

# 数据挖掘技术在财务风险分析中的应用

□张 苗

**【内容摘要】**进入 21 世纪,金融的信息化发展速度加快,各个机构已经搭建起来数据平台,逐步实现数据大集中,形成金融数据。与此同时,数据挖掘技术在过去几十年里得到长足发展,技术与方法日益完善,应用到各个领域。在金融领域运用数据挖掘技术,不但可以发挥出数据的作用,而且为企业的决策层提供了更有效的支持。为了在市场经济条件下提高竞争力,越来越多的公司对风险管理的需求更加重视。财务风险分析不仅成为企业发展过程中的一个重要课题,也是每一个企业管理者所追求的永恒目标。只有正确分析财务风险,才能够保证企业的持续运行,并为企业带来经济利润,从而促进国家经济的长效发展。从而可以看出,这项工作是很必要的。由于技术的不断进步,数据挖掘学科理论的逐渐完善,再加上数据挖掘技术在运用的过程中并不复杂,所以在对企业财务风险分析与危机预警机制中常常被人们所运用。

**【关键词】**数据挖掘; 财务风险; 交互挖掘; 时间序列; 关联规则

**【作者简介】**张苗(1993.11~),女,安徽合肥人; 贵州财经大学硕士研究生; 研究方向: 金融大数据分析

## 一、数据挖掘的相关理论

(一) 数据挖掘的定义。数据挖掘,是指从隐性的、过去未知的、潜在有用的数据中提取信息的过程。数据挖掘(DM)这个概念是最近几年来,由于人工智能的兴起和数据库的发展而出现的一门新的技术应用,它综合运用了形式识别、数据统计学、人工神经网络等一些先进技术。

(二) 数据挖掘的主要方法。数据挖掘是一门非常综合的跨学科技术,不仅在计算机领域有着广泛的应用,在数学和信息管理等领域也有着广泛的应用前景。主要方法有:决策树分类方法、神经网络法、关联规则挖掘方法、贝叶斯分类方法、聚类分析方法。

(三) 数据挖掘工具与软件。如何寻找出适合和高效的数据挖掘工具和软件,对数据挖掘项目是否能成功以及完成质量有至关重要的影响,根据使用工具的用途和方法不同,数据挖掘工具可分为普通数据挖掘工具和专用数据挖掘工具两大类。三种主要的数据挖掘工具和软件: Intelligent Miner、SPSS Clementine、WEKA。

## 二、财务风险相关理论

(一) 财务风险的概念。财务风险,是指企业对未来损失和盈利的不确定性。可能是由于企业经营管理不善,或企业资金运转等存在一定问题所导致的。其主要的特征:一是客观存在性。财务风险是固有存在,不会因人意志的变化而发生变化,无法避免也不能够消灭它,唯一的办法就是运用技术来发现风险,从而规避风险。二是全面存在性。财务风险伴随着公司财务管理的整个进程。三是不稳定性。财务风险可能在一方面通过一系列机制会发生,但同时换到另一种环境下不一定会发生。四是共存性。风险与收益是相互联系相互作用的,财务风险与收益是同步变化的,风险系数越低会导致收益相应的减少,相反的,风险系数越高收益也就越高。

(二) 财务风险的作用。考虑到经济和贸易在全球化自

由的发展,市场交易的运作更加自由,企业在扩大经营的同时,作为企业的管理层,对财务风险的意识更加强烈,往往企业会通过分析提前对财务和公司提出应对策略,想方设法地进行风险控制,然后再把这样的方法广泛地应用到其他的领域。企业对待财务风险的分析方法和应用有以下实际的作用:在国家的经济决策方面。在经济体制属于市场的情况下,我们国家的企业,能够保证经济活动处在正常范围以内,但是,国家从整体上对经济的宏观调控,却会很大程度上影响企业的发展,对企业的经营管理也会产生相应的影响,在这样的影响下,就可以看到国家的决策影响有多大。因为国家的任何一条关于经济的政策,都或大或小的影响,或者决定企业的生命和活动以及经济效益。通常国家在出台相关的政策之前,如果不会考虑到相关企业的经营范围和承受能力以及资金流现况的时候,只是从单方面的作出经济调控,这是会影响到国家内诸多企业的。

(三) 运用数据挖掘技术分析企业财务风险。

1. 风险发生的相关要素,寻找风险发生的隐性征兆。

2. 确定风险存在。

3. 确定企业财务风险分析的对象。对企业的财务风险分析对象做出归类,找出企业财务风险的对象是财务风险分析中比较重要的一个方面,由于数据挖掘结果的不可确定性,而人们对问题的分析应该是有目的的有规划的,没有目的地对数据进行处理是不可取的。

4. 准备财务风险分析数据。数据准备是数据挖掘的首要步骤,数据可以来自于现有的会计信息系统,也可以是其他业务准备企业的实际财务数据,进行目前的风险分析。

5. 数据预处理。在对可靠的数据确定之后,可以对数据进行实际的判断。从专业的角度上选择与财务和风险有关的一些变量,或者转变变量。最后,确定风险发生的概率、频率以及要计量的损失数额。

6. 数据挖掘。针对所取得的数据进行分析,并挖掘出潜

在的信息。从而,评估风险的性质和程度。并且对风险进行评价,比较风险与风险评价标准,对风险进行排序。

7. 结果评价与解释。选择企业用户的财务决策目的,并结合其挖掘的数据,对其结果进行相关的评价。

8. 对风险进行监控和审计。

### 三、关联规则挖掘和企业财务风险分析

(一) 企业的财务风险和分析指标。财务风险指标是分析过程中最为重要的部分,对指标的选择是在数据挖掘方法过程中对诸多财务指标进行比较。财务指标的具体内容有以下几方面。

1. 企业盈利能力指数。企业盈利能力指数包括: 毛利率指数、营业利润率指数、净利润指数、每股收益指数、净资产收益率指数、总资产收益率指数。

2. 经营能力指标。经营能力指标包括: 流动资产周转率、存货周转率、固定资产周转率、总资产周转率、应收账款周转率。

3. 企业增长能力指数。企业增长能力指数包括: 净资产增长率、总资产增长率、主营业务收入增长率、净利润增长率、每股收益增长率。

4. 企业偿债能力指数。企业偿债能力指数包括: 流动性比率指数、速动比率指数、现金比率指数、付息倍数指数、股东权益比率、现金净流量与负债的比率、现金净流量与净利润的比率、现金净流量与销售收入的比率。

5. 现金流动指数。现金流动指数包括: 现金流偿债比率指数、现金流利润率、现金流资产收益率、现金比率。

(二) 财务指标相关性分析。对财务分析理论的研究成果,使用传统的方法,对分析的变量和相关系数进行检测,结合分析理论对相关的财务指标进行合理的分析,但是前提是必须抛出一些高度相关的财务指标。相关系数可以表示为:

$$r_{xy} = \frac{n\sum x_i y_i - \sum x_i \sum y_i}{\sqrt{n\sum x_i^2 - (\sum x_i)^2} \sqrt{n\sum y_i^2 - (\sum y_i)^2}}$$

在公式中  $x$  和  $y$  表示两个变量。 $r_{xy}$  为它们的相关系数,  $-1 \leq |r_{xy}| \leq 1$ 。当  $|r_{xy}| = 1$  时,说明  $x$  和  $y$  是完全线性的。当  $r_{xy} = 1$  表示  $x$  与  $y$  完全正相关,当  $r_{xy} = -1$  表示  $x$  与完全负企业在利用关联规则交互挖掘算法,对所在的企业进行财务风险分析过程中,进行财务风险指标的挖掘。

(三) 根据时间间隔确定风险等级。在数据分析方法里,关联规则数据挖掘可以根据挖掘类型分为三种: 布尔关联规则、类别关联规则和数量关联规则。五个风险等级的划分是依据风险水平。当然,不同的设计师对财务风险的看法是不同的。所以,对于风险的定义中,风险划分区间也有很大的随机性和不稳定性,因此在实际的分析中要做一些适当的改变。可以利用关联规则交互挖掘算法对企业财务风险指标进行挖掘,财务风险分析是在风险识别的基础上,对财务风险的特征进行明确的定义和描述。金融风险评估应从定量和非定量两个方面进行界定。根据指标的性质,金融危机预警可分为五个层次: 危机预警、高风险预警、低风险预警、正常预警和卓越预警。

(四) 关联规则交互挖掘和建立企业财务风险分析模型。风险概念层次树的构建: 第一级表示的是风险的相关方面。第二级涉及到的是一些联系企业财务风险重要指标,包括利

润率、投资回报率、流动性周转率等。第三层次是最低层次,是指企业金融中最为代表的指标层次。相关的内容有: 毛利率、净利润、每股收益、应收账款周转率等。在财务风险的挖掘中,低层概念的挖掘便是重点。并将其扩展到相应的高层概念挖掘,在高层概念中找到财务指标的相互规律。支持阈值递减的交互挖掘策略: 策略实际是指每个层次上的最小支持阈值。通常情况下,财务指标水平越低,对应的最小支持阈值越小。假设流量和现金比率为频繁项集。在高低层次间的关联规则,根据低层的最小支持度确定最小支持度。

### 四、结语

当前社会,行业内对财务风险的分析方法有很多种。许多的研究工作人员,为了研究目标和目的,不断地为企业寻找控制企业财务分析的方法。从许多不同的角度来看,主要的解决和控制方法有以下几种: 财务报表分析、财务指标分析、德尔菲法。

随着社会的发展,企业对于公司的风险要求和管理,也越来越严格,需要找到一种有效、全面、客观评价企业风险的方法。本文中,关联规则的交互挖掘方法是适合于现代的企业,可以处理每天的有效数据,不需要复杂的数学推理。只需对不同的数据整理分析,就可挖掘出数据库里的隐性信息。没有必要建立任何假设。只要对阈值进行调整,数据中的规则就可以自己找到。通过这些方法,可以有效的减小人为的普通因素的影响,从客观角度上对企业的财务风险进行了合适的评估。此外,也一定程度上提高了对数据挖掘运行的效率。

### 【参考文献】

- [1] 马超群, 兰秋军, 陈为民. 金融数据挖掘 [M]. 北京: 科学出版社, 2007
- [2] 雷英杰, 张善文, 李续武, 周创明. MATLAB 遗传算法工具箱及应用 [D]. 西安电子科技大学, 2005
- [3] 陈燕. 数据挖掘技术与应用 [M]. 北京: 清华大学出版社, 2011
- [4] Tam K Y, Kiang M Y. Managerial applications of neural networks: the case of bank failure predictions [J]. Management Science, 1992, 38(7): 926 ~ 947
- [5] 杨钊. 浅谈数据挖掘技术在现代企业财务分析中的作用 [J]. 中国总会计师, 2010, 4: 161 ~ 163
- [6] 梁循. 数据挖掘原理与方法 [M]. 北京: 北京大学出版社, 2008
- [7] 吴慧香. 数据挖掘在财务风险预警系统中的应用 [J]. 财会通讯(综合), 2008, 2: 80
- [8] 赵春. 基于数据挖掘技术的财务风险分析与预警研究 [D]. 北京化工大学, 2012, 6
- [9] 胡世涛. 基于数据挖掘的财务风险分析方法应用研究 [D]. 华北电力大学, 2017, 12
- [10] 种镇国, 莫中杰. 基于因子分析的中国房地产上市公司绩效评价 [J]. 经济研究导刊, 2010, 12
- [11] 孙丽华. 基于灰色关联模型的中国林业上市公司绩效评价研究 [J]. 林业经济问题, 2013, 6