~1982 年间的 65 家失败公司和 64 家正常公司为样本，并将其分为训练样本和测试样本两组，采用三层前馈 ANN，使用 Altman 构建的 5 个财务比率，和多元判别分析法比较研究表明，神经网络具有更好的预测精度^[14]；Coats 和 Fant 对 47 家财务困境公司和 47 家正常公司运用神经网络模型进行判别时，对财务困境公司的预测准确率达到 91%，明显高于多元判别法 72% 的准确率^[15]。

Tam & Kiang (1992) 选用美国德克萨斯州 1985~1987 年间的 59 家破产银行和 59 家正常银行为研究对象，运用多元判别分析法、Logistic 回归法、K-近邻法、ID3 决策树法和神经网络技术分别建立模型进行分析，得出在破产前一年的预测误判率分别为 14.2%、17.7%、22.8%、22.8%、10.8%，表明神经网络技术的预测能力要好于其

他方法^[16]。

Desheng 和 Wu(2006)选取了制造业中的 48 家公司为研究样本,使用 7 个财务比率,分别建立了反向传播神经网络(BPNNs)、多元判别分析(MDA)和包含概率的神经网络预测模型(PNN)^[17]。研究结果证明 PNN 比另外两模型预测效果更好,短期和中期的预测精度分别为 87.5%和 81.25%。

支持向量机(SVM)在学习能力和泛化能力方面具有以下优良特性:SVM 基于结构风险最小原则而不是经验风险最小原则,能较好地避免过拟合;SVM 算法是一个凸优化问题,局部最优解一定是全局最优解,不容易陷入局部最优;实际应用中,SVM 能在样本容量相对较小的情况下取得较好的效果,尤其适用于样本量较少的分类预测问题。

MinandLee(2000)利用支持向量机(SVM),以韩国中小型制造企业为样本进行了财务困境的预测。该模型将支持向量机模型与神经网络、线性判别分析进行了比较,结果显示线性判别分析的预测精度最差,神经网络的预测精度优于线性判别分析,支持向量机的预测精度最高^[18]。支持向量机等机器学习方法虽然具有很高的预测性,但是,它们存在最大的缺点就是信息不透明。

Fan A et al. (2000)将改进后的支持向量机模型应用于财务预警研究,研究得出采用该模型的预测效果要好于传统方法。鉴于 SVM 有此优良特性,采用 SVM 预测模型,与之前提到判别分析模型和 logit 模型分别做实证研究,结果非常理想^[19]。

Van GeStel T et al. (2003)将最小二乘法引入到支持向量机,构建了最小二乘支持向量机(LS-SVM)模型研究财务困境预测,结果表明模型的判别准确率高达 89.91%^[20]。

Kyung-Shik Shin et al. (2005)研究了支持向量机和 BP 神经网络在破产预测中的效果差异。研究显示,在测试集的大小不变,而训练集的大小由 1856 降至 200 以下后,BP 神经网络模型在测试集上的准确率由 71.7%降至 48%,而支持向量机在测试集上的准确率由 74.6%升至 87.9%^[21]。可见,在小样本情况下 BP 神经网络的预测精度显著下降。

Melek Acar Boyacioglu, Yakpu Kara (2009)等分别运用 ANN、SVM、三种多变量模型(MDA、Logit 和 K 均值聚类分割法)进行银行的财务困境预测,结果显示作为新开发的学习算法,SVM 表现优于其他方法。

1969 年,组合预测理论由 J.M.Bates 和 C.W.J.Granger 首次提出,发现将不同的预测方法进行适当的组合,形成一种组合预测模型,可以提高预测精度。组合预测在

单项预测模型的结果基础上，确定各个模型的权重，并以某种方法进行组合。组合预测一般可分为线性组合模型与非线性组合模型^[22]。

西方国家在财务困境预测方面从二十世纪初就开始了，研究成果显著，提出了线性模型如单变量模型、多变量模型、非线性模型如神经网络模型、支持向量机模型，并提出了组合预测思想。除了上述介绍的预测方法外，国外学者还提出遗传算法(genetic algorithm)、粗集方法(rough set theory)、支持向量机(support vectormachines)等人工智能方法，以及 K 近邻法、聚类分析、贝叶斯决策法等非参数方法。

1.2.3 国内研究综述

近年来随着证券市场发展，投资者对财务困境预测产生了迫切需求，我国学者开始了对财务困境预测的研究。1986 年吴世农、黄世忠首次在我国介绍了企业破产的分析指标和预测模型。陈静(1999) 使用 1995~1997 年我国上市公司的财务数据进行研究。她选择了 27 家 ST 公司作为研究样本，并用同行业、同规模的 27 家非 ST 公司为配对样本，利用多元判别分析模型进行预测。

吴世农、卢贤义(2001)以 FiSher 判别分析、多元线性回归和逻辑回归分析三种方法，分别构建了相应的模型，结果证明逻辑回归模型的判定能力最好^[23]。

刘凤娇(2006)在 Altman 的“Z-Score”模型的基础上，选取沪、深两地 80 家上市公司作为样本进行了研究，验证了“Z-Score”模型在财务风险预警中有着准确的判别结果。文章还研究了财务失败的原因及财务失败后所采取的措施。

为了对企业财务困境预测模型进行比较，杨保安、季海(2001)提出神经网络技术在财务困境预测上的应用，并用实例进行示范性设计和验证。发现，对比统计方法中判别分析方法在预测正确率上有很大的提高。

杨淑娥，黄礼(2005)采用 BP 人工神经网络工具，选择了 120 家上市公司作为训练样本，并使用 69 家上市公司作为检测样本，建立了财务困境预测模型。取得了建模样本 90.8%和检验样本 90%的正确率^[24]。

吕长江、周现华(2005)以 1999~2002 年的制造业上市公司为样本，分别采用多元判别分析、LogiSTic 回归和人工神经网络建立上市公司财务困境预测模型。结果显示，三个模型在公司发生困境前 1 年、前 2 年、前 3 年均能进行较好的预测，但神经网络模型的预测准确率最高，LogiSTic 回归模型次之，多元判别分析法最差^[25]。

惠守博、王文杰(2006)提出了利用支持向量机建立财务预警系统的方法，筛选了财务评价指标体系，利用支持向量机的分类能力建立了财务预警的模型。然后利用

上市公司的财务数据进行检验，证明了基于支持向量机的财务预警模型的可行性和实用性，实验表明支持向量机在小样本情况下具有良好的非线性建模能力和泛化能力。

杨瑶 (2007)对 ST 前三年的财务数据进行检验，分别选取了包括偿债能力、发展能力、盈利能力、运营能力等财务指标，并采用主成分分析法对财务指标进行筛选，提取了 5 个主成分，并运用 Logit 回归模型进行预测，发现越靠近 ST 的时间，模型判别正确率越高。

2008 年，刘彦文分别用 Logit 预警模型、神经网络预警模型和神经网络预警模型构建财务困境预测模型，文章特色在于指标体系中引入了盈余管理指标和非财务信息。

2009 年，姚国芳构建了一个组合预测模型，选取了盈利能力、偿债能力、资产运营能力、成长能力和现金流量五个方面的指标体系。根据 2004-2006 年间的 88 家制造业上市公司被 ST 前 3 年财务数据，先后构建了 Logit 回归、Probit 回归、BP 神经网络三种单项预测模型及基于 Logit 回归和 Probit 回归的线性组合预测模型和基于 BP 神经网络的非线性组合预测模型。应用结果表明，组合预测模型在分类精度和稳健性两方面都较相应的单一模型有了提高，线性组合预测模型的稳健性优于利用 BP 神经网络构建的非线性组合预测模型，但分类精度低于后者。

刘静淑(2010)建立基于支持向量机财务预警模型，选取 2007—2009 年符合条件的 33 家由于“财务状况异常”被“ST”的上市公司，并且按照行业分类和资产规模相当的原则，选取 33 家财务状况良好的正常上市公司作为配对样本，并筛选 11 个财务指标建立指标体系。得到如下主要的研究结论：①支持向量机预警模型在判别上市公司财务状况有更高的准确率。②将 SVM 和 BP 神经网络模型进行自身性能稳定性比较，支持向量机有小样本研究、维数不受限制等优势，而 BP 神经网络模型自身性能欠稳定。

我国关于财务困境预测的研究近十年来得到了较大的发展，但与西方相比，国内研究还相对滞后。主要存在以下几个问题：一是预测变量的选取上，很多凭直观经验或直接借用国外的研究成果。在选择适合我国企业特点的指标上没有大的突破。二是在选取样本时，未考虑行业的因素，而国外学者针对石油、运输等行业建立了不同的行业模型。三是缺乏组合预测思想，鲜有构建出传统统计分析方法与人工智能方法相结合的组合预测模型。

1.3 研究内容、方法、创新点及框架

1.3.1 研究内容

根据研究问题的需要，本文的研究内容大致如下：

第一章为绪论，主要介绍本文的研究目的、研究意义，以及对国外和国内学者关于企业财务困境预测这一领域具有代表性的相关研究进行了述评，最后介绍了本文的研究方法、研究内容和主要的创新点。

第二章是财务困境预测模型的介绍，包括三种单项财务困境预测模型包括多元判别分析模型、逻辑回归模型、支持向量机模型和二种组合预测模型包括线性组合模型和基于人工神经网络的组合预测模型，并对它们的原理和使用方法做了详细说明。

第三章介绍了财务困境预测的指标体系的选择以及数据的预处理。本章首先介绍了指标的选取原则，财务指标分偿债能力、营运能力、盈利能力、获现能力等方面选取，再加上资产规模和审计意见两个非财务指标；然后对这些指标数据进行了例如正态检验，显著性检验，以及因子分析等一系列预处理，最终得到公共因子，作为预测模型的输入变量。

第四章为上市公司财务困境预测实证研究。对预处理后的公共因子进行实证研究，分别建立了多元判别分析、Logit 回归模型、支持向量机模型、线性组合预测模型和神经网络组合预测模型，并对实证研究的结果进行比较。

第五章是结论部分，得出本文的研究结论及不足。

1.3.2 研究方法

(1) 本文采用理论分析和实证分析相结合的方法。首先通过理论分析初步选取了影响财务困境的财务指标，然后通过因子分析提取公因子，作为预测模型的输入变量。

(2) 实证研究，运用多元判别分析、Logit 回归模型等统计方法和支持向量机和神经网络等人工智能型预测方法，模型来对公司财务业绩进行预测。

(3) 在所提出的企业财务困境预测方法与思路中，强调组合研究，包括线性组合和基于神经网络的非线性组合。

(4) 在论文中将比较研究方法贯穿始终，在对企业财务业绩预测的实证研究过程中，对多元判别分析、Logit 回归模型和支持向量机模型及线性组合模型和神经网络组合模型的预测结果，最后对各模型的结果进行比较分析。

1.3.3 创新点

本文的创新点有：

(1) 利用多种单项模型构建了一个线性组合预测模型和一个基于神经网络学习的组合预测模型，然后对预测结果的准确率进行比较分析。

(2) 分别对预测模型的训练结果和预测结果的准确性进行比较，分析模型的准确率差异，用来判断模型的稳定性。

1.3.4 财务困境预测框架

企业财务困境预测的大致过程可以用图 1-1 表示：

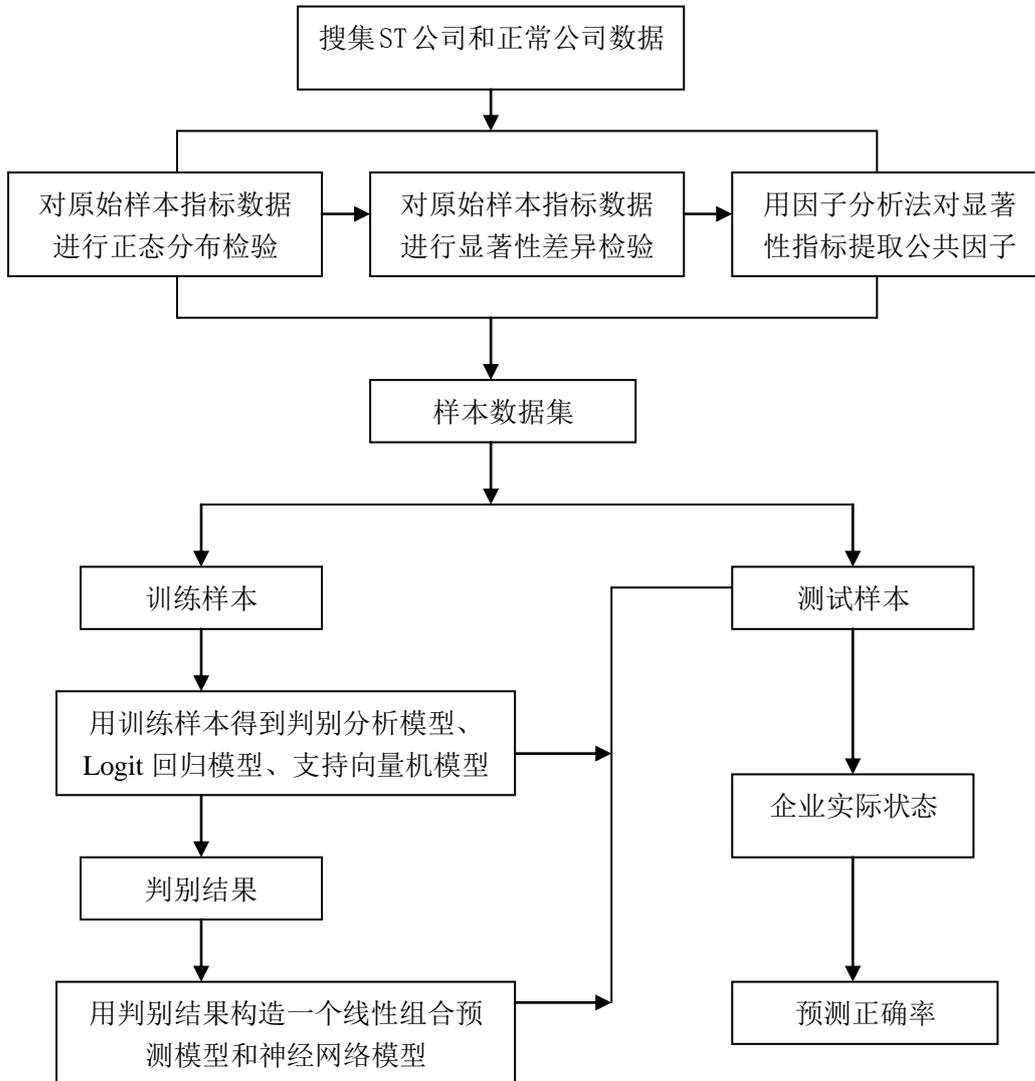


图 1-1 企业财务业绩预测过程框架

第 2 章 财务困境预测模型方法介绍

本章首先介绍了判别分析模型、Logit 回归模型、最小二乘支持向量机模型的基本原理和判别方法；然后介绍了模型的组合方法，包括线性组合和基于神经网络的组合。

2.1 基于单模型的财务困境预测方法

2.1.1 多元判别分析 (MDA)

判别分析是一种进行统计和分组的技术，根据观察或测量到若干变量值，判别样本所属类型的一种统计分析方法。它于 20 世纪 30 年代由费雪提出。近年来，在自然科学、社会学及经济管理学科中都有广泛的应用。

判别分析的特点是根据观察对象的分类和若干表明观察对象特征的变量值，总结出客观事物分类的规律性，建立判别函数，然后对待判对象进行判断其类别。然后，当遇到新的样本点时，只要根据总结出来的判别公式和判别准则，就能判别该样本点所属的类别。在判别分析中，称分类变量为因变量，而用以分类的财务指标称为判别变量或自变量。

判别分析就是希望利用已经测得的变量数据，找出一种判别函数，使得这一函数具有某种最优性质，能把属于不同类别的样本点尽可能地区别开来。为了更清楚的认识判别分析的实质，以便能灵活的应用判别分析方法解决实际问题，我们有必要了解“划分”这样概念。

设 R_1, R_2, \dots, R_k 是 p 维空间 R^p 的 k 个子集，如果它们互不相交，且它们的和集为 R^p ，则称 R_1, R_2, \dots, R_k 为 R^p 的一个划分。

在两个总体的距离判别问题中，利用 $W(\mathbf{X}) = \mathbf{a}'(\mathbf{X} - \bar{\boldsymbol{\mu}})$ 可以得到空间 R^p 的一个划分

$$\begin{cases} R_1 = \{\mathbf{X} : W(\mathbf{X}) \geq 0\} \\ R_2 = \{\mathbf{X} : W(\mathbf{X}) < 0\} \end{cases} \quad (2.1)$$

新的样品 \mathbf{X} 落入 R_1 推断 $\mathbf{X} \in G_1$ ，落入 R_2 推断 $\mathbf{X} \in G_2$ 。

这样我们将会发现，判别分析问题实质上就是在某种意义上，以最优的性质对 p 维空间 R^p 构造一个“划分”，这个“划分”就构成了一个判别规则。这一思想将在

后面的各节中体现的更加清楚。

本文旨在研究两类公司的分类问题。一类是 ST 公司，另一类是非 ST 公司。判别函数的一般形式是：

$$Z = \beta_0 + \beta_1 * X_1 + \beta_2 * X_2 + \beta_3 * X_3 + \dots + \beta_k * X_k \quad (2.2)$$

其中： Z 为被解释变量(判别值)， X_1 、 X_2 、 $X_3 \dots X_k$ 是自变量，反映研究对象的财务比率， β_1 、 β_2 、 $\beta_3 \dots \beta_k$ 为各变量的判别系数。

根据 Fisher 的线性判别模型，我们可以计算出两类样本的均值 Z_1 和 Z_2 ，“判别点”则根据对称分类原则确定：

$$Z_0 = \frac{Z_1 + Z_2}{2} \quad (2.3)$$

最后，我们将每个公司的 Z 分值与判别点 Z_0 进行比较，如果 $Z > Z_0$ 则判为财务健康公司，反之则判为财务困境公司。

2.1.2 逻辑回归模型 (Logit)

Logit 模型就是使用 Logit 逻辑回归，根据样本数据使用最大似然估计法估计出各参数值，再经过一定的数学推导运算，可以求得公司陷入财务困境的概率。

Logit 回归的假设前提是：(1) 因变量 Y_i 是二分变量；(2) 数据必须来自随机样本；(3) 变量 Y_i 被假定为 k 个自变量 $X_{ij} (j=1,2,\dots,k)$ 的函数，因变量与自变量之间的关系是非线性的；(4) 自变量之间不存在多重共线性关系。

Logit 概率函数的形式为：

$$P_i = F(Y_i) = \frac{1}{1 + e^{-Y_i}} = \frac{e^{Y_i}}{1 + e^{Y_i}} \quad (2.4)$$

其中， Y_i 是二值响应线性组合模型，即

$$Y_i = \alpha + \sum_{j=1}^k \beta_j X_{ji} \quad (2.5)$$

当 $Y_i = 0$ 时，表示公司 i 属于非财务困境公司；当 $Y_i = 1$ 时，表示公司 i 属于财务困境公司。 $X_{ji} (j=1,2,\dots,k)$ 表示第 i 家公司的第 j 个指标， α 为截距， β_j 为第 j 个指标待估计的参数。

Logit 回归拟合出的回归方程为：

$$\ln \frac{P_i}{1-P_i} = Y_i = \alpha + \sum_{j=1}^k \beta_j X_{ji} \quad (2.6)$$

P_i 是在条件下 X_{ji} 事件发生的概率。通过设定临界值作为事件发生与否的标准，如果 P_i 大于临界值，则判定事件发生；反之，判定事件不发生。

利用二项 Logit 回归模型进行财务困境预测的关键是求出回归系数 $\beta_1, \beta_2, \dots, \beta_j$ 以及常数 α ，再将回归系数带入(2.1)，(2.2)式即可求出 P_i ，即每一个样本的概率值。

在研究某个上市公司是否会发生财务困境的问题中，可以将 ST 公司的状态用数字 1 表示，而将非 ST 公司的状态用数字 0 表示，这种变量被称为虚拟变量。Logit 回归模型中因变量只能取两个值 0 和 1(虚拟因变量)，通过分类的财务指标变量(自变量)得到事件发生的概率，也就是因变量取 1 的概率。

利用 Logit 回归模型来进行财务业绩预测的关键是要求得式(2.6)中的回归系数 $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ ，然后将求得的回归系数回代入式(2.4)，求得每一个样本的概率值 p_i 。利用 SPSS 统计软件计算回归系数的过程如下：

(1)回归模型整体显著性检验。将训练样本数据录入统计软件，对回归模型整体进行显著性检验，计算卡方统计量以及相应的伴随概率 P 值。若 P 小于给定的显著性水平 α ，拒绝原假设，模型整体显著；若 P 大于 α ，模型整体不显著。

(2)回归系数的显著性检验。计算 Wald 统计量和相应的概率 P 值，若回归系数的概率 P 值小于给定的显著性水平 α ，说明该回归系数显著不为 0；反之，该回归系数显著为 0。

(3)回归模型的拟合优度评价。计算 $-2 \log likelihood$ 的值，该值越大，说明模型的拟合优度越高。计算 Nagelkerke R^2 统计量的值，该值越接近于 1，模型的拟合优度越高。

(4)计算模型的回归系数 $\beta_0, \beta_1, \beta_2, \dots, \beta_k$ ，将回归系数和测试样本数据代入式(2.4)，即可计算出测试样本的概率值 p_i 。如果 p_i 大于等于 0.5，则判断该公司为财务业绩上升公司；如果 p_i 小于 0.5，则判断该公司为业绩下降公司，据此可以求出模型的预测准确率。

多元逻辑回归模型的优点:(1)不要求解释变量服从正态分布，克服了多元判别模

型统计假设约束条件的局限性，因此其应用范围广泛;(2)预测准确度较高。缺点是计算过程较为复杂，且有较多近似处理。

2.1.3 最小二乘支持向量机(LS-SVM)

支持向量机 (Support Vector Machine, SVM) 是由 Vapnik 与其领导的贝尔实验室研究小组一起开发出来的一种新的机器学习技术，这里的“机 (machine, 机器)”实际上是一个算法。在机器学习领域，常把一些算法看做是一个机器。Vapnik 等人提出的 SVM 基于二分类问题，其核心是研究如何根据有限样本进行模式识别和回归预测，并使在对未知样本的估计过程中，期望风险达到最小。

SVM 方法的基本思想是：通过一个非线性映射 ϕ 将输入向量映射到一个高维的特征空间，从而在这个空间中寻找最优分类超平面，这个平面能够尽可能多的将两类数据点分开，同时使分开的两类数据点距离分类面最远，即在这个高维的特征空间中可以应用线性学习机的方法来解决样本空间中的高度非线性分类和回归等问题。

支持向量机最初主要用于模式识别问题中的线性可分二分类问题的研究。分类问题就是根据给定的训练集 $T = \{(x_i, y_i), i = 1, 2, \dots, l\}, x_i \in R^n, y_i \in \{+1, -1\}$ ，寻找决策函数：

$$f(x) = \text{sgn}(g(x)) \quad (2.7)$$

推断任一样本点 x 相对应的 $y = f(x)$ 的值，其中， $\text{sgn}(\cdot)$ 为符号函数。机器学习理论把寻找决策函数的方法命名为分类学习机。当 $g(x)$ 为线性函数时，由 (2.7) 确定的分类方法称为线性分类学习机；当 $g(x)$ 为非线性函数时，称为非线性分类学习机。

支持向量机分类问题主要分为三类：线性可分支持向量机、线性不可分支持向量机、非线性支持向量机。下面对三种分类支持向量机进行一一介绍。

(1) 线性可分支持向量机

支持向量机问题实质上是求解具有唯一解的二次型规划问题。

给定训练样本集： $(x_1, y_1), \dots, (x_i, y_i), \dots, (x_l, y_l), i = 1, 2, \dots, l, x_i \in R^n, l$ 为样本的个数， n 为解释变量的数目，在本文的研究中指 n 个财务指标， $y_i \in \{+1, -1\}$ 为被解释变量。在本文中，当公司财务业绩下降时， $y = -1$ ；财务业绩上升时， $y = +1$ 。

假设训练样本集能被一个分类超平面：

$$w^T x + b = 0 \quad (2.8)$$

分开。 $w = (w_1, w_2, \dots, w_l)^T$ 为垂直于分类超平面的法向量， $x = (x_1, x_2, \dots, x_l)^T$ 为训练集向量， b 为常数。

如果超平面能够把训练集中的全部向量准确的分成两类，且能够使得该超平面与离其自身最近的向量间的距离达到最大值，则称训练集向量被该超平面分开，且该超平面为最优超平面。如图 2-1 所示，下图中的五角星和四角星分别代表两类样本（本文为财务业绩上升样本和财务业绩下降样本），粗实线为分类超平面，虚线代表经过两类样本中离分类面最近的样本与分类面平行的平面，两条虚线间的距离为分类间隔，其大小为 $2/\|w\|$ （ $\|w\|$ 为向量 w 的模）。最优分类超平面是指不仅能把两类样本正确的分开，还能使得分类间隔 $2/\|w\|$ 达到最大值的分类面。其中，位于图（2-1）虚线上的训练样本点就是支持向量。

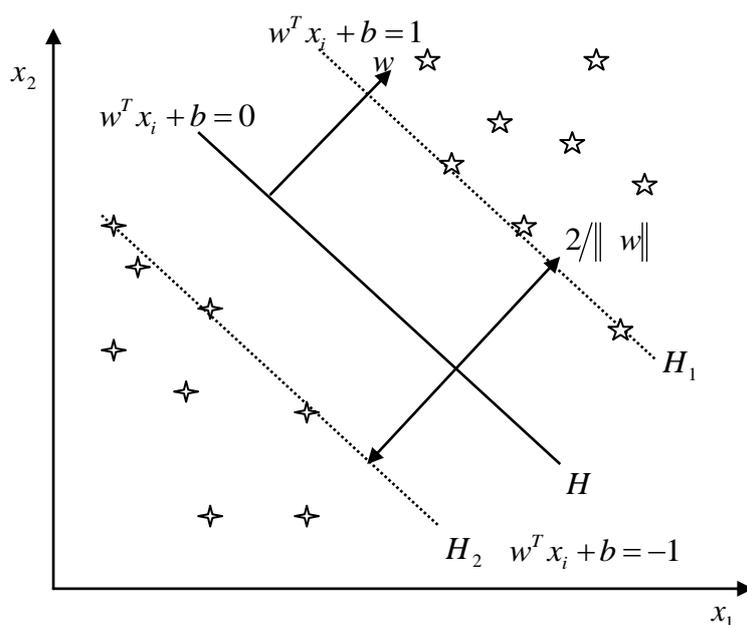


图 2-1 线性可分支持向量机示意图

为了详细的描述分类超平面，假设训练样本集数据满足以下条件：

$$w^T x_i + b \geq 1, \text{ 若 } y_i = +1 \quad (2.9)$$

$$w^T x_i + b \leq -1, \text{ 若 } y_i = -1 \quad (2.10)$$

则它们可以被组合为下式：

$$y_i(w^T x_i + b) - 1 \geq 0 \quad (2.11)$$

对于满足不等式(2.9)的样本点，一部分在超平面 $H_1: w^T x_i + b = 1$ 上，一部分不在 H_1 上；同样，对于满足不等式(2.10)的样本点，一部分在超平面 $H_2: w^T x_i + b = -1$ 上，一部分在 H_2 外。超平面 H_1 和 H_2 满足下列条件：

- ① H_1 和 H_2 平行；
- ② H_1 和 H_2 之间没有任何训练样本点；
- ③ H_1 和 H_2 之间的距离为 $2/\|w\|$ 。

最优分类超平面就是不仅能把两类样本正确的分开，还能使得分类间隔 $2/\|w\|$ 达到最大值的分类面。求 $2/\|w\|$ 的最大值等价于求 $\frac{1}{2}\|w\|^2$ 得最小值。

$$\min \Phi(w) = \frac{1}{2}\|w\|^2 \quad (2.12)$$

由(2.12)式可以看到，每一对 w 和 b 都可以确定一个超平面的方程 $w^T x_i + b = 0$ ，我们的目的就是要找到一对最优的 (w, b) ，使得对于某些样本点 x_i ，有 $y_i(w^T x_i + b) - 1 = 0$ 。在图中这些 x_i 位于虚线所在的平面上，这些 x_i 就是所谓的“支持向量”。

统计学习理论中指出，一个规范分类超平面构成的预测函数集：

$$f(x) = \text{sgn}(w^T x + b) \quad (2.13)$$

$$\text{的 VC 维 } h \text{ 满足: } h \leq \min\{[R^2 A^2], n\} + 1 \quad (2.14)$$

这里， n 是向量空间的维数， R 是覆盖所有样本向量的超球半径， $\|w\| \leq A$ 。因此，如果想要 VC 维 h 尽量小，那就要保证 A 值尽量小。这样 VC 维的降维问题就转化成最小化 $\|w\|$ 也就是增大分类间隔 $2/\|w\|$ 的问题。支持向量机方法的出发点就是在经验风险固定的条件下，把最小化期望风险问题转化为最小化 $\|w\|$ 的问题。即支持向量机就是要在满足 (2.14) 的条件下，最大化分类间隔 $2/\|w\|$ 。所以，分类间隔最大化问题可以表示为下式：

$$\min \Phi(w) = \frac{1}{2}\|w\|^2 \quad (2.15)$$

$$s.t. y_i(w^T x_i + b) \geq 1, i = 1, 2, \dots, l \quad (2.16)$$

问题(2.15)、(2.16)是数学上常见的求最优解的问题。式(2.15)、(2.16)的最优解为下面(2.15)拉格朗日函数的鞍点:

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 - \sum_{i=1}^l \alpha_i [y_i(w^T x_i + b) - 1] \quad (2.17)$$

其中 $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_l)^T \in R^l$, α_i 为对应约束条件(2.17)的拉格朗日乘子。如果 α_i 非负, 对应的输入样本变量 x_i 就是支持向量。

由于上面的拉格朗日函数在鞍点处的 w 和 b 的梯度为零, 所以:

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^l \alpha_i y_i x_i = 0 \Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i \quad (2.18)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^l \alpha_i y_i = 0 \Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.19)$$

把(2.18)和式(2.19)代入到(2.17)中, 并对它关于 α 求极大值, 就可以得到问题(2.15)、(2.16)的对偶问题:

$$\max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (2.20)$$

$$s.t. \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.21)$$

$$\alpha_i \geq 0, i, j = 1, 2, \dots, l \quad (2.22)$$

如果 α 为问题(2.20)~(2.22)解, 将 α 代入(2.18), 则:

$$w = \sum_{i=1}^l \alpha_i y_i x_i \quad (2.23)$$

根据库恩塔克定理, 最优解还应该满足下列条件:

$$\alpha_i [y_i(w^T x_i + b) - 1] = 0, i = 1, 2, \dots, l \quad (2.24)$$

大多数样本点对应的 α_i 值均为 0, 对应的 α_i 值不为 0 的样本点即为支持向量。由式(2.24)可知, 只有支持向量才对 w 有贡献, 即对最优超平面和决策函数有贡献。 α_i 和支持向量决定了最优超平面, 支持向量因此得名, 对应的学习方法称为支持向量机。

关于 b 的求解由 (2.24) 可知, $\alpha_i \neq 0$, 必有 $y_i(w^T x_i + b) = 1$ 即 $w^T x_i + b = 1$, $b = (1 - y_i w^T x_i) / y_i$, 然后将 (2.41) 代入 $b = (1 - y_i w^T x_i) / y_i$ 得:

$$b = y_i - w^T x_i = y_i - \sum_{i=1}^l \alpha_i y_i x_i^T x_i \quad (2.25)$$

实际应用中, 为了计算可靠, 通常先分别计算与所有支持向量相对应的 b 值, 然后求平均值, 即 $b = \frac{1}{N_{SV}} \sum_{x_i \in SV} (y_i - \sum_{i=1}^l \alpha_i y_i x_i^T x_i)$ 。

因此, 对于给定的未知样本点 x , 只需计算: $f(x) = \text{sgn}(w^T x + b)$ 就可以判断样本点 x 所属的类别。如果 $f(x) = 1$, 则将该样本点划分到 $y = 1$ 这一类, 本文中为财务业绩下降的公司; 如果 $f(x) = -1$, 则将样本点划分到 $y = -1$ 这一类中, 本文对应财务业绩上升的公司。

(2) 线性不可分支持向量机

在实际操作中, 多数情况下问题并不是线性可分的, 再者即使问题是线性可分的, 由于各种原因, 也会存在一些被错误分类的样本点, 这会对所构建的最优分类超平面造成严重的影响。因此, 对于线性不可分的问题, 我们通过增加一个大于等于 0 的松弛变量来放宽约束条件。也就是说对每一个训练样本点 $(x_i, y_i), i = 1, 2, \dots, l$, 引入一个非负的松弛变量 $\xi_i \geq 0, i = 1, 2, \dots, l$, 这样约束条件变为

$y_i(w^T x_i + b) \geq 1 - \xi_i$, $\sum_{i=1}^l \xi_i$ 描述了训练集被错误分类的程度。因此, 我们在解决线性

不可分问题时, 就要实现两个目的: 第一, 分类间隔 $2/\|w\|$ 尽可能大; 第二, 错分程度 $\sum_{i=1}^l \xi_i$ 尽可能小。为了完成以上两个目标, 引入一个惩罚因子 $C > 0$, 前面 (2.15)、(2.16)

的问题就转化为下面的优化问题:

$$\min_{w, b, \xi} \Phi(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad (2.26)$$

$$s.t. y_i(w^T x_i + b) \geq 1 - \xi_i \quad (2.27)$$

$$\xi_i \geq 0, i = 1, 2, \dots, l \quad (2.28)$$

其中， ξ_i 为样本点离其所属类别的距离，刻画了样本点 x_i 违反约束 $y_i(w^T x_i + b) \geq 1$ 的程度， C 可以看做是对误差的惩罚因子， C 越大，对错误的惩罚越严重。

同线性可分问题一样，(2.26) ~ (2.28) 也是一个二次型规划问题，引入拉格朗日函数，该问题的最优解即为下面拉格朗日函数的鞍点：

$$L(w, b, \alpha) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i [y_i(w^T x_i + b) + \xi_i - 1] - \sum_{i=1}^l \beta_i \xi_i \quad (2.29)$$

上式中， α_i ， β_i 为非负的拉格朗日乘子，由于在鞍点处 w 、 b 和 ξ_i 的梯度为零，可得：

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^l \alpha_i y_i x_i = 0 \Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i \quad (2.30)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^l \alpha_i y_i = 0 \Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.31)$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \beta_i = 0 \quad (2.32)$$

将式(2.30)~(2.32)代入式(2.29)，并对它关于 α 求最大值，得到(2.44)~(2.46)的对偶问题：

$$\max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (2.33)$$

$$s.t. \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.34)$$

$$0 \leq \alpha_i \leq C, i = 1, 2, \dots, l \quad (2.35)$$

根据库恩塔克定理可知，最优点处拉格朗日乘子 α_i 与约束条件的积为：

$$\alpha_i [y_i(w^T x_i + b) + \xi_i - 1] = 0, i = 1, 2, \dots, l \quad (2.36)$$

$$(C - \alpha_i) \xi_i = 0, i = 1, 2, \dots, l \quad (2.37)$$

求解 (2.33)~(2.35) 得到的 α_i 中， α_i 的取值有以下三种可能性：

① $\alpha_i = 0$ ，由(2.37)可得 $\xi_i = 0$ 。此时所有的样本点都被正确地分类；

② $0 < \alpha_i < C$ ，由(2.37)可知 $\xi_i = 0$ ，且由(2.36)可知 $y_i(w^T x_i + b) = 1$ 。此时的样本点 x_i 为标准支持向量；

③ $\alpha_i = C$ ，由(2.37)， ξ_i 取值不确定，可能为 0，也可能不为 0。当 ξ_i 不为 0 时，由(2.36)， $y_i(w^T x_i + b) < 1$ ，此时对应的样本点是误判点。当 ξ_i 为 0 时，对应的样本点为边界支持向量。

根据(2.22)，可以求得 $w = \sum_{i=1}^l \alpha_i y_i x_i$ ，再由(2.20)可知，任一支持向量满足

$$y_i(w^T x_i + b) = 1, \text{ 所以, } b = y_i - w^T x_i = y_i - \sum_{i=1}^l \alpha_i y_i x_i^T x_i。$$

同样在实际应用中，通常使用 b 的平均值，即 $b = \frac{1}{N_{SV}} \sum_{x_i \in SV} (y_i - \sum_{i=1}^l \alpha_i y_i x_i^T x_i)$ 。

从前面的讨论可以得到，支持向量机模型计算的复杂程度主要取决于样本点的个数，而与变量的空间维数没有关系，因此，支持向量机可以成功的处理高维空间的问题。

(3) 非线性支持向量机

如果训练集样本是非线性的，那么我们就需要寻找一个最优分类超曲面。但是我们对于超曲面的几何空间知识了解的比较少，且几乎所有的几何图形都是关于超平面的，所以为了解决这个问题，我们引入一个非线性映射，用非线性映射将输入空间映射到某一具有更高维数的特征空间。把原来输入空间的非线性问题经过映射到特征空间后，就可以在特征空间用线性可分支持向量机对问题进行求解了。

针对非线性分类问题，为了将输入空间映射到一个高维的特征空间，我们引入一个非线性映射函数 $\varphi(\cdot)$ ：

$$\varphi: x \in R^n \rightarrow \varphi(x) \in H \quad (2.38)$$

在特征空间对样本点进行线性分类，此时分隔超平面为：

$$w^T \varphi(x) + b = 0 \quad (2.39)$$

$b \in R$ ， $w \in H$ ，求分类超平面就转化为求解下面的最优化问题：

$$\min_{w,b,\xi} \Phi(w) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \quad (2.40)$$

$$s.t. y_i(w^T x_i + b) \geq 1 - \xi_i \quad (2.41)$$

$$\xi_i \geq 0, i=1,2,\dots,l \quad (2.42)$$

(2.40) ~ (2.42) 最优解为下面拉格朗日函数的鞍点:

$$L(w,b,\alpha) = \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i - \sum_{i=1}^l \alpha_i [y_i(w^T x_i + b) + \xi_i - 1] - \sum_{i=1}^l \beta_i \xi_i \quad (2.43)$$

$\alpha_i \geq 0$, $\beta_i \geq 0$ 为拉格朗日乘子, 由于鞍点处 w 、 b 和 ξ_i 的梯度为零, 由此可得:

$$\frac{\partial L}{\partial w} = w - \sum_{i=1}^l \alpha_i y_i x_i = 0 \Rightarrow w = \sum_{i=1}^l \alpha_i y_i x_i \quad (2.44)$$

$$\frac{\partial L}{\partial b} = \sum_{i=1}^l \alpha_i y_i = 0 \Rightarrow \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.45)$$

$$\frac{\partial L}{\partial \xi_i} = C - \alpha_i - \beta_i = 0 \quad (2.46)$$

将 (2.44)~(2.46) 代入 (2.43), 并对它关于 α 求最大值, 得到(2.58)~(2.60)的对偶问题为:

$$\max_{\alpha} \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l \alpha_i \alpha_j y_i y_j x_i^T x_j \quad (2.47)$$

$$s.t. \sum_{i=1}^l \alpha_i y_i = 0 \quad (2.48)$$

$$0 \leq \alpha_i \leq C, i=1,2,\dots,l \quad (2.49)$$

其中, $K(x_i, x_j) = \varphi^T(x_i) \varphi(x_j)$ 称为核函数。由库恩塔克定理可知, 最优点处拉格朗日乘子 α_i 与约束条件的乘积为:

$$\alpha_i [y_i(w^T x_i + b) + \xi_i - 1] = 0, i=1,2,\dots,l \quad (2.50)$$

$$(C - \alpha_i) \xi_i = 0, i=1,2,\dots,l \quad (2.51)$$

求解问题(2.46)~(2.48)得到 α_i , α_i 有三种可能的取值: $\alpha_i = 0$ 、 $0 < \alpha_i < C$ 、 $\alpha_i = C$,

与在前面线性不可分支持向量机问题的讨论一样，只有满足 $0 < \alpha_i < C$ 的样本点 x_i 才为支持变量。

根据(2.46)，可以解得 $w = \sum_{i=1}^l \alpha_i y_i \varphi(x_i)$ ，再由(2.40)可以知道，对于任一支持向量，

满足： $y_i(w^T \varphi(x_i) + b) = 1$ ，所以 b 为： $b = y_i - w^T \varphi(x_i) = y_i - \sum_{i=1}^l \alpha_i y_i \varphi^T(x_i) \varphi(x_i)$ 。

为了计算可靠，实际应用中，使用 b 的平均值，即

$b = \frac{1}{N_{SV}} \sum_{x_i \in SV} (y_i - \sum_{i=1}^l \alpha_i y_i \varphi^T(x_i) \varphi(x_i))$ ，其中， $K(x_i, x_j) = \varphi^T(x_i) \varphi(x_j)$ 为核函数。因此，

对于给定的测试集样本点 x ，只需计算：

$f(x) = \text{sgn}(w^T \varphi(x) + b) = \text{sgn}(\sum_{i=1}^l \alpha_i y_i K(x_i, x) + b)$ 就可以判断 x 的类别。

可以证明，只要函数 $K(x_i, x)$ 满足 Mercer 条件，就可以作为核函数。如果核函数 $K(x_i, x)$ 选取得适当，就可以将输入空间中的非线性分类问题转化为高维特征空间中的线性可分问题。

因此只有先选择合适的核函数，才能解决支持向量机的最优化求解问题。常见的核函数形式有：

①多项式核函数： $K(x_i, x) = (1 + x_i^T x)^d$

②Guass 径向基核函数： $K(x_i, x) = \exp(-|x_i - x|^2 / 2\sigma^2)$

③Sigmoi 核函数： $K(x_i, x) = \tanh(kx_i^T x - \delta)$

其中， $d > 0$ ， $\sigma > 0$ ， $k > 0$ ， $\delta > 0$ 称为核参数。

Suykens 和 Vandewalle 最早提出用最小二乘法(LeaST Squares)将标准支持向量机进行优化，将不等式约束条件转化为等式约束条件。

相比标准支持向量机，最小二乘支持向量机将不等式约束转化为等式约束，将二次规划问题转化为一个线性方程组的求解，极大地降低了求解难度，提高了求解效率。

(4) 支持向量机用于财务困境预测的过程

将支持向量机用于财务预警的过程，就是将上市公司财务困境预测问题看作分类问题。事实上，把的 ST 企业与非 ST 企业的各项警兆指标联系起来，就把财务困境

预测问题归结为一个两分类的分类问题。

根据上市公司财务状况,可将其分为非 ST 公司和 ST 公司,分别用 $B1$, $B2$ 表示。样本集记为 $D = \{(x_i, y_i) | i = 1, 2, \dots, N\}$ 。其中, $x_i \in R^n$, 是一个 n 维的属性向量, 用来描述第 i 个上市公司财务状况, $y_i \in (0, 1)$ 是上市公司类别, 当 $x_i \in B1$ 时代表上市公司财务状况良好, 即属于非 ST 公司, 这时记 $y_i = 0$; 当 $x_i \in B2$ 时代表上市公司出现财务困境, 即属于 ST 公司, 记 $y_i = 1$ 。现在的问题是如何根据样本集进行学习得到识别上市公司是否为 ST 公司的原则, 以对公司财务困境进行预测。也就是, 我们要根据样本集找出一个决策函数, 要找出一个决策函数, 就要从样本数据中发现一些规律或知识, 这常常要考虑数据中的相似性。从而根据这个决策函数, 推断某一新的输入 x 相对应的 y , 即推断它属于 ST 公司还是非 ST 公司。解决这个问题一个直观想法就是, 看一看新的输入 x 是与那些 ST 公司的输入相似还是与那些非 ST 公司的输入相似, 由此推断 x 的归属。如果新的输入 x 与那些 ST 公司的输入相似, 它的输出值 y 就应该是 1; 反之, 其输出值就应该是 0。这样做是合乎逻辑的, 因为相似的输入应该导致相同的输出。

2.2 基于组合预测技术的财务困境预测方法

2.2.1 线性组合预测模型

组合预测模型是指将若干种预测方法赋予不同的权重, 从而形成综合的组合预测模型。单一预测方法所用到的信息是有限的, 对同一问题只采用一种方法进行预测, 其预测精度往往不高, 预测风险较大。组合预测本质上是将各种单项预测看作代表不同信息的片段(pieces of information), 通过信息的集成分散单个预测特有的不确定性和减少总体不确定性, 从而提高预测精度。由于不同的预测方法所用到的信息是不相同的, 因此, 可以将各种单一的预测结果进行组合得到一种组合预测结果, 以达到改善预测效果的目的。目前, 根据组合预测模型的方式不同, 组合预测一般可分为线性组合预测法与非线性组合预测法。

组合预测模型的核心问题即如何计算出个预测方法的权重, 进而使得各预测方法更加有效地结合, 提高预测的准确性。

本文选择将 MDA(多元判别分析)模型、Logit 回归模型和最小二乘支持向量机模

型进行线性组合来构建混合组合模型。基本思路是:将MDA(多元判别分析)模型、Logit回归模型与最小二乘支持向量机模型的预测结果作为组合预测模型的信息片段,通过线性组合可以综合三个模型所获得的财务类型预测信息,最终提升模型的预测能力以及预测性能的稳健性。

在模型的混合组合形式确定为线性组合以后,面临的核心问题就是求出三个模型之间的权重。针对这一问题,本文选取最小方差法来求解这三个单项模型的权重系数。其具体原理如下:

假设 z 为上市公司的实际财务类型,即 1 为财务困境类型,0 为财务健康类型; Z 为组合模型所预测的上市公司财务类型; Z_1 为多元判别分析模型预测的上市公司财务类型, Z_2 为 logit 回归模型预测的上市公司财务类型, Z_3 为支持向量机模型预测的上市公司财务类型。则组合预测模型表示为:

$$Z=v_1*Z_1+v_2*Z_2+ v_3*Z_3 \quad (2.51)$$

其中, v_1 为多元判别分析模型在混合组合预测模型中的权重, v_2 为多元判别分析模型在混合组合预测模型中的权重, v_3 为 Logit 回归模型在混合组合预测模型中的权重,且 $v_1 + v_2 + v_3 = 1$ 。

2.2.2 神经网络组合预测模型

在非线性组合函数上的选择上,运用 BP 神经网络构建了上市公司财务预警的组合预测模型。BP (Back Propagation) 网络是 1986 年由 Rumelhart 和 McClland 为首的科学家小组提出,是一种按误差逆传播算法训练的多层前馈网络,是目前应用最广泛的神经网络模型之一,它的学习规则是使用最速下降法,通过反向传播来不断调整网络的权值和阈值,使网络的误差平方和最小。

BP 神经网络模型拓扑结构包括输入层 (input)、隐层 (hide layer) 和输出层 (output layer) (如图 2-2 所示)。

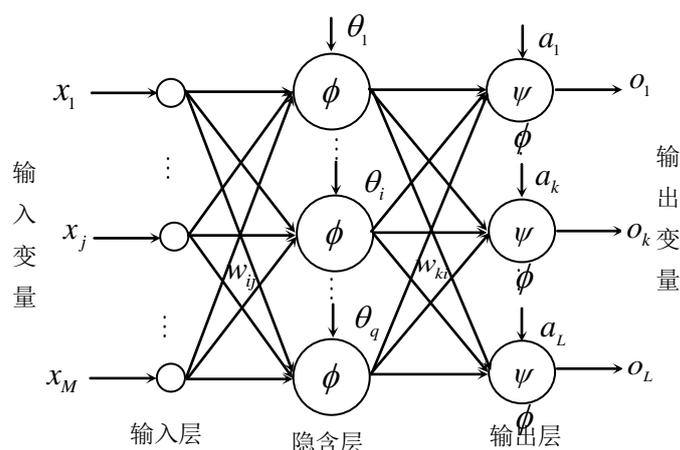


图 2-2 P 神经网络模型拓扑结构示意图

基本 BP 算法包括两个方面：信号的前向传播和误差的反向传播。即计算实际输出时按从输入到输出的方向进行，而权值和阈值的修正从输出到输入的方向进行。误差的反向传播，即首先由输出层开始逐层计算各层神经元的输出误差，然后根据误差梯度下降法来调节各层的权值和阈值，使修改后的网络的最终输出能接近期望值。

BP 网络的学习过程分为两个阶段：

第一个阶段是输入已知学习样本，通过设置的网络结构和前一次迭代的权值和阈值，从网络第一层向后计算各神经元的输出。

第二个阶段是对权值和阈值进行修改，从最后一层向前计算各权值和阈值对总误差的影响（梯度），据此对各权值和阈值进行修改。

以上两个过程反复交替，知道达到收敛位置。由于误差逐层往回传递，以修正层与层间的权值和阈值，所以称该算法为误差反向传播算法，这种误差反传学习算法可以推广到有若干个中间层的多层网络，因此该多层网络称之为 BP 网络。BP 算法程序流程如下图（2-3）

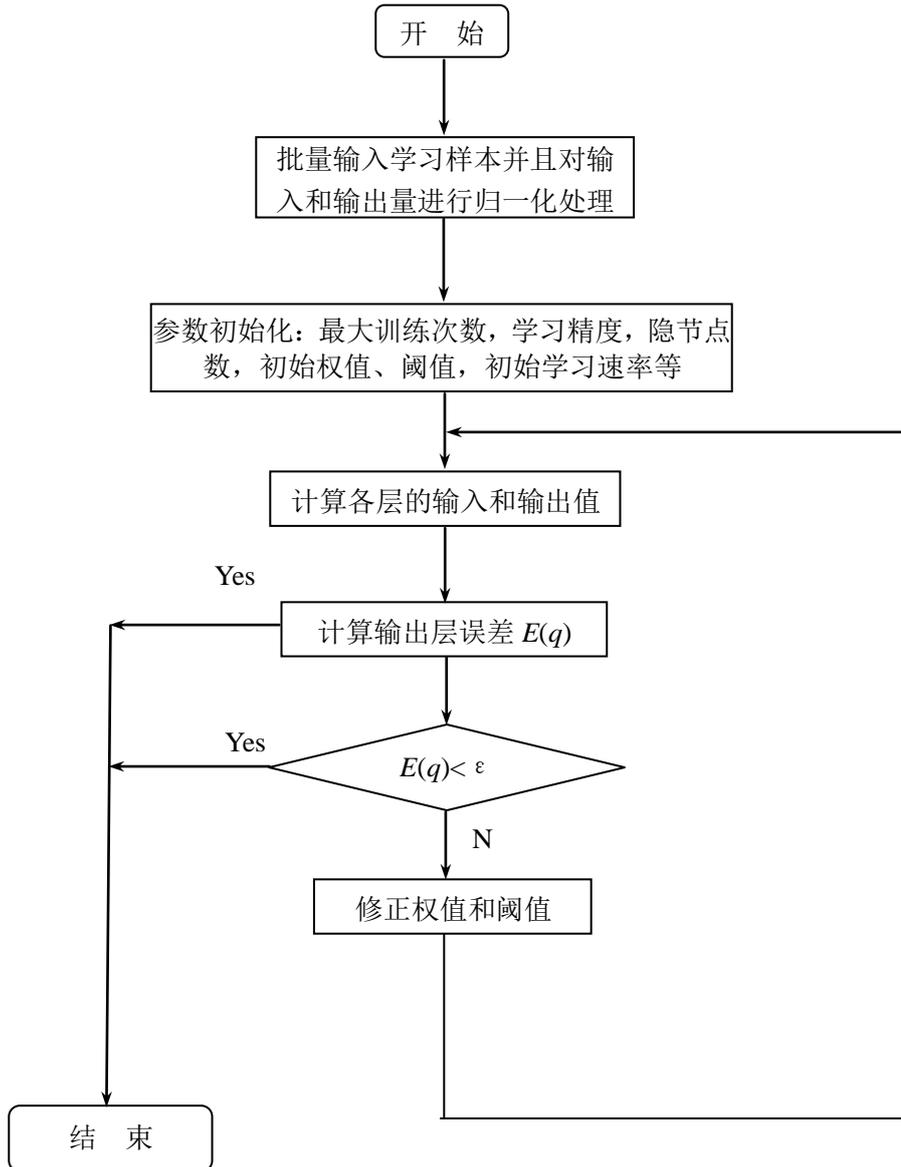


图 2-3 BP 算法程序流程图

运用 BP 神经网络建立非线性组合预测模型。输入层以多元判别分析模型的 F 值、Logit 回归模型的预测概率 P 值和支持向量机的判别结果组成的向量，作为组合预测模型的输入节点，即确定了 3 个输入节点。网络的输出层定义为一个节点，即上市公司的实际财务状况。在训练样本集中，本文定义：当为 ST 公司时，样本的输出值为 1；当为非 ST 公司时，样本的输出值为 0。BP 神经网络训练完之后，非线性函数就存储在由权值和传递函数构成的网络结构中，然后利用测试样本进行检验。

小结

本章论述了构建财务预警模型中需要需要用的方法以及它们的基本原理，包括因子分析法、多元判别分析、逻辑回归、最小二乘支持向量机等。另外特别选择了两种组合方法与单项模型进行对比，一种是线性组合模型，第二是基于神经网络的组合模型，并说明了组合的原理。

第 3 章 样本的选取和预测指标的选取及预处理

本章首先选取了 ST 公司及与之配对的正常公司最为样本；然后根据影响财务困境的因素选择财务指标，并对财务指标进行预处理。

3.1 样本选取

本文以选取的国泰安数据库的 146 家上市公司为样本，对公司的财务困境进行预测。根据我国的具体情况，以及借鉴以前文献，本文将我国上市公司中的 ST 公司确定为财务困境公司，其配对公司为非 ST 公司。配对公司的选择原则为：与所配对的 ST 公司所处的行业相同、与所配对的 ST 公司上市时间相同及资产规模相差 10% 以下。

因为中国证监会对 ST 公司进行特别处理根据上市公司(t-1)年的财务状况来决定的，如果能够获得某上市公司(t-1)年的财务报告，那么就可以确定该公司是否会因“财务状况异常”而被特别处理。所以本文选择 t-2 年的数据来预测公司的财务困境状况。

本文从沪深两市中挑选 2010 和 2011 年所有 ST 公司，剔除资料数据不全的，剩余 49 家，同时找到与其配对的 49 家非 ST 公司，作为训练样本，同时选取了 2012 年被 ST 的企业共 24 家，同时找到与其配对的 24 家非 ST 公司，作为测试样本。

3.2 预测指标体系

3.2.1 预测指标选取原则

财务困境是由于种种原因导致的企业财务状况持续恶化，财务风险加剧，出现不能清偿到期债务的信用困境，直至最终破产的一系列事件的总称。财务困境主要表现在：（1）企业总资产的账面净值小于账面记录的负债金额，即企业净资产为负值，就是所谓的“资不抵债”，表明企业已经事实上破产。（2）企业的市场销售持续萎缩，销售收入下降、产品价格下跌、主营业务收入连续负增长，存在数额巨大的未弥补亏损。

（3）企业的现金流量出现红字，企业现金流入小于现金流出，企业已经不能靠经营活动的资金维持，需要靠不断举债和其他渠道补充现金。（4）企业的应付帐款不断增加，企业经常拖欠应付款项，企业信用丧失，。应收帐款增加、存货增加都可能造成流动资金减少。

除了以上导致财务困境的财务因素外，ST 公司与正常公司在一些非财务指标上

也存在较大差异。本文尽可能建立一个完整的指标体系，寻找更多在 ST 公司与正常公司之间有显著差异的指标，以便进行实证分析。在指标的选取上主要考虑以下原则：

(1)可取得性。选取的指标数据应该容易获得，能够根据现有所公开的数据直接或间接获得。

(2)全面性。为了使财务困境预测模型更有效，在初步选取指标时，应尽量选择能够全面反映公司财务状况和经营成果的指标。

(3)相关性。所选取的指标应与财务困境的发生密切相关，能够对财务困境的发生作出必要合理的理论解释。

(4)可比性。所选择的指标应满足在同一公司不同时期具有可比性，不同公司在同一时期也应具可比性。

3.2.2 预测指标的选取

根据以上原则，考虑到我国上市公司的实际情况，本文从偿债能力、资本结构、营运能力、盈利能力、成长能力、获现能力等六个财务指标以及公司资产规模和审计意见两个非财务指标来选取财务困境的指标变量：

1. 偿债能力

企业的偿债能力和财务困境、破产风险密切相关。公司不能如期偿还到期债务是其陷入财务困境的征兆之一，因此偿债能力指标对于分析公司财务困境十分必要。本文所选取的偿债能力指标有：

流动比率=流动资产/流动负债

速动比率=(流动资产-存货)÷流动负债

现金流量债务比=经营活动现金净流量/负债合计

现金比率=现金及现金等价物期末余额/流动负债

2. 资本结构

资本结构是指企业各种资金的来源构成及其比例关系。本文选取了资产负债率，流动资产/总资产比率、流动负债/负债合计比率、产权比率等四个指标。

资产负债率=负债总额/资产总额

流动资产/总资产比率=流动资产/总资产

流动负债/负债合计比率=流动负债/负债合计

产权比率=负债总额/所有者权益总额

3. 营运能力

营运能力是指企业的经营运行能力，即企业运用各项资产以赚取利润的能力。反映营运能力的指标揭示了企业资金运营周转的情况，反映了企业对经济资源管理、运用的效率高低。企业资产周转过慢，势必造成公司陷入财务困境。本文选择应收账款周转率、存货周转率、总资产周转率来诠释营运能力。

应收账款周转率=营业收入/应收账款平均占用额

存货周转率=营业成本/存货平均占用额

总资产周转率=营业收入/平均资产总额

4. 盈利能力

盈利能力是指企业获取利润的能力，也称为企业的资金或资本增值能力，通常表现为一定时期内企业收益数额的多少及其水平的高低。经营成果的好坏也是公司能否存在和发展的基础。从投资人的角度来看，上市公司的盈利状况和盈利能力影响投资人的行为。本文选取的盈利能力指标为销售净利率、资产报酬率、净资产收益率。

销售净利率=净利润/营业收入

净资产收益率=净利润/平均股东权益

总资产报酬率=(利润总额+利息支出)/平均总资产

5. 成长能力

成长能力是企业生存的基础上，扩大规模、增加销售收入、提高公司利润的潜在能力。本文选择反映发展能力的指标为营业收入增长率、净利润增长率、总资产增长率。

营业收入增长率=(本年营业收入-上年营业收入)/上年营业收入

净利润增长率=(本年净利润-上年净利润)/上年净利润

总资产增长率=(期末总资产-期初总资产)/期初总资产

6. 获现能力

现金流在企业资产循环周转中扮演着重要的角色，贯穿于企业经营活动的全过程。从长期看，一个企业若要生存发展就必须源源不断地从经营活动中获取现金来满足其正常经营活动的各项开支，经营活动产生的净现金流才是企业偿还债务的基础和根本保障，销售现金比率指标反映每元销售得到的现金多少。

营业收入现金比率=经营活动产生现金流量净额/营业收入

除以上八个方面的财务指标外还有事务所对公司的资产规模和公司的审计意见

等非财务指标。因为审计师的审计意见对报表使用者有很大影响，而公司规模的大小也影响着公司应对困境的能力。在本文中审计意见分为两类，一种为标准无保留意见记为1，其余的审计意见记为0。而对于资产规模，由于数值较大在本文中都做lg处理。

综上所述，本文选取的财务困境预测指标如下表3-1

表3-1 上市公司财务困境预测指标体系汇总表

指标类型	序号	指标名称	计算公式
偿债能力	X1	流动比率	流动资产 ÷ 流动负债
	X2	速动比率	(流动资产 - 存货) ÷ 流动负债
	X3	现金流量债务比	经营活动现金净流量 ÷ 负债合计
	X4	现金比率	现金比率 = 现金及现金等价物期末余额 ÷ 流动负债
资本结构	X5	资产负债率	负债总额 ÷ 资产总额
	X6	流动资产/总资产	流动资产 / 总资产
	X7	流动负债/负债合计	流动负债 / 负债合计
	X8	产权比率	负债总额 / 所有者权益总额
营运能力	X9	应收账款周转率	主营业务收入 ÷ 平均应收账款
	X10	存货周转率	主营业务成本 ÷ 平均存货
	X11	总资产周转率	主营业务收入 ÷ 平均资产总额
盈利能力	X12	销售净利率	净利润 ÷ 营业收入
	X13	净资产收益率	净利润 ÷ 平均净资产
	X14	总资产报酬率	(利润总额 + 利息支出) ÷ 平均资产
成长能力	X15	营业收入增长率	(本年营业收入 - 上年营业收入) ÷ 上年营业收入
	X16	净利润增长率	(本年净利润 - 上年净利润) ÷ 上年净利润
	X17	总资产增长率	(期末总资产 - 期初总资产) ÷ 期初总资产
获现能力	X18	营业收入现金含量	经营活动现金净流量 / 营业收入
非财务指标	X19	资产规模	Log (资产规模)
	X20	审计意见	1: 标准无保留意见, 0: 非标准意见

3.3 预测指标的预处理

3.3.1 预警指标的显著性检验

通常认为，ST 公司和非 ST 公司之间应该具有显著差别，因此进入预测模型的指

标至少能显著地区别 ST 公司和非 ST 公司，这是入选指标的必要条件。显著性检验可以分为二类：一类是参数检验。一类是非参数检验。参数检验针对的是样本呈现正态分布时，可以用来进行显著性检验的统计方法。非参数检验针对样本分布不清楚，或者说样本不服从正态分布时，用来进行显著性检验的统计方法。

基于以上的分析，所以本文将预测指标的显著性检验安排如下：第一步，正态性检验。正态性检验为了对样本的分布进行判定，以此作为进行参数检验或是非参数检验的依据；第二步，独立样本的 T 检验。独立样本的 T 检验是参数检验，即对服从正态分布的预测指标样本进行独立样本 T 检验；第三步，独立样本的惠特尼检验。独立样本的惠特尼检验是非参数检验，即对不服从正态分布的预测指标样本进行独立样本的惠特尼检验。具体实施过程与结果如下：

(1) 样本分布的正态性检验

所以在初步选取指标之后，利用 SPSS16.0 统计分析软件提供的 K-S 检验，对 98 家样本的 20 个指标的样本分布进行检验。检验结果下表 3-2 所示：

表 3-2 预测指标正态性检验表

财务指标	正态参数		Z 统计量	双尾显著性 (p 值)	
	均值	标准差			
X1	1.3955	1.2268	2.1395	0.0002	
X2	1.0058	1.0597	2.4290	0.0000	
X3	0.1179	0.1972	1.7283	0.0051	
X4	0.1556	0.2432	1.6742	0.0074	
X5	55.8495	17.6202	0.8251	0.5038	*
X6	49.0478	18.5143	0.5683	0.9033	*
X7	82.9971	20.8128	2.0490	0.0005	
X8	225.5693	322.9040	2.7959	0.0000	
X9	1041.8646	10005.6743	5.0650	0.0000	
X10	4.9538	4.3442	1.4355	0.0324	
X11	0.6341	0.3361	0.7181	0.6810	*
X12	-48.1502	398.5940	4.3665	0.0000	
X13	-9.9662	25.4755	1.2479	0.0888	*
X14	-0.3867	9.1545	0.8787	0.4228	*
X15	1.6441	36.3309	0.8188	0.5138	*
X16	-640.1962	1389.3669	2.7861	0.0000	
X17	11.3182	55.9996	2.4054	0.0000	
X18	114.5864	88.8080	3.5764	0.0000	
X19	9.3322	0.5077	0.7359	0.6509	*
X20	0.8500	0.3620	5.0560	0.0000	

由表 3-2 可知, 在 0.05 的显著性水平下, X5(资产负债率)、X6(流动资产/总资产)、X11(总资产周转率)、X13(净资产收益率)、X14(资产报酬率)、X15(营业收入增长率)、X19(资产规模)的 K 统计量的概率 P 值大于 0.05, 其余 13 个指标的样本总体不符合正态分布。

根据上述结果, 本文将采用独立样本 T 检验对样本总体符合正态分布的 X5(资产负债率)、X6(流动资产/总资产)、X11(总资产周转率)、X13(净资产收益率)、X14(资产报酬率)、X15(营业收入增长率)、X19(资产规模)进行显著性检验, 然后采用两独立样本非参数检验来对其余 13 个样本指标进行显著性检验。

(2) 两独立样本的 T 检验

根据上述结果, 对服从正态分布的 7 个预测指标变量采用两独立样本 T 检验。

表 3-3 独立样本 T 检验表

财务指标	方差	方差方程的 Levene 检验		均值方程的 T 检验	
		F 值	显著性	T 值	双尾显著性 (p 值)
X5	同	0.6579	0.4193	-3.1167	0.0024
	异			-3.1167	0.0024
X6	同	0.0031	0.9559	2.6795	0.0087
	异			2.6795	0.0087
X11	同	0.5513	0.4596	2.1387	0.0350
	异			2.1387	0.0350
X13	同	20.0242	0.0000	9.7317	0.0000
	异			9.7317	0.0000
X14	同	4.5005	0.0365	9.8684	0.0000
	异			9.8684	0.0000
X15	同	0.4173	0.5198	4.8908	0.0000
	异			4.8908	0.0000
X19	同	1.3123	0.2548	4.4611	0.0000
	异			4.4611	0.0000

由表 3-3 可知, 在 $\alpha=0.05$ 的显著性水平下, X13 (净资产收益率) 的 Levene 检验的显著性 $0.000 < 0.05$, X14 (资产报酬率) 的 Levene 检验的显著性 $0.0365 < 0.05$, 因此就这两个指标来说, 两个样本的方差不齐, 再观察 t 值的双尾显著性均小于 0.05, 因此 X13 (净资产收益率)、X14 (资产报酬率) 指标通过独立样本 T 检验, 也即即两个样本中的 X13 (净资产收益率)、X14 (资产报酬率) 指标有显著性差异。

X5 (资产负债率)、X6 (流动资产/总资产)、X11 (总资产周转率)、X15 (营业

收入增长率)、X19(资产规模)的 Levene 检验的显著性 >0.05 ,因此两个样本的方差齐性,再观察 t 值的双尾显著性均小于 0.05,因此 X5(资产负债率)、X6(流动资产/总资产)、X11(总资产周转率)、X15(营业收入增长率)、X19(资产规模)未能通过独立样本 T 检验,也就是指 ST 与非 ST 样本中的五个指标是有显著性差异的。

(3) 样本分布的非参数检验

利用 SPSS16.0 版本软件对不满足正态分布的 13 个财务指标进行独立样本的均值检验,来判断 ST 公司和非 ST 公司样本的 13 个财务指标的均值是否存在显著性差异,以此来选定上市公司财务困境预测模型的判别变量。

均值差异的检验方法有很多种,其中非参数检验是指在总体不服从正态分布且分布情况不同时,用来检验数据资料是否来自同一个总体假设的一类检验方法。两独立样本非参数检验的方法有很多,本文使用了非参数检验中的 Mann-Whitney 检验来判别两组数据之间是否存在显著性的差异。

表 3-4 财务指标非参数检验

变量	Mann-Whitney U 值	Wilcoxon W 值	Z 值	双尾显著性	
X1	596.5000	1821.5000	-4.2916	0.0000	
X3	596.0000	1821.0000	-4.2951	0.0000	
X4	588.0000	1813.0000	-4.3519	0.0000	
X7	977.5000	2202.5000	-1.5859	0.1128	*
X8	760.0000	1985.0000	-3.1298	0.0017	
X9	932.0000	2157.0000	-1.9078	0.0564	*
X10	1162.0000	2387.0000	-0.2736	0.7844	*
X12	154.0000	1379.0000	-7.4356	0.0000	
X16	151.0000	1376.0000	-7.4569	0.0000	
X17	659.0000	1884.0000	-3.8475	0.0001	
X18	1087.0000	2312.0000	-0.8064	0.4200	*
X20	882.0000	2107.0000	-3.6286	0.0003	

表 3-4 是财务困境发生前 2 年各项财务指标的 Mann-Whitney 检验的检验结果。综合考虑两种检验的结果,发现: X7(流动负债/负债合计)、X9(应收账款周转率)、X10(存货周转率)、X18(营业收入现金含量)这四个预警变量的 P 值大于显著性水平,因此没有通过显著性差异检验,所以两个样本中的这 4 项指标是没有显著差异的。

综合上述样本检验结果,本文最后确定的预警指标一共包括 16 个财务指标,如表 3-5 所示:

表 3-5 通过显著性检验的指标

序号	指标名称	序号	指标名称
----	------	----	------

X1	流动比率	X12	销售净利率
X2	速动比率	X13	净资产收益率
X3	现金流量债务比	X14	总资产报酬率
X4	现金流量比率	X15	营业收入增长率
X5	资产负债率	X16	净利润增长率
X6	流动资产/总资产	X17	总资产增长率
X8	产权比率	X19	资产规模
X11	总资产周转率	X20	审计意见

3.3.2 预警指标主成分的提取

根据前面中对配对样本公司 18 个财务变量及资产规模和审计意见的显著性分析, 得到 16 个有显著性的财务指标, 但如果将这 16 个指标全部纳入到预警模型中, 会存在多重共线性, 要消除指标间的共线性、简化这些指标但又保留这些自变量的信息, 本文引入了因子分析法来给指标“降维”。

因子分析法是一种将多个指标化为少数互相无关的综合指标的统计方法, 从数学角度来看, 是一种降维处理技术。由于在实际研究中会经常遇到多指标的问题, 而且这些指标大多数是相关性的。将原来较多的指标设法通过重新组合最后变成几个新的相互无关的综合指标来取代原指标, 并且根据需要从中选取几个较少的综合指标来最大限度地反映原来指标的信息, 其中综合指标就是原来变量的公共因子。

设 p 个原始变量 X_1, X_2, \dots, X_p , m 个因子变量 F_1, F_2, \dots, F_m 两者之间的关系可以用数学模型表示为

$$\begin{cases} x_1 = a_{11}F_1 + a_{12}F_2 + \dots + a_{1m}F_m + a_1\varepsilon_1 \\ x_2 = a_{21}F_1 + a_{22}F_2 + \dots + a_{2m}F_m + a_2\varepsilon_2 \\ \dots \\ x_p = a_{p1}F_1 + a_{p2}F_2 + \dots + a_{pm}F_m + a_p\varepsilon_p \end{cases} \quad (3-1)$$

写成矩阵形式即:

$$X = AF + a\varepsilon \quad (3-2)$$

其中, A 为因子载荷系数矩阵。因子分析关键就是求出因子载荷矩阵。

因子分析的基本步骤如下:

第一步, 判断待分析的变量是否适于采用因子分析。SPSS 软件中可以通过 KMO(Kaiser-Meyer-Olkin)检验进行判断是否适合作因子分析。KMO 统计量取值在 0~1 之间。KMO 值越接近 1, 意味着所有变量间的简单相关系数的平方和远大于偏

相关系数平方和，即变量间的相关性越强，越适合进行因子分析；反之，KMO 值越接近 0，变量间的相关性越弱，越适合进行因子分析。

一般认为， $KMO < 0.5$ 不合作因子分析。表 3-6 是 KMO 检验和 Bartlett 球形检验的结果，KMO 统计量的值为 0.64，适合作因子分析。

表 3-6 KMO 检验表

KMO 和 Bartlett 的检验		
Kaiser-Meyer-Olkin 检验		.640
Bartlett 单位阵检验	近似卡方	1.173E3
	自由度	120
	显著性	.000

第二步，要确定求因子解的方法和公因子个数。表 3-7 是本文通过正交旋转提取的 7 个主成分 F1-F7，对原始指标的信息涵盖率为 84.464%，较为理想。为了对原始变量作出实际意义的解释。为了得到结果更好的因子载荷阵，初始因子实行旋转。

表 3-7 方差贡献率表

Total Variance Explained									
主成分	初始特征值			提取平方和载入			旋转平方和载入		
	特征值	占总体%	累计 %	特征值	占总体%	累计 %	特征值	占总体%	累计 %
1	4.880	30.502	30.502	4.880	30.502	30.502	2.558	15.985	15.985
2	2.598	16.236	46.738	2.598	16.236	46.738	2.239	13.994	29.979
3	1.577	9.854	56.592	1.577	9.854	56.592	1.979	12.371	42.351
4	1.372	8.576	65.168	1.372	8.576	65.168	1.925	12.033	54.383
5	1.193	7.453	72.621	1.193	7.453	72.621	1.896	11.849	66.232
6	1.008	6.298	78.919	1.008	6.298	78.919	1.612	10.073	76.305
7	.887	5.545	84.464	.887	5.545	84.464	1.305	8.159	84.464
8	.697	4.355	88.819						
9	.551	3.442	92.261						
10	.414	2.586	94.847						
11	.303	1.892	96.739						
12	.216	1.352	98.091						
13	.185	1.156	99.247						
14	.060	.372	99.619						
15	.044	.278	99.897						

16	.016	.103	100.000					
Extraction Method: Principal Component Analysis.								

第三部，进行因子旋转，对旋转后的新因子进行命名，作出实际意义的解释。

下表 3-8 是通过四次方最大法旋转后的因子载荷矩阵。

表 3-8 旋转后的因子载荷矩阵

变量	主成分						
	1	2	3	4	5	6	7
X1	.918	.237	.066	-.185	.051	.025	.044
X2	.896	.310	.049	-.123	.109	-.004	.014
X3	.197	.900	.173	-.188	-.078	.085	-.034
X4	.225	.924	.071	-.133	.046	.081	.052
X5	-.482	-.287	-.063	.711	.208	.004	.011
X6	.582	-.264	.441	.071	-.355	.318	.060
X8	-.036	-.094	.005	.960	-.055	-.024	-.008
X11	-.225	.182	.218	-.060	-.044	.648	.335
X12	.147	.048	-.049	.024	.030	.817	-.167
X13	.141	.228	.538	-.598	.304	.117	.279
X14	.202	.377	.679	-.051	.319	.141	.344
X15	.110	-.004	.278	-.131	.510	.584	.179
X16	.031	.082	.900	-.042	.161	.055	-.031
X17	.193	-.115	.131	.054	.858	-.066	-.015
X19	-.253	.111	.168	-.062	.680	.161	.230
X20	.059	-.013	.071	-.038	.138	.013	.935

从上表3-8可以看出：

①第一个公共因子 F1 在指标 X1, X2 上有较大的载荷，这些指标分别是流动比率，速动比率属于偿债能力指标，因此可将公因子 F1 命名为“偿债能力因子”。

②第二个公共因子 F2 在指标 X3, X4 上有较大的载荷，这些指标分别是代表现金比率，现金流量比率，可以将 F2 命名为“现金流量因子”。

③第三个公共因子 F3 在指标 X16 上有较大的载荷，这个指标分别是净利润增长率，因此可将公因子 F3 命名为“成长能力因子”。

④第四个公共因子 F4 在指标 X8 有较大的载荷，反映公司的产权比率，可将其命

名为“资本结构因子”。

⑤第五个公共因子 F5 在指标 X17 和 X19 上有较大的载荷，反映公司的资产规模，可将公因子 F5 命名为“资产规模因子”。

⑥第六个公共因子 F6 在指标 X12 上有最大的载荷，代表了销售净利率，因此可将其称为“盈利能力因子”。

⑦第七个公共公因子 F7 在指标 X20 即审计意见上有较大的载荷，是反映公司审计意见的指标，因此可将其命名为“审计意见因子”。

第四步，计算因子得分。因子得分即每一样本数据在不同因子上的具体数据值。

表 3-9 因子得分系数矩阵

因子得分系数矩阵							
变量	公共因子						
	1	2	3	4	5	6	7
X1	.404	-.018	-.113	.034	.063	-.020	.045
X2	.395	.043	-.126	.085	.109	-.041	.013
X3	-.069	.467	.032	.073	-.076	-.003	-.086
X4	-.027	.495	-.091	.121	.013	.001	.003
X5	-.101	.028	.028	.370	.128	.022	.022
X6	.257	-.271	.294	.085	-.297	.155	.040
X8	.102	.133	.083	.627	-.023	-.019	.070
X11	-.165	.077	-.004	-.002	-.158	.402	.216
X12	.058	-.010	-.186	.023	.043	.614	-.213
X13	-.064	-.052	.216	-.293	.045	-.054	.059
X14	-.003	.121	.314	.124	.027	-.080	.127
X15	.037	-.095	-.048	-.046	.239	.345	-.019
X16	-.102	-.040	.656	.038	-.063	-.137	-.248
X17	.152	-.096	-.056	.056	.551	-.093	-.163
X19	-.119	.060	-.058	-.016	.353	.049	.044
X20	.071	-.050	-.173	.050	-.075	-.082	.864

以 F1 为例， $F1=0.404*X1+0.395*X2-0.069*X3-0.027*X4-0.101*X5+0.257*X6+0.102*X8-0.165*X11+0.058*X12-.064*X13+0.003*X14+0.037*X15-0.102*X16+0.152*X17-0.119*X19-0.071*X20$ ，值得注意的是因子得分表达式中的各变量 X1—X20 应该是经过标准化变换后的标准变量。

小 结

综合我国学者的研究，本文以上市公司被 ST 确定为财务困境，并根据影响上市公司财务困境的因素，从偿债能力、盈利能力、成长能力、资本结构、获现能力及资产规模和审计意见等方面提取了 20 个财务指标，并用因子分析提取了 7 个公共因子，作为模型的输入变量。

第 4 章 上市公司财务困境预测模型的构建和检验

本章首先构建了多元判别分析(MDA)、多元逻辑回归模型(Logit)及支持向量机(SVM)等三种单项预测模型,然后构建了一个线性组合预测模型和一个基于神经网络学习的组合预测模型,然后对预测结果进行比较。

4.1 单模型预测结果分析

4.1.1 判别分析模型

我们把正常的上市公司定义为组合 0,财务困境公司定义为组合 1。本章采用 SPSS16.0 软件自带的 Discriminate 判别分析模块来建立财务困境预警模型。根据上述选定的 7 个建模因子,使用同时进入法进行判别,来构建 Fisher 多元判别模型。为了避免重要信息被筛选出方程,本文采用变量全部进入方程法,进行回归分析,步骤如下:

第一步:将训练样本数据(这里是 98×7 的矩阵)输入统计软件,计算 Wilks' Lambda 统计量和 Eigenvalues 统计量,判断判别模型是否整体性显著。

从表 4-1 中得出 Fisher 判别函数的特征值为 1.205,解释了 100%的方差,y 与判别函数的相关程度为 0.739,检验结果表明,模型的自变量与因变量的相关性强。

表 4-1 Eigenvalues 检验

Eigenvalues				
函数	特征值	占总体%	累计 %	典型相关
1	1.205 ^a	100.0	100.0	.739

a. First 1 canonical discriminant functions were used in the analysis.

如表 4-2 所示, Wilks' Lambda 的统计量为 0.454,卡方分布统计量为 73.143,判别函数的相伴概率为 0.000,小于显著性水平 0.05,所以拒绝各组判别函数的均值相等的零假设,表明上述判别函数达到了显著性水平,认为判别函数能较好地地区分这两类。

表 4-2 Wilks' Lambda 统计量

Wilks' Lambda				
方程检验	Wilks' Lambda	卡方统计量	自由度	显著性
1	.454	73.143	7	.000

第二步：计算模型的判别系数，根据判别系数，将训练样本数据（这里是 98×7 的矩阵）代入(5-10)式的判别模型，即可求出这 98 个样本的 Z 值，根据 Z 值的判别阈值，判断各样本是否为 ST，并求出模型的误判率。

表 4-3 显示的是未标准化的判别系数。本文财务困境预测采用非标准化系数的判别模型，因为这样我们做判别分析时可以直接将财务指标因子代入公式，不必将其标准化，这样使用起来更方便。

表 4-3 未标准化的判别系数表

Canonical Discriminant Function Coefficients	
F1	.300
F2	.469
F3	.974
F4	-.499
F5	.348
F6	.348
F7	.658
(Constant)	.000

Unstandardized coefficients

所以，可以得到线性判别模型：

$$Z=0.3 \cdot F1+0.469 \cdot F2+0.974 \cdot F3-0.499 \cdot F4+0.348 \cdot F5+0.348 \cdot F6+0.658 \cdot F7 \quad (4-1)$$

下表 4-4 给出了两组中心点的位置，显示了根据新建立的模型计算的财务困境组和财务健康组 Z 值的分布情况。通过计算得出财务困境组的重心为 1.086，财务健康组的重心为 -1.086。我们取两者的中间位置作为财务困境与财务健康组的判别标准，即 $(-1.086+1.086)/2=0$ 。根据判别模型 (4-1)，计算 Z 值，当它大于 0 时企业被划分为财务健康组，企业未来陷入财务困境的机率比较小；当它小于 0 时企业被划分为财务困境组，企业未来陷入财务困境的机率比较大。

表 4-4 两组重心位置表

Functions at Group Centroids	
0	1.086
1	-1.086

Unstandardized canonical discriminant functions evaluated at group means

原始样本的回判检验，是将样本数据代入新模型，计算出相应的判别分数-Z 值，

然后根据设定的判别准则，判断其属于哪一组别，最后将判别结果与实际结果相比较。结果如表 4-5 所示：

表 4-5 判别分析回代结果及预测

	回代结果			预测结果		
	样本个数	误判数	预测正确率	样本个数	误判数	预测正确率
正常公司	49	4	91.8%	24	6	75%
ST 公司	49	10	79.6%	24	6	75%
总体正确率			85.7%			75%

通过表 4-5 可知：

(1)回代结果：对于正常公司，49 个训练样本中有 4 个误判，多元判别模型的识别率为 91.8%；对于 ST 公司，49 个训练样本中仅有 10 个误判，预测正确率为 79.6%，得到回代结果的总体正确率为 85.7%。

(2)预测结果：24 个正常公司有 6 个误判，多元判别模型的识别率为 75%；ST 公司的测试样本中有 6 个误判误判，预测正确率为 75%；预测样本的总体正确率为 75%。

4.1.2 逻辑回归模型

使用 Spss16.0 统计软件建立 Logit 回归模型，为了避免重要信息被筛选出方程，本文采用变量全部进入方程法，进行回归分析，步骤如下：

第一步：将训练样本数据（这里是 98×7 的矩阵）输入统计软件，判断回归模型是否整体性显著。模型的拟合度是判断模型优劣的方法之一，它是判别模型和样本的拟合程度。通过已有参数，得到的观测结果可能性被叫做“似然比”。一般用对数的似然比值乘以-2 来确定模型和数据的拟合程度，记为“-2LL”，-2LL 值越小，模型拟合越好。在评价变量和模型的拟合效果时通常会采用 Cox&Snell R Square 和 Nagelkerke R Square 这两个统计量，再判别过程中，如果两个统计量的值越接近 1，则说明模型对变量的拟合程度就越高。

表 4-6 给出的是 Logit 回归模型整体显著性检验结果。从表中可以看出：表中的 -2LL 值为 0，此值很小，说明模型对数据的拟合度很理想。另外，Cox&Snell R² 的值为 0.60，Nagelkerke R² 的值为 0.8，也说明模型的拟合度很好，该模型对变量和数据的拟合程度都处于可以接受的范围内。

表 4-6 Logit 回归模型检验

Model Summary			
step	-2 Log likelihood统计量	Cox & Snell R2统计量	Nagelkerke R2统计量
1	46.020 ^a	.600	.800

第二步：计算模型的回归系数，由表 4-7 可得本文回归模型如下所示：表 4-7 从左至右分别为公共因子（模型输入变量）、回归系数的估计值、估计值的标准差、回归系数显著性检验的 Wald 统计量、自由度、显著性水平等。wald 检验值是用来检验偏回归系数的显著程度，它是偏回归系数和自由度的函数且服从卡方分布的，Wald 检验值越大自变量的作用就越显著。从 Wald 统计量以及对应的 p 值来看，因子 F1、F2、F3、F4、F5、F6、F7 的系数均小于显著性水平 0.05，说明这 7 个因子都通过了显著性检验，它们对模型的影响较显著。

表 4-7 Logit 回归模型系数检验

公共因子	回归系数	标准差	Wald统计量	自由度	相应p值
F1	-.827	.489	2.860	1	.091
F2	-1.453	.519	7.833	1	.005
F3	-4.171	1.089	14.677	1	.000
F4	2.463	1.004	6.015	1	.014
F5	-1.072	.365	8.654	1	.003
F6	-.784	.286	7.514	1	.006
F7	-2.250	.655	11.805	1	.001
ConStant	1.096	.528	4.309	1	.038

根据上表，最终得到包含 7 个主成分变量的 Logit 预警模型，构建的 Logit 回归模型： $Y = -0.827 * F1 - 1.453 * F2 - 4.171 * F3 + 2.463 * F4 - 1.072 * F5 - 0.784 * F6 - 2.250 * F7 - 1.096$

$$P = e^y / (1 + e^y) \quad (4-2)$$

根据回归系数，将测试样本数据（这里是 48×7 的矩阵）代入(5-10)式的回归模型，即可求出这 48 个样本的概率 p，若 $p \geq 0.5$ 则判断该公司是 ST 公司；否则，若 $p < 0.5$ ，则判断该公司是正常公司，并据此求出模型的误判率。

将 7 个因子代入 Logit 回归模型，可得到每一个样本的概率值 p，再根据 p 值判断其所属的类别，将测试样本回代到上述模型，可得测试样本的准确率。表 4-8 给出的是根据上面的回归模型得出的预测结果：

表 4-8 Logit 回归回代结果及预测

	回代结果			预测结果		
	样本个数	误判数	预测正确率	样本个数	误判数	预测正确率
正常公司	49	6	87.8%	24	7	70.8%
ST 公司	49	5	89.8%	24	5	79.2%
总体正确率			88.8%			75%

通过表 4-8 可知：

(1)回代结果：对于正常公司，48 个训练样本中有 6 个误判，Logit 回归模型的识别率为 87.8%；对于 ST 公司，48 个训练样本中仅有 5 个误判，预测正确率为 89.8%，得到回代结果的总体正确率为 88.8%。

(2)预测结果：24 个正常公司有 7 个误判，Logit 回归模型的识别率为 70.8%；ST 公司的测试样本中有 5 个误判误判，预测正确率为 79.2%；预测样本的总体正确率为 75%。

4.1.3 支持向量机模型

本节通过 MATLAB 分别编程，求得 LS-SVM 模型的结果，支持向量机的关键是核函数和惩罚系数的确定，不同的惩罚系数和核参数可以得到不同的支持向量机模型，但是到目前为止还没有统一的理论指导核函数的选择。鉴于以前学者的研究都表明高斯核函数的预测正确率明显高于其他核函数，所以本文也采用高斯核函数

$K(x_i, x) = \exp(-\gamma \|x_i - x\|^2)$ ，其中参数 γ 的确定采用枚举法，并将模型分为回代结果和预测结果两种情况进行分析。“预测结果”是用测试样本代入程序运行得到的结果；“回代结果”是用训练样本替代测试样本代入程序运行得到的结果。

经过反复试验，当核参数 $\gamma=5$ ，惩罚因子 $C=4$ 时，LS-SVM 程序的运行结果最好。

表 4-9 LS-SVM 回代结果及预测

	回代结果			预测结果		
	样本个数	误判数	预测正确率	样本个数	误判数	预测正确率
正常公司	48	5	89.58%	24	5	79.17%
ST 公司	48	3	93.75%	24	2	91.67%
总体正确率			91.67%			85.41%

通过表 4-9 可知：

(1)回代结果：对于正常公司，48 个训练样本中有 5 个误判，LS-SVM 模型的识别

率为 89.58%；对于 ST 公司，48 个训练样本中仅有 3 个误判，预测正确率为 93.75%，得到回代结果的总体正确率为 91.67%。

(2)预测结果：24 个正常公司有 5 个误判，LS-SVM 模型的识别率为 79.17%；ST 公司的测试样本中有 2 个误判，预测正确率为 91.67%；预测样本的总体正确率为 85.41%。

由于专业知识和自身能力的限制，核参数 γ 和惩罚因子 C 的确定是通过人工多次试算的方式得到的，最终确定的参数值可能并不是理论上的最优参数值，因而以此建立模型所得到的预测结果可能不是最佳预测结果。但是尽管如此，实证结果也验证了 LS-SVM 模型无是优于多元判别分析和 Logit 回归模型的。

4.2 组合模型预测结果分析

4.2.1 线性组合模型

假设 z 为上市公司的实际财务类型，即 1 为财务困境类型，0 为财务健康类型； Z 为组合模型所预测的上市公司财务类型； Z_1 为多元判别分析模型预测的上市公司财务类型， Z_2 为 logit 回归模型预测的上市公司财务类型， Z_3 为支持向量机模型预测的上市公司财务类型。则组合预测模型表示为：

$$Z = v_1 * Z_1 + v_2 * Z_2 + v_3 * Z_3 \quad (4-3)$$

其中， v_1 为多元判别分析模型在混合组合预测模型中的权重， v_2 为多元判别分析模型在混合组合预测模型中的权重， v_3 为 Logit 回归模型在混合组合预测模型中的权重，且 $v_1 + v_2 + v_3 = 1$ 。

在模型的混合组合形式确定为线性组合以后，面临的核心问题就是求出三个模型之间的权重。针对这一问题，本文选取最小方差法来确定多元判别分析、logit 回归模型和支持向量机模型的权重。

在确定组合预测模型的权重向量 $w = (w_1, w_2, w_3)$ 的过程中，考虑到最后预测的值应该和实际采集到的数据的残差 e 的平方和尽可能的小，为此，建立如下加权残差平方和最小的组合预测模型，通过极值的原理可以计算出确定组合权重。

本文首先求得多元判别分析的误差平方和为 0.3799，Logit 模型的误差平方和为 0.3368，支持向量机模型的误差平方和为 0.2087，，因此，考虑权重非负的线性组合预测模型的权重求解模型如下：

$$\text{Min } e_2 = (e_1 * W_1 + e_2 * W_2 + e_3 * W_3) \quad (4-5)$$

本文利用 Matlab 软件来求解，通过最小二乘法模型，可得到各指标的权重向量 $W = (0.3108, 0.3323, 0.3570)$

因此线性组合预测模型的表达式为：

$$Z = 0.3108 * Z_1 + 0.3323 * Z_2 + 0.3570 * Z_3 \quad (4-6)$$

其中： Z 为组合预测模型的预测值， Z_1 为多元判别分析模型的预测值， Z_2 为 Logit 回归模型的预测值， Z_3 为支持向量机模型的预测值，得到输出结果如表 4-10。

表 4-10 线性组合回代结果与预测

	回代结果			预测结果		
	样本个数	误判数	预测正确率	样本个数	误判数	预测正确率
正常公司	48	7	85.42%	24	9	62.5%
ST 公司	48	1	95.92%	24	1	95.83%
总体正确率			91.84%			79.17%

通过表 4-11 可知：

(1)回代结果：对于正常公司，48 个训练样本中有 7 个误判，线性组合模型的识别率为 85.42%；对于 ST 公司，48 个训练样本中仅有 1 个误判，预测正确率为 95.92%，得到回代结果的总体正确率为 91.84%。

(2)预测结果：24 个正常公司有 9 个误判，线性组合模型的识别率为 62.5%； ST 公司的测试样本中有 1 个，误判预测正确率为 95.83%；预测样本的总体正确率为 79.17%。

4.2.2 基于 ANN 的组合预测模型

在设置好网络结构和参数之后，以上述建立的多元判别分析模型、Logit 回归模型和支持向量机作为单预测模型，运用 BP 神经网络建立非线性组合预测模型。输入层以多元判别分析模型的 F 值、Logit 回归模型的预测概率 P 值和支持向量机的判别结果组成的向量 $X = (Y_1, Y_2, Y_3)$ ，作为组合预测模型的输入节点，即确定了 3 个输入节点。网络的输出层定义为一个节点，即上市公司的实际财务状况 Y。在训练样本集中，本文定义：当为 ST 公司时，样本的输出值为 1；当为非 ST 公司时，样本的输出值为 0。利用 MATLAB 语言，建立了一个 2-38-1 前向三层 BP 网络，运用 98 个训练样本作为训练集进行训练模拟，对应目标输出结果见（表 2）。BP 神经网络训练完之后，非线性函数就存储在由权值和传递函数构成的网络结构中，然后利用 48

个测试样本进行检验，把48家测试样本的多元判别分析模型的F值、Logit回归模型的预测概率P值和支持向量机的判别结果输入到组合预测模型，得到输出结果如表4-11。

表4-11 基于神经网络的组合结果

	回代结果			预测结果		
	样本个数	误判数	预测正确率	样本个数	误判数	预测正确率
正常公司	48	4	91.67%	24	3	87.5%
ST公司	48	3	93.75%	24	2	91.67%
总体正确率			92.86%			89.58%

通过表4-11可知：

(1)回代结果：对于正常公司，48个训练样本中有4个误判，神经网络组合模型的识别率为91.67%；对于ST公司，48个训练样本中仅有3个误判，预测正确率为93.75%，得到回代结果的总体正确率为92.86%。

(2)预测结果：24个正常公司有9个误判，神经网络组合模型的识别率为87.5%；ST公司的测试样本中有1个，误判预测正确率为91.67%；预测样本的总体正确率为89.58%。

4.3 财务预警模型判别结果比较分析

通过五个预测模型对我国146家上市公司进行实证分析，得出了不同的预测结果，现将各结果汇总，如下表4-12所示，从表中我们得到以下结论：

表4-12 各模型判定结果比较分析表

模型	训练样本			测试样本			正确率差异
	正常公司	ST公司	预测正确率	正常公司	ST公司	预测正确率	
判别分析模型	91.80%	79.60%	85.70%	75%	75%	75%	10.70%
逻辑回归模型	87.80%	89.80%	88.80%	70.80%	79.20%	75%	13.80%
支持向量机模型	89.58%	93.75%	91.67%	79.17%	91.67%	85.41%	6.26%
线性组合模型	85.42%	95.92%	91.84%	62.50%	95.83%	79.17%	12.67%
神经网络组合模型	91.67%	93.75%	92.86%	87.5%	91.67%	89.58%	3.28%

(1) 模型预测准确率率的比较：从上表可以看出，五个模型中神经网络组合预测模型的预测准确率最高，支持向量机模型和线性组合预测模型次之，随后是逻辑回归模型，多元判别分析排在最后。它们的预测正确率都在 75% 以上，说明它们都具有较好的预测效果。

本文所建立的线性组合预测模型的预测率高于判别分析模型和逻辑回归模型，但仍然达不到支持向量机模型的判别效果，可见把三个单项模型的结果进行线性组合预测率不会很大提高，大概取了中间值。但是，我们也可以从数据的对比中发现，线性组合预测模型在判断第二类（将 ST 公司判为正常公司）时的正确率是最高的。在实际应用中，第二类预测结果对于信贷机构而言具有更重要的意义，第二类预测结果的精确性下降一些，都会给信贷机构带来巨大的损失。

基于神经网络的非线性组合预测模型的预测正确率为 89.58%，预测正确率高于任何一种单一模型，但与支持向量机预测模型相比，变化不大

(2) 模型稳健性的比较：由于检验样本只有一组，因此，模型稳健性的判断标准以该模型在测试样本与训练样本上的分类正确率的差异作为标准。从表中的结果可以看出，预测正确率变化最小的为神经网络组合模型合预测模型，只有 3.28%，其次为支持向量机预测模型的 6.26%。因此，从模型稳健性的角度，神经网络体现出了组合模型的优势。

(3) 组合预测模型之间的比较

将两种组合模型进行对比，分别从模型的预测正确率及稳健性两方面比较其应用效果。模型预测正确率的比较，基于神经网络的非线性组合预测模型的正确率高于线性组合预测模型，但是线性组合在判断第二类（将 ST 公司判为正常公司）时的正确率比神经网络模型高。模型稳健性的比较，基于神经网络的非线性组合预测模型的预测正确率变化为 3.28%，低于线性组合预测模型。

本章小结

本章对三种单项预测模型、线性组合预测模型和基于神经网络学习的组合预测模型的预测结果进行比较，发现：单线模型中支持向量机模型的预测效果最高；组合预测模型对提高预测正确率上并不显著，但稳定性得到了提高；与线性组合预测模型相比，无论在预测正确率还是稳定性，神经网络组合都具有显著优势。

第五章 研究结论及不足

5.1 研究结论

文章通过如下五个主要步骤：

(1)本文选取了 73 家为 ST 公司和配对的 73 家为非 ST 公司为样本，对样本的财务数据预处理，分别构建了多元判别分析(MDA)、多元逻辑回归模型(Logit)及支持向量机(SVM)等三种单项预测模型，并构建了一个线性组合预测模型和一个基于神经网络学习的组合预测模型，对五个模型的结果比较分析，得出以下研究结论：

本文认为神经网络的非线性组合预测模型优于多元判别分析与逻辑回归两种统计模型，但与支持向量机模型相比，预测正确率上没有显著优势，但是稳定性提高了；与线性组合预测模型相比，无论在预测正确率还是稳定性，神经网络组合都具有显著优势。

组合预测的效果没有达到预期目标，本文认为原因是：组合模型依靠三种单项预测模型的判别结果来判断，而单项预测模型的判别结果大部分相似，经过组合很难提高判断正确率。

5.2 研究不足及未来研究方向

本文由于时间等条件的限制，也具有一定的局限性，具体表现如下：

(1)本文仅选取上市公司作为研究样本。文章基于对数据可获得性的考虑只选了上市公司，但是在我国，非上市公司所占比重较高。

(2) 上市公司的财务状况出现问题是一个连续的过程，文章只是简单的把财务状况划分为了 ST 和非 ST。企业的实际财务状况可能有三种、四种甚至五种，如何利用模型建立多状态的财务困境预测模型，对投资者、债权人和监管机构更有实际意义，这也是未来研究的一个方向。

(3) 文章在构建财务困境预测指标体系时，只是选用了直接影响财务状况的指标，影响一个上市公司财务状况还有很多因素，比如：宏观经济政策，公司治理，竞争对手情况，国际经济形势等等。由于数据搜集的困难，未将非财务指标引入模型，如把公司股权变更、企业治理结构、管理层的素质等引入，可取得模型的进一步完善。

(4)本文所选数据的有可能存在造假信息。我国上市公司监督机制尚不完善，存在

会计信息失真情况，而这种失真对于模型的准确预测是非常不利的，可能导致研究结论失真。

(5)本文研究中只选用了多元判别分析、Logit 回归模型和支持向量机三个单项模型。对于组合预测模型，不同组合预测方法对预测模型的精度有着不同的影响。在以后的研究中，可以选择不同的组合方法来看预测结果，权重可以选择其他的方法来确定。

本文在对财务困境预测模型的研究过程中，虽然取得了一些有价值的研究成果，但是应用任何一种方法进行上市公司财务困境预测都存在其不足和局限性。由于作者能力的局限，在研究财务困境问题中仍存在一些有待解决的问题。

参考文献

- [1]彭静.网络环境中企业财务困境预警研究[D].上海:上海交通大学,2008
- [2] Beaver WH. Financial ratios as predictors of failure[J]. Journal of Accounting research, 1966
- [3]Altman Edward I. Financial ratios,discriminant analysis and the prediction bankruptcy[J].Journal of Finance,1968, 23:589-609
- [4]Carmichael. Geneteal gorithms applieations in the analysis of insolveney risk[J]. Journal of Banking and Finanee,1998,22:1421-1439
- [5] Morris. Methodological Issues Related to the ESTimation of Financial DiSTress Prediction Models, Journal of Accounting Research,1984 Supplement, Vol. 22, 59-82
- [6] Ross S.A., WeSTerfield R.W., Jaffe J,Corporate Finance, New York,McGraw-Hill Companies, 1999
- [7]谷棋,刘淑莲. 财务危机企业投资行为分析与对策,会计研究,1999 (10): **28-31**
- [8] Friz patrick. A Comparison of ratios of Successful InduSTrial Enterprises with those of Failed Firms[M]. Iview York: Certified Public Accountant, 1972
- [9]Beaver,William H, Financial Ratios as Predictors of Failure,Journal of Accounting Research,1966 Supplement,Vol. 4, 71-111
- [10]Altman E. I. Financial Ratios Discriminate Analysis and Prediction of Corporate Bankruptcy,Journal of Finance, 1968(4),589-679
- [11]Altman Edward I. Prediction Financial DiSTress of Companies, Revisiting the Z Score and Zeta Models. Journal of Finance, 2000,59, 69-108.
- [12] Meyer,Pifer, Howard W,Prediction of Bank Failures. Journal of Finance, 1970, Vol. 25, 853-868
- [13] Ohlson J.Financial Ratio and the ProbabiliSTic Prediction of Bankruptcy[J].Journalof Accounting Research, 1980,18:109-131
- [14] Odom M, Sharda R.A. Neural Networks Model for Bankruptcy Prediction[J]. Proceedings of the IEEE International Conference on Neural Network,1990,2:163-168
- [15]Coats Pamela K,Franklin FL. Recognizing Financial DiSTress Patterns Using aNeural Network tool[J].Financial Management,1993,3:142-166
- [16]Kar Yan Tam,Kiang, Melody Y,Managerial Applications of Neural Networks: the Case of Bank Failure Predictions, Management Science,1992,Vol.38,926-947
- [17]Wu, Desheng(Dash),Liang, Liang and Yang, Zijiang, Analyzing the financial diSTress of Chinese public companies using probabiliSTic neural networks and multivariate discriminate analysis , Socio-Economic Planning Sciences, 2008,Vol. 42,206-220
- [18] MinandLee,Support Vector Machine for Classification of Voltage DiSTurbances, IEEE Transactions on Power Delivery,2000, 22, 1297-1303
- [19] Fan A, Palaniswami M,Selecting bankruptcy predictors using a support vector machine approach, Proceeding of the IEEE-INNS-ENNS international joint conference on neural network,2000,354-359

- [20] Van Gestel T, Baesens B, Suykens J et al., Bankruptcy prediction with least squares support vector machine classifiers, computational intelligence for financial engineering, In, IEEE international conference proceeding, 2003, 1-8
- [21] Kyung-Shik Shin, Taik Soo Lee, Hyun-jung Kim, An application of support vector machines in bankruptcy prediction model, Expert Systems with Application, 2005, Vol.28, 127-135
- [22] Van Gestel, Tony, Baesens, Bart, Martens, David. From linear to non-linear kernel based classifiers for bankruptcy prediction Neurocomputing, 2010, Vol.73, p2955-2970
- [23] 吴世农, 黄世忠. 企业破产的分析指标和预测模型. 中国经济问题, 1986, (6):23~27
- [24] 杨淑娥, 黄礼. 基于 BP 神经网络的上市公司预警模型[J]. 系统工程理论与实践, 2005, 1:12-18
- [25] 吕长江、周现华. 上市公司财务困境与财务破产的比较分析. 经济研究, 2004, 第 8 期, 64-73
- [26] 胡隽. 基于经济增加值的企业财务困境预测研究 [学位论文], 上海, 东华大学, 2006
- [27] 惠守博、王文杰. 基于支持向量机的财务预警模型与应用研究, 计算机工程与设计, 2006, 第 7 期, 第 1183-1186 页
- [28] 周元元. 基于 RS-SVM 数据挖掘技术的财务困境预测模型研究, [学位论文], 重庆, 重庆工学院, 2008
- [29] 黄真. 基于遗传算法的改进支持向量机财务预警研究, [学位论文], 北京, 北京交通大学, 2008
- [30] 宋姣. 基于支持向量机的上市公司财务危机预警研究, [学位论文], 哈尔滨, 哈尔滨工业大学, 2007
- [31] 姚宏善. 基于支持向量机的财务困境预测研究, [学位论文], 湖北, 华中科技大学, 2006
- [32] 邹睿蓉、王址道、陶艳珍. 关于上市公司财务预警机制的思考, 九江职业学院学报, 2007, 第 3 期, 第 88-89 页
- [33] 岑涌、钟萍、罗林开. 基于 GA-SVM 的企业财务困境预测, 计算机工程, 2008, 第 7 期, 第 223-225 页
- [34] 杨涛. 基于 SVM 的中国医药制造企业财务危机预警研究, [学位论文], 厦门, 厦门大学, 2009
- [35] 沈乐平、黄维民、饶天贵. 基于支持向量机的上市公司违规预警模型研究, 中大管理研究, 2008, 第 125-135 页
- [36] 李小荣. 审计意见与财务困境预测—基于中国上市公司的实证研究, 财会通讯. 综合(下), 2009, 第 6 期, 第 115-116 页
- [37] 王宏伟. 我国上市公司财务困境预测的实证研究, [学位论文], 成都, 西南财经大学, 2002
- [38] 郭亚宁、冯莎莎. 机器学习理论研究, 中国科技信息, 2010, 第 4 期, 第 208-214 页
- [39] 张学工. 关于统计学习理论与支持向量机, 自动化学报, 2000, 第 1 期, 第 32-42 页
- [40] 陈凯、朱钰. 机器学习及其相关算法综述, 统计与学习论坛, 2007, 第 5 期, 第 105-122 页
- [41] 董本清. 机器学习与支持向量机, 软件工程师, 2007, 第 7 期, 第 41-42 页
- [42] 伟东、周志华、李星. 支持向量机研究, 计算机工程, 2001, 第 1 期, 第 58-61 页
- [43] 赵冠华. 企业财务困境分析与预测方法研究, [学位论文], 天津, 天津大学, 2009

- [44]郑红军、周旭、毕笃彦.统计学习理论及支持向量机概述, 现代电子技术, 2003, 第4期, 第59-61页
- [45]王满玲、杨德礼.国外公司财务困境预测研究进展评述, 预测, 2004, 第6期, 第15-20页
- [46]卢世义.我国上市公司财务困境预测的实证研究, [学位论文], 厦门, 厦门大学, 2001
- [47] VapnikVN (著)、许建华、张学工 (译).统计学习理论, 北京, 电子工业出版社, 2004
- [48]张笑天、杨奋强.MATLAB 7.X 基础教程, 西安, 西安电子科技大学出版社,2008
- [49]杜志渊.常用统计分析方法:SPSS 应用, 济南, 山东人民出版社, 2006
- [50] 王正林、龚纯、荷倩.精通 MATLAB 科学计算, 第2版, 北京, 电子工业出版社, 2009
- [51]陈晓、陈治鸿.中国上市公司财务困境的预测, 中国会计与财务研究, 2000, 第9期, 第55-72页
- [52]闫哲.基于 logistic 模型的上市公司财务危机预测的实证研究, [学位论文], 大连, 东北财经大学, 2007
- [53]巫磊.基于 Logit 回归模型的我国上市公司财务困境预测实证研究, [学位论文], 成都, 西南财经大学, 2008

附 录

附录 A 用于实证研究的 98 个训练集样本公司

公司代码	股票名称	首次 ST 时间	配对公司代码	股票名称
000908	ST 天一	2010-02-12	600172	黄河旋风
600562	高淳陶瓷	2010-02-24	600783	鲁信创投
600077	宋都股份	2010-03-01	000565	渝三峡 A
000703	恒逸石化	2010-03-08	600078	澄星股份
000902	中国服装	2010-03-08	600086	东方金钰
000598	兴蓉投资	2010-03-10	000422	湖北宜化
000415	渤海租赁	2010-03-16	600331	宏达股份
600860	北人股份	2010-03-22	000852	江钻股份
600740	山西焦化	2010-03-31	600408	安泰集团
001896	豫能控股	2010-04-06	600070	浙江富润
600299	蓝星新材	2010-04-07	600176	中国玻纤
600885	ST 力阳	2010-04-12	600378	天科股份
000953	ST 河化	2010-04-13	600803	威远生化
000820	*ST 金城	2010-04-16	000541	佛山照明
600131	岷江水电	2010-04-16	000985	大庆华科
600179	黑化股份	2010-04-16	000159	国际实业
600444	国通管业	2010-04-19	600143	金发科技
600355	精伦电子	2010-04-20	600405	动力源
000831	*ST 关铝	2010-04-21	000581	威孚高科
000504	ST 传媒	2010-04-23	600088	中视传媒
600769	ST 祥龙	2010-04-23	000612	焦作万方
600301	南化股份	2010-04-24	000950	建峰化工
000676	ST 思达	2010-04-27	000026	飞亚达 A
000760	博盈投资	2010-04-27	600089	特变电工
000976	春晖股份	2010-04-27	000782	美达股份
600209	罗顿发展	2010-04-27	000401	冀东水泥
600455	ST 博通	2010-04-27	600063	皖维高新
000629	攀钢钒钛	2010-04-28	600010	包钢股份
000958	ST 东热	2010-04-28	000059	辽通化工
600373	中文传媒	2010-04-28	600992	贵绳股份
002072	ST 德棉	2010-04-29	000636	风华高科
600490	中科合臣	2010-04-29	600985	雷鸣科化
600538	北海国发	2010-04-29	600161	天坛生物
600634	ST 澄海	2010-04-30	600312	平高电气
600091	ST 明科	2010-05-04	600229	青岛碱业
600145	国创能源	2010-05-04	600527	江南高纤
600609	金杯汽车	2010-05-04	000338	潍柴动力
600203	福日电子	2011-01-27	600261	阳光照明
600365	通葡股份	2011-02-16	600186	莲花味精

附 录

000602	金马集团	2011-02-18	000559	万向钱潮
600335	国机汽车	2011-02-23	600416	湘电股份
000607	华智控股	2011-04-19	000400	许继电气
000595	*ST 西轴	2011-04-20	600353	旭光股份
000737	南风化工	2011-04-25	600585	海螺水泥
600074	中达股份	2011-04-26	600260	凯乐科技
600228	昌九生化	2011-04-26	000488	晨鸣纸业
900951	大化 B 股	2011-04-26	600319	亚星化学
600281	太化股份	2011-04-27	600227	赤天化
000545	*ST 吉药	2011-04-28	000989	九芝堂
002200	*ST 大地	2011-05-04	600811	东方集团
600671	ST 天目	2011-05-04	600867	通化东宝

附录 B 用于实证研究的 100 个测试集样本

公司代码	股票名称	首次 ST 时间	配对公司代码	股票名称
600894	广日股份	2012-02-14	002009	天奇股份
000420	*ST 吉纤	2012-02-28	002080	中材科技
000669	*ST 领先	2012-03-20	300003	乐普医疗
600882	*ST 大成	2012-03-20	300239	东宝生物
002114	*ST 锌电	2012-03-23	600219	南山铝业
600532	*ST 华科	2012-03-30	600311	荣华实业
000585	*ST 东电	2012-04-05	600517	置信电气
000838	*ST 国兴	2012-04-05	000671	阳光城
000899	*ST 赣能	2012-04-12	600886	国投电力
600087	*ST 长油	2012-04-13	600141	兴发集团
000056	*ST 国商	2012-04-23	600321	国栋建设
000662	*ST 索芙	2012-04-23	600135	乐凯胶片
000815	*ST 美利	2012-04-24	600257	大湖股份
000751	*ST 锌业	2012-04-26	000578	盐湖集团
000806	*ST 银河	2012-04-26	000712	锦龙股份
002102	*ST 冠福	2012-04-26	000977	浪潮信息
000155	*ST 川化	2012-04-27	600184	新华光
000767	*ST 漳电	2012-04-27	000652	泰达股份
002019	*ST 鑫富	2012-04-27	300072	三聚环保
000972	*ST 中基	2012-05-02	600148	长春一东
600392	*ST 天成	2012-05-02	000596	古井贡酒
600877	*ST 嘉陵	2012-05-02	600060	海信电器
600250	*ST 南纺	2012-05-03	600836	界龙实业
600359	*ST 新农	2012-05-03	600844	丹化科技

攻读硕士学位期间发表的学术成果

[1]李忠武. 中外股市市盈率比较研究[J]. 投资与创业, 2012(12):**3-4**

[2]王晶、李忠武. 房地产上市公司财务绩效评价研究[J]. 齐鲁珠坛, 2013(1):**36-39**

致 谢

时光荏苒，三年的研究生生活即将结束，回顾走过的三年，感慨万千。在即将踏入社会的时候，有很多的感触，在这我要感谢所有支持和帮助过我的人。

感谢我的导师赵冠华教授，本文的研究工作是在我导师赵冠华教授的精心指导和悉心关怀下完成的。导师严谨的治学态度、渊博的知识、无私的奉献精神使我深受启迪。

感谢会计学院的各位老师，你们的精彩授课拓展了我的知识视野，为我的学术研究打下了坚实的理论基础。

感谢帮助我的 425 宿舍及 525 宿舍的室友及同学们，在你们的帮助下，我顺利通过了论文答辩。

最后感谢各位评委在百忙之中评阅我的论文。论文难免有不妥之处，恳请各位老师批评指正。

李忠武
2013 年 4 月