

文章编号:1008-1534(2014)01-0039-05

基于极限学习机网络的中小企业成长性研究

高娟¹, 李玉萍¹, 龚福和²

(1. 西北工业大学管理学院, 陕西西安 710129; 2. 西北工业大学财务处, 陕西西安 710072)

摘要:根据适合中国中小企业成长性评价的指标体系,提出了一种基于极限学习机神经网络的评价模型,并以2011年前上市的30家中小企业为例,验证了这种评价方法对中小企业成长性的基本适用性。结果表明,该评价方法应用在中小企业成长性评价系统中具有很高的评价准确率。最后根据对中小企业的实证研究结果对其成长性评价作了简单总结,并对现阶段中国中小企业的成长提出了建议措施。

关键词:中小企业;成长性;极限学习机

中图分类号:F270 文献标志码:A doi: 10.7535/hbgykj.2014yx0109

Study on SMEs growth based on extreme learning machine network

GAO Juan¹, LI Yuping¹, GONG Fuhe²

(1. School of Management, Northwestern Polytechnic University, Xi'an Shaanxi 710129, China; 2. Financial Department, Northwestern Polytechnic University, Xi'an Shaanxi 710072, China)

Abstract: According to the index system suitable for SMEs growth evaluation, this paper presents a extreme learning machine network to evaluate SMEs growth. Take 30 SMEs listed in 2011 for example. It proves this evaluation method is basically applicable for SMEs growth. The results show that the evaluation method is accurate. Finally, the paper makes a brief summary based on the empirical results on SMEs, proposing measures on the growth of present Chinese SMEs.

Key words: SMEs; growth; extreme learning machine

中小企业是经济发展和社会民生的重要支柱,中小企业成长性强,死亡率也高。因此,研究中小企业的成长性问题很有必要,同时研究适合于中小企业成

长性评价的模型,客观地反映中小企业的成长状况,对于促进其稳定发展,为政府、企业、金融机构提供参考依据均具有十分重要的理论和现实意义^[1]。

中国诸多学者对中小企业成长性评价进行了研究,常见方法有模糊综合评价、灰色系统理论和因子分析等。模糊综合评价方法较全面地考虑了问题的多指标性而且计算过程简单,但各级权重难以确定且不能解决评价指标之间的相关性^[2];灰色系统理论解决了小样本带来的不确定性问题并借助关联分析方法进行综合评价,但仅靠对距离的相对测度则过度简化了问题且缺乏物理意义^[3];因子分析法利用主成分分析消除了评价指标之间的相关性,根据公共因子对企业各方面进行评价,但计算量巨大^[4]。为了利用因子分析法的优点,克服其缺点,本文采用

收稿日期:2013-06-03;修回日期:2013-07-11

责任编辑:陈书欣

基金项目:西北工业大学人文社科与管理振兴基金(GCKY2001)

作者简介:高娟(1989-),女,山西繁峙人,硕士研究生,主要从事现代会计理论与方法方面的研究。

E-mail:1191259623@qq.com

高娟,李玉萍,龚福和.基于极限学习机网络的中小企业成长性研究[J].河北工业科技,2014,31(1):39-43.

GAO Juan, LI Yuping, GONG Fuhe. Study on SMEs growth based on extreme learning machine network[J]. Hebei Journal of Industrial Science and Technology, 2014, 31(1): 39-43.

极限学习机神经网络方法来评价中小企业的成长性,具体过程见图 1。

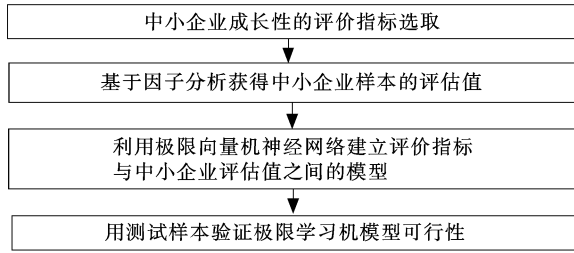


图 1 基于极限学习机的中小企业成长性研究方案流程

Fig. 1 Research plan process of extreme learning machine on growth of small and medium enterprises

1 企业成长性的评价指标选取

本着可操作性和成本-效益的原则,着重从财务状况的角度对企业的成长性进行分析。根据对中国制造业中小企业成长性的界定并结合前人对中国中小企业成长性的研究,借鉴聂敬兰^[5]的方法,构建了一套适合中国中小企业成长评价的指标模型,该评价指标体系见表 1。

表 1 企业成长评估指标体系

Tab. 1 Growth evaluation index system of the enterprise

一级指标	二级指标	指标编号	指标性质
偿债能力	资产负债率的倒数	X1	适度指标
	流动比率	X2	适度指标
	速动比率	X3	适度指标
盈利能力	净资产收益率	X4	正指标
	销售毛利率	X5	正指标
	销售净利率	X6	正指标
营运能力	存货周转率	X7	正指标
	应收账款周转率	X8	正指标
	总资产周转率	X9	正指标
成长能力	主营业务收入增长率	X10	正指标
	净资产增长率	X11	正指标
	总资产增长率	X12	正指标

2 企业成长性评价方法

极限学习机(extreme learning machine, ELM)神经网络又称基于最小二乘法的 BP 神经网络,是一种单层前馈神经网络学习算法^[6-7]。BP 神经网络输入层与隐含层及隐含层与输出层之间权值的调整是通过梯度下降法迭代获得,而极限学习机神经网络输入层与隐含层随机生成,隐含层与输出层之间权值通过最小二乘法直接获得。与 BP 神经网络相比,该方法学习速度快、泛化性能好,缺点是对电脑存储能力要求高。ELM 神经网络首先通过输入输

出训练数据获得 ELM 模型,然后利用 ELM 模型根据输入数据得到输出数据。

2.1 ELM 算法原理

极限学习机神经网络由输入层、隐含层和输出层组成。假设输入层有 n 个输入变量;隐含层有 l 个神经元;输出层有 m 个输出变量。设具有 Q 个样本的训练集输入矩阵 \mathbf{X} 和输出矩阵 \mathbf{Y} 分别为

$$\mathbf{X} = \begin{bmatrix} X_{11} & X_{12} & \cdots & X_{1Q} \\ X_{21} & X_{22} & \cdots & X_{2Q} \\ \vdots & \vdots & \vdots & \vdots \\ X_{n1} & X_{n2} & \cdots & X_{nQ} \end{bmatrix}_{n \times Q},$$

$$\mathbf{Y} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1Q} \\ y_{21} & y_{22} & \cdots & y_{2Q} \\ \vdots & \vdots & \vdots & \vdots \\ y_{m1} & y_{m2} & \cdots & y_{mQ} \end{bmatrix}_{m \times Q}.$$

设输入层与隐含层的连接权值矩阵 \mathbf{w} 和隐含层与输出层间的连接权值矩阵 $\boldsymbol{\beta}$ 为

$$\mathbf{w} = \begin{bmatrix} w_{11} & w_{12} & \cdots & w_{1n} \\ w_{21} & w_{22} & \cdots & w_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ w_{l1} & w_{l2} & \cdots & w_{ln} \end{bmatrix}_{l \times n},$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_{11} & \beta_{12} & \cdots & \beta_{1m} \\ \beta_{21} & \beta_{22} & \cdots & \beta_{2m} \\ \vdots & \vdots & \vdots & \vdots \\ \beta_{l1} & \beta_{l2} & \cdots & \beta_{lm} \end{bmatrix}_{l \times m}.$$

其中 w_{ij} 表示输入层第 i 个神经元与隐含层第 j 个神经元间的连接权值; β_{jk} 表示隐含层第 j 个神经元与输出层第 k 个神经元间的连接权值。

设隐含层神经元的阈值 $\mathbf{b} = [b_1, b_2, \dots, b_l]^T$, 隐含层神经元的激活函数为 $g(x)$, 本文隐含层激活函数 $g(x)$ 取

$$g(x) = 1/(1 + e^{-x}),$$

则极限向量机神经网络输出 \mathbf{Z} 为

$$\mathbf{Z} = [z_1, z_2, \dots, z_N]_{m \times N},$$

$$z_j = \begin{bmatrix} z_{1j} \\ z_{2j} \\ \vdots \\ z_{mj} \end{bmatrix}_{m \times 1} =$$

$$\begin{bmatrix} \sum_{j=1}^k \beta_{j1} g(W_j \cdot X_i + b_j) \\ \sum_{j=1}^k \beta_{j2} g(W_j \cdot X_i + b_j) \\ \vdots \\ \sum_{j=1}^k \beta_{jm} g(W_j \cdot X_i + b_j) \end{bmatrix}, i = 1, 2, \dots, N.$$

设 \mathbf{H} 为隐含层输出矩阵, ELM 神经网络输出公

式简化为 $H \cdot \beta = Z^T$ 。

令 $\epsilon = \|Y^T - Z^T\|$ 为逼近残差,则隐含层与输出层间的连接权值 β 可以通过求解以下方程组的最小二乘解获得:

$$\min_{\beta} \|H\beta - Y^T\|,$$

其解为 $\hat{\beta} = H^+ Y^T$,

其中 H^+ 为隐含层输出矩阵 H 的 Moore-Penrose 广义逆。

2.2 极限学习机神经网络的优化

2.2.1 模型输入参数归一化

数据归一化方法是神经网络预测前对数据常做的一种处理方法。对于企业成长性的评价指标数据^[8],将所有数据归一化到[-0.5 0.5]区间,取消各维数据间量级差别,避免因输入输出数据量级差别较大而造成 ELM 神经网络模型识别评估误差较大^[9]。本文采用最大最小法对企业成长性的评价指标进行归一化^[10]。

$$x_k = (x_k - x_{\min}) / (x_{\max} - x_{\min}) - 0.5,$$

式中: x_k 为企业成长性的评价指标数据; x_{\max}, x_{\min} 为所对应企业成长性的评价指标的上下限。

2.2.2 模型隐含层节点优化

交叉验证(CV)是模型性能评估的有效方法,可避免 ELM 神经网络模型过学习和欠学习状态。

k -CV 方法将训练集平均分为 k 组,每个子集数据分别做一次验证集,同时其余 $k-1$ 组子集数据作为训练集,这样会得到 k 个 ELM 神经网络模型,用这 k 个模型验证集输出参数的均方残差的平均值作为此隐含层节点设置的评价指标^[11],隐含层节点优化如图 2 所示。

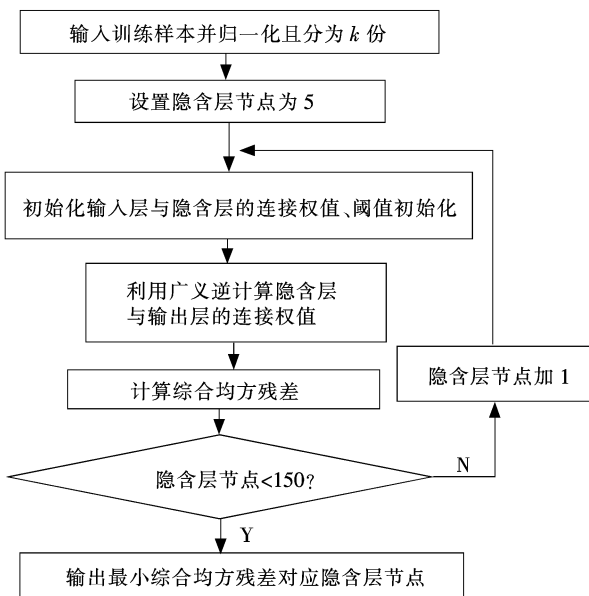


图 2 隐含层节点优化

Fig. 2 Optimization of hidden layer node

3 实例分析

3.1 样本的选择

本研究中选用中小板的中小企业作为研究样本,因为上市的中小企业财务制度比较规范,信息透明度较高。本研究数据来自新浪财经 2009 年、2010 年和 2011 年年度报告。收集了 300 家中上市企业,选出符合条件的涉及各个行业的 30 家中小企业,但由于制造业在中小企业中所占比例多,故样本中制造业企业多,其中先进制造业 15 家,医药生物 2 家,化工 3 家,房地产 2 家,科技型 5 家,航天制造业 1 家,服装类 2 家。选取样本时主要还考虑以下因素: 1) 企业上市年限超过 2 年; 2) 剔除了数据不全的样本,以保证每个样本都有自变量及因变量被合适计算; 3) 为避免异常值的影响,剔除在 2005 年至 2008 年期间出现过连续 2 年业绩亏损的公司。采用 2009 年、2010 年和 2011 年 3 年的平均值作为研究数据,较好地减少偶然性因素的影响。

3.2 中小企业评价价值确定

输出节点的选择对应于评价结果,为此需要确定期望输出。在神经网络的学习训练阶段,“样本”的期望输出值应是已知量。由于中小企业成长性是一种定性概念,很难用简单的数据或数学方法评估得出。为将这种定性描述的状态用定量化的形式来表达,本文利用新浪财经及基于因子分析法得到的值作为样本的期望输出值^[12]。

3.3 ELM 神经网络的训练与测试

利用 k -CV 方法对 ELM 模型隐含层的节点在某种意义上进行评估,以寻找到最优的隐含层节点数。本案例的计算流程如图 3 所示。

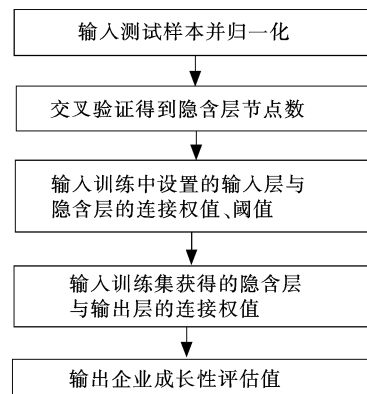


图 3 计算流程图

Fig. 3 Calculation flow chart

从因子分析法得出的期望输出值中随机抽取 70% 样本的输出值及其对应输入参数(成长性指标)作为 ELM 神经网络的训练集^[13],剩余 30% 样本的输出值及其对应输入参数作为测试集,通过交叉验

证得到 ELM 神经网络最佳隐含层节点数为 21, 此时均方差为 3.4452×10^{-4} , 这时可以避免过学习, 提高 ELM 模型的泛化性^[14]。训练集目标输出与实际输出如表 2 所示, 测试集输出对比如表 3 所示。

ELM 神经网络训练集的目标输出与实际输出一致, BP 神经网络训练集的目标输出与实际输出平均误差为 8.67%, 说明 ELM 神经网络模型比 BP 神经网络模型学习能力强。

从表 3 测试结果表明: 1) 通过分析这 9 家上市中小企业 2009 年、2010 年和 2011 年在行业中的排名, 发现这 9 家中小企业的成长基本上与现实相符。从检测结果可以看到, 网络仿真结果的误差都控制在 5% 以下, 取得了良好的评价效果。这表明神经

网络方法能够充分吸收专家的判断经验, 证实运用人工神经网络方法进行中小企业成长性评价的可用性和准确性。2) ELM 神经网络模型计算输出与实际评价结果平均误差为 11.80%, BP 神经网络模型计算输出与实际评价结果平均误差为 21.07%, 表明 ELM 神经网络泛化能力比 BP 神经网络能力强。同时据表 2 显示, 成长因子方面: 盾安环境、软控股份、新海宜、恒宝股份、沧州明珠的指标较强, 说明这些企业规模扩张速度较快, 发展前景较好, 而且企业的产品可能处于成长期, 能保持较好的增长趋势。其他指标基本也与综合指标排名保持一致。上述排名与年报显示的企业发展水平实际是基本相符的。

表 2 训练集目标输出与实际输出

Tab. 2 Target output and actual output of the training set

公司代码	002019	002012	002113	002074	002104	002111	002108
公司名称	鑫富药业	凯恩股份	天润控股	东源电器	恒宝股份	威海广泰	沧州明珠
目标输出	-0.699 0	0.315 0	-2.100 0	-0.208 0	1.055 0	0.187 0	0.450 0
ELM 实际输出	-0.699 0	0.315 0	-2.100 0	-0.208 0	1.055 0	0.187 0	0.450 0
BP 实际输出	-0.734 1	0.312 4	-2.086 1	-0.221 2	1.039 7	0.192 9	0.446 5
公司代码	002071	002057	002002	002089	002004	002011	002070
公司名称	江苏宏宝	中钢天源	鸿达兴业	新海宜	华邦颖泰	盾安环境	众和股份
目标输出	-0.172 0	0.030 0	-1.253 0	0.561 0	-0.031 0	0.492 0	-0.125 0
ELM 实际输出	-0.172 0	0.030 0	-1.253 0	0.561 0	-0.031 0	0.492 0	-0.125 0
BP 实际输出	-0.154 4	0.007 7	-1.252 4	0.561 1	-0.037 6	0.491 2	-0.115 0
公司代码	002096	002046	002026	002073	002029	002036	002049
公司名称	南岭民爆	轴研科技	山东威达	软控股份	七匹狼	宜科科技	同方国芯
目标输出	0.267 0	0.220 0	0.132 0	0.516 0	0.312 0	-0.259 0	0.295 0
ELM 实际输出	0.267 0	0.220 0	0.132 0	0.516 0	0.312 0	-0.259 0	0.295 0
BP 实际输出	0.262 8	0.255 7	0.152 4	0.497 2	0.303 1	-0.254 3	0.313 8

表 3 测试集输出对比

Tab. 3 Contrast of the test set output

公司代码	002076	002013	002009	002058	002031	002115	002016	002112	002066
公司名称	雪莱特	中航精机	天奇股份	威尔泰	巨轮股份	三维通信	世荣兆业	三变科技	瑞泰科技
目标输出	-0.010 0	-0.079 0	0.053 0	0.179 0	-0.072 0	0.335 0	-0.328 0	-0.408 0	0.346 0
ELM 输出	-0.013 2	-0.090 1	0.054 7	0.154 6	-0.066 5	0.324 2	-0.425 9	-0.416 7	0.344 3
BP 输出	-0.001 9	-0.086 4	0.042 3	0.197 0	-0.085 3	0.315 3	-0.232 6	-0.416 8	0.299 5

4 结论及建议

以中小制造型上市公司为例, 运用因子分析法研究了企业成长性的评价问题。通过实证研究, 得出如下结论。

1) 极限学习机神经网络能对大量的专家评价结果进行学习, 把专家头脑中的评价方法量化、具体化, 并把它应用到实际评价中去, 减少人的非理性因素对评价的影响, 且与传统的 BP, RBF 神经网络相比, 该方法具有学习速度快、泛化性能好等优点, 能

更准确地评价中小企业成长性。但神经网络在本文中只是简化了因子分析法复杂评价的得分过程,可以认为神经网络模型为因子分析模型的代理模型,所以本文所有的结果都是与因子分析模型对比,神经网络并不是一个单独的评价方法,它是基于因子分析而建立的。

2)选择运用极限学习机人工神经网络的原理建立中小企业成长性评价模型,希望为中小企业管理者提供长期经营决策。对股票投资者^[15],尤其是价值投资者来说,在选取创业板的股票时,不仅仅要关注它的短期成长性,更重要的是要关注其长期成长的可持续性和未来的成长潜力。

从财务层面并基于前人的研究成果,结合中小制造企业的特点运用分析法构建了中国制造业中小企业成长性评价指标体系,通过实例表明,基于极限学习机的神经网络模型具有较强的学习能力,能够得出较公平、公正的评价结果,用该评价模型对中小企业成长性综合评价是可行的。

总的来说,中小企业应加强公司治理,重视提高自己的偿债、盈利等各方面能力,从而在本质上提高企业成长能力。

参考文献/References:

- [1] 齐经民,王玉龙.基于因子分析的秦皇岛开发区中小企业成长性评价[J].价值工程,2012,29(32):69.
QI Jingmin, WANG Yulong. Evaluation of the growth of SMEs at qinhuangdao development zone based on factor analysis[J]. Value Engineering, 2012, 29(32): 69.
- [2] 刘宇泰,潘光友,刘克刚.应用性科研成果综合评价的模糊数学方法研究[J].昆明理工大学学报(理工版),2006,31(6):101-106.
LIU Yutai, PAN Guangyou, LIU Kegang. Research on fuzzy mathematics comprehensive evaluation method of applied research results[J]. Journal of Kunming University of Science and Technology (Science and Technology), 2006, 31(6): 101-106.
- [3] 刘思峰,党耀国,方志耕.灰色系统理论及其应用[M].北京:科学出版社,2010.
LIN Sifeng, DANG Yaoguo, FANG Zhigeng. Grey System Theory and Application[M]. Beijing: Science Press, 2010.
- [4] 陈思静.我国中小企业成长性影响因素研究——以中小板上市公司为例[D].镇江:江苏大学,2009.
CHEN Sijing. Factor Analysis on Growth of Small and Medium Enterprises; Take Small and Medium Listed Enterprises for Example[D]. Zhengjiang: Jiangsu University, 2009.
- [5] 聂敬兰.基于因子分析法的企业成长评价方法研究——以河南省上市公司为例[J].河南理工大学学报(社会科学版),2011,12(1):63-67.
NIE Jinglan. Based on the factor analysis of enterprise growth evaluation method research; Take listed companies in Henan for example[J]. Journal of Henan Polytechnic University (Social Sciences), 2011, 12(1): 63-67.
- [6] 丁晓剑,赵银亮.优化极限学习机的序列最小优化方法[J].西安交通大学学报,2011,45(6):7-12.
DING Xiaojian, ZHAO Yinliang. A sequential minimal optimization method for optimization extreme learning machine[J]. Journal of Xi'an Jiaotong University, 2011, 45(6): 7-12.
- [7] 史峰,王辉,郁磊,等. MATLAB智能算法30个案例分析[M].北京:北京航空航天大学出版社,2012.
SHI Feng, WANG Hui, YU Lei, et al. 30 Case of MATLAB Intelligent Algorithm [M]. Beijing: Beihang University Press, 2012.
- [8] JING Liping, NG M K, HUANG J Z. An entropy weighting k -means algorithm for subspace clustering of high-dimensional sparse data [J]. IEEE Transactions on Knowledge and Data Engineering, 2007, 19(8): 1026-1041.
- [9] DING Yuanyuan, DANG Xin, PENG Hanxiang, et al. Robust clustering in high dimensional data using statistical depths [J]. BMC Bioinformatics, 2007, 8(sup7): S8.
- [10] 陈金辉,赵雷振,杨宗宵,等.改进的BP神经网络在故障诊断中的应用[J].河北科技大学学报,2011,32(5):455-459.
CHEN Jinhui, ZHAO Leizhen, YANG Zongxiao, et al. Application of improved BP neural network in fault diagnosis[J]. Journal of Hebei University of Science and Technology, 2011, 32(5): 455-459.
- [11] 史峰,王小川,郁磊,等. MATLAB神经网络30个案例分析[M].北京:北京航空航天大学出版社,2010.
SHI Feng, WANG Xiaochuan, YU Lei, et al. 30 Case of MATLAB Intelligent Algorithm[M]. Beijing: Beihang University Press, 2010.
- [12] 龚福和,高娟.基于因子分析法的中小制造企业成长性研究[J].西安工业大学学报,2013,33(1):46-51.
GONG Fuhe, GAO Juan. Research on growth of small and medium manufacturing enterprises[J]. Journal of Xi'an Technological University, 2013, 33(1): 46-51.
- [13] 林晓松.基于BP神经网络的上市零售企业成长性评价研究[D].厦门:华侨大学,2009.
LIN Xiaosong. Based on the Research on the Growth Factors of Our Listed Retail Companies[D]. Xiamen: Huaqiao University, 2009.
- [14] 高鸿斌,张永强,张佳. BP-神经网络算法在分布式VOD系统流量预测中的应用[J].河北科技大学学报,2011,32(3):269-272.
GAO Hongbin, ZHANG Yongqiang, ZHANG Jia. Application of BP-neural network to network-flow prediction of distributed VOD system[J]. Journal of Hebei University of Science and Technology, 2011, 32(3): 269-272.
- [15] 王瑛芳.企业成长性主成分分析——以环保行业上市公司为例[J].中国管理信息化,2010,13(19):38-41.
WANG Yingfang. Enterprise growth principal component analysis; Take environmental protection listed companies for example [J]. China Management Informationization, 2010, 13(19): 38-41.