

构建基于粗糙集和BP神经网络的信用风险预警模型

王明春(博士)² 唐万生¹ 冯嘉毅² 刘鑫²

(1.天津大学自动化学院 天津 300072 2.天津工程师范学院 天津 300222)

【摘要】 本文首先建立了评价企业信用的指标体系,然后在此基础上构造了基于粗糙集和BP神经网络的企业信用风险预警模型。粗糙集用于约简指标体系中的冗余指标, BP神经网络用来构建预测企业信用的分类预警模型。最后,笔者通过实例验证了此模型的可行性和有效性。

【关键词】 风险预警 粗糙集 BP神经网络

信用风险度量是现代金融市场中风险管理的重要一环,也是商业银行信贷风险管理的关键所在。只有对信贷风险进行及时准确的度量和管理工作,帮助银行建立起具有早期预警和后期绩效评价等功能的风险管理机制,才能保障银行自身经营的安全性和盈利性,以及实现整个金融系统的稳定发展。

贷款用途、客户信用和担保效力是信贷业务风险的三个基点。贷款用途决定着一笔信贷业务所面临的环境风险,客户信用决定着具体客户运营贷款资金和承受宏观风险的能力,担保效力影响着客户的主观违约成本和客观违约损失的补偿能力。在这三个基点中,客户信用等级越高,贷款用途风险越小;客户信用等级越低,贷款用途风险越大。对于多数信贷业务而言,商业银行总是优先考量客户信用风险。

由此可见,客户信用是信贷风险结构体系中最重要、最关键的基点,也就是说准确度量客户信用风险才是信贷风险管理的关键。现有的信用风险评估模型大多只使用了一些微观财务指标,本文所构建的模型引入了一些宏观的结构指标——行业、企业的组织形式、所处区域,并结合粗糙集和BP神经网络进行了客户信用风险评估。该模型的优点是:利用粗糙集对决策表进行属性约简,克服了人工选择指标的主观性,同时减少了BP神经网络输入指标的个数,减少了BP神经网络的训练时间,并提高了评估能力。

一、信用风险预警指标体系的构建

1. 信用风险预警指标体系所涉及的微观财务指标。 银行在向企业发放贷款前首先考察的就是企业的微观财务指标。这些指标中,有反映企业偿债能力的资产负债率和利息支付倍数,有反映资金周转状况的应收账款周转率和总资产周转率,有反映流动性的流动比率、速动比率和现金比率,也有反映盈利性的息税前利润率、总资产利润率和净资产收益率。

2. 宏观一般指标对信贷风险的影响。

(1)客户组织形式对信贷风险的影响。商业银行按照客户的组织形式的不同,通常将企业划分成小企业、中小企业、单一法人、企业集团、跨国大集团。不同的客户组织形式导致信贷风险呈现出显著的差异,这种差异给商业银行的信贷风险

管理带来很大困难。如果商业银行的信贷风险管理架构是基于简单客户组织模式构建的,那么它将无法实现对复杂模式客户的有序管理,从而也无法消除或减弱信息不对称带来的显著影响,管理风险将无从谈起。

(2)客户所处行业对信贷风险的影响。行业是与企业衔接最为紧密的宏观范畴。当我们把企业置于行业这一范畴时,企业信用风险所呈现出来的特点与孤立环境下的特点就有所区别了。这种区别源于行业环境变化给企业带来的影响,以及企业参与行业竞争所带来的相对优势的变化。企业的价值取决于企业当期的和未来的收益,而行业的现状和发展趋势在很大程度上决定了行业内企业当期的和未来的收益。行业经营的基本规律,是行业内所有企业经营所必须遵循的,因而行业经营特征的差异也将造成各企业信用风险具有不同的内在结构。

(3)客户所处区域对信贷风险的影响。区域因素通常不会独立影响银行信贷风险,它一般与产业、政策构成合作群体而发挥作用。区域经济特征下的信贷风险在我国具有显著的特点,这既与产业集聚因素相关,也与地方经济发展状况相关。产业集聚理论认为,不是每个地区的经济增长速度都是相同的,经济增长往往集中于某些主导部门和某些有创新能力的行业,而这些主导部门和有创新能力的行业一般集聚在某些地区,且往往是大城市,这些地区会将这种效应向广大地区扩散。近年来,所有商业银行都把长三角、珠三角和环渤海区域作为信贷投放的重点区域,这与这些区域产业集聚带来的企业竞争优势有着密切关系。

3. 信用风险预警指标体系的建立。 由以上分析,我们建立了信用风险预警指标体系,具体如表1所示(见下页)。

二、应用粗糙集和BP神经网络研究信用风险的可行性分析

1. 粗糙集用于属性约简。 粗糙集作为一种处理不确定、不完备数据和精确问题的新的数学理论,最初是由波兰数学家 Z.Pawlak 于 1982 年提出的。粗糙集是一个功能强大的数据分析工具,它能描述和处理不完备信息;能在保留关键信

表 1 信用风险预警指标体系

指标分类		具体指标
微观财务指标	偿债能力	资产负债率
		利息支付倍数
	周转状况	应收账款周转率
		总资产周转率
	流动性	流动比率
		速动比率
		现金比率
	盈利性	息税前利润率
		总资产利润率
		净资产收益率
宏观一般指标		客户组织形式
		客户所处行业
		客户所处区域

息的前提下对数据进行化简并求得知识的最小表达；能识别并评估数据之间的依赖关系；能从经验数据中获取易于证实的规则知识，特别适于智能控制。

在评价信用风险的指标中，原始的决策表信息系统中的条件属性并不是同等重要的，甚至其中某些条件属性是冗余的。冗余属性的存在，一方面是对资源的浪费(需要占用存储空间和花费处理时间)；另一方面，也会干扰人们做出正确的决策。因此，对决策表信息系统中的条件属性进行属性约简就显得非常必要。

所谓决策表的属性约简，就是要在保持条件属性相对于决策属性的分类能力不变的条件下，删除其中不必要的或不重要的属性。粗糙集的一个重要功能就是能对决策表进行属性约简。

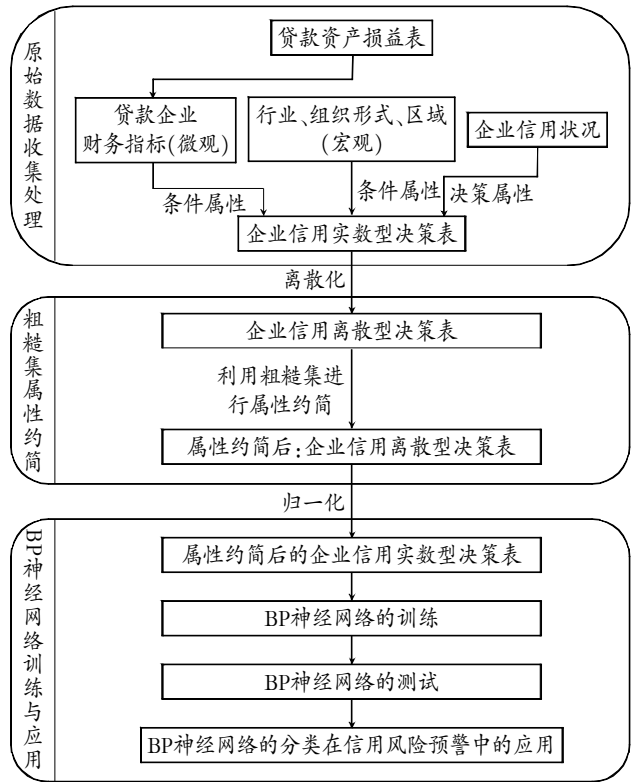
2. BP神经网络用于分类预警模型的构建。BP神经网络是应用最广泛的神经网络之一，精度高、对噪声数据不敏感、网络结构简单。BP神经网络采用最小方差学习方式，只要有足够的隐层和隐结点，它就可以逼近任意的非线性映射关系，使BP神经网络擅长处理那种规律隐含在一堆数据中的映射逼近问题。

BP神经网络克服了传统分析过程的复杂性及选择适当模型函数形式的困难。它的建立是一种自然的非线性建模过程，无须分清存在何种非线性关系，给建模与分析带来了极大的方便。

表 2 企业信用实数型决策表

行业	组织形式	资产负债率	利息支付倍数	应收账款周转率	总资产周转率	流动比率	速动比率	现金比率	息税前利润率	总资产利润率	净资产收益率	信用等级
物资	小企业	0.597 4	1.115 4	1.417 4	0.383 5	1.018 3	0.991 1	0.030 9	0.009 8	0.006 7	0.016 6	B
高科技	中小企业	0.238 7	13.500 0	4.902 4	0.126 1	2.536 1	1.988 8	1.721 1	0.017 1	0.016 0	0.021 0	AA
服装	单一法人	0.586 2	1.511 0	4.763 8	0.367 4	1.451 6	0.942 7	0.167 4	0.017 9	0.018 7	0.045 2	BB
进出口	跨国大集团	0.748 1	0.008 8	8.720 8	0.941 8	0.804 8	0.667 5	0.198 0	0.049 1	0.024 6	0.114 5	A
煤炭	企业集团	0.314 8	41.701 0	39.856 9	0.537 8	2.176 5	1.990 8	0.969 8	0.191 1	0.186 7	0.272 5	AAA
.....

3. 信用风险预警模型的构建。综上所述，我们可以建立基于粗糙集和BP神经网络的信用风险预警模型，如下图所示：



信用风险预警模型图

三、信用风险预警模型的具体应用

1. 数据获取。本文选取了某商业银行在某一时刻的108个法人的贷款资产损益表作为研究数据，通过对这些数据进行处理得到了这些企业的微观财务指标，宏观一般指标通过调查获取。受单一区域的限制，本文不再分析区域因素对客户信用风险的影响，得到由12个条件属性和1个决策属性组成的企业信用实数型决策表，其中部分数据如表2所示。

2. 离散化。本文采用专家评价法对条件属性值进行离散化处理。根据经验，专家参考离散化标准(见表3)，并依据主观判断对条件属性值进行离散化处理，在离散化处理过程中可能会根据所处行业以及产业周期不同等对评分结果进行修正，最后根据条件属性值给出评分。评分分为5个档次：1分表示很差；3分表示稍差；5分表示及格；7分表示较好；9分表示很好。离散化结果如表4所示。

表3 条件属性离散化标准

条件属性		1	3	5	7	9
一般指标	行业	专家主观判断				
	组织形式	小企业	中小企业	单一法人	企业集团	跨国大集团
偿债能力	资产负债率	>0.85	0.7~0.85	0.6~0.7	0.3~0.6	0~0.3
	利息支付倍数	<1.5	1.5~2.3	2.3~2.7	2.7~5.0	>5.0
周转状况	应收账款周转率	<1.5	1.5~2.5	2.5~3.5	3.5~5.0	>5.0
	总资产周转率	<0.5	0.5~0.7	0.7~1.0	1.0~3.0	>3.0
流动性	流动比率	<1.0	1.0~1.3	1.3~1.8	1.8~2.5	>2.5
	速动比率	<0.3	0.3~0.5	0.5~0.8	0.8~1.2	>1.2
	现金比率	<0.1	0.1~0.3	0.3~0.5	0.5~0.7	>0.7
盈利性	息税前利润率	<0	0~1.0	1.0~5.0	5.0~10.0	>10.0
	总资产利润率	<0	0~3.0	3.0~5.0	5.0~10.0	>10.0
	净资产收益率	<3.0	3.0~5.0	5.0~10.0	10.0~15.0	>15.0

表4 离散化数据表

行业	组织形式	资产负债率	利息支付倍数	应收账款周转率	总资产周转率	流动比率	速动比率	现金比率	息税前利润率	总资产利润率	净资产收益率	信用等级
9	1	7	1	1	1	3	7	1	3	3	1	B
5	3	9	9	7	1	9	9	9	3	3	1	AA
7	5	7	3	7	1	5	7	3	3	3	1	BB
5	9	3	1	9	5	1	5	3	3	3	1	A
9	7	7	9	9	3	7	9	9	3	3	1	AAA
...

3. 属性约简。在粗糙集理论中,决策表 L 可用一个四元组 $[U, CUD, V, f]$ 表示。其中, U 是对象集合, C 是条件属性集合, D 是决策属性集合, V 是属性值的集合, f 是一个信息函数, f 用于指定 U 中每一个对象的属性值。对于一个条件属性 $c_i \in C$, 可以用决策属性集合 D 的 c_i 正阈 $POS_{c_i}(D) = \bigcup_{x \in U/D} c_i(x)$ 来判

断条件属性 c_i 相对于决策属性集合 D 的分类能力。其中 $c_i(x)$ 表示集合 x 关于条件属性 c_i 的下近似。依据以上公式得到各个条件属性的分类能力大小,如表 5 所示:

表5 属性重要程度

一级指标	一般指标	偿债能力		周转状况		流动性			盈利性			
二级指标	行业 c_1	组织形式 c_2	资产负债率 c_3	利息支付倍数 c_4	应收账款周转率 c_5	总资产周转率 c_6	流动比率 c_7	速动比率 c_8	现金比率 c_9	息税前利润率 c_{10}	总资产利润率 c_{11}	净资产收益率 c_{12}
重要程度	0.541	0.016	0.148	0.082	0.295	0	0.049	0.098	0	0	0.066	0

为了进行属性约简,本文采取依据属性分类质量逐步添加属性的方法,得到一个条件属性子集 P ,使得 $POS_P(D) = POS_C(D)$ 或 $\frac{|POS_P(D)|}{|POS_C(D)|} = 1$,说明条件属性子集 P 和所有条件属性集合 C 相对于决策属性集合 D 具有相同的分类能力, P 即为条件属性集合 C 相对于决策属性集合 D 的属性约简。

属性约简的过程如表 6 所示:

表6 正阈值变化

条件属性子集	$\{c_1\}$	$\{c_1, c_5\}$	$\{c_1, c_5, c_3\}$	$\{c_1, c_5, c_3, c_8\}$	$\{c_1, c_5, c_3, c_8, c_4\}$	$\{c_1, c_5, c_3, c_8, c_4, c_{11}\}$	$\{c_1, c_5, c_3, c_8, c_4, c_{11}, c_7\}$
$\frac{ POS_P(D) }{ POS_C(D) }$	0.541 0	0.639 3	0.704 9	0.868 9	0.967 2	0.967 2	1

4. BP 神经网络的训练与应用。BP 神经网络是单向多层前向网络的一种。本文采用三层 BP 神经网络,即 1 个输入层、1 个隐含层和 1 个输出层。输入层结点数是 7 个,输出层结点数为 1 个。根据隐层结点个数的计算公式:中间结点数 $\leq \sqrt{\text{输入结点数} \times (\text{输出结点数} + 3)} + 1$,计算出隐层结点数最大为 6 个。

把属性约简后的指标作为 BP 神经网络的输入样本,选取其中 60 个样本用于网络训练,用 MATLAB 软件进行编程。第一层、第二层的激励函数分别采用 logsig 函数和 purelin 函数建立;使用 initff 函数初始化 BP 神经网络的各层权重因子;Trainbp 作为 BP 神经网络的训练函数。最大循环次数为 5 000 次,目标误差为 0.01,学习速率为 0.01。

在 BP 神经网络训练好后,利用后 48 个样本作为测试样本,得到隐层结点数 2~6 个时的评估情况,如表 7 所示:

表7 不同隐层结点个数的评估情况

	2 结点	3 结点	4 结点	5 结点	6 结点
错误率 (%)	9.32	6.09	4.64	8.27	8.64
正确率 (%)	90.68	93.91	95.36	91.73	91.36

由表 7 可知,运用该模型进行信用风险评估所得出结果的正确率在 90% 以上。

四、结论

本文通过对单笔信贷业务风险结构进行论述,综合微观层面和宏观层面建立了完善的信用风险预警指标体系,并基于粗糙集和 BP 神经网络建立了信用风险预警模型。然后利用该模型,结合某商业银行某一时点的数据,对客户信用进行了实证研究。结果表明,本文所建立的基于粗糙集和 BP 神经网络的信用风险预警模型,可有效运用于商业银行的客户信用风险预警。

【注】 本文系国家自然科学基金资助项目(编号:70471049)与天津市科技发展战略基金资助项目(编号:07ZLZLZT02900)的阶段性研究成果。

主要参考文献

- 李虹,郑丕谔.基于 BP 神经网络的个人消费信贷风险识别.哈尔滨商业大学学报(社会科学版),2007;5
- 梁琪,黄鹂皎.我国商业银行信贷风险管理体系构建探索.南开经济研究,2002;6
- 梁琪.企业信用风险的主成分判别模型及其实证研究.财经研究,2003;5
- 葛超豪,葛学健.银行信贷风险评估计量模型探讨.理论新探,2005;12