

基于关联规则的企业财务风险分析

吕 烜, 莫宏敏, 张 勇

(吉首大学 数学与统计学院, 湖南 吉首 416000)

摘要:企业对自身财务风险的管理直接影响企业的可持续性发展,具有重要意义。基于 2015 年我国上市 ST(Special Treatment)公司财务指标数据对公司财务风险进行分析。首先,对财务数据进行预处理,运用关联规则,挖掘出隐藏在财务指标中的规则,最后选取具有较高代表性的财务指标,对公司财务风险进行分析预警。实证分析表明:当上市公司财务出现危机时,总是存在一些集中出现的财务指标,它们的波动对企业财务风险有直接影响。通过分析这些指标可以对上市公司财务风险进行预警防范。

关键词:关联规则;Apriori 算法;财务风险分析

中图分类号:F275;F830 **文献标志码:**A **文章编号:**1671-1807(2017)12-0131-05

随着我国经济发展进入新常态,企业发展的环境愈发复杂,因此,企业在经营中将面临更多的不确定性,这些不确定性就给企业带来了更大的风险。企业对自身存在的财务风险进行管理直接影响到企业自身的可持续性发展。企业如果不对风险进行管理控制,就容易遭受巨大的损失,甚至破产。因此,企业财务风险分析及预防至关重要。大数据时代的核心是从数据到决策^[1],但在企业财务数据分析过程中,传统统计模型较为严格的假设条件以及复杂的计算过程大大地限制了它的实用价值。

随着数据挖掘技术的发展,企业的财务风险分析迎来了更多新的研究方法。如庄慧和饶扬胜^[2]利用 Logistic 模型对上市公司财务困境进行了预测,结果表明该模型对企业未来财务状况具有较强的预警能力,但该研究存在样本数量不足的问题。苏鲁娜^[3]利用 Z 分数模型对 A 公司进行了破产分析,列举了几个明显而重要的风险,提出相应的财务风险防范措施,但未能及时发现其他存在的风险,且未能选取相应的灵敏指标设定日常预警机制。罗旻婧^[4]运用时间序列方法对财务风险进行分析预测,具有较高准确性,但局限于有周期性规律的财务数据。徐力^[5]结合了多种数据挖掘技术,对制造业上市公司进行了财务诊断研究,结果认为不同挖掘方法各有优势,研究时

应根据需要选取具体的模型,但只选取了制造业板块,缺乏广泛性,不能很好地反映上市公司整体情况。李青^[6]利用遗传算法建立了财务危机预测模型,且模型具有较高准确率。Altman^[7]在 1968 年经过对五项财务指标加权汇总,得出 Z 值来对企业的财务风险进行预测。1977 年,Altman 与 Haldemand 等人^[8]在 Z 值公式模型的基础上,构建了 ZETA 评分模型。

基于已有的研究理论,本文采用数据挖掘中的关联规则挖掘技术,通过对我国上市 ST 公司财务指标的挖掘,探寻隐藏在各财务指标中的规则,挖掘出了净利润、销售毛利率、流动资产周转率、净资产增长率、总资产增长率、流动比率、资产的经营现金流量回报率等 7 个具有较高代表性的财务指标,然后对企业财务风险防范提出建议。

下文的结构安排是:第二部分对关联规则挖掘的原理进行了简单介绍;第三部分是基于关联规则的企业财务风险分析实证研究;第四部分给出了实证研究的结论与分析。

1 基于关联规则的企业财务风险分析模型

基于关联挖掘的企业财务风险分析主要包括以下三个阶段:

第一阶段,企业财务风险指标体系构建。

收稿日期:2017-09-12

基金项目:国家自然科学基金项目(11461027);湖南省教育厅高校科研一般项目(15C1156);吉首大学研究生科研创新项目(JGY201758)。

作者简介:吕烜(1993-),男,浙江永康人,吉首大学,硕士研究生,研究方向:经济统计;莫宏敏(1969-),男(土家族),湖南慈利人,吉首大学,副教授,博士,研究方向:矩阵理论与计算,金融数学与统计精算;通讯作者:张勇(1976-),男,湖南怀化人,吉首大学,讲师,博士,研究方向:数量金融。

第二阶段,企业财务风险分析模型构建。

第三阶段,企业财务风险防范预警。

关联规则是数据挖掘中一种重要的挖掘技术,它从大量数据中挖掘数据之间隐藏的关联关系。如果数据属性之间存在关联,那么就可以依据其他属性值对数据的某一属性进行预测。关联规则挖掘问题主要包括两个方面内容:第一个是找出数据中符合支持度阈值的数据项集;第二是利用生成频繁项集的关联规则,根据预设的置信度阈值进行筛选,得到强关联规则^[9]。

关联规则算法有多种,其中最具代表性的是 Apriori 算法。本文基于 Apriori 算法建立企业财务风险分析模型如下:

设 I 是由所有上市公司组成的集合, $D = \{T_i, i = 1, 2, \dots, n\}$ 。 D 表示上市 ST 公司组成的集合, T_i 表示上市 ST 公司,且 $T_i \in D, T_i \in I$ 。 设 X, Y 都是上市 ST 公司 T_i 的财务指标或指标集合,且 $X \cap Y = \varnothing$, 如果 T_i 同时包含 X 和 Y , 那么就可以得到关联规则:

$$X \Rightarrow Y(S, C, I) \tag{1}$$

S 为支持度,表示满足条件的公司 T_i 在上市 ST 公司集合 D 中所占的比例,计算公式如下:

$$S(X \Rightarrow Y) = \frac{S(X \cup Y)}{|D|} = \frac{P(XY)}{|\{T: X \cup Y \subseteq T, T \in D\}|} \tag{2}$$

式(2)中, $|\cdot|$ 表示集合中的元素数目。在公

司 T_i 中所有满足最小支持度的项集,称为频繁项集。

C 为置信度,反映了 X 出现的条件下, Y 出现的条件概率,表示关联规则的强度,衡量了关联规则的准确度,计算公式如下:

$$C(X \Rightarrow Y) = \frac{S(X \cup Y)}{S(X)} = P(X|Y) \tag{3}$$

通过关联规则对企业财务风险指标挖掘,根据要求调整支持度和置信度阈值,找寻出各财务指标之间的关系。选取具有高代表性的指标进行分析,对企业财务风险提出预警。

2 基于关联规则的企业财务风险分析实证研究

2.1 样本企业的选取

本文选取上市公司中的 ST 公司作为样本,主要考虑 ST 公司连续两年出现亏损,企业财务风险问题比较突出,具有较强的代表性。由于部分公司 2016 年年报暂时还未公布,所以本研究选取了 2015 年年底上市公司中的 64 家 ST 公司,剔除其中存在异常数据和缺失数据的公司,最后保留 48 家公司作为样本。样本数据来源于锐思数据库。

2.2 财务风险指标的选择

企业财务风险主要体现在盈利能力、营运能力、成长能力、偿债能力和现金流量这五个方面。在财务风险分析指标的选取上,应尽可能多地收集相关指标。本文选取了 29 个财务指标,具体指标如表 1 所示。

表 1 符号说明

符号	财务指标	符号	财务指标
X_1	每股收益	X_{16}	流动资产周转率
X_2	每股净资产	X_{17}	营业收入增长率
X_3	净资产收益率	X_{18}	净利润增长率
X_4	经营活动产生的现金流量净额	X_{19}	净资产增长率
X_5	每股净现金流量	X_{20}	总资产增长率
X_6	每股经营活动现金流量净额	X_{21}	每股收益增长率
X_7	净利润	X_{22}	股东权益增长率
X_8	销售净利率	X_{23}	流动比率
X_9	销售毛利率	X_{24}	速动比率
X_{10}	营业收入	X_{25}	资产负债率
X_{11}	每股营业收入	X_{26}	资产的经营现金流量回报率
X_{12}	应收账款周转率	X_{27}	资产的经营现金流量回报率
X_{13}	应收账款周转天数	X_{28}	经营现金净流量对负债比率
X_{14}	存货周转率	X_{29}	经营净现金流量/流动负债(现金流量比率)
X_{15}	存货周转天数		

由于财务指标之间存在一定的相关性,所以还需进一步对选取的财务指标进行相关性分析,剔除相关性较高的财务指标,以减少指标间的共线性。

通过对相关系数分析,去除相关性较高的 x_4 经营活动产生的现金流量净额、 x_{14} 存货周转率、 x_{18} 净利润增长率、 x_{22} 股东权益增长率、 x_{24} 速动比率、 x_{28}

经营现金净流量对负债比率这 6 个财务指标,保留剩下的 23 个财务指标。

2.3 风险等级划分以及财务指标数据库重构

不同设计者对财务风险的感受不一样,因此会有不同的财务风险划分标准。根据数据特征,结合赵春^[10]所定义的财务风险等级划分标准,本文将财务风险等级划分如表 2 所示。

表 2 企业财务指标风险等级定义表

财务风险指标	a 级	b 级	c 级	d 级	e 级
每股收益(元)	>1	[1,0.3]	[0.3,-0.3]	[-0.3,-1]	<-1
每股净资产(元)	>5	[5,2]	[2,1]	[1,0.5]	[0.5,0]
净资产收益率	>0.4	[0.4,0.1]	[0.1,-0.1]	[-0.1,-0.4]	<-0.4
每股净现金流量(元)	>1	[1,0.5]	[0.5,0]	[0,-0.5]	<-0.5
每股经营活动现金流量净额(元)	>3	[3,1]	[1,0]	[0,-0.5]	<-0.5
净利润(百万元)	>500	[500,100]	[100,-100]	[-100,-500]	<-500
销售净利率	>0.4	[0.4,0.1]	[0.1,-0.1]	[-0.1,-0.4]	<-0.4
销售毛利率	>0.6	[0.6,0.2]	[0.2,-0.2]	[-0.2,-0.6]	<-0.6
营业收入(百万元)	>1000	[1000,100]	[100,50]	[50,20]	<20
每股营业收入(元)	>10	[10,5]	[5,2]	[2,1]	<1
应收账款周转率(次)	>100	[100,50]	[50,20]	[20,5]	[5,0]
应收账款周转天数	[0,30]	[30,60]	[60,90]	[90,180]	>180
存货周转天数	[0,30]	[30,60]	[60,90]	[90,180]	>180
流动资产周转率(次)	>10	[10,5]	[5,2]	[2,0.5]	[0.5,0]
营业收入增长率	>0.6	[0.6,0.2]	[0.2,-0.2]	[-0.2,-0.6]	<-0.6
净资产增长率	>1	[1,0.3]	[0.3,0.1]	[0.1,0]	<0
总资产增长率	>1	[1,0.3]	[0.3,0.1]	[0.1,0]	<0
每股收益增长率	>0.5	[0.5,0.1]	[0.1,-0.1]	[-0.1,-0.5]	<-0.5
流动比率	>2	[2,1]	[1,0.5]	[0.5,0.3]	[0.3,0]
资产负债率	[0,0.2]	[0.2,0.4]	[0.4,0.6]	[0.6,0.8]	[0.8,1]
资产的经营现金流量回报率	>0.5	[0.5,0.1]	[0.1,-0.1]	[-0.1,-0.5]	<-0.5
经营现金净流量与净利润的比率	>1	[1,0.5]	[0.5,0]	[0,-0.5]	<-0.5
经营净现金流量/流动负债	>1	[1,0.5]	[0.5,0]	[0,-0.5]	<-0.5

数据库重构是为了按照定义划分的风险等级将连续的财务指标转换成离散数据,以便进行后续的关联规则数据挖掘。但样本数据中的异常值会使企业财务风险的分析产生较大偏差。因此,在数据重构前,我们对样本数据进行了数值分析,剔除样本中的异常数据,最终留下 30 家企业样本。经重构的数据库如表 3 所示。

2.4 企业财务风险指标关联规则挖掘

利用 spss modeler 对重构后的财务指标数据进行关联规则挖掘。设置了三组不同的支持度和置信度阈值,得到了不同的关联规则。在不同支持度和置信度阈值下关联规则挖掘结果如下所示:

表 3 财务指标重构表

股票代码	x_1	x_2	x_3	x_5	x_6	x_7	...
600701	c	c	c	a	c	e	...
600212	d	c	e	d	c	e	...
000962	e	b	d	d	c	e	...
002513	d	a	c	d	b	e	...
000913	c	b	c	d	c	e	...
002061	d	c	d	d	c	e	...
600721	e	c	e	d	b	e	...
002069	d	c	d	d	c	e	...
000670	c	d	c	d	c	e	...
000037	e	c	e	b	c	e	...
000950	d	b	d	c	c	e	...
...	...	...	...	...	...	...	...

当同时设置支持度和置信度阈值为 0.7 时,共计发现 81 条关联规则,涉及 x_7 净利润, x_9 销售毛利率, x_{16} 流动资产周转率, x_{19} 净资产增长率, x_{20} 总资产增长率, x_{23} 流动比率, x_{26} 资产的经营现金流量回报率等 7 个财务指标,结果如 4 所示。

表 4 $s=0.7, c=0.7$ 关联规则表

后项	前项	支持度	置信度
x_{19e}	x_{20e}	0.766 67	1
x_{7e}	x_{20e}	0.766 67	1
x_{7e}	x_{19e}	0.9	1
x_{19e}	x_{20e} 和 x_{9c}	0.733 33	1
x_{7e}	x_{20e} 和 x_{9c}	0.733 33	1
x_{7e}	x_{20e} 和 x_{19e}	0.766 67	1
x_{19e}	x_{20e} 和 x_{7e}	0.766 67	1
x_{7e}	x_{9c} 和 x_{9e}	0.833 33	1
x_{7e}	x_{26c} 和 x_{19e}	0.8	1
x_{7e}	x_{20e} 和 x_{9c} 和 x_{19e}	0.733 33	1
x_{19e}	x_{20e} 和 x_{9c} 和 x_{7e}	0.733 33	1
x_{7e}	x_{9c} 和 x_{26c} 和 x_{19e}	0.766 67	1
x_{7e}	x_{9c}	0.9	0.962 96
x_{7e}	x_{26c}	0.9	0.962 96
x_{19e}	x_{9c} 和 x_{7e}	0.866 67	0.961 54
...	...	...	...

当同时设置支持度和置信度阈值为 0.8 时,共计发现 38 条关联规则,涉及 x_7 净利润, x_9 销售毛利率, x_{19} 净资产增长率, x_{20} 总资产增长率, x_{26} 资产的经营现金流量回报率等 5 个财务指标,结果如表 5 所示。

表 5 $s=0.8, c=0.8$ 关联规则表

后项	前项	支持度	置信度
x_{7e}	x_{19e}	0.9	1
x_{7e}	x_{9c} 和 x_{19e}	0.833 33	1
x_{7e}	x_{26c} 和 x_{19e}	0.8	1
x_{7e}	x_{9c}	0.9	0.962 96
x_{7e}	x_{26c}	0.9	0.962 96
x_{7e}	x_{9c} 和 x_{7e}	0.866 67	0.961 54
x_{7e}	x_{9c} 和 x_{26c}	0.833 33	0.96
x_{9c}	x_{26c} 和 x_{19e}	0.8	0.958 33
x_{19e}	x_{9c} 和 x_{26c} 和 x_{7e}	0.8	0.958 33
x_{9c}	x_{26c} 和 x_{19e} 和 x_{7e}	0.8	0.958 33
x_{19e}	x_{7e}	0.966 67	0.931 03
x_{26c}	x_{9c}	0.9	0.925 93
x_{9c}	x_{26c}	0.9	0.925 93
x_{19e}	x_{9c}	0.9	0.925 93
x_{9c}	x_{19e}	0.9	0.925 93
x_{9c}	x_{19e} 和 x_{7e}	0.9	0.925 93
x_{26c}	x_{9c} 和 x_{7e}	0.866 67	0.923 08
x_{9c}	x_{26c} 和 x_{7e}	0.866 67	0.923 08
...	...	...	...

当同时设置支持度和置信度阈值为 0.9 时,共计发现 9 条关联规则,涉及 x_7 净利润, x_9 销售毛利率, x_{19} 净资产增长率, x_{26} 资产的经营现金流量回报率等 4 个财务指标,结果如 6 所示：

表 6 $s=0.9, c=0.9$ 关联规则表

后项	前项	支持度	置信度
x_{7e}	x_{19e}	0.9	1
x_{7e}	x_{9c}	0.9	0.962 96
x_{7e}	x_{26c}	0.9	0.962 96
x_{19e}	x_{7e}	0.966 67	0.931 03
x_{26c}	x_{9c}	0.9	0.925 93
x_{9c}	x_{26c}	0.9	0.925 93
x_{19e}	x_{9c}	0.9	0.925 93
x_{9c}	x_{19e}	0.9	0.925 93
x_{9c}	x_{19e} 和 x_{7e}	0.9	0.925 93

3 结论

通过对上市 ST 公司财务指标的关联规则挖掘,我们发现当上市公司出现财务风险时,存在部分财务指标总是频繁出现,这些指标分别为: x_7 净利润, x_9 销售毛利率, x_{16} 流动资产周转率, x_{19} 净资产增长率, x_{20} 总资产增长率, x_{23} 流动比率, x_{26} 资产的经营现金流量回报率等 7 个指标。对以上 7 个指标,通过我国上市公司的实际数据进行测试分析,证明这些关键指标的波动直接影响公司的财务风险程度。

我们可以通过对这 7 个财务指标分析对企业财务风险进行如下分析：

在企业盈利能力方面,可以分析 x_7 净利润、 x_9 销售毛利率两个指标,它们直观体现了企业的盈利能力。净利润高,则说明企业的效益好,如果净利润低,则企业的效益就差。销售毛利率低,表明企业销售毛利占净利比率低,就表明了企业通过营业赚取利润的水平不高。

在企业营运能力方面,可以分析 x_{16} 流动资产周转率,它反映了企业流动资产的利用效率。该指标越高,表明流动资产周转速度越快,则企业营运能力越强。

在企业成长能力方面,可以分析 x_{19} 净资产增长率、 x_{20} 总资产增长率两个指标,这两个指标都能直观的看出企业在一段时间内的资产规模扩张的速度。

在企业偿债能力方面,可以分析 x_{23} 流动比率。流动比率体现了企业的变现能力和短期偿债能力,比率越高,能力越强。

在企业现金流量方面,可以分析 x_{26} 资产的经营现金流量回报率,它是经营活动产生的现金流量净额

与总资产的比值。该指标显示了企业的收现能力,回报率越高,收现能力越强。

参考文献

[1] 陈文伟,陈晟.从数据到决策的大数据时代[J].吉首大学学报:自然科学版,2014(3):31—36.

[2] 庄慧,饶扬胜.数据挖掘技术在上市公司财务困境预测中的应用[J].现代经济信息,2016(19):143—144.

[3] 苏鲁娜.A公司财务风险分析与防范研究[D].北京:首都经济贸易大学,2016.

[4] 罗旻婧.基于时间序列的财务风险预测研究[D].广州:广东财经大学,2016.

[5] 徐力.基于数据挖掘的上市公司财务诊断研究[D].杭州:浙江财经大学,2016.

[6] 李清.基于适度财务指标和遗传算法的财务危机预测模型研究[J].统计与信息论坛,2010(2):3—9.

[7] ALTMAN E I. Financial ratios,discriminant analysis,and the prediction of corporate bankruptcy[J]. The Journal of Finance,1968,23(4):589—609.

[8] ALTMAN E I, HALDEMAN R G, NARAYANAN P. ZETA TM analysis a new model to identify bankruptcy risk of corporations [J]. Journal of Banking & Finance,1977,1(1):29—54.

[9] 王振武,徐慧.数据挖掘算法原理与实现[M].北京:清华大学出版社,2015:35—53.

[10] 赵春.基于数据挖掘技术的财务风险分析与预警研究[D].北京:北京化工大学,2012.

Enterprise Financial Risk Analysis Based on Association Rules

LV Xuan, MO Hong-min, ZHANG Yong

(College of Mathematics and Statistics,Jishou University,Jishou Hunan 416000,China)

Abstract: It is of great significance for enterprises to manage their own financial risks, which affects the sustainable development directly. This paper is to analyze the financial risk of the company based on the financial indicators of China's listed companies under special treatment in 2015. At first, utilize the association rules and preprocess the financial data, then Dig out the rules hidden in the financial indicators, at last select a higher representative financial indicator and pre-warning the financial risk. Empirical analysis shows that: there is always some financial indicators when the listed company was in financial crisis. Their movement has a direct impact on enterprise financial risk. By analyzing these indicators, the listed companies financial risk can be early warned and prevented.

Key words: association rules; Apriori algorithm; financial risk analysis

(上接第 6 页)

A Empirical Analysis on Coordination Development of Economy-Technology-Environment System of Energy Intensive Enterprises

ZHENG Ji-liang, ZHOU Xuan

(Faculty of Management & Economics,Kunming University of Science & Technology,Kunming 650093,China)

Abstract: Under the situation of Chinese economic construction to be driven by science and technology, the coordination development of economy, technology (technology innovation), and environment system (energy conservation & emissions reduction) of an energy-intensive enterprise will be the key factor. An evaluation model and indicator system of economy-technology-environment system about an energy-intensive enterprise is been built up then. The one, two and three dimension coordination degree of economy-technology-environment system of three famous steel firms in China are analyzed empirically and comparatively. the results show that, in the research period, although the economic effect of three firms was fluctuate greatly, the coordination degree of economy-technology-environment system of three firms appeared the trend of steady growth, and the coordination degree of technology-environment system all are good.

Key words: energy-intensive enterprise;economy-technology-environment system;coordination development