

改进后的神经网络模型在财务预警中的应用

周 游

(江苏经贸职业技术学院 南京 210007)

企业财务预警是以企业财务信息为基础,通过分析一些敏感性财务指标的变化,对企业可能面临的财务危机进行预测。BP神经网络模型是目前在财务预警领域应用较多也是较成功的一种神经网络模型。但BP神经网络模型在应用中也存在明显的缺陷:①描述财务比率的指标过多,如果将这些指标都作为模型输入将导致网络结构过于庞大;另一方面也将导致神经网络的泛化能力下降,使其预测性能变坏。因此很多研究中只选取部分指标作为模型输入,这样又将导致分析信息的不完整;②由于BP算法采用的是梯度下降搜索方法,在训练过程中容易陷入局部最优,陷入局部最优将直接导致分析结论的偏差。因此,如何在财务预警分析中改善神经网络的性能是一个关键问题。本文提出首先采用主分量分析对高维指标数据进行降维,在保留绝大多数指标信息的前提下缩小网络规模,然后采用具有全局最优搜索能力的粒子群优化算法实现对神经网络连接权值的优化训练,以避免BP神经网络陷入局部最优,提高BP神经网络模型预警结果的科学性和可靠性。

一、主分量分析的原理及实现

在进行数据分析和处理的过程中,设输入原始数据 x 的维数是 N ,希望通过预处理得到 $M(M < N)$ 维数据 y ,如果不加任何限制条件,仅对 x 进行简单的截断,那么所引起的均方误差将等于舍弃的各分量方差之和。为此,希望找到一个线性变换 W ,使得对 Wx 的截断在最小均方误差下为最优,这就要求被舍弃的分量具有较低的方差,而保留的分量具有较高的方差。PCA正是寻找这个线性变换的方法。

从几何变换的观点看,PCA的基本思想是寻找一个最佳子空间,当高维数据 x 在该子空间进行投影后,所得分量具有最大方差。同时,在子空间用新分量对原始数据进行重建时,在均方误差最小的情况下逼近效果最优,即使式(1)最小化:

$$E\left\{\|x - \sum_{j=1}^M (W_j^T x) W_j\|^2\right\} \quad (1)$$

PCA的目的就是找到一个正交变换矩阵 $W^T = [w_1, w_2, \dots, w_M]$ 对 N 维向量 x 进行正交变换,使得变换结果 y 的各分量 $\zeta_i (i=1, 2, \dots, M)$ 间互不相关。并且当所有观测数据 x 沿 w_1 方向投影时,PCA将使得到的分量 ζ_1 能量最大,即方差 $E\{\zeta_1^2\}$ 最大,这时 ζ_1 称为第一主分量(PC1);在与 w_1 正交的条件下,观测数据 x 在 w_2 上投影,使 ζ_2 能量最大,这时 ζ_2 称为第二主分量(PC2)。对于 N 维向量 x ,由于投影后的维数

$M \leq N$,因此最多可以得到 N 个分量。在实际应用中通过截取其中 $d(d < N)$ 个主分量实现信息提取和降维。

二、粒子群优化神经网络训练方法

粒子群优化算法是一种群智能演化计算技术,与遗传算法比较,PSO没有遗传算法的交叉以及变异操作,因此算法简单高效,反映了群智能算法的最新进展。

PSO中,每个优化问题的潜在解都是搜索空间中的一个“粒子”。所有的粒子都有一个由被优化的函数决定的适应值。在每一次迭代中,粒子通过跟踪两个“极值”来更新自己。第一个极值就是粒子本身所找到的最优解(个体极值 $pBest$);另一个极值是整个种群目前找到的最优解(全局极值 $gBest$),粒子根据如下公式更新自己的速度和位置:

$$v(t+1) = W \cdot v(t) + C_1 \cdot \text{rand}() \cdot [pBest(t) - x(t)] + C_2 \cdot \text{rand}() \cdot [gBest(t) - x(t)] \quad (2)$$

$$x(t+1) = x(t) + v(t+1) \quad (3)$$

其中: $v(t+1)$ 是粒子下一时刻的速度; $x(t+1)$ 是粒子下一时刻的位置; C_1 和 C_2 是学习因子,通常 $C_1 = C_2 = 2$; W 为惯性权重,一般 $W \in [0, 1]$ 。为了避免基本PSO在搜索过程中陷入局部最优,研究者们陆续提出很多PSO改进算法,这些新算法都使PSO的优化性能得到进一步提高。

采用PSO算法训练神经网络时,首先将特定结构中所有神经元间的连接权值编码成实数码串表示的个体。假设网络中包含 M 个待优化权值(包括阈值在内),则每个粒子将由 M 个权值参数组成的一个 M 维向量来表示。根据微粒群规模,按照上述个体结构随机产生一定数目的个体(微粒)组成种群,其中不同的个体代表神经网络的一组不同权值,将微粒群中每一个体映射为神经网络中的权值的解,网络权值的优化过程是一个全局寻优的过程。

计算网络在训练集上产生的均方误差,并以此构造适应度函数,用来计算个体的适应度。

$$E(X_p) = \left[\sum_{p=1}^n \sum_{k=0}^c (Y(X_p) - T_{p,c}) \right]^2 \quad (4)$$

其中: $T_{p,c}$ 为训练样本 p 在 c 输出端的给定输出; $Y(X_p) = f(X_p, W)$ 为网络输出, W 为网络连接权。

当适应度函数取得最小值时,对应神经网络输出误差最小,所对应的权值 W^* 为所求神经网络的训练结果。

三、财务指标体系

关于财务困境的描述,Altman等综合了学术界定义财务

困境的四种情形:经营失败、无偿付能力、违约和破产。我国证券市场上被特别处理(ST)的股票大多是由于“连续两年亏损或每股净资产低于股票面值”,即财务指标的恶化是上市公司被特别处理的主要原因。由于数据披露等问题,国内学者一般将ST的上市公司作为财务困境公司。基于相同的原因,本文中财务困境的界定以是否被ST为准,如果上市公司被ST,则该公司陷入财务困境;反之,则没有陷入财务困境。

有关财务困境预测模型的分析中,常用的财务比率主要有30个(见表1)。如果将30个财务比率指标全部作为神经网络的输入(即30维向量),则输入节点多达30个,过大的网络结构一方面要求巨大的训练样本集支持,另一方面也将导致神经网络的泛化能力下降;如果简单地从30个指标中选取部分指标作为神经网络输入,则因为输入的信息不完整将会使预测结果的客观性和可靠性降低。

表1 常用财务困境预测比率指标

序号	指 标	序号	指 标
1	现金/流动负债	16	速动资产/净销售收入
2	总负债/权益	17	流动资产/总资产
3	现金/净销售收入	18	速动资产/总资产
4	净收入/总资产	19	流动负债/权益
5	现金/总资产	20	普通股收益率
6	速动资产/存货	21	权益/固定资产
7	现金流/流动负债	22	留存收益/总资产
8	净销售收入/总资产	23	权益/净销售收入
9	现金流/总负债	24	总负债/总资产
10	营业收益/总资产	25	存货/净销售收入
11	现金流/总资产	26	流动资金/净销售收入
12	息税前利润/利息总额	27	长期负债/权益
13	流动资产/流动负债	28	流动资金/总资产
14	速动资产/流动负债	29	权益市场价值/负债
15	流动资产/净销售收入	30	流动资金/总资产

为此,本文利用PCA的降维与信息提取能力对原来的30个财务比率指标的高维向量进行处理。对于K组财务指标向量 x_i ,其协方差矩阵为:

$$C_x = \frac{1}{K} \sum_{i=1}^K (x_i - \bar{x})(x_i - \bar{x})^T \quad (5)$$

其中: \bar{x} 是向量 x_i 的均值,协方差矩阵 C_x 的30个递减特征值对应的特征向量 w_1, w_2, \dots, w_{30} 构成特征空间的一组样本集。实验证明当 $d=5$ 时,主分量方差已达到总方差的98.7%。将原来的30维样本 x_i 在这个($d=5$)维空间投影就得到降维后的样本 $x'_i = W_d x_i$ 。

神经网络的输出用来描述分析结果,即公司处于财务困境或未处于财务困境两种状态,因此可以构造单隐含层神经网络结构:6个输入层节点(第6个节点输入始终为-1,作为阈值使用),8个隐含层节点,2个输出层节点(描述ST/非ST二种状态,处于某状态,则对应的输出节点为1,反之为0),由此建立神经网络预测模型。

四、改进神经网络财务预警实例分析

为了使财务指标具有可比性,实验中选择的140家规模相当的制造业上市公司作为研究对象,其中财务困境类(ST)公司56家和非财务困境类(非ST)公司84家。

分别选取40家ST企业和40家非ST企业作为训练样本集,剩下的16家ST企业和44家非ST企业作为测试样本集。所有数据来自于网上公开披露数据。

实验采用上节建立的模型,以高斯变异粒子群优化方法对神经网络进行训练,步骤如下:

步骤1:设置粒子种群数为30,根据网络规模设置粒子维数($6 \times 8 + 8 \times 2 = 64$),随机初始化各粒子的起始位置和起始速度,初始化的各个粒子构成最优神经网络权值初始候选解集。

步骤2:粒子群算法迭代开始,计算各粒子在(4)式准则下的适应度函数值。

步骤3:个体极值的更新。比较每个粒子的当前适应度 $fitness(t)$ 与迭代前的个体极值 $fitnesspBest(t-1)$,若粒子的当前适应度优于迭代前的个体极值,则进行个体极值的更新,否则保留原个体极值。

步骤4:从种群中选取个体 m_i ,进行高斯变异操作,产生一个新个体 m'_i ,按Altman的算法判断是否保留变异粒子。

步骤5:全局极值的更新。在所有的个体极值中,适应度最优的个体极值即为全局极值。

步骤6:判断是否满足迭代停止条件,即是否达到设定的迭代次数($=1000$)。“否”返回步骤2,“是”转到步骤7。

步骤7:全局极值对应的粒子为种群的最优解,即神经网络权值的最优解 W^* 。

实际应用中,模型的预测能力可用两类错误来度量,第一类错误为将ST公司判为非ST公司(即将困境公司误判为正常公司),第二类错误为将非ST公司判为ST公司(即将正常公司误判为困境公司)。为了验证方法的有效性,分别采用Altman预测模型中的5维财务指标(营运资金/总资产、留存收益/总资产、息税前盈利/总资产、资本化市值/总资产、销售收入/总资产)作为神经网络输入以及本文中PCA降维后的指标作为神经网络输入,模型训练结束后,用测试样本集对神经网络模型进行测试。

在有些研究中,将训练过的模型对训练样本集的测试结果也作为一项性能指标,但从统计学习理论可知:只要构造一个足够复杂的“学习机”,总可以使其对训练样本具有理想的测试结果,但这对提高模型的“预测能力”并无帮助。

表2 测试样本集测试结果(总样本数60)

	I类错误	II类错误	错误加总
Altman5维指标	5(8.3%)	8(13.3%)	13(21.6%)
本文方法	3(5.0%)	5(8.3%)	7(13.3%)

两种方法对测试样本集测试结果存在明显的差异,因为Altman选择的5维指标只利用了原来30维指标的部分信息,而本文方法充分利用了财务指标的有效信息并通过神经网络优化训练避免了局部最优,因此预警能力显著提高。○