

# 基于数据包络和数据挖掘的财务危机预测模型研究

赵智繁 曹 倩

(北京工商大学计算机与信息工程学院 北京 100048)

**摘 要** 以往的企业财务危机预测研究只能预测企业是否具有财务危机,无法预测企业财务危机的程度,这是由于在界定企业财务危机时,只依据了企业是否为 ST 企业的分类方式。鉴于此,通过数据包络分析法,进一步细化了企业财务危机的分类,再使用关联规则算法筛选出重要的预测变量,最后使用决策树技术构建企业财务危机预测模型,并对分类的有效性和预测的准确率进行了验证。实证结果表明,基于数据包络和数据挖掘的财务危机预测模型既能保持较高的准确率,又能预测企业财务危机的程度,使得预测结果更具有参考价值。

**关键词** 数据包络,财务危机,关联规则,决策树

中图法分类号 TP391 文献标识码 A

## Study on Financial Crisis Prediction Model with Data Envelopment Analysis and Data Mining

ZHAO Zhi-fan CAO Qian

(School of Computer and Information Engineering, Beijing Technology and Business University, Beijing 100048, China)

**Abstract** Previous studies on corporate financial crisis prediction can only predict whether the enterprise has the financial crisis, can not predict what the degree of enterprise financial crisis is. This is due to the definition of corporate financial crisis, only on the basis of whether the enterprise is ST Enterprise. In view of this, this paper used data envelopment analysis to refine the financial crisis of enterprises, then selected important predictive variable by using the association rule technology, and finally constructed prediction model of enterprise financial crisis by using the decision tree technique, and the validity of the classification and the accuracy of prediction were verified. Empirical study indicates that the financial crisis prediction model based on data envelopment analysis and data mining can not only maintain a high accuracy, but also can predict the degree of enterprise financial crisis. Its prediction result is more valuable.

**Keywords** Data envelope, Financial crisis, Association rules, Decision tree

## 1 引言

对财务危机预测的定量研究,目前是管理科学领域研究的热点问题。国外的财务危机预测研究起源于 20 世纪 30 年代,随后,众多学者在该领域进行了相关的探索,涌现出大量的研究成果。国内的相关研究始于 20 世纪 90 年代末。目前,随着企业财务危机问题的不断加剧,财务危机预测研究正处在一个快速发展的时期。

纵观现有的研究成果,可以发现,相关研究的创新点大部分集中在财务危机预测建模方法的研究和预测变量的选取这两个方面,而在研究财务危机程度划分方面的工作很少。就现有的文献来看,国内外的研究样本大致可分为 2 组,即“财务困境企业”和“非财务困境企业”。在实证研究中,将财务困境企业直接等同于上市企业中的 ST 企业,相应地,非财务困境企业等同于非 ST 企业。采取此种分类方式的原因主要是为了迎合我国上市企业的实际情况、提升建模实验的可操作性以及便于不同学者的研究相互比较<sup>[1]</sup>。

随着研究地不断深入,此种分类方式出现了很多弊端。

首先,在实际的应用中,不仅需要预测企业是否具有财务危机,还需要预测企业具有的财务危机的程度,以目前这种分类方式为基础构建的预测模型的预测结果将只有两种,即预测企业将成为“ST 企业”或者将成为“非 ST 企业”,因此其应用价值并不高;其次,非 ST 企业成为 ST 企业的必要条件之一为“连续两年净利润为负”,这使得 ST 企业的财务指标数据本身就具有一定的特点,即使在预测模型中挖掘到与“净利润”相关的预测变量的一些规则,这些规则也很有可能没有意义的,在检验模型效果时,这可能会高估预测模型的准确率;最后,每年的 ST 企业数量较少,其中拥有相似行业、相近规模的企业样本数就更加稀少,目前绝大多数相关研究的训练样本集只能控制在 30 到 50 个左右,训练样本集的不足将降低预测模型的准确率与适用性。

部分学者在财务危机程度划分方面进行了积极的探索,在原有的依据 ST 与非 ST 进行分类的基础上,设置了新的划分规则。但危机程度的界定问题相对定性,难以使用数值衡量,而且新设置的划分规则具有较大的主观性,这些都为在财务危机预测模型中进行财务危机程度划分的工作造成了影响。

本文受北京市教委科学研究面上项目(KM201410011005),北京市优秀人才培养资助项目(2015000020124G029),北京工商大学教育教学改革项目(jg155225)资助。

赵智繁(1991—),男,硕士,主要研究方向为数据仓库、数据挖掘,E-mail:zhifanzhao@163.com;曹倩(1983—),女,讲师,硕士生导师,主要研究方向为高性能数据挖掘、大数据研究,E-mail:caoqian125@126.com。

目前,数据包络技术应用于财务状况评价的研究越来越成熟。数据包络技术以效率指数为核心,不必自定义权重,具有很强的客观性,特别适合对多投入、多产出的企业型决策单元进行综合评价。依据评价结果可以将企业财务危机程度细分为“安全”、“较安全”、“一般”、“危机”等不同等级。另外,以往的研究样本中必须选取一半的ST企业样本,而数据包络技术是对所有相近行业、相似规模的企业样本进行的综合对比评价,因此所有历史数据完整的企业样本都可以参与到研究中,这大大扩充了研究样本容量,提升了预测模型的稳定性和适用性。由此可见,通过使用数据包络技术来划分财务危机程度,可以弥补许多以往依据ST划分的不足。

## 2 国内外研究现状

在财务危机预测研究中,国外最早是以企业破产作为界定企业财务危机的标准,其中最具代表性的研究是Altman以33家提出破产申请的公司和与之配对的健康公司为研究样本,并使用多元线性判别方法创立了Z值模型<sup>[2]</sup>。随后,随着应用需求的提升以及研究的深入,部分学者提出了要对企业财务危机程度进行细化,以产生更合理、更实用的财务危机预测模型。Lau提出了企业陷入财务危机是具有阶段性的<sup>[3]</sup>,他将这个过程分成了4个阶段:第一阶段是财务状况开始恶化的阶段,主要体现在股利或未支付股利较上年减少40%以上;第二阶段是财务状况继续恶化的阶段,主要体现在企业无力偿还债务;第三阶段是企业受破产法保护阶段;第四阶段是破产阶段。Ohlson则提出当前依据企业是否破产的二分类方法将导致实验样本的缺少,并首次使用了非配对抽样方式弥补了破产企业样本数量不足的问题<sup>[4]</sup>。Aziz的基于现金流量的财务困境预测模型(CFB)可以以一个定量指标——企业预期现金流量的净现值,作为企业财务危机程度划分的标准<sup>[5]</sup>。

国内的大多数研究则是以上市企业被标记为ST作为界定企业财务危机的标准,同样是一种二分类方式。也有一些学者使用其他的分类方法,来弥补目前二分类方式的局限性。周敏使用基于模糊优选模型将企业样本财务危机状况分成了4个不同程度的级别<sup>[6]</sup>,并使用神经网络进行学习和预测。还有学者自行设定了一些分类标准,如吕长江的三指标聚类分析判别标准<sup>[7]</sup>、文拥军的依据连续两年净利润情况的分类标准<sup>[8]</sup>,这些分类标准均实现对企业财务危机程度的多分类。陈磊则将传统的ST、非ST的分类方法扩展为ST、\*ST、非ST3类<sup>[9]</sup>。另外也有部分研究使用了多分类报警器,使二分类预测方法应用于多分类的预测研究中<sup>[10]</sup>。总的来说,自定义一些量化的分类标准是我国学者常用的改进方法,尽管这些方法能体现出不同的财务危机程度,但依然具有很大的主观性。

目前,使用数据包络分析法对企业财务状况进行综合评价的研究成果有很多,如柏云燕使用数据包络分析法对高校财务绩效进行综合评价<sup>[11]</sup>,李小程则应用数据包络对房地产投资项目财务效率进行综合评价<sup>[12]</sup>,这些研究均取得了良好的评价效果。

本文将数据包络技术应用到企业财务危机预测模型中,通过应用数据包络分析软件,实现对企业财务危机程度的多分类,并使用关联规则对预测变量进行筛选,保留关联度较大的预测变量,剔除弱关联变量,最后使用决策树构建预测模型。

实验结果验证了财务危机程度多分类的有效性和预测模型的准确性。

## 3 研究设计

### 3.1 技术路线

财务危机预测模型预测结果的分类,与样本的财务危机程度的分类是一一对应的,如:样本的财务危机程度分为ST企业和非ST企业,则财务危机预测模型的预测结果也分为这两类。同样地,如果样本的财务危机程度分为“安全”、“较安全”、“一般”、“危机”4类,则财务危机预测模型的预测结果同样分为这4类。将使用数据包络技术对样本的财务危机程度进行评价,依据评价结果进行多分类,从而实现对预测结果的多分类,使得预测模型不但可以预测企业是否具有财务危机,还能预测企业具有的财务危机的程度。

假设企业第T年年末的财务危机程度评价结果为“危机”,由于企业财务数据的发布具有滞后性,因此企业第T年的财务数据对构建财务危机预测模型没有任何意义,应当使用企业第T-1年的财务数据,这之间相差的一年代表预测模型能提前一年预测企业的财务危机程度,没有这一年的提前期则不能称为是预测模型,这是整个财务危机预测模型的关键。需要注意的是,传统的二分类预测模型使用的是第T-2年的财务数据,本文使用的是第T-1年的财务数据,原因在于,企业被标记为ST往往发生在第T+1年的年初,而本文的财务危机程度评价结果发生在第T年的年末,这之间虽然只有几天的间隔,但在数值表示上依旧存在1年的差距。

最后使用数据挖掘技术构建财务危机预测模型,为了检验预测模型的准确性,在进行数据挖掘前应当将样本划分为训练样本和测试样本。同时,要对数据包络多分类结果的有效性和数据挖掘预测模型的准确率进行验证,以判断其是否达到了良好的细分效果。

### 3.2 预测变量

#### 3.2.1 备选的预测变量

本文数据来源于上海证券交易所官方网站中,是上市公司XBRL专栏披露的财务数据,样本为62家化学原料和化学制品企业2007年至2014年的财务数据,行业的分类标准为CSRC。

在财务危机预测模型中,一般选取财务指标作为预测变量,财务指标可以通过财务分析中的比率分析法计算得出。本文备选的财务指标预测变量源自2006年9月国资委发布的《企业综合绩效评价实施细则》中财务绩效定量评价指标规定的8个基本指标与10个修正指标,相对而言该指标更能全面衡量我国上市企业的财务状况,如表1所列。

表1 备选预测变量

A:盈利能力	B:资产质量	C:债务风险	D:经营增长
A1:净资产收益率	B1:总资产周转率	C1:资产负债率	D1:销售增长率
A2:总资产报酬率	B2:应收账款周转率	C2:已获利息倍数	D2:资本保值增值率
A3:营业利润率	B3:流动资产周转率	C3:速动比率	D3:销售利润增长率
A4:盈余现金保障倍数	B4:资产现金回收率	C4:现金流量负债比率	D4:总资产增长率
A5:成本费用利润率			
A6:资本收益率			

### 3.2.2 预测变量的计算

目前,通过比率分析计算得出的预测变量的主要问题是其无法正确代表企业的实际财务情况。例如在计算企业净资产收益率时,当分子净利润与分母平均净资产均为负数时,得到的结果反而为正,发生这种情况时,数字无法表达出“已经没有净资产的企业又获得了亏损”这种极端恶劣的情况。因

此本文结合此类数据的发生特点以及相关财务指标的实际含义,提出了一种新的分类计算方法,如表 2 所列,使得计算后的值能更好地代表企业当前的实际财务状况。表 2 中使用的处理方式“置为 MIN”,是使其与全部样本中该指标的最小值相等,即使它等于该行业下的最低水准,“置为 MAX”同理。

表 2 指标数据分类计算方法

指标名称	产生异常的原因	判断条件	处理方式	描述
A1:净资产收益率	平均净资产<0	若净利润<=0	置为 MIN	已经没有净资产的企业再次受到了亏损,此类财务状况是最糟糕的。
		若净利润>0	置为 0	已经没有净资产的企业实现了盈利,优于亏损的企业(A1<0),但劣于正常的企业(A1>0)。
A4:盈余现金保障倍数	净利润<0	若经营现金净流量<=0	置为 MIN	企业没有利润收益,且没有现金保障。
		若经营现金净流量>0	置为 1	企业虽然没有利润收益,但现金流情况优于现金流为负的企业(A4<0),劣于正常企业(A4>0)。
C2:已获利息倍数	利息支出<0	若利润总额>0	置为 MAX	企业不但拥有销售利润,且拥有利息收入,因此没有债务危机。
		若利润总额<=0	置为 1	企业的利润大部分来源于利息收入,该情况优于利润不足以支付利息的企业(C2<1),但劣于正常企业(C2>1)。
D2:资本保值	去年所有者权益<0	若今年所有者权益<=0	置为 MIN	连续两年所有者权益为负,此类财务状况很差。
		若今年所有者权益>0	置为 1	实现了财务状况的逆转,优于亏损的企业(D2<1),但劣于正常企业(D2>1)。
D3:销售利润增长率	去年销售利润<0	若今年销售利润增长<=0	置为 MIN	连续两年销售利润为负,此类财务状况很不乐观。
		若今年销售利润增长>0	置为 0	销售利润相对去年有所好转,优于连续没有销售利润的企业(D3<0),但劣于正常企业(D3>0)。

### 3.3 数据包络对财务危机程度的多分类

#### 3.3.1 数据包络输入/输出项

数据包络模型是直接使用输入、输出项数据建立的非参数的经济数学模型。应用数据包络技术的关键之一是选取合适的输入/输出数据项,其中,输入项应该是决策单元用以产生输出项的投入状况或是规模状况,输出项应该是决策单元所产生的价值状况。因此,将与企业投入资产情况及成本费用情况相关的数据项作为数据包络的输入项,主要涉及资产负债表与利润表中的会计期间费用,其中包括:总资产、股本、资本公积金、所有者权益、营业成本、销售费用、管理费用、财务费用 8 个项目;将与企业获得的收益相关的数据项作为数据包络的输出项,主要涉及利润表中的各类收入项与现金流量表中的现金流项,其中包括:营业收入、营业利润、利润总额、净利润、经营活动现金净流量 5 个项目。

#### 3.3.2 效率指数的计算

CCR 模型是一种最具代表性的数据包络分析模型,通过对多投入、多产出的决策单元进行综合评价,得出每个决策单元的效率指数,本文将企业的效率评价指数作为划分企业所处财务危机程度的标准。

假设数据包络系统结构中共有  $n$  个决策单元(DMU),每个决策单元都有  $m$  个输入变量(记为  $x$ )与  $s$  个输出变量(记为  $y$ )。每个输入/输出变量各有一个权系数,权系数是对应的输入/输出变量的一种度量,分别记为  $v$  与  $u$ ,所有权系数的取值均不小于 0。其中,决策单元(DMU)对应某个企业某一个年份的财务数据,输入变量  $x$  与输出变量  $y$  分别对应 3.2.1 节中提到的输入/输出项,权系数  $v$  与  $u$  为系统设定而非使用者主观设定。

设  $h$  为每个决策单元的效率评价指数,则  $h$  的计算方法为:

$$h_j = \frac{u^T y_j}{v^T x_j} = \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}}, j=1, 2, \dots, n \quad (1)$$

可以将每个决策单元的效率指数理解为每个决策单元的输出项数值与其对应权系数的乘积的总和除以每个决策单元的输入项数值与其对应权系数的乘积的总和。由于权系数不

是主观设定的,因此系统总是能取得许多组合合适的权系数,使得:

$$h_j \leq 1, j=1, \dots, n \quad (2)$$

需要尽可能多地考虑取不同的权系数时,每个决策单元的  $h$  值最大是多少。以第  $j_0$  个决策单元的效率指数为目标,以所有决策单元的效率指数为约束,则产生的数据包络的 CCR 模型为:

$$\begin{cases} \max h_{j_0} = \frac{\sum_{r=1}^s u_r y_{rj_0}}{\sum_{i=1}^m v_i x_{ij_0}} \\ \text{s. t. } \frac{\sum_{r=1}^s u_r y_{rj}}{\sum_{i=1}^m v_i x_{ij}} \leq 1, j=1, 2, \dots, n \\ u \geq 0, v \geq 0 \end{cases} \quad (3)$$

通常认为,决策单元的最大效率指数等于 1 时,该决策单元为 DEA 有效或者弱有效,其小于 1 时,该决策单元没有达到 DEA 有效<sup>[13,14]</sup>。

将根据每个企业每个年份的效率指数,将财务危机程度细分为“安全”、“较安全”、“一般”、“危机”4 类,一般以等深原则划分效果最好。显然,当效率指数较大时,意味着该企业本年度在财务方面以更小的投入换取了更大的回报,财务状况较好,因此财务危机程度较低。随着效率指数的降低,财务危机程度将逐渐加重。

在得到所有样本的财务危机程度后,要将第  $T$  年的财务危机程度与第  $T-1$  年的财务历史数据进行对应,以实现财务危机预测的目的。最好的方法是,在数据库中将企业第  $T$  年的财务危机程度作为一个新的属性加入到该企业第  $T-1$  年的记录中去。

#### 3.4 数据挖掘构建预测模型

##### 3.4.1 预测变量的筛选

目前,使用决策树构建企业财务危机预测模型,当预测变量较多时,容易造成树的过度生长及过拟合问题,因此需要筛选预测变量。以往研究常用的配对 T 检验、主成分分析等方法需要样本数据近似服从正态分布并满足相关性检验,但是

财务数据因其本身的特点,往往无法满足这些条件。

关联规则对样本数据是否连续、是否服从正态分布、是否满足相关性检验等均无要求,具有更强的适用性。同时,通过关联规则筛选出的预测变量与财务危机程度关联度较大,在降低预测变量个数的同时,可以保证预测模型一定的准确率。因此,可以使用  $T-1$  年的财务历史数据和关联规则技术筛选出重要的预测变量。

在关联规则中, $X$  推导出  $Y$  的支持度代表所有记录中同时包含  $X$  与  $Y$  的比率,公式为:

$$support(X \rightarrow Y) = \frac{count(X \cup Y)}{|D|} \quad (4)$$

$X$  推导出  $Y$  的置信度代表了所有同时包含  $X$  与  $Y$  的记录数与包含  $X$  的记录数之比,意义是当一条记录具有  $X$  属性时,也具有  $Y$  属性的概率是多少,公式为:

$$confidence(X \rightarrow Y) = \frac{support(X \rightarrow Y)}{support(X)} \quad (5)$$

如果“预测变量  $A1 < x$ ”推导出“危机”的支持度满足最小支持度,置信度满足最小置信度,则说明当预测变量  $A1 < x$  时,企业明年的财务危机程度为“危机”的可能性较大,因此预测变量  $A1$  将成为一个合格的预测变量。

### 3.4.2 预测模型的构建

通过关联规则只能筛选出合适的预测变量,并不能形成可读的、可验证的预测模型。因此,本文使用决策树技术来构建财务危机预测模型。决策树将通过计算每个属性的信息增益量来确定优先分裂属性。信息增益量是针对某一个变量属性而言的,即系统在携带该变量时与不携带该变量时的信息量差值。因此在决策树中,对信息量的计算特别重要。假设有一个变量  $X$ ,它可能的取值为:

$$(X_1, X_2, \dots, X_n) \quad (6)$$

每一种取到的概率分别是:

$$(P_1, P_2, \dots, P_n) \quad (7)$$

则它的信息量(熵)为:

$$H(X) = - \sum_{i=1}^n P_i \log_2 P_i \quad (8)$$

决策树总是会优先分裂信息增益最大的属性变量。决策树模型的预测变量是企业第  $T-1$  年的财务指标数据,从全部训练样本出发,根据不同的判别条件形成多个分支,最后的分类结果则是企业第  $T$  年的财务危机程度。在得到树模型后,当拥有某企业第  $S$  年的财务数据时,代入树模型后得到的分类结果就是该企业第  $S+1$  年的财务危机程度的预测结果。

## 4 实证研究与结果验证

### 4.1 实证过程

本文使用的数据包络分析软件是 MaxDEA Version 5.0,经数据包络分析法处理后,根据效率指数的大小,按照等深划分原则,将企业财务危机程度划分为“安全”、“较安全”、“一般”、“危机”,部分分析结果如表 3 所列。

将原始财务报表数据导入 SQLSERVER 数据库,经过加工计算得出财务指标数据,并在每个企业的第  $T$  年的财务指标数据里添加一个新的属性(列),称为“结果变量”,同时将该企业第  $T+1$  年的财务危机程度对应地添加进去,其中没有第  $T+1$  年的财务危机程度的元组将被删除。最终形成了 345 组历史数据,选取其中 230 组作为挖掘算法的训练样本集,剩余 115 组作为测试集。在 Business Intelligence Development Studio 中连接数据源与数据源视图,建立关联规则挖掘模型。将最小概率设为 0.6,最低重要性设为 0.2,可以得到 8 条规则,以此可以筛选出与财务危机程度相关性较大的

7 个指标,分别是:A2,A6,B1,C1,C2,C3,D2。重要的关联规则如表 4 所列。

表 3 部分 DEA 分析结果

DMU 编码	企业代码	时间	效率指数	财务危机程度
1	600075	2014/12/31	0.741865	一般
2	600075	2013/12/31	0.704105	危机
3	600075	2012/12/31	0.705499	危机
4	600075	2011/12/31	0.718443	危机
5	600078	2014/12/31	0.700721	危机
6	600078	2013/12/31	0.701828	危机
7	600078	2012/12/31	0.729745	一般
8	600078	2011/12/31	0.720041	一般
9	600091	2014/12/31	0.845934	较安全
10	600091	2013/12/31	0.7929	较安全
11	600091	2012/12/31	0.792234	较安全
12	600091	2011/12/31	1	安全

表 4 重要关联规则

概率	重要性	规则
1.000	0.356	A2 ≥ 0.214 → 安全
1.000	0.387	C2 < 7.129 → 危机
0.889	0.447	B1 ≥ 1.507 → 安全
0.778	0.353	C1 = 0.435 - 0.612 → 一般
0.650	0.305	D2 ≥ 0.632 → 安全
0.645	0.321	B1 = 1.085 - 1.507 → 危机
0.636	0.278	C3 = 3.245 - 6.232 → 危机
0.614	0.365	A6 ≥ 0.251 → 一般

建立决策树挖掘模型,以关联规则挖掘模型得出的 7 个重要指标作为预测变量,以“结果变量”字段作为结果变量,可以得到图 1 的结果。

### 4.2 结果验证

#### 4.2.1 财务危机程度分类结果的有效性验证

本文研究的核心工作是,使用数据包络技术对财务危机程度进行多分类,这与传统的基于 ST 的二分类方法相比,不但可以预测企业是否存在财务危机,还能预测企业具有的财务危机的程度,显然具有更好的应用价值。但是,由于这是一种细化分类的工作,还需要对这种细分方法的有效性进行检验,以考察新的分类结果与传统的分类结果之间有无映射关系。选取研究样本中的 ST 记录,共 33 条,另外配对并选取 33 条非 ST 记录,来分析它们经过数据包络分析后的财务危机程度分类结果,如表 5 所列。

表 5 分类结果

	危机	一般	较安全	安全
ST	30	1	2	0
非 ST	4	12	8	9

根据分类结果得出,原分类中的 ST 记录使用数据包络评价分类后,绝大部分属于“危机”的一类原分类中的非 ST 记录,绝大部分属于另外 3 类。因此,原分类结果与现分类结果之间存在一定的映射关系,现分类对原分类具有良好的细分效果。

#### 4.2.2 财务危机预测模型的准确性验证

根据挖掘准确性图表,使用 115 组测试数据进行准确性检验,通过决策树模型得到的结果分类矩阵如表 6 所列。

表 6 决策树模型验证结果分类矩阵

	危机 (实际)	一般 (实际)	较安全 (实际)	安全 (实际)
危机(预测)	28	8	3	1
一般(预测)	0	17	4	0
较安全(预测)	0	2	19	4
安全(预测)	0	2	3	24
准确率	100%	58.6%	65.5%	82.8%

由于本文的研究样本没有按照以往研究的配对抽样法进行抽样,因此在验证预测准确率方面,不适合建立以传统 ST 方法为基础的预测模型进行对比验证。但是本文的预测结果包含 4 种不同的财务危机程度,与以往只能预测是否具有财务危机的模型相比,具有更好的应用价值。而且根据目前的研究文献来看,以往大部分的财务危机预测模型的准确率在 70%到 85%之间。根据表 6 可以看出,本文方法在预测“危机”与“安全”时,具有较高的准确率,尽管在预测“较安全”与“一般”时准确率不够,但这对该方法应用价值的影响相对较小。因此本文的研究在保证较高的准确率的基础上,不但可以预测企业是否具有财务危机,还能预测财务危机的程度。

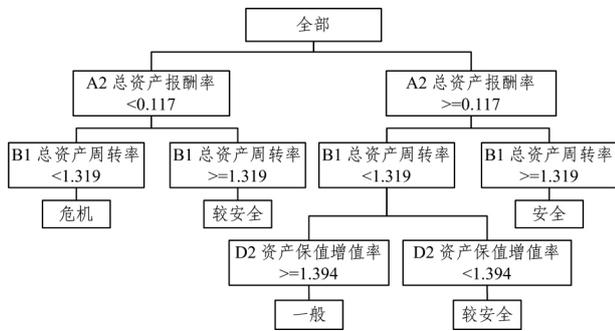


图 1 决策树结果

在企业财务危机预测中,错误主要有两类:第一类是预测的结果比实际结果要好,第二类是预测的结果比实际结果要差。Sinkey 证实了,在预测模型中,总体预测正确率高但是其中第一类错误率高的错误成本,要高于总体预测正确率低而其中第一类错误率低的错误成本<sup>[13]</sup>。本文的预测模型的错误主要来源于第二类错误,此类错误要优于上面提到的第一类错误。

结束语 本文将数据包络引入财务危机预测模型,对财务危机程度进行了细分,取得了良好的效果。就目前研究的结果来看,总资产报酬率是一个相当稳健的指标,该指标不但是反映企业总体盈利能力的重要指标,同时对企业财务危机的预测也有着重要的参考意义。无论是管理者,还是外部投资者,都应着重关注企业总资产报酬率的情况。而相关的企

业财务综合评价模型,也可适当地调高该指标的权重,以获取更为精准的评价结果。

## 参考文献

- [1] 边海容,万常选,刘德喜,等.考虑 Web 金融信息的上市企业财务危机预测模型研究[J]. 计算机科学,2013,40(11): 295-298
- [2] Altman E I. Discriminate Analysis and Prediction of Corporate Bankruptcy[J]. Financial Ratios,1968(9):589-609
- [3] Lau A H. A five state financial distress prediction model[J]. Journal of Accounting Research,1987,25(1):127-138
- [4] Ohlson J S. Financial Ratios and the Probabilistic Prediction of Bankruptcy[J]. Journal of Accountant Research,1980,18(19): 103-131
- [5] Aziz A D. Emanuel and G Lawsom. Bank Prediction: An Investigation of Cash Flow Based Models[J]. Journal of Management Studies,1988(25):419-437
- [6] 周敏,王新宇.基于模糊优选和神经网络的企业财务危机预警[J]. 管理科学学报,2002(3):86-90
- [7] 吕长江,韩慧博.上市公司财务困境预测方法的比较研究[J]. 吉林大学社会科学学报,2005(6):99-109
- [8] 文拥军.基于多类 Logistic 的财务困境预警模型实证研究[J]. 中国海洋大学学报,2009(2):46-50
- [9] 陈磊,任若恩,曹汉平.公司多阶段财务危机动态预警研究[J]. 系统工程理论与实践,2008,28(11):29-35
- [10] 邱玉莲,朱琴.基于支持向量机的财务预警方法[J]. 统计与决策,2006(16):153-155
- [11] 柏云燕.数据包络分析法在高校财务绩效评价中的应用[J]. 电子商务,2012(9):58-62
- [12] 李小程.基于数据包络分析的房地产投资项目财务效率评价研究[J]. 湖南财政经济学院学报,2011(8):94-98
- [13] 张坤.基于数据包络分析:决策树的企业财务评价与预警——以西部四省上市企业为例[J]. 财务与金融,2014(3):36-40
- [14] 徐凤江.数据包络分析法分析企业财务[J]. 现代商贸工业,2009,21(4):276-278
- [15] Sinkey. A Multivariate Statistical Analysis of the Characteristics of Problem Bank[J]. Journal of Finance,1975,3(1):21-38

(上接第 460 页)

- [14] Zhou Jian-heng, Shu Ling. Strategy for supply chain coordination based on fashion life cycle[C]//2010 International Conference on Management and Service Science, Wuhan, 2010:1-5
- [15] 陈希孺,柴根象.非参数统计教程[M].上海:华东师范大学出版社,1993
- [16] 程秩.常用预测方法及评价综述[J]. 四川师范大学学报,2002,22:5-6
- [17] 胡觉亮,何秋霞,韩曙光,等.基于改进的 BASS 模型的服装产品生命周期研究[J]. 浙江理工大学学报,2010,1(27):69-73
- [18] 胡觉亮,孔云鹏,韩曙光.需求信息更新下服装供应链的最优订货决策研究[J]. 纺织学报,2012,33(11):136-140
- [19] 胡敏敏.析 SPA 型服装品牌零售模式——以 ZARA 为例[J]. 浙江理工大学学报,2007,7:404-406
- [20] 刘咏梅,王朝晖,王咏.基于产品生命周期的服装库存管理智能决策[J]. 东华大学学报,2004,30(2):34-38
- [21] 刘铮,徐琪.基于服务一购买意愿的双渠道短生命周期产品需求

预测模型——以快时尚服装行业为例[J]. 东华大学学报,2013,5(39):668-681

- [22] 乔亮亮.多元线性回归预测法在服装制造中的应用[J]. 长江大学学报,2010,1(7):145-160
- [23] 盛方正,季建华.需求预测信息更新的零售商三阶段最优采购策略[J]. 管理工程学报,2009,2:156-159
- [24] 徐贤浩,陈雯,廖丽平,等.基于需求预测的短生命周期产品订货策略研究[J]. 管理科学学报,2013,4:22-23
- [25] 楼俊钢,蒋云良,申情,等.软件可靠性预测中不同核函数的预测能力评估[J]. 计算机学报,2013,36(6):1303-1311
- [26] 王静,蔡勇,蒋刚.基于多尺度核函数的铆接件腐蚀疲劳预测[J]. 计算机应用研究,2015,32(4):1074-1077
- [27] 汪廷华,陈峻婷.核函数的选择研究综述[J]. 计算机工程与设计,2012,33(3):1181-1186
- [28] 梅玲,张亚丽.结构化数据核函数综述[J]. 计算机工程,2010,36(23):199-201