

文章编号: 1001-0920(2004)03-0295-04

模块小波神经网络在工业产品质量控制中的应用

李换琴¹, 万百五²

(1. 西安交通大学 理学院, 陕西 西安 710049; 2. 西安交通大学 系统工程研究所, 陕西 西安 710049)

摘 要: 针对输入空间包含多种类型的数据时, 以单一的神经网络为模型, 其收敛很困难的问题, 提出一种基于模块小波神经网络的建模方法. 利用分而治之的思想, 模块神经网络通过一个门控网络进行分类和协调, 可以将一个复杂任务分解成几个简单的子任务, 每个子任务由一个局部专家网络学习. 与传统的模块网络不同, 这里的专家网络是小波网络而不是 BP 网络. 将所提出的网络模型用于热连轧产品质量建模, 并与单一的神经网络建模结果进行比较. 建模结果表明, 模块小波神经网络模型优于单一神经网络模型.

关键词: 模块小波网络; 高维输入; 质量模型; 热连轧机

中图分类号: TP18 **文献标识码:** A

Application of modular wavelet neural networks in industries product quality control

LI Huan-qin¹, WAN Bai-wu²

(1. Faculty of Science, Xi'an Jiaotong University, Xi'an 710049, China; 2. Institute Systems Engineering, Xi'an Jiaotong University, Xi'an 710049, China. Correspondent: WAN Bai-wu, E-mail: wanbw@mail.xjtu.edu.cn)

Abstract: When the input space consists of several different classes of input data, it becomes very difficult to converge the network during the training phase. A modular wavelet neural network is presented to overcome this difficulty. Based on the divide-and-conquer concept, a modular network is capable of dividing a complex task into subtasks, and modeling each subtasks with an expert network. To model such activities, a gating network is used for the classification and allocation of the input data to the corresponding expert network. Different from traditional modular networks, here each expert network is a wavelet network. The performance of such networks in modeling product quality is examined and compared with that of singular networks. Modeling results demonstrate that the proposed method takes a significant improvement over the general singular network model.

Key words: modular wavelet neural network; high-dimension input; quality model; hot rolling mill

1 引 言

一个企业要在市场中立于不败之地, 必须生产出高质量的产品, 并能根据客户的要求设计新产品. 这就要求企业开发产品质量模型和产品质量控制模型.

热连轧过程是钢铁制造的一个重要过程, 它包

含炼钢和轧制两个阶段. 炼钢过程中各化学成分的含量、各阶段炉温的高低、卷曲阶段钢卷温度等, 对成品钢的质量都有很大影响. 某钢厂 3 年多来详实记录的实测数据有 15 000 多组, 经过数据挖掘, 从中得出覆盖 12 个钢种的 9 026 个样本. 根据理论分析和实际操作经验, 经模型抽象和变量大量筛选, 认为

收稿日期: 2003-01-20; 修回日期: 2003-05-14

基金项目: 国家 863 计划资助项目 (863-51-011); 国家自然科学基金资助项目 (60274055); 西安交通大学自然科学基金资助项目 (0900-5-73024).

作者简介: 李换琴 (1964—), 女, 陕西韩城人, 副教授, 博士, 从事大规模复杂系统建模与控制的神经网络方法研究; 万百五 (1928—), 男, 江苏南京人, 教授, 博士生导师, 从事大工业过程控制、优化与辨识等研究.

成品钢的质量主要与 23 个化学元素及轧制过程中加热炉出口温度、板材厚度等 9 个参数的变化有关。质量指标主要考虑板材的金属物理性能,如断裂延伸率、抗拉强度和屈强比等。可以看出,钢的质量指标与影响它的参数之间是一种多变量复杂非线性关系,采用通常的统计与回归方法难以建立其模型。

神经网络理论的研究与发展为大规模复杂系统的建模提供了强有力的工具^[1]。近几年,神经网络已越来越多地应用于钢铁工业过程建模^[2~6]。但是,所有这些应用只局限于输入变量不超过 10 维的情形。从原理上说,神经网络能逼近任意复杂非线性函数,但随着问题复杂程度的增加,训练一个复杂网络所需样本和学习时间都急剧增加,而且所得到的复杂网络并不一定能揭示出问题的层次和结构。

针对热连轧过程的特点,人们相继提出了一些适合高维输入的各种改进算法的前馈神经网络建模方法^[7~9],然而建模效果却不甚理想。建模精度难以提高的原因有两个方面:一是输入变量维数(32 维)高,数据之间存在耦合现象,学习输入-输出关系网络需要大量隐层神经元;二是训练样本覆盖了 12 个钢种,用它训练多层前馈神经网络建立的产品质量模型,实际上是 12 个钢种质量指标在某种意义下的平均值。

为解决上述问题,本文利用分而治之的技术,同时鉴于对该问题小波网络优于普通 BP 网络^[9],提出一种基于模块小波神经网络的建模方法。该方法将复杂问题分为几个较为简单的子问题,通过一个门控网络进行协调,一个局部专家网络解决一个子问题。与必须学习整个原始任务的单一神经网络(NN)相比,各 NN 上的平均负载降低了,从而克服了单一 NN 的限制。

2 模块网络

模块网络也称为组合网络,可以学习不同区域函数表达式不同的复杂函数。它通过将一个任务分解成若干个功能独立的子任务,不同的任务分配到不同的子网络来实现这一目的。

模块网络结构如图 1 所示,它由两部分组成:专家网络和门控网络。输入数据同时加到专家网络和门控网络,其基本思想是用门控网络将一个任务分配给一个或几个专家网络。如果模块网络的输出不正确,与期望值有误差,就改变这些专家网络和门控网络的权值,使这些子网络局部化,以适应所分配的子任务。所有局部专家的输出通过门控网络进行调

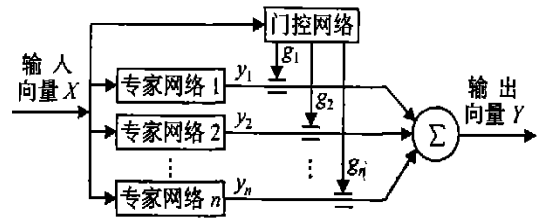


图 1 模块神经网络的结构

和,模块化网络的最终输出为

$$Y = \sum_{i=1}^n g_i y_i \quad (1)$$

式中: n 为专家网络的个数, y_i 为第 i 个专家网络的输出, g_i 为门控网络第 i 个输出神经元的激励。

文献[10, 11]从竞争混合的角度描述了模块化网络,在门控网络中使用 softmax 激励函数。更精确地说,门控网络对第 i 个输出单元使用 softmax 激励 g_i ,即

$$g_i = \frac{\exp(u_i)}{\sum_j \exp(u_j)} \quad (2)$$

式中 u_i 是流到门控网络第 i 个输出神经元的输入加权和。图 2 表示了门控网络及 g_i 的形成。由式(2)知,产生较小激励 u_i 的第 i 个局部专家网络的输出 y_i 不会对最终输出产生较大的影响。所以,在模块化网络中使用 softmax 激励函数可以提供一种“竞争”混合机制。

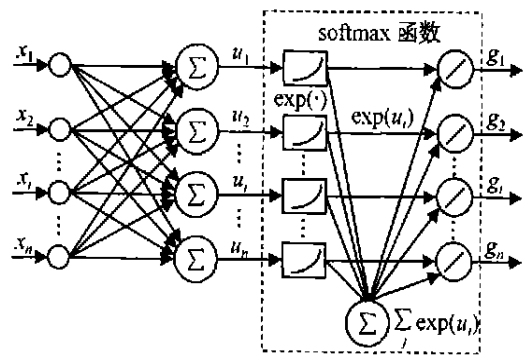


图 2 门控网络

模块网络学习中的关键问题是如何选择期望函数。Jacobs 提出了竞争学习机制^[11],在目标函数中加入惩罚项,各专家网络相互竞争而不合作,最终使得一个专家子网络激活。期望函数定义为

$$E^c = - \ln \sum_i \frac{g_i}{\sigma_i} \exp \left[- \frac{1}{2\sigma_i^2} (d^c - y_i^c)^2 \right] \quad (3)$$

式中: y_i^c 是专家网络 i 在情况 c 时的输出, σ_i 是专家网络 i 对输出所做贡献的比例参数, d^c 是 c 时的期望输出。式(3)要求各个专家网络产生整个输出向量

而不是一部分,因此在给定的训练中,局部专家网络的目标不直接受其他子网络权值的影响 因为在训练中,当其他某些子网络的权值改变时,门控网络会改变分配所选网络的责任 当一个专家网络产生的误差小于其他专家网络所产生的平均加权误差时,其责任会增加;反之,责任减少.

因为 y_i 是各专家网络的权系数 w_i 的函数, g_i 为门控网络第 i 个输出神经元的激励,是权系数 u_i 的函数,所以要学习的参数是 $2n$ 个权矩阵 w_1, w_2, \dots, w_n ; u_1, u_2, \dots, u_n 以及参数 σ_i , 分别记为

$$\begin{aligned} W &= [w_1, w_2, \dots, w_n]^T, \\ U &= [u_1, u_2, \dots, u_n]^T, \\ \sigma &= [\sigma_1, \sigma_2, \dots, \sigma_n] \end{aligned} \quad (4)$$

网络的学习过程就是调节这些参数,使

$$f(W, U, \sigma) = - \ln \prod_i \frac{g_i}{\sigma_i} \exp \left(- \frac{1}{2\sigma_i^2} d^c - y_i^c \right) \quad (5)$$

达到最小

3 热连轧过程

热连轧过程包含炼钢和轧制两个阶段 炼钢产生钢坯,钢坯从加热炉出来,经高压水除磷后进入粗轧机、精轧机,再经层流冷却机构将热轧带钢冷却至卷取温度后进入卷取机成卷 由连铸机至卷取机的连续工艺过程如图 3 所示

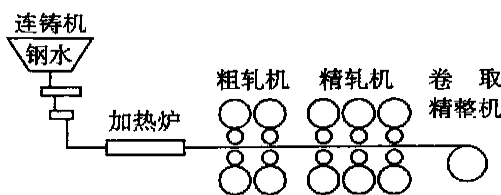


图 3 热连轧过程示意图

可见现代化钢铁企业生产的一个特点是工序繁多,热连轧产品生产要历经炼铁、炼钢、连铸、热轧等多个环节,每一步都对最终产品(板材)的质量产生影响 多辊热连轧机是一个串联型工业大系统,针对具体的产品质量要求,课题组把多辊热连轧生产线处理为一个静态系统,它的输入是分阶段加入的,具体包括化学元素含量、温度控制变量和时间变量等,而输出则表现为产品的质量指标 产品质量控制就是通过控制系统的输入变量来达到系统优良的输出,这需要多道工序的通力合作才能实现

衡量热连轧产品(板材)的质量指标一般有两类:一类是机械性能指标,如断裂延伸率、屈服强度、抗拉强度等;另一类是几何性能指标,如厚度、宽度、

板形等 本文要解决的是对机械性能指标的质量控制 输出变量包括断裂延伸率、屈服强度及抗拉强度等指标,输入变量包括化学元素成分、温度控制和时间等,而厚度、宽度等指标则作为扰动输入 经过模型抽象和变量筛选,最终确定输入变量为 32 个,输出变量为 3 个.

4 基于模块网络的热连轧产品质量模型

本文利用模块网络建立热连轧机生产线产品质量模型来进行质量预测 为了缩短训练时间及加快收敛速度,一般将神经网络质量模型分解成单个质量指标为输出的子质量模型 下面以断裂延伸率为输出建立其质量模型

4.1 样本数据归一化

由于实际测量的样本数据的各个分量取值范围不同,有的甚至相差很大,不能直接用实测数据训练神经网络 建模过程中,首先对数据进行归一化

用 $(X(i), e\ell(i))$ ($i = 1, 2, \dots, 9\ 026$) 表示实际测量的 9 026 个样本,其中 $X(i) = (x_1(i), x_2(i), \dots, x_{32}(i))$, 相应的归一化样本表示为 $(\underline{X}(i), \underline{e\ell}(i))$ ($i = 1, 2, \dots, 9\ 026$), 其中

$$\begin{aligned} \underline{X}(i) &= (x_1(i), x_2(i), \dots, x_{32}(i)), \\ x_j(i) &= \frac{x_j(i) - m_{\min}(x_j(k))}{m_{\max}(x_j(k)) - m_{\min}(x_j(k))}, \\ \underline{e\ell}(i) &= \frac{e\ell(i) - m_{\min}(e\ell(k))}{m_{\max}(e\ell(k)) - m_{\min}(e\ell(k))}. \end{aligned}$$

以下为表示方便, $(X(i), e\ell(i))$ 归一化后的样本 $(\underline{X}(i), \underline{e\ell}(i))$ 仍用 $(X(i), e\ell(i))$ 表示,即假定 $(X(i), e\ell(i))$ 已被归一化

4.2 网络结构确定

样本数据包含 12 个钢种,每一个钢种用一个专家子网络学习,因此,选择含有 12 个专家网络和 12 个输出的门控网络的组合网络为模型 门控网络为两层线性输出网络,每一个专家网络为 3 层前向小波网络^[9].

4.3 建模过程与结果

将 9 026 个样本分成两组:随机选取 5 000 组作为建模样本,其余 4 026 组作为检验样本 专家网络隐层神经元的激励函数取为 B 样条小波函数^[9] 隐层神经元个数开始时设置为 25,反复试验进行调节,当取值为 37 时,建模效果达到最好 这时,网络的拟合均方误差为 0.001 1,拟合命中率为 93.3%,检验误差为 0.001 3,检验命中率为 90.5%.

为比较起见,本文同时选择两隐层前馈小波神

神经网络建立此模型 隐层神经元的个数开始时分别取为15和10 为找到最优结果,反复试验 当隐层神经元的个数开始时分别取为30和10时,建模效果比较好,模型的拟合误差为0.0025,拟合命中率为83.3%,检验误差为0.0034,检验命中率为81.5% 这里,命中率指的是网络输出值与实际值相对误差小于0.05的样本占总样本的百分比

图4和图5分别给出了模块小波神经网络建模结果和双隐层小波神经网络建模结果,纵坐标表示钢板材质量指标——断裂延伸率的期望输出值,横坐标表示所建立的神经网络模型的输出值 图中给出的是随机抽取500个样本点的模型测试结果

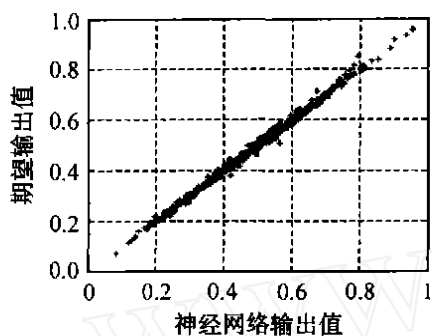


图4 模块小波神经网络建模结果

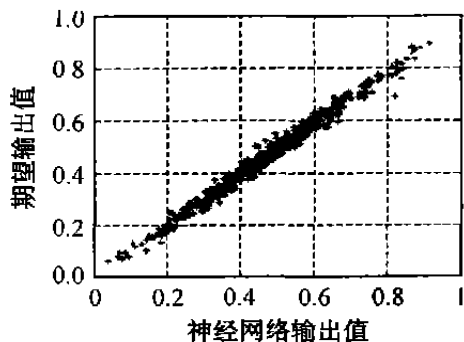


图5 双隐层小波神经网络建模结果

5 结 语

本文从企业多年记录的数据中挖掘其内在联系,建立了工业生产的产品的质量模型 该模型可用于产品质量控制和新产品的设计

本文提出了一种模块小波神经网络建模方法 这种由多个神经网络组合成的网络结构,建模精度远高于单一神经网络

参考文献(References):

[1] Parlos A, Chong K, Atiya A. Application of recurrent multilayer perceptron in modeling complex process

dynamics[J]. *IEEE Trans on Neural Networks*, 1994, 5(2): 255-266

[2] Chun M S, Yi J J, Moon Y H. Application of neural networks to predict the width variation in a plate mill [J]. *J of Materials Processing Technology*, 2001, 111 (1-3): 146-149.

[3] Liu Z Y, Wang W D, Gao W. Prediction of mechanical properties of hot-rolled C-Mn steels using artificial neural networks [J]. *J of Materials Processing Technology*, 1996, 57(3-4): 332-336

[4] Aistleitner K, Mattersdorfer L G, Haas W, et al Neural network for identification of roll eccentricity in rolling mills [J]. *J of Materials Processing Technology*, 1996, 60(1-4): 387-392

[5] Dukman Lee, Yongsug Lee. Application of neural-network for improving accuracy of roll-force model in hot-rolling mill [J]. *Control Engineering Practice*, 2002, 10(4): 473-478

[6] Pican N, Alexandre F. Artificial neural networks for the presetting of a steel temper mill [J]. *IEEE Expert*, 1996, 11(1): 22-27.

[7] 贾磊, 万百五, 冯祖仁. 以高维输入神经网络作为生产线产品质量模型 [J]. *控制与决策*, 2000, 15(5): 569-572 (Jia L, Wan B W, Feng Z R. Neural network quality model with high dimension inputs for production-line products [J]. *Control and Decision*, 2000, 15(5): 569-572)

[8] 邢进生, 万百五. 两阶段混合算法在轧钢数据建模中的应用 [J]. *西安交通大学学报*, 2000, 34(12): 105-107. (Xing J S, Wan B W. Mixed algorithm with two stages and its application in modeling for hot steel rolling data [J]. *J of Xi'an Jiaotong University*, 2000, 34(12): 105-107.)

[9] 李换琴, 万百五. 基于小波神经网络的大型多辊热连轧机产品质量模型 [J]. *系统工程*, 2002, 20(5): 55-58 (Li H Q, Wan B W. A product quality model for large-scale hot rolling mill based on high-dimension input wavelet neural network [J]. *Systems Engineering*, 2002, 20(5): 55-58)

[10] Jacobs R A, Jordan M I, Nowlan S J, et al Adaptive mixtures of local experts [J]. *Neural Computation*, 1991, 3(1): 79-87.

[11] Jacobs R A, Jordan M I, Nowlan, et al Task decomposition through competition in a modular connectionist architecture: The what and where vision task [J]. *Cognitive Science*, 1991, 15(2): 219-250