

# 基于神经网络技术的商业银行信用风险评估<sup>\*</sup>

王春峰, 万海晖, 张 维

(天津大学系统工程研究所, 天津 300072)

**摘要:** 研究了神经网络技术在商业银行信用风险评估中的应用。实证结果表明, 与传统统计方法(判别分析)相比, 神经网络技术具有更高的预测精度和更强的鲁棒性。

**关键词:** 神经网络; 判别分析; 信用风险; 信用风险评估

## Credit Risk Assessment in Commercial Banks Using Neural Networks

WANG Chun-feng, WAN Hai-hui, ZHANG Wei

(Institute of Systems Engineering, Tianjin University, Tianjin 300072)

**Abstract** In this paper, the problem of credit risk assessment in commercial banks is investigated by using the method of neural network. Empirical results show that neural network is a promising method of evaluating credit risk in commercial banks in terms of predictive accuracy, adaptability and robustness.

**Keywords** neural network; discriminant analysis; credit risk; credit risk assessment

### 0 引言

近年来,随着金融的全球化趋势及金融市场的波动性加剧,商业银行的风险管理一直是国际国内金融界关注的焦点<sup>[1]</sup>。商业银行在营运过程中面临的金融风险主要有信用风险、利率风险、汇率风险、流动性风险和操作风险等,其中信用风险占有特殊的地位。信用风险是指,借款人由于种种原因,不愿或无力偿还银行贷款本息,使银行贷款无法收回,形成呆帐损失的可能性。在现代商业银行经营中,信用风险是影响其安全高效运营的主要原因。从国际上看,在许多发达国家,由于金融机构贷款信用质量低下,如对不发达国家的银行贷款、银行住宅贷款、农业抵押贷款等,导致呆帐和不良贷款不断增加,造成流动性危机,最终诱发倒闭,给金融业和整个国民经济造成严重损失<sup>[2]</sup>。世界银行对全球银行业危机的研究表明,导致银行破产的最常见原因,就是信用风险<sup>[3]</sup>。因此,国际金融界近年来对信用风险的关注日益加强,如旨在加强信用风险管理的《巴塞尔协议》已在西方主要发达国家全面实施。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,加强信用风险的管理尤为重要。其原因在于,第一,由于历史原因,不良资产一直是影响我国银行业有效经营的主要因素;第二,在国有专业银行向商业银行转轨过程中,面临的主要问题突出表现为比例较大的不良资产,呆坏帐的负担是我国商业银行进一步发展的障碍,加强信用风险管理是解决这一问题的关键;第三,由于历史原因,我国商业银行的各项资本资产比例与巴塞尔协议的要求尚有相当大的差距,因此加强信用风险管理是我国商业银行与国际金融业接轨的关键措施之一。

在西方发达国家,商业银行的信用风险管理技术已比较成熟,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,统计方法得到广泛的应用,如判别分析和Logistic回归等。自从80年

<sup>\*</sup> 收稿日期: 1998-02-04

资助项目: 国家自然科学基金委 95 重大项目 (79790130) 及国家教委跨世纪优秀人才基金 (96-170) 共同资助

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风险管理中(详见本文第二部分)。

在我国,由于现代商业银行体制刚刚建立,其信用风险管理技术较为落后,特别是作为信用风险管理的核心技术——信用分析与评估技术仍处于传统的比例分析阶段,远不能满足商业银行对各种形式贷款安全性的准确测量。目前我国银行机构主要使用计算贷款风险度的方法进行信用风险评估——在对企业进行信用等级评定的基础上,考虑贷款方式、期限以及形式因素,进而确定贷款的风险度。其中作为核心的信用等级评定,是通过对企业的某些单一财务指标进行评价,而后加权平均确定的<sup>[4]</sup>。该方法的重大缺陷在于指标和权重的确定带有很大的主观性,使得评级结果与企业的实际信用状况有很大出入,因此需要引入科学方法来确定有效指标,并建立准确的定量模型来解决信用评估问题。

针对这种情况,本文的目的在于,根据我国商业银行的具体情况,结合国际上目前较为流行的人工神经网络技术,研究我国商业银行风险评估问题。

本文的结构安排如下:第一部分首先介绍了商业银行信用风险的管理方法与程序;然后综述了信用风险评估的各种模型;第三部分对神经网络模型和反向传播学习算法进行了介绍;第四部分是模型的具体构造,包括指标体系的确定和样本数据的处理;第五部分利用神经网络技术,进行了实证研究,并与判别分析模型的结果进行了比较;最后总结了全文。

## 1 信用风险管理方法与程序

商业银行信用风险管理一般包括,信用风险分析——风险的测量与评估,信用风险管理与控制——贷款定价、贷款发放、贷款检查。

信用风险分析是指,对可能引起贷款风险的因素进行定性分析、定量计算,以测量借款人的违约概率,为贷款决策提供依据。如是否发放、价格确定、发放条件和形式等。显然,信用分析是信用风险管理的首要工作和关键。

在信用风险分析时,商业银行典型的作法是,首先依据借款人的资料,遵循5C原则(即借款人品质(Character)、能力(Capacity)、资本(Capital)、担保(Collateral)、环境(Condition)),或LAPP原则(流动性(Liquidity)、活动性(Activity)、盈利性(Profitability)、潜力(Potentialities)),定性分析借款人目前的财务状况、管理水平、宏观与行业经济状况;确定反映借款人还款能力的指标体系;根据指标体系,利用一定的定量技术、模型(评估模型)判别借款人的违约概率。

## 2 信用风险评估模型

企业的信用风险问题——能否如期还本付息,受很多因素的影响如企业财务状况、管理水平、行业状况、宏观经济形势等,目前的金融理论尚不能对这一问题在理论上作出满意的解释。国外商业银行处理这一问题通常采用所谓的“经验主义方法论”,即把信用风险分析看成是模式识别中的一类分类问题——将企业划分为能够按期还本付息和违约两类。其具体做法是根据历史上每个类别(如期还本付息、违约)的若干样本,从已知的数据中发现其规律,从而总结出分类的规则,建立判别模型,用于对新样本的判别。尽管有人将这种方法称为“粗暴的经验主义方法”<sup>[5]</sup>,但在目前的金融理论状况下,它可能是最有效的方法,也是国际金融界和学术界视为主流的方法。

国际上,通常将商业银行信用风险的测度转化为企业财务状况的衡量问题,因为信用风险的形成——企业是否能如期还本付息主要取决于企业财务状况。影响企业财务状况的因素很多,我们不能仅根据企业的某些单一指标,而应根据影响企业财务状况的多维指标来评估企业的财务状况。因此,这一方法的关键步骤和难点在于,指标体系的确立和评估模型的选择——如何将多维指标综合起来。目前采用的方法有统计方法、专家系统、神经网络技术等。

### 2.1 统计方法

国外在对信用风险的评估中,广泛采用了基于统计判别方法的预测模型。这些方法和模型都是在

Fisher 于 1936 年作出的启发性研究之后提出来的<sup>[6]</sup>。总的来说, 这些模型都被表述为一类分类问题, 它们接受定义在已选变量集合上的一个随机观测值样本, 建立判别函数, 进行分类。根据判别函数形式和样本分布的假定不同, 常见的模型有: 回归分析法<sup>[7]</sup>、多元判别分析法<sup>[8]</sup>、Logit 法<sup>[9]</sup>、Probit 法<sup>[10]</sup>、因子 Logistic 法<sup>[11]</sup>和近邻法<sup>[12]</sup>等。尽管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<sup>[13]</sup>。

### 2.1.1 判别分析(Discriminant Analysis)

在针对分类问题的多种不同的统计方法中, 最常用的是多元判别分析法(MDA: Multivariate Discriminant Analysis)。MDA 是根据观察到的一些统计数字特征, 对客观事物进行分类, 以确定事物的类别。它的特点是已经掌握了历史上每个类别的若干样本, 从中总结出分类的规律性, 建立判别公式, 当遇到新的事物时, 只要根据总结出来的判别公式, 就能判别事物所属的类别。

MDA 的关键在于建立判别函数, 目前, 统计学中建立判别函数常用方法有: 一是已知总体分布的前提下求得平均出错概率最小的分类判别函数, 通常称为贝叶斯(Bayes)判别函数; 二是未知总体分布或未知总体分布函数时, 费歇(Fisher)准则下的最优线性判别函数<sup>[14]</sup>。在这里, 判别分析模型是通过采用极大化组间比和组内方差的 Fisher 方法来建立的。在满足如下条件时, 可以证明, 由 Fisher 方法得出的判别规则最优极小化误判代价。

- a 每组均服从多元正态分布;
- b 每组的协方差矩阵相同;
- c 每组的均值向量、协方差矩阵、先验概率和误判代价是已知的

MDA 有两个总体与多个总体判别之分。在两个总体判别时, 判别公式为:

$$DF(X) = X^T \sum^{-1} (\mu_1 - \mu_2) - 1/2 (\mu_1 - \mu_2)^T \sum^{-1} (\mu_1 + \mu_2)$$

其中,  $\mu_1, \mu_2$  和  $\sum$  分别是各组均值和共同协方差矩阵, 依据一定判别规则即可对原始样本进行分类。显然, 判别函数  $DF(X)$  是各独立变量(或指标)的线性函数。在满足上述 a、b、c 假设下, 它能使误判达到最小。

Altman 应用判别分析法建立了著名的 Z-score 模型和在此基础上改进的 ZETA 模型, 这两个模型的目的是预测企业破产的概率, 为银行贷款决策提供支持。目前 ZETA 模型已商业化, 广泛应用于美国、意大利等国的商业银行, 取得了巨大的经济效益<sup>[15]</sup>。这一方法的最大缺陷在于其过于严格的假定条件, 实际中均有不同程度的违背。

### 2.1.2 Logit 方法

Logit 分析与判别分析法的本质差异在于前者不要求满足正态分布或等方差, 其模型采用 Logistic 函数,

$$Y = \frac{1}{1 + e^{-\eta}}, \quad \eta = c_0 + \sum_{i=1}^p c_i X_i$$

其中  $X_i (1 \leq i \leq p)$  表示第  $i$  个指标,  $c_i$  是第  $i$  个指标的系数,  $Y$  是一表示企业财务状况好坏的概率。

由于 Logistic 回归不假定任何概率分布, 不满足正态情况下其判别正确率高于判别分析法的结果。

### 2.1.3 近邻法

近邻法是一种非参数方法, 当已知总体表现为显著非正态分布时, 特别是当属于同一类的样本在变量空间形成聚类时, 近邻法十分有效。与参数类方法相比, 近邻法用于对总体分布施加很少约束的情况, 是一种十分灵活的方法。

近邻法不仅放松了正态性假定, 也避免了传统技术对模型函数形式设定的困难。任何一个样本到底划归哪一类是由其  $k$  个近邻划归类型所确定。任意两个样本之间的距离可定义为  $d(x, y) = (x - y)^T \text{cov}^{-1}(x - y)$ , 其中  $\text{cov}^{-1}$  是合并协方差的逆。这样, 一个样本划归为它的  $k$  个近邻的多数(即当一个样本的  $k$  个近邻的大多数划归 1 类, 则该样本也应划属 1 类)。

## 2.2 分类树<sup>[12]</sup>

80 年代末期, 有学者提出一种利用机器学习技术发展起来的符号方法——分类树。该方法不象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。此前,曾有学者采用了一种叫做递归分割的类似技术生成判别树。两种方法都采用了一种非返回跟踪的分割方法将样本集递归分割成不相交的子集。它们的差别只是在分割准则上,分类树方法旨在极大化分割子集的熵,而递归分割方法则旨在使期望误判损失达到最小。

上述四种方法在国外已有大量应用,实证结果发现<sup>[12,13]</sup>: 1) 企业财务状况的评价可以看作是一类基于一系列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比率的关系常常是非线性的; 3) 预测变量(财务比率)可能是高度相关的; 4) 大量实证结果表明,许多指标不成正态分布。

因此,传统的分类方法不能很好地解决这些问题。作为研究复杂性的有力工具,神经网络技术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,特别是能处理任意类数据,这是许多传统方法所无法比拟的。通过不断学习,能够从未知模式的大量的复杂数据中发现其规律。神经网络方法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,它是一种自然的非线性建模过程,毋需分清存在何种非线性关系,给建模与分析带来极大的方便。该方法用于企业财务状况研究时,一方面利用其映射能力,另一方面主要利用其泛化能力,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>[16]</sup>。

针对这种情况,本文的目的在于,根据我国商业银行的具体情况,利用人工神经网络技术,研究我国商业银行的信用风险评估问题。

### 3 人工神经网络基本原理

人工神经网络(Artificial Neural Network)是由大量简单的处理单元广泛连接组成的复杂网络,用于模拟人类大脑神经网络的结构和行为。它反映了人脑功能的许多基本特性,但它并不是人脑全部的真实写照,而只是对其作某种简化、抽象和模拟<sup>[17]</sup>。

#### 3.1 前向网络

神经网络是由按层进行组织的神经元和连接弧所组成的,这些神经元以不同的方式组织起来形成不同的网络结构。本文所应用的是一种称为前向网络的特殊的网络结构。

在前向网络中有三种神经元:输入神经元、输出神经元和隐含神经元。输入神经元接收环境的信息,位于网络的最底层;输出神经元把信息传递到环境中去,位于网络的最高层;隐含神经元不与环境发生交互作用,因此是不可见的。其中每个神经元都是一个简单的计算装置,其特性由简单的数学函数所描述。神经元*i*接收其它神经元传递来的输入信息,根据和函数 $net_i$ 进行加权平均,然后根据传递函数 $f_i$ 产生输出信息,输出信息又按照网络的拓扑结构传递到下一个神经元。我们应用Rumelhart于1986年提出的如下函数:

$$I_i = \sum_j w_{ij}x_j + \phi, \quad x_i = f_i = \frac{1}{1 + e^{-net_i}}$$

其中, $I_i$ 为神经元*i*的层输入; $x_i$ 为神经元*i*的输出; $w_{ij}$ 为神经元*i, j*之间的连接权; $\phi$ 为神经元*i*的偏置。

每一条连接弧都被赋予一定的数值来表示连接弧的连接强度。正的权值表示影响的增加,负的权值表示影响的减弱。在前向网络中,神经元间前向连接,同层神经元互不连接,信息只能向着一个方向传播。前向网络的连接模式是用权值向量 $w$ 来表示的。在网络中,权值向量决定着网络如何对环境中的任意输入作出反应。同样,网络也是通过不断调整权值来完成整个学习过程的。

#### 3.2 反向传播学习算法(Back-propagation Learning Algorithm)

如何针对一个特定的任务设定一个合适的权值是非常困难的,特别是不知道总体分布的时候。通常的解决办法是通过对样本的训练来完成网络的学习任务的。典型的学习算法是搜索权值空间以找到最适合给定样本的权值。在各种学习算法中,多层网络的反向传播算法(简称BP算法)应用最为广泛。BP算法最早是由Werbos在1974年提出来的,Rumelhart等人于1985年发展了该理论,提出了清晰而又严格的算法。BP算法适用于前向网络,它采用有导师学习的训练形式,提供输入矢量集的同时提供输出矢量集,通

过反向传播学习算法, 调整网络的连接权值, 以使网络输出在最小均方差意义下, 尽量向期望输出接近, 反向学习的进程由正向传播和反向传播组成. 在正向传播过程中, 输入信息经隐含神经元逐层处理并传向输出层, 如果输出层不能得到期望的输出, 则转入反向传播过程, 将实际输出与期望输出之间的误差沿原来的连接通路返回, 通过修改各层神经元的连接权值, 使误差减小, 然后转入正向传播过程, 反复循环, 直至误差小于给定的值为止

设有  $N$  个训练对组成的训练集, 每一个训练对用输入矢量  $X_i = (x_{i1}, x_{i2}, \dots, x_{in})$  和输出矢量  $D_i = (d_{i1}, d_{i2}, \dots, d_{in})$ ,  $1 \leq i \leq N$ . 在前向传播中, 把  $X_i$  作为网络的输入, 根据现有的  $W$  计算网络的输出  $Y_i = (y_{i1}, y_{i2}, \dots, y_{in})$ . 比较实际输出  $Y_i$  与期望输出  $D_i$  之间的差异, 计算每一个输出单元的平方误差  $(y_{ij} - d_{ij})^2$ ,  $1 \leq j \leq n$ . 把这些误差进行加总得到误差函数

$$E = \sum_{i=1}^N \sum_{j=1}^n \frac{(y_{ij} - d_{ij})^2}{2}$$

我们所要做的就是通过改变  $W$  来减小  $E$ , 以使得所有的输入矢量都尽量与相应的输出矢量相匹配. 因此学习的过程就转化为定义在权值空间上的目标函数  $E$  的极小化问题

在训练过程中总是以尽可能快的减小  $E$  的方式进行. 一般它依赖于在权值空间中是否沿梯度方向搜索, 所以采用梯度下降法来训练权值. 每一个权值  $w_{ij}$  的变化量  $\Delta w_{ij}$  按如下方式计算

$$\Delta w_{ij} = -\eta \frac{\partial E}{\partial w_{ij}}$$

其中  $\eta$  为学习率, 是控制算法收敛速度的参数

在第一阶段得到的总误差平方和在第二阶段被一层一层地反向传播回去, 从输出单元到输入单元. 权值的调整决定于传播过程中的每一步. 由于  $I_{kf}$  和  $E$  都是连续可微的, 因此, 可以应用以下公式计算  $E / \Delta w_{ij}$  的值

$$\frac{\partial E}{\partial w_{ij}} = \frac{\partial E}{\partial f_i} \frac{\partial f_i}{\partial a_i} \frac{\partial a_i}{\partial w_{ij}}$$

$W$  的修改可以有两种方式, 一是对于每一训练对  $(X_i, D_i)$  都修改一次  $W$ , 另一种方式是输入全部的训练对后再加总  $\Delta w_{ij}$  并进行修改. 训练矢量集中训练对的数目称为一个 epoch. 当 epoch 不是非常大的时候, 后一种方式能够加快收敛的速度. 因为第一种方式只能针对某一特定的训练对减小误差函数, 而可能增大其它训练对的误差函数; 第二种方式总是以减小总体误差函数为目标的. 所以我们采用第二种方式

## 4 样本数据和模型构造

### 4.1 指标与样本数据

本文以某国有商业银行的信用风险评估为例, 选择其九十余家企业客户为对象, 考察它们的短期贷款偿还情况. 所用的财务比率是参照国内财政部考核企业财务状况、国外相关财务比率指标, 对反映企业流动性、盈利性、增长性、速动性、偿债性等方面指标, 利用 STATGRAPHICS 软件包进行主因子分析, 得到以下五个比率:

- $X_1 =$  运营资金 / 资产总额
- $X_2 =$  保留盈余 / 资产总额
- $X_3 =$  息税前利润 / 资产总额
- $X_4 =$  普通股、优先股市场价值总额 / 负债账面价值总额
- $X_5 =$  销售收入 / 资产总额

调查取得的样本总数为 95 家, 经过对样本的初步审查, 删除不合格的样本 8 个. 因此有效样本总数为 87, 其中能够偿还银行信用贷款的企业 44 家, 剩下 43 家不能按时偿还贷款. 作为一种控制的手段, 每一家能够偿还贷款的企业都按照资产规模、行业范围等与一家不能偿还贷款的企业相匹配. 我们将这 87 家企业划分为两个子样本集合, 根据经验的做法, 训练子样本集合由 54 家企业构成, 包括 27 家能够偿还贷款的企业与 27 家不能偿还贷款的企业; 剩下的 33 家企业组成测试子样本集合, 包括 17 家能够偿还贷款的企业

和 16 家不能偿还贷款的企业

4.2 模型构造

4.2.1 线性判别模型

我们利用上述五个比率来构造判别分析模型 对根据样本数据计算得到的五个比率进行 Kolmogorov-Smirnov 检验以考察是否服从正态分布的假设 在删除异常数据后,  $X_1, X_2, X_3, X_5$  均服从正态分布, 只有  $X_4$  不服从正态分布(注: 取对数处理后仍不满足正态分布), 这必然会对判别分析的结果产生一定的影响 判别分析模型是根据 STATGRAPHICS 软件包计算而得到的

4.2.2 神经网络模型

神经网络模型的建立主要需考虑两方面的问题: 一是确定网络结构; 二是学习参数的调整 首先考虑网络结构的确定 网络结构主要包括连接方式、网络层次数和各层节点数 网络的连接方式代表了网络的拓扑结构, 本文应用的前向网络是目前应用最广的神经网络 可以证明<sup>[18]</sup>, 在一定条件下, 一个三层的 BP 网可以以任意精度去逼近任意映射关系 而且经过实验发现, 与一个隐含层相比, 用两个隐含层的网络训练并无助于提高预测的准确率 各隐含层节点数的选择较为复杂, 并无确定的法则, 只能根据一些经验法则, 通过实验来确定 一般来说, 可考虑的经验法则有: 1) 隐含层节点数不能是各层中节点数最少的, 也不是最多的; 2) 较好的隐含层节点数介于输入节点和输出节点数之和的 50%至 75%之间; 3) 隐含层节点数的理论上限由其训练样本数据所限定

根据以上几条, 通过试值法测试隐含层节点数对网络性能的影响 表 1 显示了不同的隐含层节点数对应的训练集合和测试集合的误判值

表 1 隐含层节点数的影响

隐含层 节点数	第一类错误		第二类错误		总误判	
	训练集	测试集	训练集	测试集	训练集	测试集
0*	5	8	2	5	7	13
1**	4	7	3	6	6	13
2	3	7	1	5	4	12
3	3	8	0	5	3	13
4	3	7	0	5	3	12
5	3	6	0	4	3	10
6	3	7	0	5	3	12
7	3	9	0	6	3	15
8	2	11	0	8	2	19

注: \* 当误差平方和达到 0.031 时陷入训练瘫痪; \*\* 当误差平方和达到 0.028 时陷入训练瘫痪

我们把误判分为两类: 第一类错误与第二类错误 在统计学中, 第一类错误称为“拒真”, 第二类错误称为“纳伪” 这里, 我们把第一类错误定义为把不能偿还贷款的企业误判为能偿还贷款的企业的错误, 第二类错误定义为把能够偿还贷款的企业误判为不能偿还贷款的企业的错误 显然, 第一类错误比第二类错误严重的多 犯第二类错误至多是损失一笔利息收入, 而犯第一类错误则会造成贷款不能收回, 形成呆帐 因此我们应尽量避免第一类错误 表 1 显示当隐含层节点数为 5 时, 误差是最低的 输出层单元数在本例中可以选择一个(输出结果为 0/1)或两个(00/01), 试验发现, 两个输出单元的结构提高了判别的准确率 最终我们选定的网络结构包含 5 个输入节点、5 个隐含层节点以及 2 个输出节点, 如图 1 所示

Stornetta 和 Huberman 曾经讨论过把输入、输出的值调整到  $[-0.5, 0.5]$  之间对训练过程的影响<sup>[19]</sup>, 结果表明收敛时间缩短了 30~ 50%. 修改后的传递函数如下

$$f_i = \frac{1}{2} \frac{e^x - e^{-x}}{e^x + e^{-x}}$$

我们的结果表明, 采用这种修改后的传递函数, 收敛时间缩短了 45%.



为了快速计算, Rumelhart, Hinton 和 Williams 1986 年对 BP 算法进行修整以改进训练时间, 即在前面的公式上再加一个称为“惯性项”的调整量, 则调整公式为

$$w_{ij(n+1)} = -\eta \frac{\partial E}{\partial w_{ij(n)}} + \alpha w_{ij(n)}$$

其中  $\alpha$  为惯性参数,  $0 < \alpha < 1$ .

该方法将前次计算的权重引入学习过程, 用惯性项保持其变化速度于一定程度, 以抑制收敛过程中振荡的发生

学习参数  $\alpha$   $\beta$  的调整亦只能采用试值法 首先置初值为  $\alpha = 0.90$ ,  $\beta = 0.70$  由于在网络训练过程中并未出现振荡, 所以保持学习率  $\alpha$  不变, 保证网络以较快的速度收敛; 逐渐增加  $\beta$  值, 调试网络性能, 发现当  $\beta = 0.66$  时, 网络收敛最快 若再增大  $\beta$  值, 又开始出现振荡 因此, 最后选定的学习参数为  $\alpha = 0.90$   $\beta = 0.66$

本网络的训练是应用 NeuraWorks Explorer 软件包在 486 微机上实现的 每次运行的最大训练次数设定为 4000 次

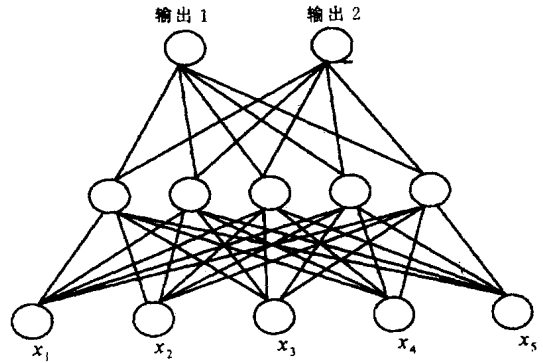


图 1 神经网络结构

## 5 结果分析

应用 STATGRAPHICS 软件包, 最终得到线性判别模型的判别系数如表 2 所示 五个比率中的四个在 1% 的显著水平下服从正态分布 通过考察相关系数矩阵来检验解释变量的线性相关性, 发现没有变量是其它变量的线性组合

表 2 判别分析结果

变量	$X_1^*$	$X_2^*$	$X_3^*$	$X_4$	$X_5^*$	常量
系数	0.32054	1.78868	5.16075	0.07098	-0.06743	-0.52731

\* 在 1% 的显著水平下服从正态分布

每一个模型的预测精度都同时用训练样本和测试样本表示出来 通常精度的度量用两类错误来度量 结果如表 4 所示 几乎从所有的方面来看, 神经网络都优于判别分析法, 除了在测试样本中神经网络的第二类错误多了一个误判 且在两类错误中, 第一类错误显然更为严重, 因而它的误判代价也就更高一些

表 3 训练样本和测试样本的误判

模型	训练样本			测试样本		
	第一类错误	第二类错误	总误判	第一类错误	第二类错误	总误判
LDA	5(13.16%)	0(0.00%)	5(6.76%)	11(40.74%)	3(10.71%)	14(25.45%)
NN	3(7.89%)	0(0.00%)	3(4.05%)	6(22.22%)	4(14.29%)	10(18.18%)

## 6 总结

本文将神经网络技术应用于商业银行信用风险评估中, 结果表明, 神经网络技术具有广泛的应用前景 总体来讲, 神经网络具有一些其他方法无法比拟的优点:

1) 神经网络方法是一种稳健的、非参数的方法, 具有很强的非线性映射能力, 其学习经验的能力强, 分类精度高;

2) 神经网络采用分布式存储结构, 容错能力强 网络中少量单元的局部缺损不会造成网络的瘫痪, 影响全局, 反映了神经网络的鲁棒性;

3) 神经网络进行大规模并行处理, 反映为同一层所有神经元同时进行计算, 且每一神经元存储的信息也同时参与计算;

4) 神经网络具有在新环境下的泛化能力, 即在经过一定数量带噪声样本训练之后, 网络通过学习来抽取规则或记忆知识, 抽取样本隐含的关系并记忆, 并对新情况下数据进行内插或外推;

5) 神经网络能不断接受新样本、新经验并不断调整模型, 自适应能力强, 具有动态特性

尽管神经网络已经得到了广泛的应用, 但也存在以下缺点:

**网络结构:** 建立神经网络模型中最重要的问题就是确定网络的拓扑结构。对于一个特定的分类问题并没有一个规范的模式来构造网络的结构, 通常都是通过实验的方法来得到一个适合的网络结构。Lippman (1987) 建议, 对于任意复杂的判别函数, 一个隐含层就足够了<sup>[19]</sup>。虽然如此, 隐含层单元数的选择余地仍然是非常大的。隐含层单元数的增多, 有助于对训练样本的拟合, 但同时降低了模型的自由度, 削弱了模型的泛化能力, 即过度训练问题。

**训练效率:** 除了上述网络结构中存在的问题之外, 还有训练的复杂性问题。

a. **收敛问题:** Rumelhart 等人在 1986 年曾用微分方程作了证明, 仅当网络权值调整步无限小时, 收敛才有效。这意味着, 训练要用无限的收敛时间。经验观察也说明了训练过程中收敛是不可预测的, 且时间是相当长的。

b. **局部极小问题:** BP 学习算法用梯度下降法调整网络权值, 这对凸状的差错曲面将是有效的, 但对实际问题中经常遇到的严重卷绕的非凸状曲面却不能取得最佳效果;

c. **训练瘫痪问题:** 在某些情况下, 网络训练过程中权值的修改实际上是停顿了。一旦进入这种状态, 它使训练时间按数量级增长, 这种网络瘫痪是一个严重问题;

d. 系统在训练过程中可能产生振荡行为。

以上几点都是在网络训练过程中出现的若干问题。通过采用一些改进的算法, 大部分问题都能得到一定程度的解决。

**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。人们无法根据网络的权值得到输入变量的相关重要性。近年来, 在这个研究方向上取得了不少的进展。Miller, Todd 和 Hedge (1989) 建议在构造网络结构时使用遗传算法<sup>[17]</sup>。还有一些其它学者使用启发方法来解释输入变量的相关重要性。如 Tam 和 Kiang (1990) 应用误差函数的偏导数来解决问题<sup>[12]</sup>, 还有人则把模糊逻辑与神经网络的解释能力联系起来。

## 参考文献:

- [1] Freeman A. A survey of international banking. The Economist, 1993: 1~ 37.
- [2] Saunders A. Financial Institutions Management. Boston: Irwin, Inc., 1994.
- [3] 世界银行. 新兴市场经济中的商业银行. 北京: 中国财政经济出版社, 1994.
- [4] 曾国坚, 何五星. 银行风险论. 北京: 中国计划出版社, 1995.
- [5] Freedman R S, et al. AI in the Capital Markets. Chicago: Probus Publishers, 1995.
- [6] Fisher R A. The use of multiple measurements in taxonomic problems. Ann. Eugenics, 1936, 7: 179 ~ 188.
- [7] Meyer P A, Pifer H. Prediction of bank failures. Journal of Finance, 1970, 25: 853~ 868.
- [8] Altman E I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. Journal of Finance, 1968, 23: 589~ 609.
- [9] Press S J, Wilson S. Choosing between Logistic regression and discriminant analysis. J. Amer. Statist. Assoc., 1978, 73: 699~ 705.
- [10] Barth J R, et al. Thrift institution failures: estimating the regulator's closure rule. G. G Kaufman (Ed.), Research in Financial Services, Greenwich, CT: JAI Press, 1989.



- [12] Tam K Y, Kiang M. Managerial applications of neural networks: the case of bank failure predictions. *Management Sciences*, 1992, 38(1): 926~ 947.
- [13] Eisenbeis R A. Pitfalls in the application of discriminant analysis in business, finance, and economics. *Journal of Finance*, 1977, 32: 875~ 900.
- [14] 张尧庭, 方开泰. 多元统计分析引论. 北京: 科学出版社, 1982.
- [15] Altman A I. *Corporate Financial Distress*. New York: John Wiley & Sons, 1983.
- [16] Odom M D, Sharda R. A neural network for bankruptcy prediction. *International Joint Conference on Neural Network*, June 17~ 20, 1990, vol II (San Diego, Calif.), 163~ 168.
- [17] 施鸿宝. 神经网络及其应用. 西安: 西安交通大学出版社, 1993.
- [18] Lippman R P. Introduction to computing with neural nets. *IEEE ASSP May*, 1987, 4: 4~ 22.
- [19] Stornetta W S, Huberman B A. An improved three layered backprop algorithm. *Proceedings of the IEEE First International Conference on Neural Networks*, eds M. Caudill and C. Butler (San Diego, Calif: SOS Printing), 1988.

---

(上接第 15 页)

#### 参考文献:

- [1] Vijay Grubaxani. The New World of Information Technology Outsourcing. *Communication of The ACM*, 1996, 39(7): 45~ 46.
- [2] Richard L Huber. How Continental Bank Outsourced Its "Crown Jewels." *Harvard Business Review*, Jan-Feb, 1993: 121~ 129.
- [3] 《微电脑世界》. Trust Bank 将 IT 运作全部外包. *微电脑世界*, 1997.
- [4] Lacity M C, Willcocks L R and Feeny D F. IT Outsourcing: Maximize Flexibility and Control. *Harvard Business Review*, May-June, 1995.
- [5] Willcocks L P and Currie W L. Information Technology in Public Services: Towards the Contractual Organization? *British Journal of Management*, June, 1997, Vol Special Issue: 107~ 120.
- [6] Lance Eliot. From Downsizing to Upsizing: Doing IT Right. *Decision Line*, July 1996: 9~ 10.
- [7] Pranlahad C and Hamel G. The Core Competence of the Corporation. *Harvard Business Review*, 1990, 63(3): 79~ 91.
- [8] Quinn J B and Himer F G. Strategic Outsourcing. *Sloan Management Review*. Summer, 1994: 43~ 55.
- [9] Leslie Willcock and Guy Fitzgerald. *Business Guide to IT Outsourcing*. Business Intelligence, 1994.
- [10] 李敏强, 纪仕光, 寇纪淞. 信息技术对现代企业结构和管理模式变革的影响. *系统工程与市场经济*. 中国系统工程学会编著, 北京: 科学技术文献出版社, 1996.