

## 研究简报

基于神经网络的企业信用等级评估<sup>①</sup>

陈雄华, 林成德, 叶武

(厦门大学自动化系, 厦门 361005)

**摘要:** 企业信用等级评估是金融领域重要的问题, 论文采用人工神经网络模型研究企业信用等级的评估问题。按照企业样本在信用等级的分布状况来抽样, 然后, 根据企业样本性质的不同, 将其分为制造业和非制造业两大类。利用偏相关分析方法建立了企业信用评级的指标体系。此外, 还介绍了几种企业信用评级常用的评估模型, 并将神经网络评估模型的性能和其他的信用评估模型作了比较, 实验结果表明神经网络模型具有更好的预测准确性。

**关键词:** 神经网络; 信用评级; 信用风险评估模型

中图分类号: TP389; TP183

文献标识码: A

文章编号: 1000-5781(2002)06-0570-06

## Credit risk assessment of enterprise basing on neural network

CHEN Xiong-hua, LIN Cheng-de, YE Wu

(Department of Automation, Xiamen University, Xiamen 361005, China)

**Abstract:** Credit rating of the enterprise is very important problem in the financial field. In this article, we researched this problem using artificial neural network model. We got the samples basing on the distribution of credit rating of the all samples. And then, we divided the samples of the enterprises into two sets, one is manufacture and the other is non-manufacture, we also developed an indicators' system using partial correlation method. Furthermore, we introduced many credit assessment models in common use and compared the result of our model with that of some of them. The result from the experiment shows that our model is more applicable and suitable.

**Key words:** neural network; credit risk assessment; credit risk assessment model

## 0 引言

商业银行对客户的信用评价是银行贷款的核心内容, 对银行客户的信用评估是否合理、科学、准确关系着银行贷款的成败。

银行的风险管理一直是国际国内金融界关注的焦点<sup>[1,2]</sup>, 在西方发达国家, 商业银行的信用风险管理技术已经比较成熟, 许多定量技术和支持工具、软件已付诸商业应用, 继传统的比例分析之

后, 统计方法得到了广泛的应用, 如判别分析和 Logistic 回归等。自从 20 世纪 90 年代末期以来, 人工智能技术如神经网络、专家系统、分类树也被应用于商业银行信用的管理中<sup>[3,4]</sup>。

从国内看, 对处于新兴市场和转轨型经济环境下的我国商业银行而言, 加强信用风险管理尤为重要。这是因为我国的商业银行刚刚建立, 其信用风险管理技术较为落后, 特别是客户信用分析与评估技术仍处于传统的比例分析阶段。目

①. 收稿日期: 2001-04-16; 修订日期: 2002-06-24.

前我国银行机构对企业的信用等级评定是通过企业的某些单一财务指标进行加权平均确定的<sup>[4]</sup>. 该方法的重大缺陷在于指标和加权值的确定带有很大的主观性, 使得评估结果和实际状况有很大的出入. 因此需要引入科学方法来确定有效评估指标, 并建立准确的定量模型来解决信用等级评估的问题.

## 1 传统的信用评估模型和神经网络

国际上, 对企业信用的评估, 通常将商业银行对企业信用风险的测度转化为对企业财务状况的衡量问题, 因为信用风险的形成, 主要取决于企业的财务状况. 具体的作法是根据历史上每个类别(如信用等级 AAA、AA、A、B 等)的若干样本, 从已知的数据中发现规律, 从而总结出分类的规则, 建立判别模型, 用于新样本的判别. 国外在对银行客户信用评价中广泛采用了基于统计判别方法的模型, 这些方法都是在 Fisher 于 1936 年所作的研究之后提出来的<sup>[4]</sup>. 常用的模型还有: 多元判别分析法(MDA)、Logit 模型、分类树法等.

MDA 是除美国外其他国家用得最多的统计方法. MDA 可以具体分为一般判别分析和逐步判别分析, 前者不考虑变量筛选, 后者考虑变量筛选. 基于 MDA 建立的判别模型有 Chesser 判别模型、ZETA 判别模型、逐步判别分析模型等. 它们的共同特点就是通过对一些财务指标进行加权得到一个输出结果, 再对这个输出的结果进行分段判别, 从而得到对应的一个信用等级.

Logit 分析与判别分析的本质差异在于不要财务指标满足正态分布或等方差, 由于一般判别分析方法的局限, Logit 分析在预测中得到了相当广泛的应用, 1981 年以后的研究绝大多数都用 Logit 分析, 在一些国家建立了许多相应的模型<sup>[5-9]</sup>.

20 世纪 80 年代末期, 有学者提出一种基于机器学习技术发展起来的符号方法, 即分类树方法. 该方法不像传统方法那样通过判别函数形成决策规则来判别样本所属的类型, 而是创立了一个对原始样本进行最佳分类判别的分类树.

以上介绍的这些模型已经得到了广泛的应用, 但它们仍存在着许多缺陷, 实证结果发现:

1) 企业财务状况的评价可以看作是在基于独立变量基础上的分类问题;

2) 企业财务状况的好坏与财务指标的关系是非线性的;

3) 许多财务指标可能是高度相关的;

4) 许多财务指标不成正态分布.

因此, 传统的分类方法不能很好地解决这些问题. 神经网络技术作为研究复杂性问题的强有力的工具, 近年来在模式识别与分类、识别滤波、自动控制、预测等方面已展示了其非凡的优越性, 特别是它能处理任意类型的数据, 这一特点是许多其他方法所无法比拟的. 神经网络通过不断学习, 能够从特定模式的大量的数据中发现其潜在的规律. 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的困难, 它是一种自然非线性建模过程. 勿须分清是何种非线性关系, 给建模与分析带来了极大的方便. 该方法用于企业财务状况研究时, 一方面利用其映射能力, 另一方面利用其泛化能力, 即在经过一定数量的带噪声的样本的训练之后, 网络可以抽取样本所隐含的特征关系, 并对新情况下的数据进行内插和外推以推断其属性. 它在分类问题中的出现, 最早是用于对银行破产的预估<sup>[4]</sup>.

神经网络人工智能方法能处理一系列信息输入(如财务比率等的数据), 并能产生相应的输出, 网络通过学习能够产生一个成功反映所有输入输出变量相应关系的模式. 除此之外神经网络不需要变量之间必须线性相关或是相互独立这一假设. 变量之间存在的微妙联系, 在数据不连续或有噪声的情况下, 也可被系统辨识并生成定性估评(如信用等级). 简而言之, 除部分不明确的结果之外, 神经网络能够在相似点和类似点上给出有价值的结论.

在美国和欧洲, 已有一些实际工作<sup>[10, 11]</sup>, 尝试将此技术应用到金融服务领域, 尤其是信用分析领域. 在我国这方面的实际工作还较少, 还处于研究探索阶段<sup>[12]</sup>.

## 2 样本取样和变量的筛选

### 2.1 样本的选取

本文所用的数据来源于福建省某银行对中小

企业(非上市公司)信用评估时所使用的数据.在样本收集过程中,要考虑的主要问题是样本的代表性、可靠性、和抽样误差.不同行业不同企业的财务特性是存在差异的.由于其经营生产流程、经营周期、经营性质等均有不同,反映其财务状况和生产经营状况必然有所不同.本文将行业分为制造类和非制造类.

客户信用等级是反映客户偿还债务能力和价值大小的相对评价标准,目前我国银行一般将客户分为AAA级、AA级、A级、BBB级、BB级、B级6个信用等级.由于本文中银行提供的数据只有少量BBB级、BB级、B级的客户,故把这3个等级合并为A级以下这一个等级.所使用的样本按银行对企业的信用评估等级AAA级、AA级、A级、A级以下4个等级的数据分布进行取样.

根据以上的分析,借鉴国外研究取样的经验并结合我国的实际,确定以下的取样原则:1)随机取样;2)主要按银行提供的已评估样本的分布来取样,尽量考虑行业、资产规模的样本的分布;3)尽量选取足够多的样本,以保证样本的代表性.

这样,把企业分成制造类和非制造类两大类,再依等级AAA、AA、A、A以下4级分类划分,根据4类的分布情况,同时考虑资产规模,进行随机抽样.选取了每个企业的财务数据,每年的数据作为一个样本,由于银行提供的数据中制造类的比例为8:4:2:1,所以按比例进行抽样,结果AAA、AA、A、A以下的样本比例为126:68:34:12,其中150

个取为学习样本,而余下的90个作为检验样本;对于非制造类,比例为3:2:1:1,抽样样本数为66:43:21:20,其中90个为学习样本,其余的60个为检验样本.

## 2.2 财务比率分析与变量筛选

企业的信用状况如何,要依据一些标准的财务指标加以规定.现根据实际情况筛选那些对信用等级的评估起重要作用的财务指标.财务指标的选取依据两个原则:一是指标最好能比较直接地反映偿债能力;二是指标的设计应该避开所谓的行业标准,即尽量使这些指标的值得在不同行业 and 不同企业有相同或相似的意义.

在参考国家财政统计评价司关于企业信用评级指导思想<sup>[13]</sup>的基础上并参照银行所提供的数据将信用等级分为G1(营运资本/总资产)、G2(资产负债率)、G3(流动比率)、G4(速动比率)、G5(存货周转率)、G6(应收账款周转率)、G7(销售利润率)、G8(资产报酬率)、G9(息税前收益/总资产)、G10(销售收入/总资产)、G11(负债总额/总资产)、G12(固定资产/总资产)、G13(息税前收益/总负债)、G14(息税前收益/营运资本)、G15(存货/销售收入)、G16(流动负债/净资本)、G17(利润/销售收入)这些财务比率进行偏相关分析.

将随机选取的制造类240条样本和非制造类150条样本用SPSS统计软件进行偏相关分析.得到的结果见表1.

表1 制造类样本偏相关分析

项目	G1	G2	G3	G4	G5	G6	G7	G8	G9
S	0.694 6	0.893 2	0.665 3	0.671 0	0.033 6	-0.026 8	0.698 4	0.897 8	0.899 6
P	0.000	0.000	0.000	0.000	0.412	0.513	0.000	0.000	0.000
项目	G10	G11	G12	G13	G14	G15	G16	G17	
S	0.817 0	0.893 2	-0.104 8	0.777 9	-0.075 8	-0.037 2	0.751 5	0.063 0	
P	0.078	0.000	0.000	0.000	0.064	0.364	0.000	0.123	

表1是制造类各财务比率与信用等级分S的零阶相关,这里对应的数据表示相关系数,P表示对于相关系数为0的假设成立的概率,参与相关

系数计算的观测量数都是240.非制造类与制造类进行同样的处理,各财务比率与信用等级的零阶相关见表2.

表 2 非制造类样本偏相关分析

项目	G1	G2	G3	G4	G5	G6	G7	G8	G9
S	0.727 6	0.788 2	0.788 2	0.709 3	-0.012 9	0.058 1	0.699 9	0.620 2	0.603 7
P	0.005	0.000	0.000	0.010	0.875	0.480	0.000	0.000	0.000
项目	G10	G11	G12	G13	G14	G15	G16	G17	
S	0.625 5	-0.268 5	-0.062 7	0.498 0	0.074 6	0.250 7	-0.353 1	-0.053 3	
P	0.057	0.000	0.002	0.000	0.364	0.002	0.000	0.000	

注: 表 2 变量 S 和 P 的意义同表 1, 参与相关系数计算的观测量数都是 150.

从各变量间的零阶相关系数可以看出信用等级分与财务比率的相关系数, 以及显著性检验结果的不相关概率. 可以知道信用等级分与以下财务比率有较明显线性关系:

对制造类: G1、G2、G3、G4、G7、G8、G9、G10、G11、G13、G16.

对非制造类: G1、G2、G3、G4、G7、G8、G9、G10.

由于各财务比率间的相互影响, 再根据信用等级分与各财务比率单独的偏相关进行分析得出结果见表 3、4.

表 3 制造类各财务比率和信用等级分独立的偏相关分析

项目	G1	G2	G3	G4	G7	G8
S	0.681 9	0.638 0	0.794 0	0.807 9	0.612 6	0.650 0
P	0.000	0.059	0.023	0.009	0.000	0.000
项目	G9	G10	G11	G13	G16	
S	0.845 4	0.615 8	0.436 1	-0.355 4	0.371 0	
P	0.000	0.003	0.894	0.703	0.614	

表 4 非制造类各财务比率和信用等级分独立的偏相关分析

项目	G1	G2	G3	G4	G7	G8	G9	G10
S	0.598 9	0.772 7	0.714 2	0.707 8	0.623 2	0.559 0	0.679 2	0.784 8
P	0.024	0.039	0.003	0.000	0.000	0.002	0.032	0.027

表 3 和表 4 对应数据表示该变量在排除其它变量的影响后, 与信用等级分 S 的相关系数, P 表示不相关的概率.

由表 3 可得出结论: 对制造类的信用等级分与 G11(负债总额 / 总资产) 的相关系数为 0.436 1, 不相关的概率为 0.894, 可不考虑; 对 G13(息税前收益 / 总负债) 的相关系数为 -0.355 4, 不相关的概率为 0.703, 也可不考虑; 对 G16(流动负债 / 净资本) 的相关系数为 0.371 0, 不相关的概率为 0.614, 也可不考虑. 而对非制造类, 信用等级分与各财务比率的偏相关性都比较强, 对这 8 个比率应给予考虑.

综上所述, 对制造类或非制造类, 在下面的信用等级判别模型中采用的财务比率变量统一为: G1、G2、G3、G4、G7、G8、G9 和 G10 这 8 个量.

### 3 实际应用

#### 3.1 网络构建

根据文[ 14, 15] 提到的建模方法, 建立了一个 3 层 BP 神经网络信用等级评判模型: 其输入层为 8 个神经元, 输出为 1 个神经元, 而隐含层神经元的个数, 用文[ 15] 提供的方法: 即逐步增加结点, 每增加 1 个结点后, 通过学习训练, 如果引起判断准确性的下降, 则采用未增加前的结点数, 本文对模型进行试验后发现, 当结点增加到 18 的时候, 评判的准确率最高, 在以后就有下降的趋势, 故模型的隐结点数定为 18. 将 AAA、AA、A、A 以下 4 级的期望输出分别设为 0.9, 0.7, 0.5, 0.3.

网络模型确定后, 公司该年的 8 个财务比率

作为一个矢量输入,在各个比率输入之前将数据先进行归一化处理

$$x' = \frac{x - \min(\circ)}{m\alpha_x(\circ) - \min(\circ)}$$

归一化后,  $x' \in [0, 1]$ .

### 3.2 网络的训练

采用动量法和学习率自适应调整的改进型BP算法来训练网络<sup>[14]</sup>,将输入信号和期望输出一同输给网络,网络在学习过程中将不断地调整网络的权值和阈值,当输出值和期望的输出间的平均误差  $e$  小于目标误差  $\epsilon$  时,学习过程结束,在本实验中,经过了多次实验,发现目标误差设为 0.006 时,评判的准确性最高,特别应指出,这里的误差精度不宜取得过高,因为样本集的输入输出夹杂有噪音信号,学习开始时,网络权值的调整主

要只针对实际信号进行,但随着学习的继续,网络权值的结构也反应了噪音的某些特征,因此学习过度将会引入噪音信号.

### 3.3 模型评估应用

当网络学习完成后,就可以利用这一模型对样本进行预测,即对银行客户进行信用等级的评价,建立的评判规则如表 5 所示.

表 5 评判规则

级别	AAA	AA	A	A 以下
模型的输出 $y$	$y \geq 0.8$	$0.6 \leq y < 0.8$	$0.4 \leq y < 0.6$	$y < 0.4$

现将制造类和非制造类的学习样本和检验样本均输入到网络中进行评估,并和一般回归模型和逐步判别模型作比较得出表 6 的结果.

表 6 各种模型评判结果

类别	模型	不同类别判断准确性(判断正确/总个数)				学习 样本/ %	检验 样本/ %
		AAA	AA	A	A 以下		
制造类	一般回归	56/ 79	28/ 39	17/ 22	5/ 10	70. 67	65. 56
		26/ 41	19/ 29	10/ 12	4/ 8	(106/ 150)	(59/ 90)
	逐步判别	59/ 79	27/ 39	18/ 22	5/ 10	72. 67	68. 89
30/ 41		20/ 29	9/ 12	3/ 8	(109/ 150)	(62/ 90)	
非制造类	神经网络	69/ 79	31/ 39	17/ 22	10/ 10	84. 67	73. 33
		32/ 41	21/ 29	8/ 12	6/ 8	(127/ 150)	(66/ 90)
	一般回归	31/ 42	19/ 26	10/ 12	3/ 10	76. 67	68. 33
17/ 24		14/ 17	6/ 9	4/ 10	(69/ 90)	(41/ 60)	
非制造类	逐步判别	34/ 42	19/ 26	10/ 12	3/ 10	78. 89	68. 33
		17/ 24	16/ 17	6/ 9	4/ 10	(71/ 90)	(43/ 60)
	神经网络	38/ 42	24/ 26	10/ 12	9/ 10	90	76. 67
19/ 24		13/ 17	7/ 9	7/ 10	(81/ 90)	(46/ 60)	

注:表中 AAA、AA、A、A 以下 4 栏中的每一单元格均有两组数据,第 1 组数据为对应信用等级学习样本评判正确的样本数/该信用等级学习样本总数,第 2 组数据为对应信用等级检验样本评判正确的样本数/该信用等级检验样本总数,学习样本和检验样本栏中给出了评判准确率和对应的评判正确的样本数/对应样本总数.

从表 6 可以看出基于 BP 神经网络的信用评价模型,在判别的准确性方面优于其他常用的模型,虽然得到的结果没有像国外同等实验那么显著,但对那些不是经过专家筛选出来的,且在可靠性、准确性方面也存在着疑问的学习样本,有这样的结果已经说明 BP 神经网络提供了更好的银行客户信用等级评判模型.

## 4 展望和思索

在以上研究基础上,对工作中存在的不足和

欠缺作以下归纳,并对完善工作提出了几点思考.

因为样本量还不够大,因此在实用性和普及性上尚待进一步研究和探索.其次因为样本数量的限制,不可能对不同行业进行具体的模型构建,只能以一种统一的眼光来构造一个通用的模型,而在经济领域,行业间的差异较大,如果强制地把它划归到统一的模型中去研究,必定使模型出现失真,所以在样本数据量足够的情况下,应该对一些特殊的行业进行特别的处理.

因为神经网络的输入和输出是一种映射的关系,这种映射关系不能处理一些特殊的情况,如在

信用评估中, 当某个指标满足不了要求, 则不管信用分多高都不能评为 AAA. 这样的硬性规则在信用评估中是很多的, 如果能够总结出所有的硬性规则, 并用这些规则对神经网络的输出作二次修正, 则应该可以提高模型评判的准确性.

万事万物都有其矛盾的统一性和特殊性, 对于我国的企业, 则存在行业性质、企业规模、政策倾向、经济地域等诸多方面的差异, 如何解决这些差异对建模的影响, 应该由人工智能和经济领域的专家们共同探讨, 共同研究. 建议由经济领域的专家根据企业的各种属性和特点划分出有较强可比性的经济客体的集合, 再由人工智能领域的专家针对各个不同的客体集合进行人工神经网络的建模. 由于在每个集合内部有较强的可比性, 它们存在的规则也较明显, 而神经网络通过学习就可以捕捉到这种潜在的规则.

#### 参考文献:

- [ 1 ] Freeman A. A survey of international banking [ J ]. *The Economist*, 1993, (6): 1—37
- [ 2 ] 曾国坚, 何五星. 银行风险论 [ M ]. 北京: 中国计划出版社, 1995
- [ 3 ] 高业培. 企业失败理论与实证研究 [ D ]. 厦门: 厦门大学, 1999
- [ 4 ] 王春峰, 万海晖, 张 维. 基于神经网络技术的商业银行信用风险评估 [ J ]. *系统工程理论与实践*, 1999, 9: 24—32
- [ 5 ] Fisher R A. The use of multiple measurements in taxonomic problems [ J ]. *Ann. Eugenics*, 1936, (7): 179—188
- [ 6 ] Altman E I. Financial ratios discriminant analysis and the predication of corporate bankruptcy [ J ]. *Journal of Finance*, 1968, 23(9): 589—609
- [ 7 ] Elam E. The effect of lease data on the predictive ability of financial ration analysis for small business failure prediction [ J ]. *Journal of Financial and Quantitative Analysis* 1972 (3): 65—76
- [ 8 ] Press S J, Wilson S. Choosing between Logistic regression and discriminant analysis [ J ]. *J. Amer Statist Assoc*, 1978, 73(7): 699—705
- [ 9 ] Eisenbeis R A. Pitfalls in the application of discriminant analysis in business finance, and economics [ J ]. *Journal of Finance*, 1977, 32: 875—900
- [ 10 ] Odom M D, Sharda R A . Neural network for bankruptcy prediction [ J ]. *International Joint Conference on Neural Networks*, 1990, 2 (6): 163—168
- [ 11 ] Hopfield J J, Tank D W. Neural computation of decisions in optimization problem [ J ]. *Biological Cybernetics*, 1985, 52: 123—145
- [ 12 ] 王世彤. 企业财务综合分析方法—互异代表系与人工神经网络 [ J ]. *系统工程理论与实践*, 1998 (12): 103—109
- [ 13 ] 财政部统计评价司. 企业效绩评价问答 [ M ]. 北京: 经济科学出版社, 1999
- [ 14 ] Dutta S, *et al.* Bond rating: A non-conservative application of neural networks [ J ]. *Proc. IEEE Int. Conf. on Neural Networks*, 1992, (2): 154—173
- [ 15 ] Poddig T. Bankruptcy prediction: A Comparison with Discriminant Analysis [ M ]. New York: Apostolos-Paul Refenes, 1993. 137—159
- [ 16 ] 陈雄华, 林成德. 基于改进型 BP 算法的外债风险指标预测 [ J ]. *厦门大学学报(自然科学版)*, 2001, 40(5): 1017—1021

#### 作者简介:

陈雄华(1976—), 男, 福建政和人, 硕士. 研究方向: 计算机技术在经济领域的应用.

林成德(1948—), 男, 浙江定海人, 博士, 教授. 研究方向: 随机系统的估计和应用, 控制在经济领域的应用.

叶 武(1975—), 男, 福建莆田人, 硕士. 研究方向: 计算机技术在经济领域的应用.

## 5 结束语

将 BP 神经网络技术应用用于商业银行客户信用等级评价中, 结果表明, 神经网络模型具有广泛的应用前景.

当然, 基于神经网络的银行信用等级评估还有一些尚待解决的问题, 比如对判别结果的解释能力, 一般情况下神经网络都被认为没有解释能力, 这是神经网络作为一种技术方法的主要缺陷, 人们无法根据神经网络的权值得到输入变量的相关重要性. 近年来, 国外一些专家学者提出在构造网络结构的同时使用遗传算法, 还有一些学者使用启发式的方法来解释输入变量的相关重要性, 有人则把模糊逻辑与神经网络的解释能力联系起来, 在这个研究方向上已取得了不少的进展.