

# Adaboost-BP 方法在企业信用评级中的应用研究

各机构应独立于其他经济利益，加强自我监督、明确自身职业道德标准、培养社会责任感，运用科学的方法做出客观评判。

◎ 王会昌 高 静 / 文

## 企业信用评级指标变量的选取

信用评级指标体系是进行企业信用评级工作的依据，也是衡量信用评级结果是否客观公正的标尺。本文对国内外几大企业信用评级机构的指标体系进行归纳汇总，结果显示各指标体系在资产结构、资本流动性、经济效益、现金流量以及企业的发展前景等方面都有所涉及。但具体来看仍存在着诸多差异，各机构均自立体系进行评级，还未形成规范统一的指标体系。

企业信用评级方法研究的开创者 Altman 在建立 Z 值模型时，曾提出最初指标的选取应遵循两个原则：一是该指标在以前的研究中出现频率；二是该指标与所研究问题的潜在相关性。在考虑了这两个基本原则的基础上，本文参考国内外主要评级指标体系，并结合我国证券市场的实际情况，选取企业信用评级指标体系（如表 1 所示）。

## Adaboost-BP 方法及其适用性分析

在过去的企业信用评级相关研究中，可以发现企业的信用水平与企业信用评价指标之间的关系存在以下几个特点：一是企业信用评级实质上是一种分类问题；二是现实中一个企业信用水平的高低与我们所选取的评价指标之间常常体现出非线性的关系；三是企业信用评价指标的分布在现实中并不呈现正态分布；四是我们所选取的企业信用评级指标之间并非相互独立，它们所反映的经济内容存在很大程度上的相关性。

根据以上特点，显而易见，传统的评级方法在实际评级过程

中是很难适用的。相较之，Adaboost-BP 方法则体现了其非凡的优越性。首先，该方法是一种非线性模型，其自然的学习过程对非线性关系具有很好的模拟能力且给构建模型与研究分析提供了很大的方便。其次，该方法在数据分布未知或是数据不连续、存在相关关系等问题下都不会影响其辨识和学习。再次，该方法具有强大的自适应能力，它只通过预先提供的一批相互对应的输入——输出数据，即可自动调节训练样本的权重并对得到的弱分类器进行最后的整合，找到两者之间的潜在规律。最后，Adaboost-BP 方法可以保持良好的泛化能力，即便在训练次数很大的情况下也很少会出现过度拟合现象，对少样本、高维数据具有很好的适用性，其精度高的优点也格外突出。

具体来讲，Adaboost-BP 方法是将 BP 神经网络作为基础的弱分类器，之后通过 Adaboost 算法将其升级，构造出一种组合式的强分类器作为最后的决策分类器。其核心思想在于通过调整训练集中每个样本的权重，即改变数据分布来实现由弱分类器到强分类器的跨越。其运用流程如图 1 所示。

该方法的具体算法步骤如下：

第一步，数据选择和网络初始化。从样本空间中随机选择  $m$  组训练数据，初始化测试数据的分布权值  $D_i(i)=1/m$ ，根据样本输入输出维数确定神经网络结构，初始 BP 神经网络权值和阈值。

第二步，弱分类器预测。训练第  $t$  个弱分类器时，用训练数据训练 BP 神经网络并且预测训练数据输出，得到预测序列  $g(t)$  的预测误差和  $e_t$ ，误差和  $e_t$  的计算公式为：

$$e_t = \sum_i D_i(i) \quad i = 1, 2, \dots, m(g(t) \neq y)$$

表1 企业信用评级指标体系

考查内容	偿债能力				营运能力				现金流量			获利能力				发展能力			
指标名称	流动比率	速动比率	资产负债率	产权比率	应收账款周转率	存货周转率	固定资产周转率	总资产周转率	现金债务总额比	销售现金比率	总资产现金回报率	营业利润率	成本费用利润率	资产报酬率	净资产收益率	营业收入增长率	净资产增长率	总资产增长率	营业利润增长率

式中,  $g(t)$  为预测分类结果,  $y$  为期望分类结果。

第三步, 计算预测序列权重。根据预测序列的  $g(t)$  预测误差  $e_t$  计算序列的权重  $a_t$ , 权重计算公式为:

$$a_t = \frac{1}{2} \ln\left(\frac{1 - e_t}{e_t}\right)$$

第四步, 测试数据权重调整。根据预测序列权重  $a_t$  调整下一轮训练样本的权重, 调整公式为:

$$D_{t+1}(i) = \frac{D_t(i)}{B_t} * \exp[-a_t y_i g_t(x_i)] \quad i = 1, 2, \dots, m$$

上式中,  $B_t$  是归一化因子, 目的是在权重比例不变的情况下使分布权值和为 1。

第五步, 强分类函数。训练  $T$  轮后得到  $T$  组弱分类函数  $f(g_t, a_t)$ , 由组弱分类函数  $f(g_t, a_t)$  组合得到了强分类函数  $h(x)$ 。

$$h(x) = \text{sign}\left[\sum_{t=1}^T a_t * f(g_t, a_t)\right]$$

## 企业信用评级的 Adaboost-BP 模型构建

### (一) 样本数据选取

鉴于上市公司的财务数据比较容易获得, 本文选取在我国上海证券交易所 A 股市场上市的制造业企业作为研究对象, 选取了 100 家“正常上市”状态企业, 30 家“ST”企业。在这里, 本文定义上市公司被实施特别处理的年份为  $T$ , 根据中国上市公司的年报披露制度, 本文将采用  $T-2$  和  $T-3$  年的财务数据作为样本数据, 最终形成了容量为 260 的面板数据。

需要说明的是, 本文沿用前人研究方法, 由于“正常上市”状态公司属于正常经营的公司, 其盈利能力正常、资产整体质量

图1 Adaboost-BP 方法流程图

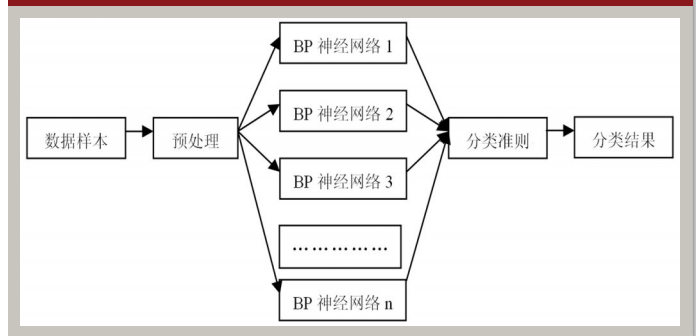


表2 样本选取数量一览表

	训练集	测试集	合计
正常上市公司	60	40	100
ST 公司	17	13	30
合计	77	53	130

表3 Mann-Whitney U 检验结果

	应收账款 周转率	存货 周转率	营业收入 增长率	营业利润 增长率	销售现金 比率
Mann-Whitney U	1171.000	1423.000	1321.000	1407.000	1178.000
Wilcoxon W	1636.000	6473.000	1786.000	1872.000	1643.000
Z	-1.818	-.425	-.989	-.514	-1.779
Asymp.Sig.	.069	.670	.323	.607	.0755

好、偿债能力强、抗风险能力强, 这样的上市公司信用状况良好, 所以定义其为“信用好”企业。与此相对应的, 由于“ST”公司是被特殊处理的连续亏损公司, 公司财务状况不佳。这样的公司抗风险能力差, 偿债能力弱, 将其视为处于财务困境、存在违约行为的公司是合理的, 所以本文定义其为“信用差”企业。

### (二) 数据预处理

#### 1. Mann-Whitney U 检验

在本文中，首先运用两样本秩和检验对样本数据进行分析，以判断出信用好与信用差的企业的各财务指标是否存在真正统计学意义上的差异，从而起到筛选指标的目的。

检验结果显示（见表3）：在0.05的显著性水平下，应收账款周转率、存货周转率、营业收入增长率、营业利润增长率及销售现金比率这五个指标不能显著地区分不同信用水平的上市公司，将其删去。

2. 信息增益方法

由于决策树的特征选择比较稳健，本文选择采用决策树的信息增益原理对样本数据进行二次处理，即根据指标载有信息量的大小对其做出选择。本文最终删去排名最靠后的四个指标，速动比率、产权比率、现金债务总额比、总资产现金回报率。

3. 主成分分析

主成分分析法可使数据主特征明朗化，降低输入变量的维数，减少网络结构中的节点数和网络的复杂度，进而提高模型的学习和预测效率。本文采用主成分分析方法对原始变量之间的高度相似性形成的冗余信息进行剔除。

（1）KMO 检验及 Bartlett 检验

由检验结果可以看出，KMO 值大于0.5且 Bartlett 检验显著性水平小于0.05，综合看来本文经前两步处理后的样本数据适合进行因子分析。

（2）主成分提取过程

利用SPSS软件进行主成分分析，得到变量的相关系数矩阵的特征值及贡献率（见表4）。

由特征值和累积贡献率，我们根据累积贡献率达到80%的原则提取5个重要主成分F1、F2、F3、F4和F5。然后，我们对这5个主成分变量建立原始因子载荷矩阵，利用正交旋转法对其进行结构调整简化，计算得出旋转后因子载荷矩阵（见表5）。

从该表可以概括出这五个主成分分别为获利能力因子、偿债能力因子、营运能力因子、成长能力因子和自有资产管理能力因子。之后根据因子得分系数矩阵及相关线性组合方程将原始指标数据转化成为新的指标体系数据，并作为后面所建模型的输入数据。

（三）模型输出结果

利用MATLAB软件构建企业信用评级的Adaboost-BP模型，模型对测试样本的测试输出结果如表6所示。

图2 KMO 检验及 Bartlett 检验结果

Kaiser-Meyer-Olkin Measure of Sampling Adequacy.		.595
Bartlett's Test of Sphericity	Approx. Chi-Square	1029.071
	df	45
	Sig.	.000

表4 方差贡献率

主成分	特征值	方差贡献率（%）	累积方差贡献率（%）
1	2.843	28.428	28.428
2	1.831	18.309	46.737
3	1.361	13.615	60.352
4	1.127	11.267	71.619
5	.854	8.539	80.158
6	.778	7.783	87.941
7	.537	5.366	93.307
8	.355	3.549	96.856
9	.242	2.425	99.281
10	.072	.719	100.000

表5 旋转后的因子载荷矩阵

	F1	F2	F3	F4	F5
流动比率( $X_1$ )	-.037	.937	-.047	.014	.013
资产负债率( $X_2$ )	.074	-.384	.090	.018	-.121
营业利润率( $X_3$ )	.960	-.011	.047	.082	.036
成本费用利润率( $X_4$ )	.973	-.037	.028	.065	.030
资产报酬率( $X_5$ )	.435	.010	.195	.226	.178
净资产收益率( $X_6$ )	.059	.019	.024	.039	.961
固定资产周转率( $X_7$ )	.006	.081	.156	.068	.013
总资产周转率( $X_8$ )	.061	-.046	.982	.081	.023
净资产增长率( $X_9$ )	.044	.016	.050	.141	.219
总资产增长率( $X_{10}$ )	.126	.012	.085	.965	.039

表6 模型输出结果汇总表

	Adaboost-BP 方法	BP 神经网络方法
第一类错误	6 ( 23.08% )	17 ( 65.38% )
第二类错误	14 ( 17.50% )	13 ( 16.25% )
合计	20 ( 18.87% )	30 ( 28.30% )

第一类错误即将“信用差”企业误判为“信用好”企业，第二类错误即将“信用好”企业误判为“信用差”企业。

#### (四) 输出结果分析

##### 1. 预测准确率分析

预测准确率是考量一个企业信用评级模型好坏的重要测度。从总体上看，基于 Adaboost-BP 方法和 BP 神经网络方法构建的信用评级模型对测试样本的预测准确率都达到了 70% 以上。但具体来讲，Adaboost-BP 模型比 BP 模型的预测准确率要高出近 10 个百分点。这说明，将 Adaboost 与 BP 神经网络方法相结合在客观上能够提高企业信用评级模型的判断准确率。因此在这一角度上，企业信用评级的 Adaboost-BP 模型是明显优于 BP 神经网络模型的。

##### 2. 评级成本比较分析

下面这个方程描述了在考虑两类误判成本时的期望误判成本：

$$EC = C(1|2) \times P(1|2) \times P(2) + C(2|1) \times P(2|1) \times P(1)$$

式中， $EC$  为期望误判成本； $P(1)$  和  $P(2)$  分别代表正常上市状态企业和 ST 企业的先验概率； $P(1|2)$  和  $P(2|1)$  分别代表犯第一类错误的概率和犯第二类错误的概率； $C(1|2)$  和  $C(2|1)$  则相应地代表一类错误成本和二类错误成本。

在这里，本文将两类公司各自占上证 A 股制造业类上市公司的比例作为先验概率，则先验概率  $P(1)$  和  $P(2)$  分别为 0.89 和 0.11。由于一类错误的成本会远远高于二类错误成本，在这里，本文采用相对成本比值来计算两类错误的成本，即假设一类错误成本是二类错误成本的  $n$  倍，将  $n$  分别取值为 1.5, 10。具体分析结果显示当两类错误成本相等时，基于 Adaboost-BP 方法的评级模型期望误判成本略高于 BP 神经网络模型，但差异不大。当相对成本比值为 5 和 10 时，前者的期望误判成本则明显低于后者。同时，我们可以计算出当  $n$  约等于 1.89 时，两个模型的期望误判成本相等。 $n > 1.89$  时，前者期望误判成本低于后者； $n < 1.89$  时，前者期望误判成本高于后者。这个分析结果对我们进行信用评级的实践具备一定的指导意义。

#### 相关建议

我国的企业信用评级行业目前正处于初级发展阶段，评级活动的诸多内容还未形成统一的标准与规范，造成了评级机构良莠

不齐、评级标准缺失、评级方法随意性大、评级结果混乱的现象。结合本文研究内容，现遵循评级工作流程对企业信用评级机构提出相关建议。

第一，在评级指标的选择上，评级机构应当全面考量。这里主要强调受评企业的行业差异。以电力企业、房地产企业与商业银行为例，电力企业因受政府较强监管，其市场地位稳定且经营风险偏低。而房地产企业是典型的资本密集型企业，资本结构特殊，同时受政府的政策导向影响比较大。商业银行为金融机构其最大特点是负债经营，它不仅具有一般企业的外部经营风险，同时承担着极大的内部风险。综合而言，这三类企业的行业信用环境存在巨大差异。因此，在进行信用评级之前，评级机构应当对此有深刻的认识，从而在评价指标的选择上有所侧重，进一步才能对受评企业做出客观公正的判断。

第二，在评级方法的选择上，评级机构应当结合不同方法的特点及受评对象的实际情况进行选择。举例来说，若受评对象历史数据缺乏，则以多元统计判别分析方法为代表的评级方法则难以适用。若评级活动成本受到严格控制且对评级结果的一致性要求较高，则以专家打分法为主的主观判别方法的使用会受到很大程度上的限制。若需要呈现评级结果的具体过程解释，则神经网络类的方法难以满足。所以，不同的受评企业其评级方法的选择应具有针对性。

第三，企业信用评级是一个动态过程，各评级机构应当做好评级数据的积累，开展动态追踪评级。这样做一是为自身评级结果的质量检验打好基础，二是为提高自身评级业务能力提供参考。

第四，我国的信用评级行业刚刚起步，目前尚未设立统一的机构对评级业务进行有效管理。具体的，信用评级理论、标准和方法都还缺乏科学性和规范性。这就要求各评级机构加强横向联系与交流，共同构建行业规范，逐步推进我国信用评级业的发展。

企业信用评级作为一种中介服务其最终产品将作为影响相关投资者投资决策的因素，各机构应独立于其他经济利益，加强自我监督、明确自身职业道德标准、培养社会责任感，运用科学的方法做出客观评判。

作者单位：中国人民银行酒泉市中心支行