

Credit Risk Assessment of Small and Medium-sized Listed Companies Based on Multilayer Perceptron

Li Dong, Mu Zhang

School of Economics and Finance, Guizhou University of Finance and Economics, Guiyang 550025, China

基于多层感知器的中小上市公司信用风险评估

董礼, 张目

贵州财经大学金融学院, 贵阳, 550025, 中国

Abstract

According to the proportion of "non ST" and "ST" in training, this paper used the principle of multilayer perceptron (MLP) to establish some small and medium-sized enterprise credit risk assessment model based on neural network, and finally, this model was chosen to evaluate the credit risk of 114 small and medium listed companies in 2018. According to the operating conditions of listed companies, the samples are divided into two categories----"good" and "bad". For each listed company, the six first level indicators that are profitability, solvency, growth capacity, operating capacity, cash flow and capital and its sixteen two level indicators are mainly considered in this paper. The simulation results show that the correct rate of the training sample is as high as 88.5% and the correct rate of the test sample is up to 94.4%. The results are ideal.

Keywords: multilayer perceptron; credit risk assessment; small and medium-listed companies

摘要

本文根据训练集中不同“非 ST”与“被 ST”的比例, 利用多层感知器 (MLP) 原理建立了多个中小企业神经网络信用风险评估模型, 最终选择最佳模型对我国 2018 年 114 家中小上市公司进行信用风险评估。按照各上市公司经营状况把样本分为“好”、“差”两类。本文对于每一家上市公司, 主要考虑其盈利能力、偿债能力、成长能力、营运能力、现金流量和资本结构六个一级指标及相应的 16 个二级指标。仿真结果表明, 训练样本的正确率高达 88.5%; 测试样本的正确率高达 94.4%; 其结果较为理想。

关键词: 多层感知器; 信用风险评估; 中小上市公司

1. 引言

最近几年, 不管是次贷危机还是欧债危机, 它们都有一个共同的特征: 因企业财务状况的恶化而导致信用风险加剧, 进而导致资本市场的波动, 同时也给投资者、金融机构等相关经济市场参与者带来了巨大的损失, 也给相关金融机构带来了巨大的冲击。为了能更好的评判企业等贷款客户的信用等级, 进而能够提高自身的盈利能力、达到控制风险所带来的损失以及能

够做出更佳决策的目的。所以，如何能够提高对贷款客户信用等级的评价是商业银行等相关金融机构重点关注的问题之一。

目前，学术界对信用风险评估的方法主要有传统的信用风险评价方法、多元统计分析方法和人工智能方法。在传统的信用风险评价方法中，最具有代表性的有“5C”要素分析法（张玲和张佳林（2000）[1]）和 LAPP 原则等（柯孔林和周春喜（2005）[2]）；这一类信用风险评价方法具有直观明了、操作简单等优点，但其缺点也显而易见[2]，例如，以定性分析为主，其主观性较强，缺乏客观的评价基础；借款人财务数据的真实性和准确性直接影响着信用风险评价结果的可靠性，而现实中的财务数据常常失真；综合分析能力较差，对财务状况缺乏整体的概括。在多元统计分析方法中，具有代表性的有 Altman（1968）[3]率先运用到公司破产和违约风险评估中的多元判别分析法、Logistic 模型（Ohlson（1980）[4]；石晓军（2006）[5]）、Probit 模型（郑昱（2009）[6]）和数学规划方法（Sueyoshi（2001）[7]；张忠志等（2004）[8]）；这一类方法科学实用、操作简便，可为相关银行科学评估中小企业信用风险、合理制定授信额度提供重要的参考依据，具有一定的实际应用价值。而对于人工智能方法在信用风险评估中的应用，主要有专家系统（杨保安和朱明（2006）[9]；张鹏，曾永泉，岳超源（2013）[10]）、支持向量机（Xu et al（2008）[11]）和神经网络等（Poddig（1995）[12]；Yurt（1995）[13]）。随着国内外研究学者在信用风险评估领域的深入研究，神经网络模型最近几年越来越受国内外学者的青睐（郭英见和吴冲（2009）[14]等）。

不管采用哪种方法对企业的信用状况进行评估，都无法回避一个问题，即如何选择最佳指标？事实上对信用风险评估结果的影响变量因素很多，且各变量之间可能存在高度的相关性，如果对变量的处理欠佳，最终的结果是预测模型效果大幅度下降。而人工智能中的神经网络模型对与

指标变量没有任何的假设等要求。所以运用神经网络模型对企业信用风险状况进行评估越来越受相关研究学者的青睐（张新红、王瑞晓（2011）[15]、江训艳（2014）[16]、姚德权、王文进（2015）[17]、肖斌卿、杨旻、余哲、沈才胜（2016）[18]）。上述引用神经网络模型对信用风险的评估取得了很好的效果，因此，本文在他们的研究基础之上，着重研究怎样提高 SPSS 中的多层感知器对中小上市公司信用评价的准确率，即模型的推广能力。并针对实证结果提出相应的风险防范建议。

2. 多层感知器模型简介

多层感知器主要由输入层、隐藏层和输出层三部分所构成。当把多层感知器用于两类的模式分类时，其原理相当于用超平面把两类样本分开。它的权值算法一般采用误差反向传播算法（BP 算法）学习。笔者利用 SPSS19.0 中的多层感知器构建了中小企业上市公司信用风险评估模型，SPSS 相对于 MATLAB 和 Python 等软件来说，在构建网络模型的时候不需要编程，使用起来比较简单，在给定训练样本的情况下，SPSS 能够自动的选择最佳的模型结构。

（一）人工神经网络结构

多层感知器的隐含层可以为一层或多层，与之相关的另一个概念是 BP 神经网络。有大量的研究表明，对于实际工作中的大部分问题都可以通过一个只含有一个隐含层的 BP 网络结构来逼近（只要神经元节点数足够多）。所以一个三层的 BP 网络已经可以完成任意复杂关联的分析问题。在网络结构内部，层与层之间全连接，同一层之间的神经元无连接。多层的网络设计使得网络结构能够从输入中挖掘更多的信息，完成更复杂的任务。其中网络结构内部的传递函数必须可微，BP 网络一般使用 Sigmoid 函数或者线性函数作为传递函数。

（二）反向传播算法的学习规则

在确定好层数和每一层的神经元个数后，还需要确定各层之间的权值系数才能根据输入给出正确的输出。误差反向传播算法是以一种具有监督方式进行的算法，其中的学习过程主要由正向传播和反向传

播两种方式组成。首先人为的设定一个期望输出值。然后在输入层输入训练样本，在这一过程中所输入的样本数据经过隐含层的激励函数向输出层输出结果，这一过程称为“正向传播”。输出的实际结果与期望输出的结果差即是误差。然后按照误差平方最小这一规则，误差经过误差函数转入反向传播，此过程称为“误差反向传播”。误差反向传播过程中不断地对权值和阈值进行修改调整，所以误差反向传播神经网络也简称 BP 网。随着两种传播过程的不断交替进行。实际输出的结果不断地向所对应的期望输出结果逼近，这一不断反复进行的过程提高了网络结构对输入样本数据响应正确率。

3. 实证分析

3.1 中小上市公司样本数据的选取

基于上文所构建的网络模型，下面将对中小板、创业板和 A 股等 2018 年的 114 家随机中小上市公司进行分类，即根据公司财务状况将企业分为信用状“好”和信用状况“差”的企业。由于“非 ST”的企业，其财务状况良好，违约的可能性较小，信用风险较低，这类企业被划分为信用状况“好”的企业；其余的“被 ST”的企业通常财务状况较差，违约的可能性较大，所以被划分为信用状况“差”的企业^[19]。到目前为止，由于对中小企业概念的界定还很模糊，所以本文将参照陈晓红^[20]等对中小上市公司的筛选原则筛选本文的研究样本：(1) 2015 年 12 月 31 日之前在沪深证券交易所上市的公司，这样保证能有一年年以上的分析时段。(2) 在 2016 年至 2018 年间的流通股小于一亿股。研究中以被“ST”的企业作为信用状况“差”的企业；非“ST”的企业作为信用状况“好”企业。

经过筛选后，考虑到被“ST”的企业数量一般小于非“ST”企业，所以最终以 2:1 的比例确定了非“ST”和被“ST”企业数量共 114 家^[21]。其中深圳 A 股 13 家、中小板 66 家、上证 A 股 25 家、创业板 9 家以及 B 股 1 家。

3.2 指标体系的构建

本文在考虑到数据的有效性、客观性、易于观察和测量性后，最终以上市公司财务指标即定量指标作为本文的研究指标，其中主要选取了盈利能力、偿债能力、成长能力、营运能力、现金流量和资本结构六个一级指标，相应的选取了 16 个二级指标，分别是盈利能力下的 3 个二级指标、偿债能力下的 3 个二级指标、成长能力下的 3 个二级指标、营运能力下的 3 个二级指标、现金流量下的 2 个二级指标和资本结构下的 2 个二级指标（见表 1）。

表 1 中小上市公司财务指标

一级指标	二级指标
盈利能力	X ₁ =资产报酬率（%）
	X ₂ =营业利润率（%）
	X ₃ =成本费用利润率（%）
偿债能力	X ₄ =流动比率（%）
	X ₅ =速动比率（%）
	X ₆ =现金流动负债比
成长能力	X ₇ =每股收益增长率（%）
	X ₈ =营业利润增长率（%）
	X ₉ =净资产增长率（%）
营运能力	X ₁₀ =营业周期（天/次）
	X ₁₁ =固定资产周转率（次）
	X ₁₂ =流动总资产周转率（次）
现金流量	X ₁₃ =销售现金比率（%）
	X ₁₄ =总资产现金回收率（%）
资本结构	X ₁₅ =资产负债率（%）
	X ₁₆ =股东权益比率（%）

3.3 实证分析

在实证分析部分，本文首先以 2:1 的比例确定好训练集和测试集个案数量，然后再针对训练集中“非 ST”和“被 ST”的不同比例来训练网络模型，最终确定不同类别的最佳比例在什么范围内模型的预测效果最佳。其中“非 ST”和“被 ST”的不同比例是随机的从样本中抽取的。从表 2 可以看出训练集中不同比例的“非 ST”和“被 ST”情况下训练出的模型所得出的训练集和测试集的正确总计百分比情况。

从表 2 可以看出，当训练集中的“非 ST”和“被 ST”的比例在 24:12 到 24:15 范围内时，模型的预测准确率最高，随着比例的不断变化，不同比例训练集下的模型预测率在不断地下降。图 1 能够很好地

表 2 不同比例下的总计百分比

训练集		测试集		正确总计百分比	
非 ST	ST	非 ST	ST	训练集	测试集
42	36	34	2	89.7%	88.9%
44	34	32	4	82.1%	86.1%
46	32	30	6	89.7%	86.1%
48	30	28	8	88.5%	94.4%
50	28	26	10	85.9%	91.7%
52	26	24	12	88.5%	88.9%
54	24	22	14	92.3%	83.3%
56	22	20	16	88.5%	80.6%
58	20	18	18	92.3%	83.3%
60	18	16	20	91.0%	80.6%
62	16	14	22	91.0%	75.0%
64	14	12	24	92.3%	72.2%
66	12	10	26	93.6%	72.2%
68	10	8	28	87.2%	66.7%
70	8	6	30	92.3%	66.7%
72	6	4	32	87.2%	50.0%
74	4	2	34	82.1%	69.4%

反映出模型预测准确率与测试集中“非 ST”和“被 ST”的比例之间的函数关系图。如图 1 所示。

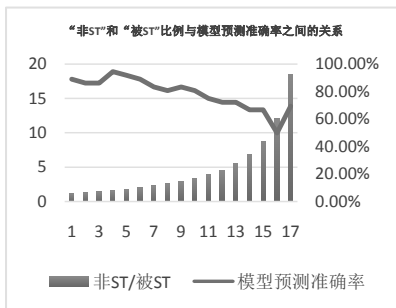


图 1 “非 ST”和“被 ST”的不同比例与模型预测准确率之间的关系

左轴表示训练集中“非 ST”和“被 ST”的比例，右轴表示模型对测试集的正确预测百分比。底轴表示模型个数。从图 1 可以看出总的有 17 个网络结构模型，这 17 个模型的训练集中的“非 ST”与“被 ST”之比从左至右不断增加；与此同时，网络模型对测试集中的正确预测率不断降低，很明显当“非 ST”与“被 ST”之比为 12 时，网络模型对测试集的正确预测百分比最低，为 50%。当“非 ST”与“被 ST”之

比为 1.6 时，网络模型对测试集的正确预测百分比最高，为 94.4%。在建模过程中，由于样本数量的限制，所以本文在对训练集中的不同类别进行随机配比时也受到一定的限制。但是从本文中的表 2 或者图 1 可以看出，训练集中的不同类别的比例所训练的模型对测试集的正确预测率的影响不可忽视。通过对以上的结果进行分析，本文最终选择第四个模型作为本文的模型。所以下文将对训练集中“非 ST”和“被 ST”的比例为 48:30 时所训练的模型即第四个模型所得出的结果进行简要的分析。从图 2 可知该模型的网络结构图，从图可知，输入层共有 16 个神经元节点，该模型只有一个隐含层，且隐含层的神经元节点数为 4。

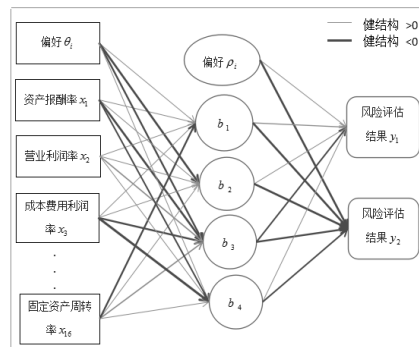


图 2 多层感知器的网络结构

表 3 显示了有关神经网络的信息，从表 3 可知，输入层单位数（神经元节点数）为 16，在指标数据输入之前已对数据进行标准化；其中隐含层中的单位数为 4，只有一个隐含层，激活函数是双曲正切函数；输出层只有一个因变量，单位数为 2，激活函数为 Softmax。

在把 78 个训练样本输入网络后，隐层节点数是从 1 开始逐个增加，当隐含层的节点数为 4 时，网络模型的误判率最低，故隐含层节点数取 4。此时我们所选取的权值系数及偏好值系数如表 4 所示，如输入层第一个神经元节点与隐含层第二个神经元节点的权值系数为 -0.738，且偏好值为 -0.224。

表 3 网络信息

输入层	因子	1	X_1	9	X_9
		2	X_2	10	X_{10}
		3	X_3	11	X_{11}
		4	X_4	12	X_{12}
		5	X_5	13	X_{13}
		6	X_6	14	X_{14}
		7	X_7	15	X_{15}
		8	X_8	16	X_{16}
隐含层		隐含层数		1	
		隐含层中的单位数		4	
		激活函数		双曲正切函数	
输出层		因变量		风险评估结果	
		单位数		2	
		激活函数		Softmax	

表 4 参数估值

预测器		预测值					
		隐含层				输出层	
		b_1	b_2	b_3	b_4	0	1
输入层	偏好	-0.42	-0.22	-0.48	-0.14		
	X_1	-0.27	-0.74	0.54	0.78		
	X_2	0.21	-0.94	0.59	0.64		
	X_3	0.185	-0.05	-0.15	0.80		
	X_4	0.08	-0.28	-0.47	-0.29		
	X_5	0.37	-0.32	0.23	0.28		
	X_6	0.43	0.17	-0.42	-0.03		
	X_7	0.30	-0.95	-0.28	-0.14		
	X_8	0.11	-0.05	-0.25	-0.01		
	X_9	-0.43	-0.18	0.59	-0.10		
	X_{10}	0.16	-0.35	-0.73	0.02		
	X_{11}	0.06	-0.26	0.11	0.28		
	X_{12}	0.02	-0.48	0.23	0.47		
	X_{13}	-0.54	0.06	0.33	-0.29		
	X_{14}	0.29	-0.08	0.28	0.02		
	X_{15}	0.07	0.05	0.04	0.07		
X_{16}	-0.08	0.41	0.67	-0.56			
	偏好				0.20	-0.37	
	b_1				0.16	0.11	
	b_2				-1.0	0.60	
	b_3				0.19	0.32	
	b_4				0.83	-0.56	

将两大样本的正确率进行交叉对比，从表 5 可知，训练样本中有 48 个是非 ST 样本，30 个被 ST 的样本；测试样本中有 28 个是非 ST 样本，8 个被 ST 的样本；网络模型对训练集分区中信用状况“好”的企业正确预测率达到 93.8%，测试集达到

100%；对训练集分区中信用状况“差”的企业正确预测率达到 80.0%，测试集达到 75.0%；预测样本的总体预测准确率高达 94.4%，说明该模型具有较好的预测能力。

表 5 分类

样本	已观测	已预测		
		0	1.0	百分比
训练	0	45	3	93.8%
	1.0	6	24	80.0%
	百分比	65.4%	34.6%	88.5%
测试	0	28	0	100.0%
	1.0	2	6	75.0%
	百分比	83.3%	16.7%	94.4%

4. 结论

本文通过建立多层感知器网络模型对训练样本中不同比例的中小上市公司的信用状况进行“好”和“差”评估，仿真结果表明，当训练集和测试集中的“非 ST”和“被 ST”的比例在 24:12:12:6 到 24:15:14:4 范围内时，模型的预测准确率最高，并随着比例的不断变化，模型的预测率在不断地下降。通过对比不同比例下的模型，本文选择训练集中“非 ST”和“被 ST”的比例在 48:30 时的模型作为本文的预测模型，该模型的训练样本的正确率高达 88.5%；测试样本的正确率高达 94.4%。

但是，整个模型的构建也是存在不足的，从主观角度讲，由于对网络模型进行训练的样本可能出现选择性偏差，最终可能影响该网络模型在中小上市公司信用风险评估中的准确性和稳健性；其次，由于指标在整个计算过程中是非常重要的，本文在选取的这 16 个二级指标中具有一定的缺陷，因为这 16 个二级指标中全是定量指标，没有选取一个定性指标。从客观方面讲，网络模型固有的局限性，该局限性主要是指网络模型黑箱的问题。“黑箱”是指，对于网络内部参数的含义不能进行具体的解释，难以进行一个逻辑的分析，是一个不透明的过程。

致谢

本研究得到了国家社会科学基金项目(71263011)和 2017 年度第二批贵州省基础研究计划(软科学类别)项目(黔科合基础(2017)1516-1)的资助。

参考文献

- [1] 张玲,张佳林.信用风险评估方法发展趋势.预测,2000,(4):72-75.
- [2] 柯孔林,周春喜.商业银行信用风险评估方法研究述评.商业经济与管理,2005,(6):55-60.
- [3] Altman E I. Financial ratios, discriminant analysis and the prediction of corporate bankruptcy. *Journal of Finance*, 1968, 23(4): 589-609.
- [4] Ohlson J A. Financial Ratios and the Probabilistic Prediction of Bankruptcy. *Journal of Accounting Research*, 1980, 18(1): 109-131.
- [5] 石晓军. Logistic 违约率模型最优样本配比与分界点的模拟分析.数理统计与管理,2006,25(6):675-682.
- [6] 郑昱. 基于 Probit 模型的个人信用风险实证研究.上海金融,2009,(10):85-89.
- [7] Sueyoshi T. Extended DEA-discriminant analysis. *European Journal of Operational Research*, 2001,131:324-351.
- [8] 张忠志,唐焕文,荣莉莉.数据包络分析在信用评估中的应用.运筹与管理,2004,13(1):112-117.
- [9] 杨保安,朱明.神经网络与专家系统相结合的银行贷款风险管理决策研究——国家自然科学基金项目 79770086 回溯.管理学报,2006,3(4):387-390.
- [10] 张鹏,曾永泉,岳超源.信用风险度量的新专家方法.科技进步与对策, 2013,7:52-54.
- [11] Xu X J, Zhou C G, Wang Z. Credit scoring algorithm based on link analysis ranking with support vector machine. *Expert Systems with Applications*, 2008, 36(2):1-8.
- [12] Poddig T. Bankruptcy prediction: a comparison with discriminant analysis. *Neural Networks in the Capital Market*. John Wiley and Sons Ltd, 1995.
- [13] Yurt A. Neural networks in corporate failure prediction: the UK experience. *Proceedings of the Third International Conference on Neural Networks in the Capital Market*,1995.
- [14] 郭英见, 吴冲. 基于信息融合的商业银行信用风险评估模型研究. 金融研究, 2009(01):95-106.
- [15] 张新红,王瑞晓.我国上市公司信用风险预警研究.宏观经济研究,2011,(1):50-54.
- [16] 江训艳. 基于 BP 神经网络的商业银行信用风险预警研究. 财经问题研究, 2014,(S1):46-48.
- [17] 姚德权, 王文进. 基于风险资产结构不确定性的商业银行整合风险度量研究. 财经理论与实践,2015,(6):8-13.
- [18] 肖斌卿, 杨旸,余哲, 沈才胜. 小微企业信用评级模型及比较研究.系统工程学报,2016,(6):798-807.
- [19] 张目, 周宗放. 基于多目标规划和支持向量机的企业信用评估模型. 中国软科学, 2009(04):185-190.
- [20] 陈晓红,张泽京,王傅强.基于 KMV 模型的我国中小上市公司信用风险研究.数理统计与管理,2008(01):164-175.
- [21] 边海容, 万常选, 刘德喜, 江腾蛟. 考虑 Web 金融信息的上市企业财务危机预测模型研究.计算机科学, 2013, 40(11): 295 298+315.