

文章编号: 2095-2163(2019)05-0287-04

中图分类号: TP391.1

文献标志码: A

基于 C5.0 的制造企业财务预警研究

张丹, 曹红苹

(上海工程技术大学 管理学院, 上海 201620)

摘要: 随着制造业公司面临的环境愈加复杂多变, 公司为保持长远的稳定发展态势必要加强自身的风险管控。财务预警作为风险分析的一种工具, 可为预警提供有力的基础, 与管控的有效性息息相关。C5.0 决策树通过训练数据可得到清晰的分类规则, 这些规则可为风险分析提供有效依据。通过 2016 年 90 家制造企业的 31 项财务指标构建出 C5.0, 得出基本每股收益、总资产增长率、营业毛利率、现金比率和全部现金回收率对企业风险预测的影响程度较高, 企业应当加强该层面的监管。

关键词: 财务预警; C5.0; 制造企业

Research on financial early warning of manufacturing enterprises based on C5.0

ZHANG Dan, CAO Hongping

(School of Management, Shanghai University of Engineering Science, Shanghai 201620, China)

[Abstract] The manufacturing companies are exposed to an increasingly complex and changeable environment, which means that some effective measures should be taken to improve their risk management so as to get a sustainable future. As a tool for risk analysis, financial early warning can provide a strong basis for early warning, which is closely related to the effectiveness of management and control. The C5.0 decision tree can provide clear classification rules through training data and provide criteria for risk management. C5.0 is constructed through 31 financial indicators of 90 manufacturing companies in 2016, and the basic earnings per share, total asset growth rate, operating gross profit margin, cash ratio and total cash recovery rate have a higher impact on corporate risk prediction. For these companies, supervision on the five points is supposed to be reinforced.

[Key words] financial warning; C5.0; manufacturing enterprise

0 引言

在国家整体经济中,制造业占主体地位,是国民经济的基础。近年来,中国制造业成品走向国门,面向世界,整体呈现良好态势。然而需要指出,制造业的前进路上也将面临机遇与挑战并存的前景,使得制造业的发展中必然存在着一定的不确定性。在此环境背景下,企业为了谋求长远的发展,加强风险预测及管理势在必行,这对企业的财务预警机制是一种考验。

财务预警是一种分析预测问题,是通过选取一定的指标,并且建立模型达到推测的目的,从而为决策提供依据。对企业而言,其陷入财务困境的界定向来有很多标准,如破产清算、资产折现出现严重问题、失去偿还银行贷款及债券的能力、被证券交易所进行特别处理(ST)。在研究过程中,学者们会从研究出发点选择相应的标准。预警分析最初只是通过单个的变量来预测,如 Fitzpatrick 选择财务比率用于预测,Beaver 提出了基于 5 个财务指标的预警模

型。然而通过单个指标虽然较为方便,但是考虑到企业经营涉及到多个层面,产生广泛影响,单个指标也容易出现被操纵的可能性,多变量的研究应运而生。当预警研究转入多变量分析阶段后,模型建立即成为预警提升的热点和焦点。其中,Altman 的 Z-Score 模型堪称经典,随后又陆续推出了 F-score、Probit、Fisher、Bayes、Logit 等模型,并且随着人工智能的涌现,神经网络、决策树等也已开始用作分析工具^[1]。

通过对国内外学者的研究文献综合探讨后可知,预警研究多是侧重于有效指标筛选以及模型建立方面。然而,不少预警模型在构建上都是只会显示指标提取和结果输出,而关于内部详尽的分类过程以及哪些因素的影响程度却仍不得而知,因而可将其视作一种黑箱问题。随着信息化时代的来临,数据信息量越来越大。在各类数据挖掘方法中,决策树算法由于可以产生清晰明确的分类规则,正逐渐大范围地应用到分析预测研究中。对于企业陷入危机、出现财务风险时有各种表现,是否被证监会特

作者简介: 张丹(1995-),女,硕士研究生,主要研究方向:财务预警、分类预测;曹红苹(1968-),女,副教授,硕士生导师,主要研究方向:网络编程、市场营销。

收稿日期: 2019-05-22

哈尔滨工业大学主办 ◆ 科技创新与应用

别处理(ST)则是其最直观的特征形式。本文中,认定企业陷入财务困境的标准为被证券交易所 ST,而被 ST 的公司则处于风险状态,未被标示的即是无风险的。选取制造业 2016 年 45 家被 ST 的公司相关财务指标,同时选 45 家正常企业作为对照,融合风险水平指标和现金流分析,利用 SPSS Modeler 构建 C5.0 决策树对制造业公司的风险预警构建模型,探索对企业风险影响较大的指标。对此,本文拟展开如下研究论述。

1 样本数据的选取

作为国民经济基础的制造业,其基数众多,处于危机状态的也并不鲜见,制造业被 ST 的公司占有行业被 ST 的将近 60%。从重要性和实用性角度出发,文章选取制造业公司来进行预警分析。通过国泰安数据库进行样本的筛选,定位为 A 股的公司,

剔除指标不健全的和因非经营因素导致公司出现财务危机的,在系统内筛选出 45 家 ST 公司。为研究企业财务预警的影响因素,须设定对照组,文献中常用的是 1:1 或者 1:2,考虑到风险组的公司有 45 家,样本较为充足,本文按照 1:1 的原则选取了 45 家正常经营的公司。为增强可比性,正常公司的选择按照制造业中经营范围相同或类似的、且资产规模不超过样本组公司组 15% 获取。样本公司选定后的重点在于评价指标的筛选,指标的筛选上通常涵盖盈利能力、经营能力、偿债能力、发展能力、比率结构这 5 个方面在内。为全面评估企业的财务状况,除上述常用的衡量指标外,考虑到企业发生危机很大程度与资金运作有关,引入现金流分析指标。此外,企业风险水平指标以及可反映企业整体价值的基本每股收益也被引入评价体系。文中有关财务指标的整体选取结果详见表 1。

表 1 财务指标选取表

Tab. 1 Financial indicators for selection

盈利能力	现金流分析	经营能力	风险水平	发展能力	偿债能力	比率结构
资产报酬率 (X_1)	净利润现金净含量 (X_7)	流动资产周转率 (X_{12})	财务杠杆(X_{14})	资本积累率(X_{16})	流动比率(X_{22})	流动资产比率 (X_{26})
总资产净利润率 (X_2)	营业收入现金净含量(X_8)	总资产周转率 (X_{13})	经营杠杆(X_{15})	固定资产增长率 (X_{17})	速动比率(X_{23})	营运资金比率 (X_{27})
流动资产净利润率(X_3)	全部现金回收率 (X_9)			总资产增长率(X_{18})	现金比率(X_{24})	固定资产比率 (X_{28})
净资产收益率 (X_4)	现金再投资比率 (X_{10})			营业收入增长率 (X_{19})	资产负债率 (X_{25})	流动负债比率 (X_{29})
营业毛利率 (X_5)	基本每股收益 (X_{11})			经营活动产生的净流量增长率(X_{20})		综合税率(X_{30})
成本费用利润率 (X_6)				可持续增长率 (X_{21})		净利润综合收益占比(X_{31})

2 C5.0 模型理论

决策树是一种归纳性学习算法,基于已经选择的模糊数据以递归形式分割数据,使得子树集能够得到最好的划分,过程中推理出一系列的分类规则,该规则通常以树形展现,直观简洁,便于理解。决策树算法来自于 CLS(概念学习系统),常见的有基于 ID3、C4.5 和 C5.0 算法的构建的决策树^[2]。各个组别之间的差异是分组考虑的重要因素,也决定了不同种决策树算法各自的独特性。最初的决策树研究基于 ID3 算法,该算法在处理离散数据方面具有良好的效果,但却可能过于拟合,这将不利于后续的研究分析。针对这一不足,在分类规则中融入了信

息增益率,标准调整为有最大信息增益率的分割变量,这是 C4.5 与 ID3 相比存在着不同差异的地方。同时,从 ID3 局限于离散型数据这一问题入手,C5.0 算法则对此做了进一步的拓展,新的算法可以将连续属性的数据转换为离散化,使得原始数据的应用类型更为多样。在大数据处理层面,Boosting 的使用既可以增加预测的准确力度,也可以降低运行时内存的占用率,而与 C4.5 相比,预测的效率和效果都有一定范围内的提升。C5.0 构建决策树基于 2 个层面,对此可阐述如下。

(1) 通过最大属性信息增益率选择属性。

(2) 自上而下的递归方式。构建过程中,数据的分类模式类同于一棵倒立的树木,原始数据的整

体由根而发散,开叉点即为分类节点,通过分析找到最优的开叉点,层层向下分开,直至达到所有的数据都在同类的叶节点上而无需再分为止。

3 实证分析

本文用于实证分析软件为 SPSS Modeler 18.0, 样本数据为 90 家制造业上市公司的 31 项财务指标。在仿真研究中,选取的样本中 ST 的风险状态设置为“是”,将正常公司的设置为“否”,所有的公司按公司代码、相关指标、风险状态整合到 Excel 文件内。

利用 Modeler 源下的 Excel 文件项读取整合文件,再通过字段选项下的类型选项对读取的变量进行设定,内容详情分述如下。

(1) 因为股票代码仅仅是为了识别不同公司,对于预警问题无关紧要,故而将股票代码的角色设定为“无”。

(2) 相关财务指标在预警研究中会对分析产生重大影响,且为分析的基础,其角色设定为“输入”。

(3) 预警研究的重点在于企业的风险状态,也是预期从模型中输出的结果,所以把企业风险状态的角色设定为“目标”,通过文件的读取和变量的设定这两步完成了数据的初始获取及设定。

对于预测分析而言,通常是经过对训练集的模拟输出、得到规则模型后,再将测试集信息代入以实现最终的分析预测。所以训练集的选取关系着规则模型的建立。行业分析的惯例是 70% 的训练集,30% 的测试集,本文仍沿用此法,在字段选项下的分区选项中划分好训练和测试的比例。在此基础上,则会利用模型中的 C5.0 构建模型,为方便查看及理解过程,输出类型择定为生成规则集,同时为提升模型评估性能,运用交叉检验,折叠次数为 10,在 SPSS Modeler 18.0 中构建出设计流程分析图,具体如图 1 所示。

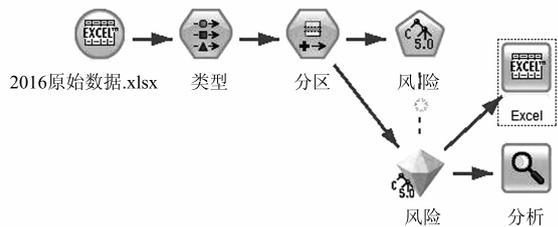


图 1 基于 C5.0 的制造业风险预测数据流图

Fig. 1 Manufacturing risk forecast data flow based on C5.0

流程运行后,得出 27 个测试样本中有 22 个预测符合事实,准确率达到 81.48%,交叉验证的均值

85.7,标准误差为 5.1,预测效果良好,准确率超过 80%,仿真后得到的风险预测分析结果见表 2。

表 2 基于 C5.0 的制造业风险预测分析表

Tab. 2 Manufacturing risk forecast analysis based on C5.0

预测判定	训练集		测试集	
	数目	比例/%	数目	比例/%
正确	63	100	22	81.48
错误	0	0	5	18.52

在预测风险时,影响程度较高的指标有基本每股收益 X_{11} 、总资产增长率 X_{18} 、营业毛利率 X_5 、现金比率 X_{24} 、全部现金回收率 X_9 ,重要性分别为:0.33、0.28、0.16、0.16、0.16。且推演得到的判定规则见表 3。

表 3 企业风险判定规则表

Tab. 3 Rules for companies' risk judgment

风险类型	判定规则
风险组	$X_{11} \leq -0.110$ 或 $X_9 \leq 0.055$
正常组	(1) $X_5 \leq 0.327$ 且 $X_{11} \geq -0.110$ 且 $X_{18} \leq 0.813$ 且 $X_{24} > 0.039$
	(2) $X_9 > 0.055$ 且 $X_{11} > -0.110$ 且 $X_{18} \leq 0.813$

规则集中指标主要分布在偿债能力、发展能力、盈利能力及现金回拢层面,这些本质上均为资金流的把控。关于企业在今后的日常运营中亟需酌情关注的要点,可做阐释叙述如下。

(1) 提高流动资产的变现能力,可适当抽取部分资金用以清算短期债务。针对货币资金、应收账款设置风险预警标准,而对交易的数量、单价、日期一定要准确记录并加强监管核对。赊销时要提高对购买商的信用额度审核力度,同时对金额较大的应收款项的收回也要加强跟进,必要时还可将金额收回度纳入员工考核标准。

(2) 优化产品的结构,淘汰收益率低的产品,减少同质产品的生产销售,加强产品的研发创新力度,实行产品差异化的竞争。

(3) 减少债务融资的占比,企业可以将股权变为多元化,如职工入股,或者构建第三方金融平台,募集个人资金。

此外,针对向银行的贷款,可实行债转股,公司也可以通过非公开发行股票方式(定向增发)募集配套资金。

4 结束语

文章通过对 90 家制造业上市公司 31 项财务指 (下转第 292 页)