

宝钢 IF 钢大生产产品性能预测

初元璋 祁鹏 张娅

北京科技大学材料科学与工程学院,北京 100083

摘要 以 BP 算法为基础开发了 ANN 学习预测系统,用于宝钢 IF 钢大生产产品性能预测。同时,应用在宝钢 IF 钢大生产数据对该系统进行了测试和分析,并与多元线性回归结果进行预测精度比较。结果表明,ANN 学习预测系统,除 σ_2 误差较高(9.0%)外, σ_1 , δ , r 和 n 值均 $<5.0\%$,且比多线性回归方法精度高。

关键词 BP 算法; IF 钢; 性能预测

分类号 TG335.5

自 1995 年宝钢 IF 钢投入大生产以来,已达到年产 10 万 t 的水平。但是在 IF 钢生产中还存在着一定问题,主要表现是产品的质量不稳定。由于 IF 钢的生产过程中影响 IF 钢最终产品性能的因素较多,而且各影响因素对性能的影响和各因素之间的相关作用复杂,目前还没有得到可靠的数学模型,无法在生产过程中准确地预测产品性能和产品可能会出现的问题,进而在后序生产过程中应采取相应的补救措施。

本文研究开发了预测精度较高的 ANN 系统,采用较为合理的方法对这些数据进行处理,实现 IF 钢大生产产品性能的预测,用于生产和产品质量的控制。

1 BP 网络的设计与实现

BP 网络作为最常用的一种人工神经网络算法,具有很强的函数逼近能力和自学习、自组织能力,可以在没有数学模型条件下,根据可靠的数据建立起自变量与因变量间的映射关系,因而比较适合于 IF 钢大生产产品性能的预测。同时, BP 网络是人工神经网络中应用最为成熟和广泛的一种算法^[1],应用该算法解决 IF 钢大生产产品性能预测的问题是可行的。

IF 钢大生产中,影响最终产品性能的因素(包括:化学成分、各个工序工艺参数、产品规格等)很多,且关系复杂。随着对 IF 钢生产过程的

冶金学研究和 IF 钢大生产数据分析的深入,以及 IF 钢系列新品种的不断开发,对影响 IF 钢性能的因素和所关心的 IF 钢的性能认识会有所变化^[2],这将会导致 BP 网络的输入、输出节点的个数和各个节点所代表的物理意义的变化。因此,在设计 BP 网络的拓扑结构和数据结构时,应使各层的节点数可以随所处理数据组成的变化而变化。

IF 钢大生产过程中采集的数据量大,且不断有新数据加入。因此,在设计 BP 网络时,应使其具有随时添加可学习样本的能力。此外,为了尽可能地减小 BP 算法学习和预测的误差,在网络的数据结构和拓扑结构设计除了要便于误差的计算和权值的修正外,还要设法使样本数据的归一、样本随机排序简单易行。

为了实现上述目的,在实现 BP 网络时,采用了以下的处理方法:

(1) 所有样本数据和网络节点均采用链表结构。例如,输入层每个节点就是输入层链表的一个节点,其数据结构如图 1 所示。

这样,每个节点中记录了可以找到下一个节点的信息。如果在程序中记下了输入层的第一个节点,就可以按顺序找到每一个输入层节点。同样,输入数据、输出数据、隐层节点、输出层节点均采用类似的数据结构。这样既可以方便、可靠地找到每一个数据,又可以方便地保证各项数据信息的正确传递。图 2 所示就是从输入数据到隐层节点的数据对应关系。从图中可以看出,尽管各层和数据层面之间的对应关系

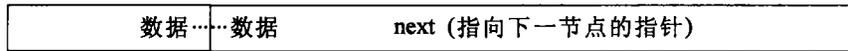


图 1 BP 网络单个输入层节点的数据结构

Fig.1 Data structure of input layer node of BP neural networks

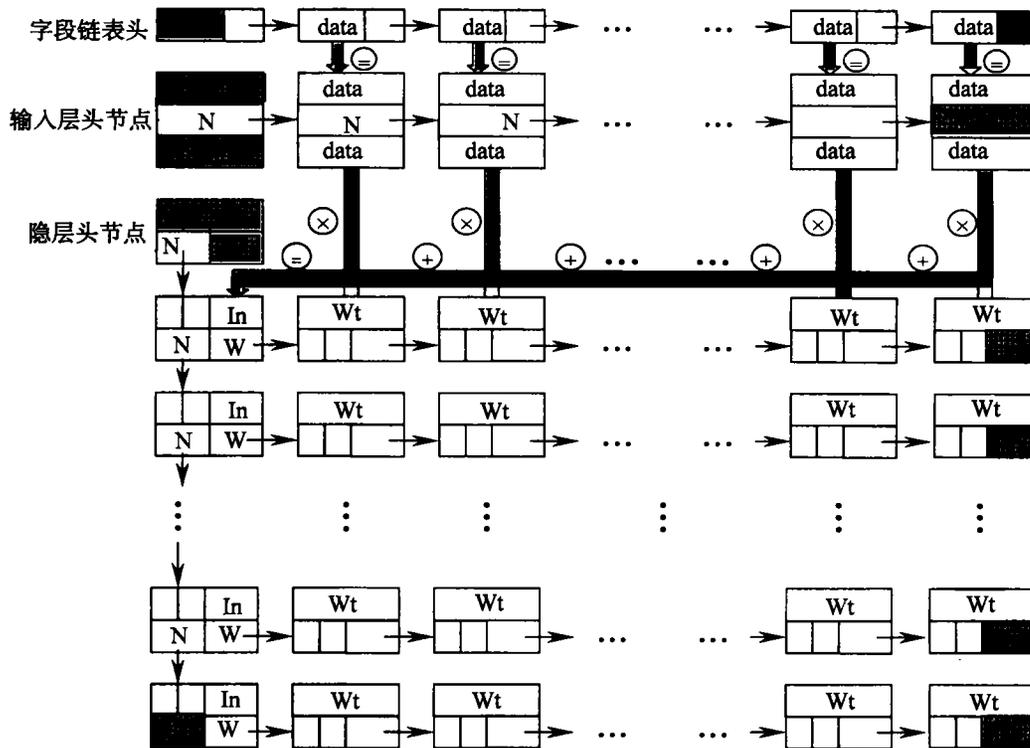


图 2 样本记录中字段、输入层节点、隐层间节点对应关系

图中阴影框表示相应数据为空, 细线箭头为指针, 带纹理的粗箭头为数据流向, 其一侧圆圈内运算符是与其所连接数据的关系; 图中 data 代表字段链表中节点保存的数据, In 代表输入值, N 代表 next 指针, W 代表权值链表的头指针, Wt 代表权值链表中节点保存的权值。

Fig.2 Corresponding relationship among fields in simple records, nodes in input layer and nodes in hidden layer

很复杂, 数据间的运算也比较繁琐, 但程序实现过程中的数据对应关系清晰明了。采用这样的数据结构的另一个优点是, 不必知道样本中数据的个数(即 BP 网络输入、输出层节点的个数), 只要找到样本中的第一个数据项, 就可以根据样本数据直接构造网络的输入、输出层, 进而构造出隐层, 从而建立一个拓扑结构, 它可根据所处理数据自动生成的 BP 网络。在对样本进行学习或预测时, 也不必知道样本数, 而是只要找到第 1 个样本即可。

(2) 网络的拓扑结构是根据所学习的样本的数据构成确定的。即根据 1 个样本数据的链表结构中节点个数确定相应的 BP 网络的输入、输出层的节点个数。由于 BP 网络的隐层节点数据对网络学习和预测的结果有很大影响, 为了确保隐层节点能适应样本数据的结构, 程序中采用隐层节点数随输入层节点变化的策略, 实际

上就是隐层节点数由输入数据的项数确定, 可以保证网络的拓扑结构较好地适应样本的数据构成。

(3) 在 BP 算法的实现中, 还采用了输入、输出数据归一化、样本随机排序等措施来减小学习和预测过程中的误差, 使得网络取得较好的运算结果。

由于采用上述方法, 使得本文所建立 ANN 学习预测系统有较好的学习和预测精度。应用该系统能够对宝钢 IF 钢大生产中实际数据进行处理和分析, 预测宝钢 IF 钢大生产产品的性能, 为产品的质量控制在提供依据。

2 实例分析

2.1 IF 钢大生产数据的初步分析

收集了宝钢 IF 钢大生产过程中收集了大量

的数据,其中包含了化学成分、生产工艺参数、产品规格、产品性能及其在各个工序中的编号等.

根据IF钢的用途和特点,反映IF钢产品优劣的基本性能有5项,分别是屈服强度($\sigma_{0.2}$)、抗拉强度(σ_b)、延伸率(δ)、各向异性塑性应变比(r)和加工硬化指数(n).通过对IF钢大生产过程进行深入的冶金学原理分析和对IF钢大生产数据的数据分析,确定了影响IF钢产品性能的17项主要因素包括:

(1)化学成分(11项):碳、氮、钛、磷、硫、硅、锰、镍、铜、铝、铬的含量(无含铌IF钢样本,故未取铌含量,此外,碳、氮、钛、硫4个参数中隐含了有效钛碳比).

(2)4个温度参数:开轧(加热)温度SRT、粗轧终轧温度RT4、精轧平均温度FT,卷取平均温度CT(精轧、卷取的最高、最低温度因为不能较好反映整个冷轧卷的状况而舍去;另外由于现场条件等的限制,未能取到对性能的影响更为显著的退火工艺参数).

(3)2个尺寸参数:热轧后产品厚度HPR、冷轧后产品厚度HPC(这两个参数实际上相当于冷轧压下率).

2.2 应用人工神经网络的结果及对该系统的检验

为了检验ANN学习预测系统预测IF钢大生产产品性能的精度,本文选取了873组包含了前述所有参数的、宝钢IF钢大生产中采集的数据,按照冷轧号把它分为两部分:冷轧号在前的823组数据,用于网络的学习;其余50组数据用于预测,即40组数据的值在已学习过数据的极值所确定的范围内(称为A类样本),其他10组数据中至少有1项或2项数据的值超出了已学习数据的范围(称为B类样本).

同时,用多元线性回归结果(回归时用到的所有873组数据,分别采用SPSS和MATLAB 2种数据分析软件进行多元线性回归)对上述

50组数据进行预测,并将其结果与人工神经网络的预测结果进行了对比.结果如表1所示.

由表中数据可以知道:ANN学习预测系统的预测精度要高于多元线性回归方程的预测精度,尤其对于所有数据的值全部落在由已学习过样本的极值确定的闭区间内的样本(A类样本),ANN学习预测系统的预测结果有较高的精度,5项性能指标中只有屈服强度的平均预测误差较大为9.0%,其余4项的预测误差均在5.0%左右,5项平均误差为4.9%,可满足工程使用的要求;当被预测样本中有数据的值超出由已学习过样本的极值确定的闭区间(B类样本)时,屈服强度和 r 值的预测误差急剧增大,分别达到了14.0%和21.7%,已没有参考价值,但此时抗拉强度、延伸率和 n 值的预测误差还在5%左右,说明这时神经网络的预测精度已不确定.这是因为应用神经网络逼近函数时,实质上是通过多次对输入值较简单的处理(但不是线性处理)和对每次处理结果的加权迭加,在学习过的范围内接近被逼近函数映射关系的,在进行权值修正时没有(也不可能)考虑学习范围以外逼近的误差大小,相当于用大量曲线的加权迭加,在一定范围内拟合另一条复杂曲线,但不考虑在此范围外与实际曲线间的差别,因此对未学习过的范围内的数据进行预测时,其误差将是不可预测的.这也说明应用神经网络进行函数逼近时,比如这里的性能预测,不具备数据定义域上的推广能力.

2.3 ANN学习预测系统特点

(1)预测精度高,高于多元线性回归所得方程的预测精度.

(2)完全不用数学或物理模型,因而无须知道影响因素与被影响因素间的任何确定关系,只要确定影响因素与被影响因素即可,故可用于尚未研究透彻的领域.

(3)适应范围广,ANN学习预测系统的输

表1 ANN学习预测系统对50组测试数据预测的平均误差
Table The Results of processing 50 samples by ANN Learn-Forecast System %

预测方法		$\sigma_{0.2}$	σ_b	δ	r	n	样本数
人工神经网络预测结果	A类样本	9.0	2.1	4.1	5.7	3.5	4.9
	B类样本	14.0	4.3	6.3	21.7	6.5	10.6
	全部样本	10.0	2.5	4.5	8.9	4.1	6.0
SPSS 预测结果		12.5	7.3	6.1	9.0	17.7	10.5
MATLAB 预测结果		6.5	11.5	3.7	9.2	19.5	10.1

入、输出节点可自动适应数据的结构,只需将数据以.dbf表或.txt文本格式保存,即可用该程序处理.

(4)对新样本具有敏感性,当遇到工艺变化等情况,可通过对新样本的反复学习迅速达到较高的预测精度.

3 ANN 学习预测系统的应用

ANN 学习预测系统可以根据 IF 钢的化学成分、生产工艺及产品规格预测产品的最终性能,从而为质量控制和生产工艺的合理选择提供依据.在生产过程中要预测某一产品的最终性能时,先将产品已知的化学成分和工艺参数、规格参数作为输入值输入网络,其他未知的参数取其工艺设定值,即可以预测出该产品在已

有条件下根据预定工艺进行后序生产时产品的最终性能,这样就可以预测按照预定的后序工艺能否生产出满足要求的产品,进而根据预测的结果调整工艺参数、化学成分等影响性能的因素.同时,该系统还可应用于解决其他输入、输出关系不明确,但可收集到一定量的数据的复杂问题的结果预测,因而具有一定的实用价值和推广性.

参考文献

- 1 祁鹏. 宝钢 IF 钢大生产数据分析及产品性能预测: [硕士学位论文]. 北京:北京科技大学, 2000
- 2 江建华, 马衍伟, 王先进, 等. 用人工神经元方法预报 IF 钢性能. 北京科技大学学报, 1997, 19(5): 467
- 3 张乃尧, 阎平凡. 神经网络与模糊控制. 北京:清华大学出版社, 1998

Performance Forecast of IF Steel Mass-Produced in Bao Steel

CHUN Yuanzhang, QI Peng, ZHANG Ya

Material Science and Engineering School, UST Beijing, Beijing 100083, China

ABSTRACT Develop ANN learn-forecast system by employing BP algorithm to forecast the performance of IF steel, test and analyze the system by using data collected from BAO Steel, and compare the precision of forecasted data with that of the multivariate linear regression model. The results show that the relative errors of ANN learn-forecast system on σ_b , δ_1 , r and n are all less than 5.0% except that on $\sigma_{0.2}$ is 9.0%. It is concluded that this system has a higher forecast precision than the multivariate linear regression model.

KEY WORDS BP algorithm; IF steel; performance forecast