

基于灰色 BP 神经网络的实验材料供应预测*

丁 硕¹ 巫庆辉¹ 常晓恒¹ 王 东² 张 放¹

(1. 渤海大学工学院 锦州 121013; 2. 渤海大学实验管理中心 锦州 121013)

摘 要:针对单一 BP 神经网络对实验器材需求量预测准确度偏低的情况,提出了一种将灰关联分析与 BP 神经网络相结合的实验材料需求预测模型。先用灰关联分析法计算出影响需求量的各因子之间的关联度,然后选择关联度较大的 3 个优势因子作为 BP 神经网络的训练样本,建立了 3 层 BP 网络预测模型。以某实验材料的实际需求量为实例进行算法检验,对比分析了灰色 BP 网络模型和单一 BP 网络模型的预测准确性。实验结果表明:灰色 BP 网络模型将原有 6-10-1 的 BP 网络结构简化为 3-6-1 结构,灰色 BP 网络模型预测的最大相对误差仅为 -1.36%,而单一 BP 网络模型的预测最大相对误差为 -4.18%,灰色 BP 模型比单一 BP 模型的预测精度更高,结构更简单。

关键词:灰关联分析;BP 神经网络;实验材料;预测

中图分类号: TN957.52+9 TP183 **文献标识码:** A **国家标准学科分类代码:** 520.2060

Prediction of laboratory equipment support based on grey relation analysis and BP neural network

Ding Shuo¹ Wu Qinghui¹ Chang Xiaoheng¹ Wang Dong² Zhang Fang¹

(1. College of Engineering, Bohai University, Jinzhou 121013, China; 2. Experiment Management Center, Bohai University, Jinzhou 121013, China)

Abstract: To deal with low accuracy of single BP neural network in the prediction of laboratory equipment demand, a prediction model based on grey relation analysis and BP neural network is proposed. The correlation degrees between influencing factors were first calculated using grey relation analysis method. Three major factors with higher correlation degrees were chosen as training samples for the BP neural network and a three-layered BP network prediction model was established. The real laboratory equipment demand of an experiment was taken as an example to test the algorithm. And a comparative analysis of the prediction accuracy of the grey BP network model and the single BP network model was done. The result shows that the grey BP network model reduces the topology of BP network from 6-10-1 to 3-6-1. Its maximum relative error is -1.36%, while the one of the single BP network model is -4.18%. The prediction accuracy of the grey BP network model is higher and its structure is also simpler.

Keywords: grey relation analysis; BP neural network; laboratory equipment; prediction

1 引 言

高校常规实验器材的储备量过大会增加储存空间和保管费用,还会使实验器材陈旧老化,储备量过小又无法满足正常教学任务,所以,对实验器材的需求量做出准确预测可以为实验材料的科学供给提供重要依据。由于影响实验器材需求量的因素较多,统计数据呈现高维、非线性特性,传统方法的预测结果与实际相差较大,无法满足实际需要。人工神经网络技术的快速发展为实验器材的

需求量预测提供了新的方法。其中,BP 神经网络因为其自身算法简单、非线性映射能力强^[1-5],被广泛应用于各个领域的预测。但是,当训练样本数量较小时,BP 网络逼近能力较差;当训练样本数量足够大时,其逼近能力虽有所加强,但是此时 BP 网络又存在收敛速度慢、易陷入局部极小值等缺点,最终导致网络不能收敛^[6-11]。文献[12]用 BP 网络对编号为 001 的实验器材的需求量做出预测,但是,对于作为预测对象主体的影响因子之间的关联度考虑不够,即未考虑各因子之间存在的信息重叠

收稿日期:2016-09

* 基金项目:国家自然科学基金资助项目(61401044)、渤海大学教学改革项目(BDJG-15-YB-C-018)资助

现象,从而导致预测准确度不高。所以,本文采用灰关联分析方法,定量计算出多影响因子之间的关联关系,将各影响因子按相对实验器材需求量的重要性进行排序,提取关联度较大的3个优势因子作为BP网络的输入,实现网络模型输入特征向量的降维,构建了3层灰关联BP网络的预测模型,并与文献[12]中所建立的单一BP网络模型的预测结果进行对比,验证改进的预测模型的有效性。

2 实验器材需求量影响因子的灰关联分析

2.1 灰关联系数的计算

灰关联分析的本质是通过比较数据序列曲线几何形

状的相似度来判断序列之间的联系是否紧密,相似度越大,序列之间的关联度越大,反之关联度越小。相似程度用关联系数和关联度来描述,关联度反映了各影响因子对结果的影响程度,关联度的数值大小与影响程度成正比关系^[13-15]。

根据对实验室实验材料的实际供应情况分析并参考相关文献[12],选用6个指标作为实验器材需求量的影响因子,以文献[12]中的样本集为例对6个影响因子进行灰关联分析,编号为001的实验材料需求原始数据如表1所示,其中: X_1 —使用时间、 X_2 —使用次数、 X_3 —使用人员中业务素质差的使用人员所占比例、 X_4 —故障率、 X_5 —平均故障间隔时间、 X_6 —采购难度系数。

表1 实验材料需求原始数据表

年份	X_1 /kh	X_2 /次	X_3 /%	X_4 /(次/kh)	X_5 /h	X_6	实际需求/个
2002	5.363	855	0.090	0.003 3	259	0.36	5
2003	6.547	920	0.101	0.004 4	360	0.40	7
2004	6.301	876	0.098	0.004 2	312	0.31	6
2005	6.861	959	0.131	0.005 1	430	0.53	8
2006	6.927	1 204	0.150	0.006 2	535	0.70	10
2007	6.890	1 180	0.128	0.005 9	522	0.63	9
2008	6.721	974	0.094	0.005 7	499	0.59	8
2009	6.799	1 024	0.143	0.006 3	577	0.71	9
2010	7.131	1 265	0.169	0.007 4	630	0.84	11
2011	8.642	1 524	0.192	0.007 9	679	0.88	13

令 X 为灰关联因子矩阵,将实验器材的实际需求量作为参考序列 $x_0(k), k=(1,2,\dots,10)$,将6个影响因子作为比较序列 $x_i(k), i=(1,2,\dots,6), k=(1,2,\dots,10)$,可得灰关联因子矩阵如式(1)所示。按照式(2)方法计算可以得到 $x_i(k)$ 对 $x_0(k)$ 在第 k 点的灰关联系数 $\xi_{0i(k)}$,式(2)中 λ 为分辨系数, $\lambda \in (0,1)$, λ 越小表示关联系数间的差异越大,本文取0.5,比较序列与参考序列之间的关联系数如表2所示。

$$X = (X_0(k), X_1(k), \dots, X_6(k)) = \begin{pmatrix} x_0(1) & x_1(1) & x_2(1) & \dots & x_6(1) \\ x_0(2) & x_1(2) & x_2(2) & \dots & x_6(2) \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ x_0(10) & x_1(10) & x_2(10) & \dots & x_6(10) \end{pmatrix}_{10 \times (6+1)} \quad (1)$$

$$\xi_{0i(k)} = \frac{\min_i \min_j |x_0(k) - x_i(k)| + \lambda \max_i \max_j |x_0(k) - x_i(k)|}{|x_0(k) - x_i(k)| + \lambda \max_i \max_j |x_0(k) - x_i(k)|} \quad (2)$$

表2 比较序列与参考序列之间的关联系数

ξ_{01}	ξ_{02}	ξ_{03}	ξ_{04}	ξ_{05}	ξ_{06}
0.395 7	0.969 2	0.567 2	0.605 1	0.682 1	0.712 7
0.477 5	0.556 4	0.676 7	0.961 0	0.760 2	0.697 4
0.386 5	0.573 3	0.664 1	0.879 2	0.639 4	0.600 6
0.636 5	0.531 5	0.550 9	0.539 9	0.927 0	0.905 1
0.480 9	0.573 8	0.724 8	0.708 3	1.000 0	0.712 5
0.797 2	0.619 2	0.582 8	0.754 7	0.756 8	0.786 2
0.703 3	0.613 8	0.425 2	0.333 6	0.735 9	0.705 1
0.737 0	0.738 9	0.476 0	0.810 4	0.659 5	0.994 5
0.368 2	0.516 3	0.516 6	0.928 1	0.864 7	0.591 1
0.357 8	0.481 0	0.631 7	0.683 0	0.920 5	0.719 4

2.2 灰关联度的计算

令 r_i 为比较序列 $x_i(k)$ 对参考序列 $x_0(k)$ 灰关联度, r_i 计算方法如式(3)所示。

$$r_i = \frac{1}{6} \cdot \sum_{k=1}^{10} \xi_{0i(k)} \quad (i = 1, 2, \dots, 6) \quad (3)$$

2.3 灰关联度排序

由式(3)计算得出各个比较序列与参考序列之间的关联度,各影响因子的关联度排序如表3所示。关联度数值表明比较序列对参考序列的影响程度^[16],由式(3)可以看出,原始6个影响因子对实验器材需求量的影响程度的顺序为 $X_5 > X_6 > X_1 > X_2 > X_3 > X_4$ 。

表3 各影响因子的关联度排序

影响因子	关联度	排序
X_1	0.593 4	6
X_2	0.685 9	4
X_3	0.646 2	5
X_4	0.800 4	3
X_5	0.882 9	1
X_6	0.824 9	2

3 基于灰关联 BP 网络的实验材料需求预测模型的建立

以表1中编号为001的实验器材从2002—2011年的需求量相关数据进行网络建模,以2002—2010年的数据作为训练样本,2011年的数据作为测试样本,灰关联BP网络的实验材料需求预测模型的建立过程如下。

3.1 网络的拓补结构

输出层节点:输出层神经元数为实验材料的需求量,所以输出层节点数为1;输入层节点:输入层节点数为经过灰关联分析后提取的优势因子数量,合理选取影响因子对于预测模型的准确性起着至关重要的作用。为了尽可能地避免输入特征向量提取过程中的主观性,依据2.3节计算结果,按照各影响因子的关联度由大到小顺序分别提取前1~5个因子进行实验;隐含层节点:每种情况下隐含层节点数目依据经验公式(4)并经过反复多次试验最终确定,式(4)中 N_i 为输入层节点数, O_i 为输出层节点数, H_i 为隐含层节点数^[17-18]。6种网络的拓补结构如表4所示,当因子数量为6时,即为未经过灰关联分析的单一BP网络。

$$H_i = \sqrt{N_i + O_i} + \alpha, 1 < \alpha < 10 \quad (4)$$

表4 不同拓补结构网络训练误差对比

因子数量	拓补结构	训练步数	R 值	均方误差
1	1-12-1	12	0.980 68	0.007 2
2	2-12-1	23	0.990 01	0.003 6
3	3-6-1	35	0.995 86	0.000 8
4	4-8-1	132	0.991 03	0.002 4
5	5-9-1	1965	0.982 39	0.005 6
6	6-10-1	2877	0.980 13	0.009 8

3.2 网络的训练结果与分析

根据表4依次建立6种拓补结构的BP网络,将表1中的数据分别作归一化处理输入到BP网络中,BP网络相关参数设置如表5所示。利用LM算法训练6种网络,不同拓补结构网络训练误差对比如表4所示。

表5 BP网络参数

参数名称	参数值	参数名称	参数值
训练函数	trainglm	误差函数	均方误差(mse)
转换函数	Sigmoid	学习速率	0.1
学习函数	learnngdm	预测精度	0.01
隐层传递函数	tansig	最大迭代次数	5 000
输出层传递函数	purelin		

由表4可以看出,在相同的BP网络参数配置条件下,与未经过灰关联分析的BP网络相比(因子数量为6时),灰色BP网络模型,可以不同程度地减少单一BP网络的训练步数和降低网络预测的均方误差,这是因为灰色BP网络模型降低了原始特征向量维数,较大幅度地降低了原始特征向量间的信息重叠,故将灰关联分析与BP网络相结合用于实验材料的需求量预测是可行的。但是,进一步查看表4可知,提取的优势因子数量越少,网络结构越简单,训练周期越短,但预测误差较大;增大优势因子数量,会在一定程度上减小预测误差,但网络结构复杂,导致训练周期过长。所以,优势因子数量是影响灰色BP网络模型的预测精度和建模效率的重要因素。

为了进一步说明优势因子数量的确定过程,在提取前1~5个因子5种情况下,分别对网络实际输出值与期望输出值进行线性回归分析,线性回归系数R值越接近于1,表明网络的训练输出结果与目标值越接近。限于篇幅,只给出提取3个和5个优势因子时训练样本集的线性回归结果,分别如图1、图2所示。由图1和图2可以看出,当提取 X_5 、 X_6 、 X_4 这3个指标作为优势因子时,回归系数 $R=0.995 86$,实际拟合直线与理论拟合直线几乎重合,各个样本点紧密分布于拟合直线周围,网络经过35次迭代即可收敛,训练输出结果与目标值较为接近。当提取 X_5 、 X_6 、 X_4 、 X_2 、 X_3 这5个指标作为优势因子时, $R=0.982 39$,实际拟合直线与理论拟合直线位置不重合,各个样本点较为分散地分布在拟合直线周围,网络经过1 965次迭代才可收敛,训练输出结果与目标值相差较大。进一步观察表4可知,当提取前1、2、3个优势因子时,网络的均方误差值呈现下降趋势,训练时间变化不大,但当提取前4、5个优势因子时,网络的均方误差值反而呈现上升趋势,训练时间也明显延长,这是因为提取过少的优势因子不能完整地反映原始预测指标的特征信息,而提取过多的优势因子会增大信息冗余量,彼此之间存在大量的信息干扰,会弱化原始预测指标的特征信息,所以,最终确定

提取 X_5 、 X_6 、 X_1 这 3 个指标作为优势因子作为 BP 网络的输入特征向量,建立了 3—6—1 结构的实验材料需求预测模型。

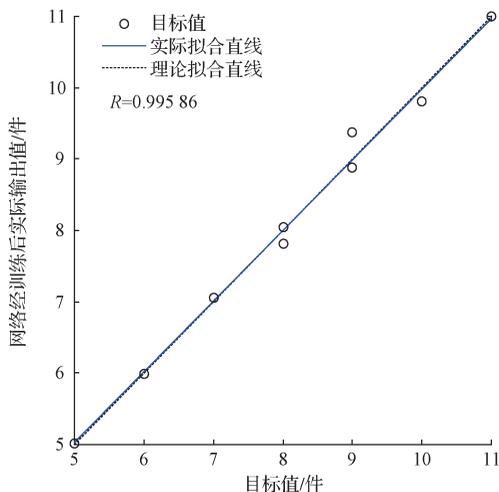


图 1 训练样本集的线性回归结果(3 个优势因子)

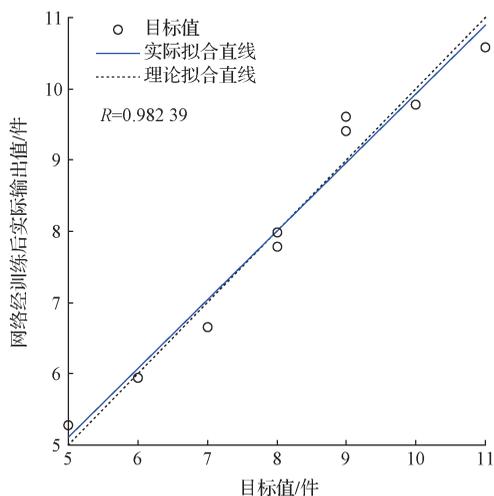


图 2 训练样本集的线性回归结果(5 个优势因子)

4 仿真实验

为了检验灰关联 BP 网络预测模型的泛化能力,以文献[12]中编号为 001 的实验器材 2002—2010 年的数据作为训练样本,以 2011 年的数据作为测试样本,应用灰关联 BP 网络模型进行需求量预测,并与文献[12]中建立的单一 BP 网络预测模型作对比,2 种网络的预测结果对比如图 3 所示。可以看出,灰关联 BP 网络对编号为 001 的实验器材在 2011 年预测值为 12.823 2 个,2011 年实际值为 13 个,相对误差仅为 -1.36%;单一 BP 网络预测值为 12.457 0 个,相对误差为 -4.18%。在各个训练样本点,灰关联 BP 网络的预测结果与实际值几乎完全一致,且绝对误差波动很小,绝对误差明显小于单一 BP 网络的预测

结果。这是因为对原始指标进行了灰关联分析,提取了 3 个优势因子,使单一 BP 网络结构得以简化,在较大程度上降低了输入向量之间的信息重叠,BP 网络的迭代次数由 2 877 次降至 35 次,收敛速度明显加快,提高了建模效率和预测准确度。

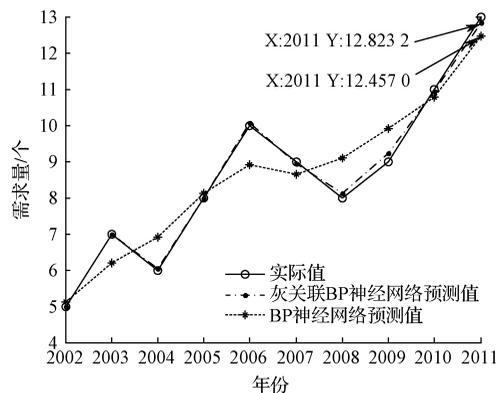


图 3 2 种网络的预测结果对比

5 结论

采用灰色关联分析法计算出各影响因子对实验材料需求量的影响程度,将各因子对需求量影响程度的大小进行排序,选择影响程度较大的 3 个优势因子作为输入数据,有效减少了 BP 网络的输入特征向量维数,简化了 BP 网络的拓补结构,提高了预测模型的建模效率。选用文献[12]中编号为 001 的实验器材需求量作为实例验证,仿真结果表明,建立的灰关联 BP 网络预测模型的预测结果精度明显高于单一 BP 网络模型,并且模型拓补结构简单、所需样本数量少、训练效率明显提高,验证了所建立的预测模型的有效性。

参考文献

- [1] 韩韬,陈晓辉.基于预测模型的 BP_Adaboost 算法改进[J]. 桂林理工大学学报,2014,34(3):589-594.
- [2] 赵会敏;雒江涛;杨军超,等.集成 BP 神经网络预测模型的研究与应用[J]. 电信科学,2016,32(2):60-67.
- [3] 丁硕,常晓恒,巫庆辉,等. Elman 和 BP 神经网络在模式分类领域内的对比研究[J]. 现代电子技术,2014,37(8):12-15.
- [4] 胡帅,顾艳,姜华,等. 基于 PCA-BPNN 的学生写作成绩预测模型研究[J]. 国外电子测量技术,2015,34(12):35-38.
- [5] 刘春,马颖. 遗传算法和神经网络结合的 PSD 非线性校正[J]. 电子测量与仪器学报,2015,29(8):1157-1163.
- [6] 庄育锋,胡晓瑾,翟宇. 基于 BP 神经网络的微量药品动态称重系统非线性补偿[J]. 仪器仪表学报,2014,35(8):1914-1920.

- [7] 丁硕,常晓恒,巫庆辉,等. DGA 与 GRNN 的联合变压器故障诊断研究[J]. 电子测量技术, 2014, 37(5): 142-146.
- [8] 胡帅,顾艳,曲巍巍. 主成分分析与 BP 网络结合的教学质量评价[J]. 自动化技术与应用, 2016, 35(4): 10-14.
- [9] DING S, WU Q H. Research on inverse model based on ANN and analytic method for induction motor[J]. Automation and Control, 2011, 5(4): 356-370.
- [10] LIU D T, ZHOU J B, PENG Y. Data-driven Prognostics and Remaining Useful Life Estimation for Lithium-ion Battery: A Review [J]. Instrumentation, 2014, 1(1): 59-70.
- [11] 丁硕,常晓恒,巫庆辉,等. 数值优化改进的 BP 神经网络逼近性能对比研究[J]. 山东科学, 2014, 27(1): 68-72.
- [12] 黄启来,汲万峰,赵明. 基于神经网络技术的实验器材保障应用研究[J]. 实验技术与管理, 2014, 31(7): 235-236, 240.
- [13] 梁娜,张吉刚. 基于灰关联分析-神经网络的组合预测模型[J]. 江汉大学学报:自然科学版, 2009, 37(4): 57-59.
- [14] 于红斌,齐娜,赵永进. 灰关联神经网络在葡萄酒评价中的应用[J]. 微处理机, 2013(4): 49-51.
- [15] 王劲松,李宗玉,隋雷. 基于灰色神经网络的战场态势分析及预测[J]. 电光与控制, 2015, 22(12): 15-19.
- [16] 杨永明,王燕,范秀君,等. 基于灰关联-神经网络的电力工程造价估算[J]. 重庆大学学报, 2013, 36(11): 15-20.
- [17] 丁硕,巫庆辉. 基于改进 BP 神经网络的函数逼近性能对比研究[J]. 计算机与现代化, 2012(11): 10-13.
- [18] 丁硕,常晓恒,巫庆辉. 数值优化改进的 BP 网络的模式分类对比[J]. 计算机系统应用, 2014, 23(5): 139-144.

作者简介

丁硕, 1979 年出生, 硕士, 副教授, 主要研究方向为人工智能、神经网络理论及应用。
E-mail: dingshuo2004@sina.com

(上接第 74 页)

- [10] 吴秀梅,赵连玉. 力传感器动态标定装置的研究[J]. 天津理工大学学报, 2007, 23(2): 75-77.
- [11] 范兵,陈锡侯,肖朝晖,等. 磁场式位移传感器误差的数理模型研究[J]. 仪器仪表学报, 2016, 37(5): 985-992.
- [12] 张德福,葛川,李显凌,等. 高精度位移传感器线性度标定方法研究[J]. 仪器仪表学报, 2015, 36(5): 982-988.
- [13] 沈春丰,陈光柱,韩振铎. 基于虚拟仪器的力传感器动态标定技术研究[J]. 煤矿机械, 2008, 29(6): 65-67.
- [14] WANG ZH Q, WANG H R, MENG Q M. RLS and LMS blind adaptive multi-user detection method and comparison in acoustic communication[J]. Instrumentation, 2015, 2(2): 47-54.
- [15] 易大义,陈道琦. 数值分析引论[M]. 杭州:浙江大学出版社, 1998.

作者简介

杜天军,男,1974 年出生,工学博士,高级工程师,主要研究方向为自动测试系统和机载设备故障诊断。
E-mail: dtj2000@163.com.

黄世超,男,1977 年出生,工学学士,高级工程师,主要研究方向为航空维修工程及管理。

刘书国,男,1982 年出生,工学博士,主要研究方向为自动测试系统和机载设备故障诊断。

官利,男,1981 年出生,工程师,主要研究方向为自动测试系统和机载设备故障诊断。