

煤炭企业 ERP中基于神经网络的 BOM数量模型

夏 鑫^{1,2},姜海虹²,陈哲明³

(1. 厦门大学工商管理博士后流动站,福建 厦门 361005; 2 山东工商学院 管理科学与工程学院,山东 烟台 264005; 3. 东营职业学院,山东 东营 257091)

[摘 要] 根据煤炭企业影响物料消耗因素的特点,进行了煤炭企业 BOM神经网络模型的选择及算法改进,并以乳化油为例进行实证研究,得出了相应结论。

[关键词] 煤炭企业;物料清单(BOM);企业资源计划(ERP);神经网络(ANN)

[中图分类号] F426.21 [文献标识码] A [文章编号] 1672-5956(2007)02-0056-05

BOM Quantity Model of Artificial Neural Network Based
on ERP in Coal Mining EnterprisesXIA Xin^{1,2}, JIANG Hai-hong², CHEN Zhe-ming³

(1. Postdoctoral Research Station of Business Administration, Xiamen University, Xiamen 361005, China; 2 Shandong Institute of Business and Technology, Yantai 264005, China; 3 Dongying Vocational Institute, Dongying 257091, China)

Abstract: According to the characteristics and the influence factors of materials consumption in coal industry, the author chooses a BP artificial neural network model and improves its algorithm for coal industry BOM. It takes Emulsible Oil as an example and draws a corresponding conclusion.**Key words:** coal industry; Bill of Material (BOM); Enterprise Resource Planning (ERP); Artificial Neural Network (ANN)

BOM (Bill of Material, 物料清单)模型包括 BOM结构模型和 BOM数量模型。BOM结构模型是确定产品消耗物料的种类;BOM数量模型就是采用一定的方法确定 BOM结构模型中的各种物料的消耗数量。煤炭企业 BOM的结构模型层次比较简单,而其 BOM的数量模型相对比较复杂,如何合理、准确确定 BOM数量模型是煤炭企业成功实施 ERP的关键。

一、煤炭企业 BOM 的 BP神经网络模型的选择及算法改进

在煤炭企业 ERP系统中,由于物料消耗不构成产品实体,其 BOM的各物料消耗数量是一个取决于地质条件、开采工艺方式、管理水平等等诸多因素的变量,这些因素之间相互作用复杂,作用机制不明确,判定决策又无条理可循。但实际处理时,我们并不需要知道影响物料消耗的每一个因

[收稿日期] 2006-12-09

[作者简介] 夏鑫,1969年生,男,湖南安化人,山东工商学院副教授,博士,厦门大学在读博士后,研究领域为管理信息系统和成本管理,重点研究作业(成本)管理、ERP及其在煤炭企业应用的理论和方法,(电子信箱)xiaxin0302@126.com。

素,也不需要弄清楚决定物料消耗的作用机理。我们只需要做到当影响物料消耗的各因素发生变化时,可以根据相应的变化对物料消耗数据进行调整,从而可以更有效地进行物料定额、物料需求计划和物料采购计划的制定,同时更有利于进行物料控制和库存管理。而神经网络由于其自学习、自组织和自适应的特性,所以对处理煤炭企业物料消耗这种影响因素复杂、知识背景不清楚和推理机制不明确的非结构化问题非常有效。

BP网络是逆传播(back propagation)学习算法的简称,其具体学习算法分为累积误差逆传播算法和标准误差逆传播算法。最基本的BP网络是三层前馈网络,即输入层、隐层和输出层,节点

之间前向连接。本文利用神经网络确定物料消耗的基本思想是:按作业找出影响物料消耗的因素,作为神经网络的输入节点;利用企业以前的相同作业的物料消耗数据,或者同行业先进企业相同作业的物料消耗数据作为样本,训练神经网络。网络经过学习训练后,当以后影响物料消耗的各因素发生变化时,可以通过网络输入变化后的因素参数确定物料的消耗量。可见,煤炭企业BOM的BP神经网络的输入层是影响某物料消耗的各种因素(a_1, a_2, \dots, a_q),输出层为该种物料的消耗量(c_1),其拓扑结构如图1所示。

恰当地选择网络的输入变量和输出变量、合

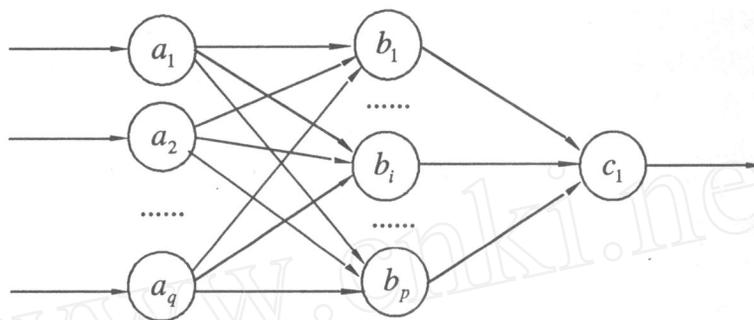


图1 煤炭企业BOM的BP网络拓扑结构

理地确定网络拓扑结构、充分利用建模样本集信息和恰当地终止网络的训练过程是建立实用BP网络预测模型的关键。从国内外有关神经网络预测模型的研究看,应用BP神经网络建立各种预测模型的方法主要存在以下问题:一是将建模样本简单地分成训练样本子集和验证样本子集的做法,在某种程度上会导致所建网络对验证样本子集的“过渡拟合(overfitting)”;二是BP网络的拓扑结构对所建网络的预测能力的影响认识不足、重视不够;三是在基于BP网络的预测建模过程中,仅限于BP网络对训练样本的拟合或网络在训练样本子集上的系统误差很小是片面的。

针对BP神经网络模型的上述问题,对影响BP网络模型的几个关键问题进行确定或改进:

第一,由于网络拓扑结构问题的复杂性,目前还缺少通用的确定优化网络拓扑结构的算法,本文仍然采用经验试错法来确定优化的网络拓扑结构。第二,针对将建模样本简单分成训练样本和验证样本所造成的过渡拟合问题,采用交互验证法(cross-validation)来评判所选的网络拓扑结构及相应的训练参数是否满足要求。第三,为了改善

预测模型的泛化能力,确定优化的或合理的网络拓扑结构和建立实用的BP网络模型,将一组训练后的BP网络整体视为网络预测模型,并将这组训练后的BP网络的输出值进行平均,将所得的值作为网络预测模型的预测值。很容易证明,与单个网络相比,网络组平均函数能更好地拟合真正的回归函数,网络组平均函数对新样本有更好的处理能力,可以改善所建网络预测模型对新样本的预测效果。第四,由于标准BP算法存在收敛速度慢、局部极值等主要缺点,采用动量-自适应学习率调整算法进行网络的训练。动量法降低了网络对于误差曲面局部细节的敏感性,有效地抑制网络陷入局部极小,而自适应学习率调整算法有利于缩短时间。

二、煤炭企业BOM数量模型实例及结果分析

1. 影响物料消耗因素的确定和数据获取

煤炭企业影响物料消耗因素很多,主要有作业量(原煤产量、月推进度、掘进进尺等)、地质构造(断层、褶曲、岩岩岩侵入等)、煤层赋存状况(煤

层硬度、顶板和底板强度、倾角、煤层稳定性、夹矸系数、煤的容重等)、工作面设计情况(采高、工作面走向长、工作面长、生产能力、工作面回采率等)、采煤工艺或支护方式、水文和物料管理水平等因素。对于不同物料其影响因素也有所不同。通过对有关煤炭生产现场专家的调查和分析,我们从中选择了8个因素作为乳化油消耗神经网络模型的输入,以乳化油单位消耗量作为网络输出。下面以某矿采煤作业的乳化油单位消耗数据,筛选其中的51个样本数据,作为算例建模数据。有关原始数据资料略。

2 样本设计及预测过程

在神经网络结构中,参数的选择十分重要。输入层和隐含层个数的增多会增加神经网络的表达能力,但个数的增多同时会影响网络的收敛速度并加大噪声干扰。为了使模型既具有理论价值,又具有现实可操作性,应选择典型并具有代表性的指标。通过分析采煤生产过程中影响乳化油单位消耗的各种因素,选取了地质构造、煤层厚度、煤层倾角等8个指标构成输入层。其输出层只有一个指标,即乳化油单位消耗数量。激活函数选取sigmoid函数 $f(x) = 1 / (1 + \exp(-x))$ 。

输入层中各输入因素的说明: a_1 煤层厚度(单位:米); a_2 煤层倾角(单位:度); a_3 平均采高(单位:米); a_4 工作面长(单位:米); a_5 煤层硬度; a_6 夹矸系数; a_7 地质构造情况; a_8 瓦斯水文情况。

(1) 指标评价价值的确定和无量纲化处理。在8个指标中有定性指标和定量指标之分,而且由于不同的指标是从不同的侧面影响回采工作面乳化油消耗情况,因此指标之间无法进行比较,并考虑神经网络训练的收敛问题,需对各指标进行无量

纲化处理。

定性指标。在影响回采工作面乳化油消耗的各因素中,定性指标有地质构造情况 a_7 、瓦斯水文情况 a_8 等。这些指标的评价我们采用专家打分的办法进行评价,其取值为0.1~1之间。评分标准为:地质构造情况按1~5类依次评分为[0.1]、[0.25]、[0.50]、[0.75]和[1.00];瓦斯水文情况按1~4类依次评分为[0.1]、[0.4]、[0.7]和[1.00]。

定量指标。通过调查分析,将煤层倾角 a_2 、煤层硬度 a_5 、平均采高 a_3 、夹矸系数 a_6 等确定为正指标,将煤层厚度 a_1 、工作面长 a_4 等确定为负指标。

正指标预处理函数为

$$f(x) = \begin{cases} 0.1 & x = x_{min} \\ 0.1 + 0.9 \times \frac{x_i - x_{min}}{x_{max} - x_{min}} & x_{min} < x < x_{max} \\ 1 & x > x_{max} \end{cases}$$

负指标预处理函数为

$$f(x) = \begin{cases} 1 & x = x_{min} \\ 0.1 + 0.9 \times \frac{x_{max} - x_i}{x_{max} - x_{min}} & x_{min} < x < x_{max} \\ 0.1 & x > x_{max} \end{cases}$$

(2) 训练过程。将样本分成不相交的6组数据,前45个样本按每组9个分成5组,后6个样本作为新样本。前5组依次以1组为验证样本,其余4组作为训练样本。

以第5组样本作为验证样本,以第1~4组样本进行训练,网络隐层为2层,网络拓扑结构为8-12-7-1,训练5238次,验证样本平均误差和为0.0039,训练误差随训练次数变化以及验证样本目标值与预测值逼近情况分别如图2和下页图3。

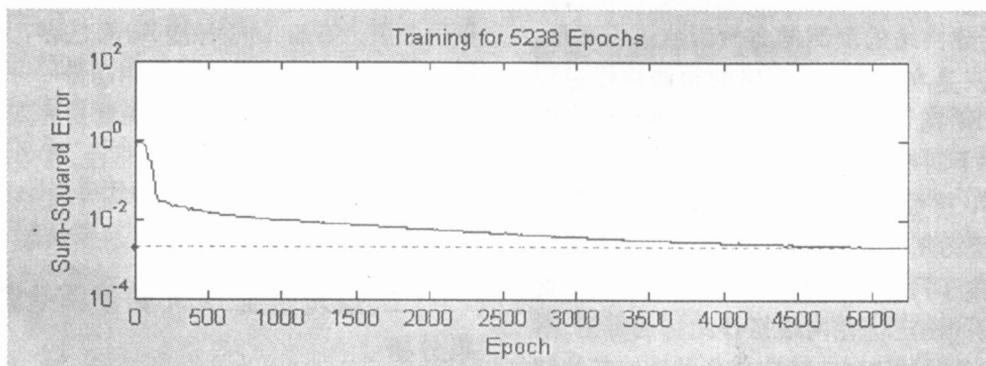


图2 训练误差随训练次数变化示意图

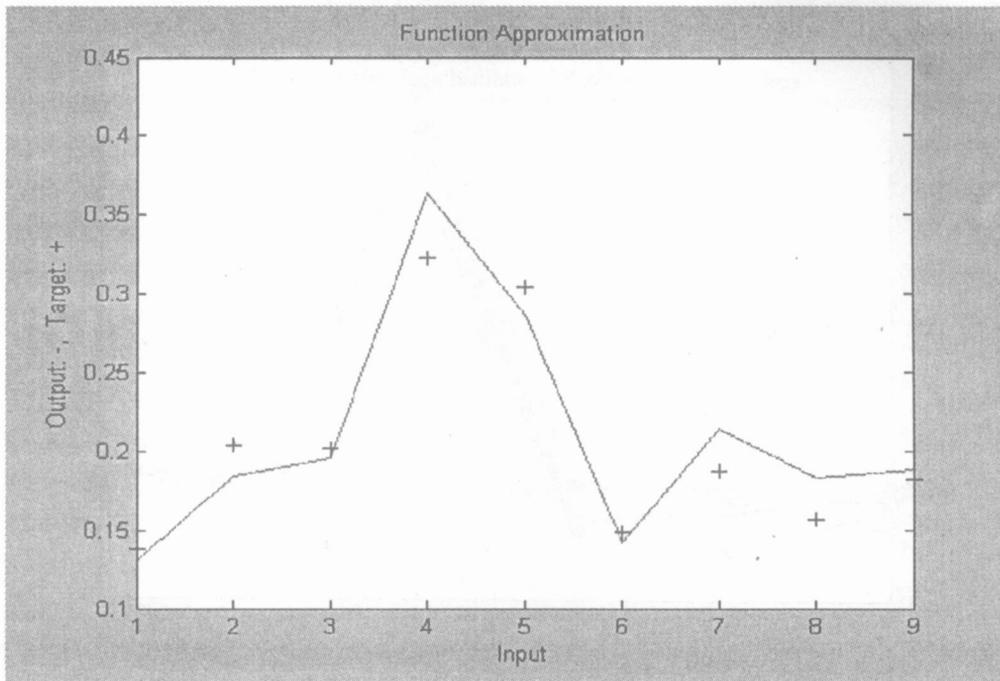


图 3 验证样本目标值与预测值比较

以第 4 组样本作为验证样本,以第 1、2、3、5 组样本进行训练,网络隐层为 2 层,网络拓扑结构为 8 12 6 1,训练 7 259 次,验证样本平均误差和为 0.005 5。

以第 3 组样本作为验证样本,以第 1、2、4、5 组样本进行训练,网络隐层为 2 层,网络拓扑结构为 8 12 7 1,训练 7 324 次,验证样本平均误差和为 0.003 1。

以第 2 组样本作为验证样本,以第 1、3、4、5 组样本进行训练,网络隐层为 2 层,网络拓扑结构为 8 11 7 1,训练 3 768 次,验证样本平均误差和为 0.004 7。

以第 1 组样本作为验证样本,以第 2~5 组样本进行训练,网络隐层为 2 层,网络拓扑结构为 8 12 7 1,训练 5 776 次,验证样本平均误差和为 0.004 2。

将最后 6 个样本作为新样本,提供给上述训练好的 5 个网络,可以得到网络预测值。各单个网络预测值和平均预测值与目标值的逼近如下页图 4 所示,目标值和平均预测值在图中标明,其他为单个网络预测值。预测值结果及与目标值的误差计算如表 1 所示。网络预测的平均误差和为 0.004 056,大于各网络平均误差和的平均值 0.004 28。

3. 实例分析

一般认为相对误差小于或等于 10% 的验证样本数在验证样本子集中所占比例大于 60%,所建预测模型就有实用价值。由表 1 可知,6 个测试样本中有 1 个样本的相对误差大于 10%,相对误差小于或等于 10% 的测试样本数在测试样本子集中所占比例为 83%,因此所建预测模型具有一定的实用价值。

表 1 人工神经网络模型预测值结果及误差计算

No	目标值	预测值	误差	相对误差	还原目标值	还原预测值	误差
46	0.199 4	0.184 1	- 0.015 3	7.67%	119.247 3	104.557 2	- 14.690 0
47	0.156 1	0.169 5	0.013 4	8.58%	77.673 5	90.5392 7	12.865 8
48	0.140 9	0.143 1	0.002 2	1.56%	63.079 5	65.1917 5	2.112 3
49	0.854 8	0.807 6	- 0.047 2	5.52%	748.518 6	703.200 3	- 45.318 3
50	0.229 7	0.199 2	- 0.030 5	13.28%	148.339 3	119.055 2	- 29.284 1
51	0.220 8	0.242 7	0.021 9	9.92%	139.794 1	160.821 0	21.026 9

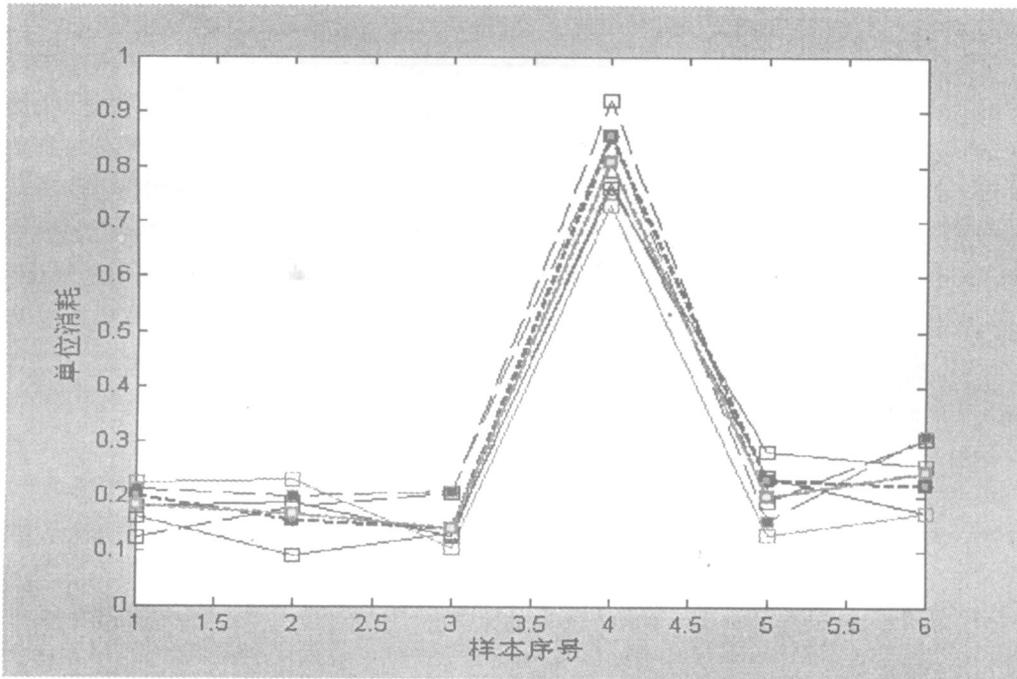


图 4 各网络预测值和目标值比较

三、结论

首先,从上述预测值和目标值的结果比较和分析可以看出,应用神经网络模型可以比较有效地确定不同生产条件的物料消耗量。当生产和管理中的影响物料消耗量各因素发生变化时,可以输入变化后的因素参数,利用经过训练后的神经网络模型来确定新条件的物料消耗量,从而使得物料消耗量的确定有一定的理论依据,也更加接近于物料实际消耗的现实值,更有利于物料消耗、库存及成本的控制。

其次,由于神经网络是一个动态系统,可以在系统中不断地通过学习修正自我,使得 ERP 系统中物料消耗量的确定不但可以根据其影响因素的变化而不断调整,还可以在实施过程中通过不断的学习训练而不断完善自我结构和连接,从而使物料消耗量的确定更加精确、更为准确有效。

最后,从还原后的预测值与原目标值的误差计算来看,网络预测结果不是很精确,仍存在一定误差。在实际应用时,可以通过搜集更多样本并改进数据预处理方法来提高其适用性。

[参考文献]

- [1] 夏鑫. 煤炭企业基于作业的 ERP 研究 [J]. 煤炭经济研究, 2004, (12): 32 - 34.
- [2] 吴忠. 基于神经网络动态确定 ERP/MRP 系统提前期的研究 [J]. 计算机应用与软件, 2003, (7): 5 - 6.
- [3] 李英龙. 采选工业系统综合建模方法研究 [M]. 昆明: 云南大学出版社, 2001: 10.
- [4] 蒋宗礼. 人工神经网络导论 [M]. 北京: 高等教育出版社, 2003: 2.
- [5] 夏鑫. 煤炭企业物料管理策略研究 [J]. 山东工商学院学报, 2006, (3): 24 - 27.

[责任编辑:刘丽娟]