

Financial Risk Early Warning Model of Decision Tree Research Based on Z-Score Indicators

Yunyi Yang, Zhonghe Sun, Jing Lu

IT Center, China National Audit Office, Beijing

Email: peteryoung@audit.gov.cn, sunzhonghe@audit.gov.cn, lujing@audit.gov.cn

Received: Aug. 16th, 2017; accepted: Aug. 31st, 2017; published: Sep. 7th, 2017

Abstract

Financial risk early warning model is one of the primary means of forecasting financial crisis. Z-score model has been applied widely since it has precise formula and good practicability. However, as regard to the adequacy of using Z-score model to forecast financial crisis for domestic companies, there is a dispute. Therefore, we collect financial statements data from 2007-2013 of Shanghai listed companies, and prove that Z scores between ST companies and non-ST companies are linearly inseparable. It reveals that Z-score model with linear separable principle in nature, has its own limitation. Hence, this paper brings up a financial risk early warning model of decision tree based on Z-score indicators and it solves the problem of Z-score model. The empirical research shows that the decision model given by this paper can forecast financial crisis three years before it happens with 75.37% accuracy, and with 95.45% accuracy one year ahead of crisis occurrence.

Keywords

Financial Risk Early Warning, Z-Score, Decision Tree

基于Z-Score指标决策树的财务风险预警模型研究

杨蕴毅, 孙中和, 卢靖

审计署计算机技术中心, 北京

Email: peteryoung@audit.gov.cn, sunzhonghe@audit.gov.cn, lujing@audit.gov.cn

收稿日期: 2017年8月16日; 录用日期: 2017年8月31日; 发布日期: 2017年9月7日

摘要

财务风险预警模型是预测财务危机的主要手段之一。Z分数模型以其公式严谨、可操作性强等特点，被广泛应用。然而，Z分数模型是否适用于预警中国企业的财务危机仍存在争议。为此，归集了沪市上市公司2007~2013年的财务报表数据，实证发现ST公司与非ST公司的Z值线性不可分。揭示了：本质上为线性区分方法的Z分数模型存在局限性。因此，利用Z分数模型原有指标，提出了基于决策树的财务风险预警模型，解决Z分数模型失效问题。实证研究表明：所提出的预警模型提前3年预测ST公司财务危机的准确率达75.37%，提前1年预测的准确率达95.45%。

关键词

财务风险预警，Z-score，决策树

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

国内学者胡剑(2014)提出财务风险是在公司各项财务活动中,由于内外部环境及难以预料或无法控制的因素使财务系统运行偏离预期目标而形成的经济损失的机会性或可能性[1]。财务风险客观存在,虽不能完全消除,但可预防和降低。

财务风险预警的主要手段之一是建立财务风险预警模型。一个好的财务风险预警模型不仅能给投资者及早带来预警信号,及时做出决策规避风险,还能帮助企业管理层提早发现企业财务恶化的征兆,预防财务危机。财务风险预警模型有单变量和多变量两种模式。早在20世纪三十年代,西方经济学家帕特里克就运用单个财务比率对公司财务风险进行预警研究。在多变量模式中,有日本开发银行的多变量预测模型,中国学者周首华、杨济华的F分数模型,以及引入现金流量因素的预警模型等。

Z分数模型以其严谨的公式和较强的操作性被广泛应用。美国学者Edward I. Altman (1968)选取1946~1965年间的33家破产公司和33家正常经营的公司,提取出5种最具共同预测能力的财务比率,建立起类似于回归方程式的鉴别函数——Z分数模型[2]。其中,Z分数为五项财务指标的加权和,推荐以1.81为是否有破产危机的临界值,如果Z值低于1.81,则公司存在很大的破产危险。

目前我国学者对Z分数模型是否适用于国内企业存在争议。一派学者通过小规模数据验证ST公司的Z值小于1.81,认为该模型有效;另一派学者发现破产公司与非ST公司的Z值无明显分界,对Z分数模型持否定态度。但是这两派研究都存在数据规模较小、验证方式不够全面、未能从方法原理上进行分析等缺陷。

为此,本文首先归集了沪市上市公司2007~2013年913家财务报表数据,提取Z分数指标,按Z分数模型计算这些公司的Z值,对ST公司和非ST公司的Z值进行比对分析。研究发现ST公司和非ST公司的Z值呈现混杂分布的形态,两类公司间不具备线性可分的条件。这一结果表明:本质上为线性区分方法的Z分数模型在分析我国公司财务风险时存在方法上的缺陷。为了解决这一问题,本文在Z分数模型原有指标的基础上,提出了预测财务风险的决策树方法。研究发现,影响企业财务风险预警的主要

财务指标有留存收益/总资产, 营运资产/总资产和销售收入/总资产。实验结果表明, 在非平衡数据状态下, 提前三年预测公司是财务破产风险的准确率为 75.37%, 提前两年的预测准确率为 88.54%; 提前一年的预测准确率为 95.45%。

本文共分为五个部分, 第一部分引言, 介绍背景意义。第二部分综述了国内学者对 Z 分数模型有效性的分歧。第三部分对 Z 分数模型有效性进行分析, 实证 Z 分数模型对我国上市公司财务风险预警失效。第四部提出基于决策树的上市公司财务风险预警模型, 并验证模型。第五部分, 总结结论。

2. Z 分数预警模型适用性分歧

Z 分数模型的基本原理是用基于多重指标的线性多元函数识别有财务风险的公司。经典 Z 分数模型的表达式为: $Z = 1.2X_1 + 1.4X_2 + 3.3X_3 + 0.6X_4 + 0.99X_5$ 。其中: Z 为判别函数值; $X_1 =$ 营运资产/总资产; $X_2 =$ 留存收益/总资产; $X_3 =$ 息税前利润/总资产; $X_4 =$ 权益市 Z 场价值/负债账面总价值; $X_5 =$ 销售收入/总资产。奥特曼提出判断公司破产的临界值: 若 $Z > 2.675$, 表明公司的财务状况良好, 发生破产的可能性较小; 若 $Z < 1.81$, 则公司存在很大的破产危险, 即 1.81 是公司破产发生的定量界定。

目前, 国内部分学者就 Z 分数模型对我国上市公司的财务风险预警是否有效进行了分析, 主要存在两种观点: 一些学者认为传统的 Z 分数模型仍适用于我国公司的财务风险预警, 而另一些学者则对此持否定态度。

支持 Z 分数模型有效的研究方法有多种。向德伟(2002), 张蔚虹, 朱海霞(2012), 王芳云(2005)通过计算几十家研究样本中 ST 公司的平均 Z 值和非 ST 公司的平均 Z 值, 并与 Z 分数模型警戒值进行比较, 发现 ST 公司平均 Z 值小于 1.81, 非 ST 公司平均 Z 值大于 1.81, 从而认为该实证研究大体上支持 Z 分数模型的有效性[3] [4] [5]。李艳(2012), 常树春(2013)选取个别 ST 公司和非 ST 公司作为样本数据, 对小规模样本数据 Z 值进行计算, 发现 3 到 5 家非 ST 公司 Z 值大都低于 1.81, 得出 Z 分数模型有效结论[6] [7]。徐秀渠(2010)修改原 Z 分数模型的警戒值, 使得样本数据的 Z 值落到修改后的相应预警区间内[8]; 叶海英(2011)关注 ST 公司 Z 值, 却忽略了非 ST 公司 Z 值是否落在相应预警区域内, 仅从样本 ST 公司 Z 值小于 1.81 推出 Z 分数模型有效[9]。

同样, 反对 Z 分数模型有效研究方法也有多种。周剑涛(2009), 陈珍静(2011)通过研究几十家 ST 和非 ST 公司的财务数据, 发现 Z 分数模型在验证我国上市公司中的破产公司时有较高的准确度, 而作为预测非 ST 公司财务状况模型时, 具有很大的误判性[10] [11]。张楚堂, 杨志强(2009)把公司分成业绩优秀、业绩良好、业绩一般、业绩很差 4 类, 用每一类公司整体的平均 Z 值判断, 发现 Z 分数模型对业绩优良企业解释效果最好, 对破产公司的解释效果最差[12]。邵丽娜(2011)通过选取沪深股市共 193 家公司 2007 年~2010 年的财务数据, 对每一年 ST 公司和非 ST 公司的 Z 值进行分析, 发现 Z 模型对 ST 公司和非 ST 公司的准确率都不理想, 得出 Z 分数模型并不适用于我国上市公司的财务风险预警[13]。

然而上述两派的研究方法都存在缺陷。首先, 支持 Z 分数模型的主流研究方法是以 ST 公司和非 ST 公司的 Z 均值与预警临界值比较得出结论。虽然 ST 公司 Z 均值与非 ST 公司 Z 均值有所区别, 但 ST 公司的 Z 值区间和非 ST 公司的 Z 值区间存在高度重叠, 因此用 Z 均值的方法难以预判具体公司的财务风险水平。其次, 在支持派别的研究方法中, 个别样本数据选择过小, 使得结论的可靠性降低。比如只选取 3 到 5 家 ST 公司验证, 推出 Z 分数模型有效的结论。此外, 有的研究方法只验证 ST 公司的 Z 值, 而非 ST 公司的 Z 值能否与破产公司 Z 值区分开, 无从得知。否定 Z 分数模型派别的主要理由是非 ST 公司和 ST 公司的 Z 值并没有明确的分界线, 因此预测失效。否定派别的缺陷在于实验数据规模小, 不具代表性, 且仅呈现 ST 公司与非 ST 公司 Z 值无明显分界这一实验现象, 既未从原理上解释原因, 也未总结出规律和提出有效的解决方法。

3. 基于上市公司财报的 Z 分数模型有效性分析

在前人研究的基础上, 本节以较大规模的样本数据集分析 Z 分数模型的有效性。最终获取了 2007~2013 年 913 家沪市上市公司的财务报表数据, 同时判定 ST 公司和非 ST 公司的 Z 值, 并进行分组区分和综合比较。分析结果表明: ST 产公司与非 ST 公司的 Z 值线性不可分, 因此采用 Z 分数模型预测我国公司财务风险时存在方法上的缺陷。

3.1. 数据准备

选取沪市 913 家上市公司 2007 年至 2013 年 7 年中每年被评为 ST 公司及非 ST 公司的财务数据作为数据样本, 通过网上爬取的方式获得数据。为计算 Z 分数模型中的 X_1 到 X_5 五个指标, 需要利润分配表、资产负债表、利润分配表摘要、股票成交概况、股本结构五大类报表。为获取每个公司在每一年是否被评为 ST 公司, 需知每个公司每年的股票简称, 该信息可通过金融界上企业的半年报列表获得。最终从金融界网站上爬取利润分配表、资产负债表、利润分配表摘要、半年报列表, 从新浪财经上爬取股本结构, 并从上海证券交易所爬取股票成交概况。爬取的这些报表均为网页, 以 html 文件的形式存放在本地。通过分析不同网站不同类型财务报表数据的网页代码结构, 编写数据抽取工具从本地的 html 文件中自动解析和抽取每类报表中的财务指标数据存放到数据库中。最后, 从这 6 张数据表中选取所需的财务指标计算 X_1 到 X_5 五个指标, 并获取每个企业每年的 ST 标识, 关联后得到建模数据表, 至此数据准备过程完成。详细的数据基础如表 1 所示。

3.2. Z 分数模型有效性实证分析

基于上述数据的分析结果表明: 根据 Z 值大小判断一个企业是否有财务风险, 并不能取得很好的效果。以 2011 年至 2012 年国内 800 多家沪市上市公司中每年被评为 ST 公司的企业和当年的非 ST 公司企业的财务数据作为样本, 通过计算 Z 值发现 Z 分数模型并不能很好的区分 ST 公司和非 ST 公司企业。因 2011 年结果与 2012 年类似, 图 1 仅展示了 2012 年 Z 分数模型的预测结果。图 1 中的 X 轴表示公司编号, Y 轴表示 Z-score 值, 红色表示 ST 公司, 黑色表示非 ST 公司。

图 1 中完整呈现 800 多家上市公司的 Z 值, 其中右上角子图进一步放大显示了 Z 值取值 -1 到 5 时相关企业的分布情况。可以看出 ST 公司和非 ST 公司在特征空间中交叉混杂分布, 并不存在明确的分界线, 例如按照 $Z = 1.81$ 并不能区分 ST 公司和非 ST 公司。这一结果表明: ST 公司与非 ST 公司的 Z 值是线性不可分的, 本质上为线性区分方法的 Z 分数模型不能够分析国内企业的财务风险。

表 2 进一步对在 Z 值各区域 ST 公司和非 ST 公司的分布情况给出统计, 更加完整地反映了 Z 分数模

Table 1. Overview of adopted data sets

表 1. 数据基础统计表

年份	非破产公司数	破产公司数	原始数据表	抽取指标
2007	667	79	利润分配表	营业收入
2008	675	72	资产负债表	流动资产合计、流动负债合计、 资产总计、盈余公积金、未分配利润、负债合计
2009	688	64	利润分配表摘要	息税前利润
2010	660	77	股票成交概况	年末收市价
2011	682	72	股本结构	流通 A 股、流通 B 股、流通 H 股、总股本
2012	720	61	半年报列表	企业简称
2013	742	26		

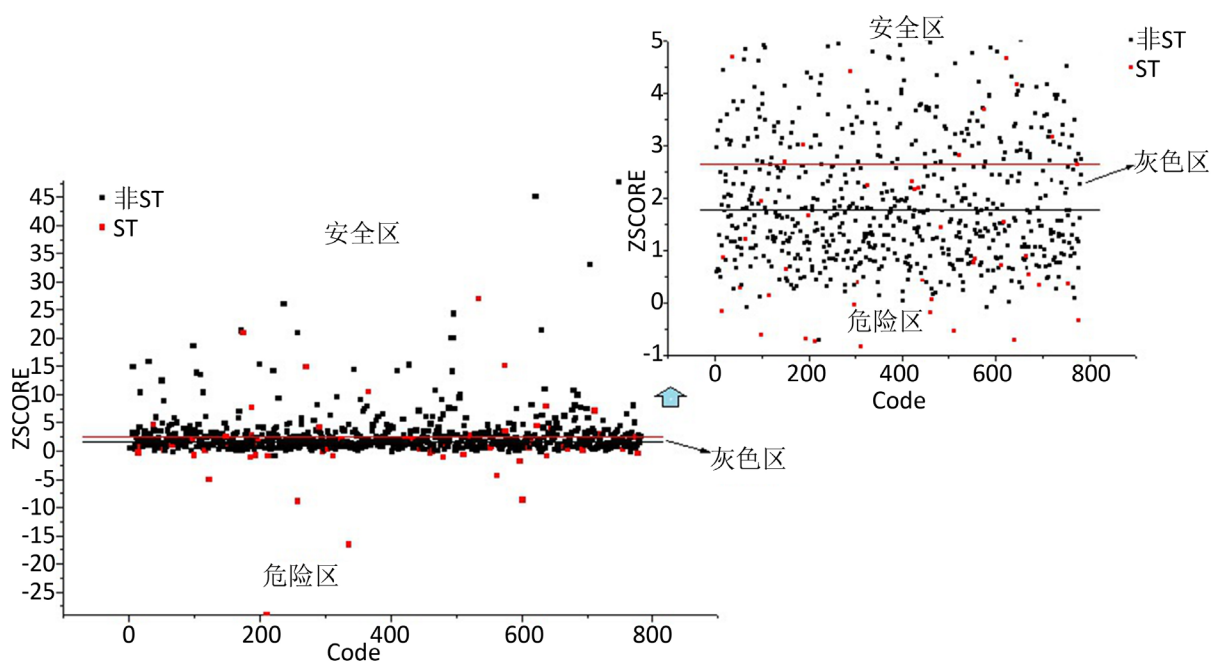


Figure 1. Prediction result of enterprise financial risk in 2012

图 1. 2012 年企业财务风险预测结果图

Table 2. Overview of prediction result of Z-Score model in 2012

表 2. 2012 年 Z 分数模型预测结果统计

真实值/预测值	破产区	灰色区	安全区	合计	准确率	覆盖率
破产公司	38	8	15	61	9.92%	62.30%
非破产公司	345	158	217	720	93.53%	30.14%
合计	383	166	232	781		

型的失效情况。这里只对 2012 年做出说明，其他年份类似。从表 2 中统计得出，Z 分数财务风险预警模型对我国上市公司的预测结果并不理想，要么准确率低要么覆盖率低，或者两者同时都低。2011 年预测破产公司的准确率为 13.38%，覆盖率为 52.78%，2012 年预测 ST 公司的准确率仅为 9.92%，覆盖率为 62.30%。

4. 基于决策树的财务风险预警模型

前述分析表明：传统的 Z 分数模型并不适用于我国上市公司的财务风险预警。因此，有必要探索更好的方法预测企业的财务风险。

数据挖掘中的分类算法能够从已知类别的历史数据中学习到一个模型，然后用该模型对未来未知类别的数据进行类别预测，符合对公司进行财务风险预警的场景，即通过对历史已知的是否有财务风险的公司样本的学习，预测未来的未知财务风险的公司样本是否有财务风险。同时，分类算法中有很多算法都能解决样本的线性不可分问题，这也弥补了前面提到的 Z-score 模型本质上只能解决线性可分问题的缺陷。

分类算法常用的有贝叶斯网络、支持向量机、决策树等。这些算法都能解决样本线性不可分的问题。但是贝叶斯、支持向量机等这些算法的模型结果的可解释性较差。J. Han. M. Kamber (2000) 提出决策树算法从一组无次序、无规则的实例中推理出以决策树表示的分类规则，用这些规则来预测未知类别记录的

类别[14], 这些规则可以将线性不可分的实例区分开。而且, 通过这些规则使用者可以知道每一类的样本有什么特征, 并可以从财务角度识别模型的合理性及指标的意义, 解释性较好。

4.1. 基于决策树的财务风险预警模型

本节选择决策树算法对我国上市公司的财务风险进行预测。由于 X_1 到 X_5 这 5 个指标仍然能够反映公司的财务状况, 因此构造决策树判定模型的目的是找出五项 Z 分数指标和企业是否存在财务风险间的关系, 用它来预测风险未知的企业的类别。它将全部已知是否有财务风险的企业作为训练样本, 将全部样本集作为初始节点 S_0 , 采用自顶向下的递归分裂方法, 层层筛选出最优的 Z-分数指标作为当前节点的分裂依据, 形成新的分支节点 S_i , 直至每个叶节点中的企业都属于同一类别为止。构建决策树判定模型的具体算法流程如下:

输入: 已知破产与否的企业样本, 即训练集

1. 训练集中的所有记录看作一个节点 S_0 ;

2. 遍历 X_1 到 X_5 每个指标的每种分裂方式, 找到每个指标最好的分裂点, 再比较每个指标最好的分裂点, 选取最好的分裂指标;

3. 根据第 2 步选取的分裂指标的分割点, 将训练集分成两个节点 S_1 和 S_2 ;

4. 对 S_1 和 S_2 分别执行 2~3 步, 直到每个节点中的实例都属于一类为止。

输出: 决策树, 该树的每个分支是一个判定条件, 叶子节点表示所属类别。

上市公司数据中 ST 与非 ST 样本数量存在明显的不平衡性, 即正常公司的数量远多于破产公司的数量, 例如 2007~2013 年国内上市公司财务数据中非 ST 公司是 ST 公司的 11.7 倍。这会对基于决策树算法的结果模型产生影响, 使模型具有偏向性, 具体表现为: 对少数类的识别率远低于多数类的识别率。因此, 我们采用非平衡数据和数据平衡 2 种方式构建决策树判定模型, 然后通过对比两种方式的实验结果, 选取比较合理且比较实用的方式。

第一种方式是在模型中保持数据原样, 不对其进行平衡处理。建模选择 2007 年至 2010 年被评为 ST 的公司和当年的非 ST 公司作为训练集来训练决策树模型, 用 2011 年至 2013 年被评为 ST 的公司和当年的非 ST 公司作为测试集来评价模型的好坏。而数据平衡方式是指对训练集数据作适当参数平衡, 测试集数据仍保持原样。参数平衡是指在训练阶段, 通过对不同类别的实例权重设置来改变数据的分布情况, 在此模型中, 就是通过复制较少类别(ST 公司)的实例, 以减少数据的不平衡程度。在训练阶段, 经过多次试验得出, 当 ST 公司的权重设为 2, 非 ST 公司的权重设为 1 时, 得到的模型效果较好。用 2007 年至 2010 年被评为 ST 公司的企业和当年的非 ST 公司企业进行 2 倍参数平衡后作为训练集来训练决策树模型, 用 2011 年至 2013 年被评为 ST 公司的企业和当年的非 ST 公司企业进行测试。非平衡数据下的决策数模型如图 2 所示, 参数 2 倍平衡数据下的决策树模型与图 2 类似, 不再赘述。

4.2. 基于决策树的财务风险预警模型规则解释

通过前述的决策树模型, 可以得到判定破产的重要规则, 模型规则如表 3 所示。

从未平衡数据和参数 2 倍平衡的两种模型结果来看, 规则几乎是一致的。两种方式下的决策树的顶点都是以 X_2 作为第一次的分裂指标, (如决策树图), 下一个分裂指标为 X_1 。由于未平衡数据状态下的规则与参数 2 倍平衡状态下的规则类似, 因此仅对未平衡数据状态下的规则进行解释:

规则(1) “ $X_2 \leq -0.095$ 且 $X_1 \leq -0.23 \rightarrow$ 破产危机 = 是(188,90.0%,1)”表明: 如果一个公司满足 $X_2 \leq -0.095$ 并且 $X_1 \leq -0.23$, 那么此公司会被判断有破产危机, 即有较高的财务风险; 此次实验中 188 家破产公司中的 90% 都满足上述条件。规则中的 $X_2 =$ 留存收益/总资产, 留存收益是指企业从历年实现的利润中提取或留存于企业的内部积累, 它来源于企业的生产经营活动所实现的净利润, 该指标反映企业累积获利能力; $X_1 =$ 营运资产/总资产, 营运资产 = 流动资产 - 流动负债, 该指标反映企业的流动性, 流动

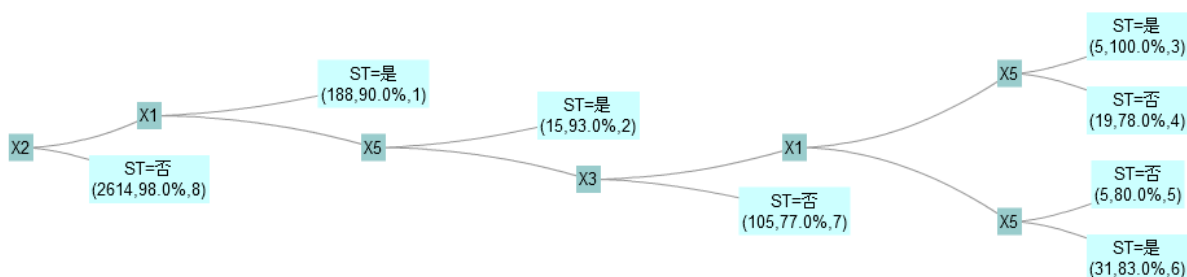


Figure 2. Decision tree under the imbalanced data

图 2. 非平衡数据状态下的决策树

Table 3. Rules of model for financial risk judgment

表 3. 判定财务风险的模型规则

未平衡 数据训练规则	(1) $X_2 \leq -0.095$ 且 $X_1 \leq -0.23 \rightarrow$ 破产危机 = 是(188,90.0%,1)
	(2) $X_2 \leq -0.095$ 且 $X_1 > -0.23 \rightarrow X_5 \leq 0.10 \rightarrow$ 破产危机 = 是(15,93.0%,2)
	(3) $X_2 \leq -0.095$ 且 $X_1 > -0.23$ 且 $X_5 > 0.10$ 且 $X_3 \leq 0.016$ 且 $X_1 > -0.054$ 且 $X_5 > 0.26 \rightarrow$ 破产危机 = 是(31,83.0%,6)
参数2倍 平衡训练规则	(1) $X_2 \leq -0.078$ 且 $X_1 \leq -0.23 \rightarrow$ 破产危机 = 是(362,94.0%,1)
	(2) $X_2 \leq -0.078$ 且 $X_1 > -0.23$ 且 $X_5 \leq 0.10 \rightarrow$ 破产危机 = 是(33,96.0%,2)
	(3) $X_2 \leq -0.078$ 且 $X_1 > -0.23$ 且 $X_5 > 0.10$ 且 $X_3 \leq 0.016$ 且 $X_1 > -0.054$ 且 $X_5 > 0.26 \rightarrow$ 破产危机 = 是(65,86.0%,8)

资本越多,说明不能偿债的风险越小。也就是说,在国内现行条件下,当一个公司留存收益/总资产 $\leq -0.95\%$ 时,即积累利润已经亏空并且亏空大于总资产的 9.5%,同时企业在短期经营中已经资不抵债,流动负债超过流动资产的金額达到总资产的 23%时,这个企业很可能面临财务危机。

规则(2) “ $X_2 \leq -0.095$ 且 $X_1 > -0.23$ 且 $X_5 \leq 0.10 \rightarrow$ 破产危机 = 是(15,93.0%,2)”表明:如果一个公司满足 $X_2 \leq -0.095$, $X_1 > -0.23$, 并且 $X_5 \leq 0.10$, 那么此公司会被判定有破产危机。当一个公司留存收益/总资产 $\leq -0.95\%$ 时,即积累利润已经亏空并且亏空大于总资产的 9.5%,即便 X_1 (营运资产/总资产) $>$ 临界值-23%时,仍要继续判断 $X_5 =$ 销售收入/总资产是否小于 10%。也就是说即便一个公司在亏空不超过总资产 9.5%,流动性也还可以的时候,如果销售收入跟不上(不到总资产 10%),那么这个公司也有很大可能陷入较高的财务风险。

其他规则以此类推。由上述模型结果得出:最重要的判定指标是 $X_2 =$ 留存收益/总资产, $X_1 =$ 营运资产/总资产, 和 $X_5 =$ 销售收入/总资产。并可得出,决策树模型判定的规律是,首先判断公司留存收益的亏空程度,再判断其短期流动性,继而判断其销售收入与公司规模的符合程度,一般情况下,亏空不高于总资产的 10%,流动负债超过流动资产的金額不高于总资产的 23%,销售收入不低于总资产的 10%。如果一个公司亏空大,且流动性差,那么很有可能面临财务危机;如果一个公司亏空大,销售收入跟不上,即便流动性较好,也很可能陷入财务危机。

4.3. 基于决策树的财务风险预警模型验证

对于非平衡数据模型和数据平衡模型,我们都选择相同的测试集,即 2011~2013 年上市公司的财务数据来检验模型,检验结果如表 4 所示。

在非平衡数据状态下,从模型验证结果表 4 可以得出,提前三年预测无破产危机公司的准确率和覆盖率都达 95% 以上。对于有破产危机公司的预测,在准确率上表现较理想,维持在 75% 以上的水平,在覆盖率上稍差一些,维持在 60% 左右。提前一年预测破产公司的财务风险的准确率为 95.45%,覆盖率为

Table 4. Prediction result of decision tree under balanced and imbalanced data with the duplicated rare instances
表 4. 参数 2 倍平衡与非平衡数据状态下决策树预测结果

预测年份		未来一年内									
训练年份	真实值\预测值	非平衡数据状态				参数平衡数据状态					
		是	否	合计	准确率	覆盖率	是	否	合计	准确率	覆盖率
	是	42	30	72	95.45%	58.33%	53	19	72	75.71%	73.61%
	否	2	680	682	95.77%	99.71%	17	665	682	97.22%	97.51%
	合计	44	710	754			70	684	754		
		未来两年内									
训练年份	真实值\预测值	非平衡数据状态				参数平衡数据状态					
		是	否	合计	准确率	覆盖率	是	否	合计	准确率	覆盖率
2007~2010	是	85	48	133	88.54%	63.91%	103	30	133	75.74%	77.44%
	否	11	1391	1402	96.66%	99.22%	33	1369	1402	97.86%	97.65%
	合计	96	1439	1535			136	1399	1535		
		未来三年内									
训练年份	真实值\预测值	非平衡数据状态				参数平衡数据状态					
		是	否	合计	准确率	覆盖率	是	否	合计	准确率	覆盖率
	是	101	58	159	75.37%	63.52%	119	40	159	62.63%	74.84%
	否	33	2111	2144	97.33%	98.46%	71	2073	2144	98.11%	96.69%
	合计	134	2169	2303			190	2113	2303		

58.33%；提前 2 年预测的准确率为 88.54%，覆盖率为 63.91%；提前 3 年预测的准确率为 75.37%，覆盖率为 63.52%。在参数平衡数据状态下，从表 4 的模型验证结果看出，结果与非平衡数据状态下类似，在未来三年内的无破产危机公司的预测都比较准确，准确率和覆盖率都达 95% 以上。而对于有破产危机公司的预测，数据 2 倍平衡下的结果比数据非平衡下的结果略胜一筹。虽然准确率要比非平衡数据的结果稍差一些，但在覆盖率上能达到 70% 以上。提前 1 年预测破产的准确率为 75.71%，覆盖率为 73.61%；提前两年预测破产的准确率为 75.74%，覆盖率为 77.44%；提前三年预测破产的准确率为 62.63%，为 74.84%。总体看来，在参数 2 倍平衡状态下的模型在准确率和覆盖率两方面表现更均衡一些，但值得注意的是，随着预测年份的推移，两种模型的准确度都有所下降。

5. 总结

通过对传统 Z 分数模型有效性实证分析，发现国内有破产危机的公司与正常公司 Z 值本身的非线性可分性，因此基于多元线性回归函数的 Z 分数模型的预测效果并不理想，已不能完全反映当代国内企业的财务风险与危机。

基于决策树算法的财务风险预警模型为我国公司的财务风险预警提供了一种有效的方法。它有能力解决有破产危机的公司与正常公司的非线性可分的特点。它可根据所要预测的年份选择最近的数据进行训练，以保证模型跟进现行市场变化，也可根据模型规则判断财务风险的关键领域及风险程度。该模型在非平衡数据与参数 2 倍平衡数据状态下的模型都表现良好，后者在准确率和覆盖率两项的表现上更为均衡。在两种方式下，对正常公司的判定准确率和覆盖率都能达 95% 以上，对有破产危机的公司的判断

准确率和覆盖率几乎都能达 75% 以上。值得注意的是, 虽然决策树模型能较为准确的预判公司是否面临财务风险, 并且通过指标可以判断所面临财务风险的程度, 但随着时间的推移, 模型准确率会有所下降。因此预判公司是否有财务风险时, 需要选择近几年的数据做训练。这样训练出的模型才符合实情, 也能提高模型的精确性。

基金项目

国家科技支撑计划项目《面向国家审计“免疫系统”的审计模拟与仿真平台研发及应用示范》(批准号: 2012BAH08B02), 国家自然科学基金项目《基于社会计算的社会保障国家审计政策功能研究》(批准号: 71272216)。

参考文献 (References)

- [1] 常树春, 范文雨. 基于 Z-score 模型的汽车行业上市公司财务风险实证分析[J]. 商业会计, 2013(13): 89-91.
- [2] 陈珍静. Z-score 模型在我国上市公司财务预警中适用性的探讨——基于交通运输设备制造业的实证分析[J]. 国际商务财会, 2011(4): 60-63.
- [3] 胡剑. 论企业财务风险的防范[J]. 会计师, 2014(4x): 47-48.
- [4] 李艳. 基于 Z-Score 模型对我国社会服务业上市公司破产预测的实证分析[J]. 商, 2012(7): 84, 109.
- [5] 邵丽娜. Altman's Z-score 模型对我国上市公司的适用性分析[J]. 财会研究, 2011(24): 49-51.
- [6] 王芳云. 运用 Z 计分模型对上市公司财务风险的实证研究[J]. 山东纺织经济, 2005(5): 42-44.
- [7] 向德伟. 运用“Z 记分法”评价上市公司经营风险的实证研究[J]. 会计研究, 2002(11): 53-57.
- [8] 徐秀渠. Altman's Z-Score 模型在企业风险管理中的应用研究[J]. 经济经纬, 2010(4): 103-106.
- [9] 叶海英. 对上市公司财务预警的实证研究——Z 计分模型[J]. 中国外资, 2011(24): 84-85.
- [10] 张楚堂, 杨志强. 基于 Z-score 模型的国有公司财务分析[J]. 武汉理工大学学报, 2009(15): 145-149.
- [11] 张蔚虹, 朱海霞. Z-Score 模型对科技型上市公司财务风险预警的适用性检验[J]. 科技管理研究, 2012, 32(14): 228-231.
- [12] 周剑涛, 庞文凤. Z-SCORE 财务预警模型在上市公司应用的实证研究[J]. 财会通讯, 2009(24): 25-27.
- [13] Altman, I.E. (1968) Financial Ratios Discriminant Analysis and the Prediction of Corporate Bankruptcy. *Journal of Finance*, 4, 589-609.
- [14] Han, J. and Kamber, M. (2000) Data Mining: Concepts and Techniques.

知网检索的两种方式:

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择: [ISSN], 输入期刊 ISSN: 2160-7311, 即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入, 输入文章标题, 即可查询

投稿请点击: <http://www.hanspub.org/Submission.aspx>

期刊邮箱: mm@hanspub.org