

高炉透气性指数的改进多层超限学习机预测模型

苏晓莉, 尹怡欣, 张 森[†]

(北京科技大学 自动化学院, 北京 100083; 北京科技大学 钢铁流程先进控制教育部重点实验室, 北京 100083)

摘要: 高炉透气性指数是高炉操作者衡量高炉顺行状态的指标之一。针对传统透气性指数测量模型的缺陷, 本文提出了一种基于改进的多层超限学习机(multi-layer extreme learning machine, ML-ELM)的高炉透气性指数预测模型。首先分析影响高炉透气性指数的相关操作参数, 考虑到高炉生产数据含有大量噪声, 运用小波去噪方法消除数据的噪声干扰。然后建立高炉透气性指数预测模型。在建模过程中, 将偏最小二乘(partial least square, PLS)与多层超限学习机算法结合, 消除多层超限学习机最后一层隐藏层的多重共线性, 提高了模型预测精度。并且所提出的改进算法称为PLS-ML-ELM。最后使用现场生产数据对该模型训练和测试, 预测结果表明所提出模型能够快速、精确地预测高炉透气性指数, 并且为高炉的后续操作提供有效的决策与支持。

关键词: 高炉; 透气性指数; 建模; 多层超限学习机; 预测

中图分类号: TP273 **文献标识码:** A

Prediction model of improved multi-layer extreme learning machine for permeability index of blast furnace

SU Xiao-li, YIN Yi-xin, ZHANG Sen[†]

(School of Automation, University of Science and Technology Beijing, Beijing 100083, China;
Key Laboratory of Advanced Control of Iron and Steel Process (Ministry of Education),
University of Science and Technology Beijing, Beijing 100083, China)

Abstract: Permeability index of blast furnace is one of significant indicators of measuring the anterograde state of blast furnace for operators. Aiming at the defects of traditional permeability index measurement model, this paper proposes a prediction model for permeability index based on improved multi-layer extreme learning machine algorithm (ML-ELM). Firstly, relevant operation parameters are chosen through analyzing the mechanism of blast furnace. Given to blast furnace production data contain noise, wavelet transform is adopted to get rid of interference. Secondly, the prediction model of permeability index is established. Multi-layer extreme learning machine and partial least square method (PLS) are combined to overcome output matrix multicollinearity of the last hidden layer for ML-ELM and prediction accuracy is improved. And the improved algorithm is named as PLS-ML-ELM. Finally, practical production data are used to train and test this model. Simulation results indicate that the model can quickly and accurately predict permeability index and can offer efficient decision for sequent blast furnace operation.

Key words: blast furnace; permeability index; model; multi-layer extreme learning machine; prediction

1 引言(Introduction)

高炉炼铁是一个化学反应众多、强耦合、多变量、非线性的复杂工业过程^[1-2]。及时掌握高炉运行状况和根据其进行合理操作是实现高炉炼铁高产低耗的关键所在。目前, 高炉透气性指数是综合反映高炉运行状态的重要指标之一。根据透气性指数, 高炉操作人员可以尽早发现和避免高炉悬料、崩料等高炉炉

况失常现象, 从而判断高炉是否处于顺行状态。由此可见, 对于高炉操作者来说, 高炉透气性指数的准确预测是至关重要的。

传统的透气性指数测量模型通常用高炉的冷风流量与全压差(热风压力与炉顶压力之差)之比来表征高炉透气性指数^[3], 数学表达式为

$$T_z = \frac{Q}{(p_1 - p_2)} = \frac{Q}{\Delta p}, \quad (1)$$

收稿日期: 2016-05-07; 录用日期: 2016-10-18.

[†]通信作者. E-mail: zhangsen@ustb.edu.cn.

本文责任编辑: 吴敏.

国家自然科学基金重点项目(61333002), 国家自然科学基金项目(61673056)资助.

Supported by Key Program of National Natural Science Foundation of China (61333002) and Program of National Natural Science Foundation of China (61673056).

其中: T_z 表示高炉透气性指数, Q 表示冷风流量, p_1 表示热风压力, p_2 表示炉顶压力。

由于高炉内部反应复杂、因素众多, 仅使用冷风流量、热风压力和炉顶压力这3个高炉参数难以全面表征高炉透气性指数, 而且不能实现准确预测。这直接导致高炉操作人员不能依据其进行后续的高炉生产优化控制, 进而无法满足高炉现场的需求。目前, 国内外关于高炉透气性指数的预测模型十分少见。

多层超限学习机算法(multilayer extreme learning machine, ML-ELM)是在超限学习机-自动编码器(extreme learning machine-autoencoder, ELM-AE)的基础上提出的。ML-ELM的训练速度和测试速度比现有的深层网络更快, 它的泛化能力可以达到甚至优于其他深层网络的精度。而且, ML-ELM不需要微调。因此, ML-ELM可以被用来建立预测模型。

针对高炉透气性指数的相关分析, 本文选取透气性指数作为输出量, 使用高炉生产数据, 结合ML-ELM算法, 建立高炉透气性指数预测模型。考虑到ML-ELM最后一层隐藏层存在多重共线性问题, 本文运用偏最小二乘方法(partial least square, PLS)优化ML-ELM, 从而实现高炉透气性指数的高效、精确预测。为了保证模型的精准度和可信度, 本文根据高炉相关机理, 分析影响透气性指数的因素, 并且运用小波变换方法去除现场生产数据噪声。实验结果表明, 采用改进ML-ELM算法建立的预测模型能够实时精准预测高炉透气性指数, 为高炉顺行和后续操作提供支持。

本文结构如下: 第2部分介绍影响透气性指数的因素及高炉生产数据噪声处理; 第3部分简要介绍ML-ELM和与其相关的算法, 并详细介绍改进ML-ELM算法, 最后给出透气性指数预测模型; 本文的仿真实验和仿真结果在第4部分进行介绍; 第5部分对本文进行总结。

2 透气性指数影响因素分析及数据预处理(Permeability index factor analysis and data preparation)

测量高炉透气性指数的传统模型以冷风流量、热风压力和炉顶压力为根据, 但是高炉内部环境复杂, 仅以这3个参量不足以全面反映高炉透气性指数的变化, 从而不能准确判断高炉炉况。由于高炉内部环境复杂, 故影响高炉透气性指数的因素众多。因此, 在建立高炉透气性指数预测模型之前, 首先确定影响透气性指数的相关因素作为预测模型的输入参量。同时, 高炉操作现场环境复杂, 影响透气性指数的相关参数的数据带有噪声, 需要进行数据去除噪声处理, 从而提高预测模型的性能, 实现精准预测的目的。

2.1 透气性指数影响因素分析(Permeability index factor analysis)

本文除了采用冷风流量、热风压力和炉顶压力这3个与高炉透气性指数相关的参数作为输入参量外, 还结合高炉内部机理分析影响其他高炉透气性指数的相关因素。其中, 风温表示高炉鼓风的温度。风温越高, 鼓风带入炉内的热量越多, 进而产生高炉下部的热波动; 风速是指鼓风在高炉风口出风处的速度, 风速可以影响高炉炉缸的活跃程度, 并且影响高炉下部的煤气流分布情况, 因此也影响高炉下部的热波动; 炉顶温度可以反映高炉内部的热交换状态。由高炉的热制度、送风制度可以影响高炉透气性指数可知, 风温、风速和顶温这3个参数均影响高炉透气性指数。

综合上述高炉内部机理分析, 本文选取风量、风温、风压、风速、顶温、顶压这6个高炉参数。结合Pearson相关系数方法, 分别计算这6个参数与透气性指数之间的相关程度。设参数 X 和 Y 有如下两组数值: $X: x_1, x_2, \dots, x_n; Y: y_1, y_2, \dots, y_n$, 则参数 X 和 Y 的Pearson相关系数定义如式(2)所示:

$$\rho = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{\sum_{i=1}^n (x_i - \bar{x})^2} \sqrt{\sum_{i=1}^n (y_i - \bar{y})^2}}, \quad (2)$$

其中: 相关系数 ρ 的绝对值($|\rho|$)表示两个参数之间的相关程度, 当 $0.8 \leq |\rho| \leq 1.0$ 时, 表示两个参数有着极强的相关性。通过选取1500组高炉现场生产数据, 根据式(2)计算得到每个参数与透气性指数的相关系数, 结果如表1所示。

表1 高炉参数与透气性指数的Pearson相关系数的计算结果

Table 1 Results of Pearson correlation coefficient between parameters of blast furnace and permeability index

高炉参数	Pearson系数
风量	0.8641
风温	0.8532
风压	0.9219
风速	0.9085
顶温	0.8276
顶压	0.8132

由表1可知, 所选取的6个高炉参数分别与透气性指数有着极强的相关性。

根据高炉内部机理分析和Pearson相关系数分析, 本文选取风量、风温、风压、风速、顶温、顶压这6个高炉参数作为预测模型的输入参量。由于本文对高炉透气性指数进行预测, 故将其作为预测模型的输出参量。

2.2 基于小波变换的高炉现场数据预处理 (Preparation of blast furnace field data based on wavelet transform)

由于高炉内部环境复杂、换炉操作等因素的影响,从高炉现场所获得的高炉生产数据包含大量噪声^[4].并且,使用伴有噪声的高炉数据建立预测模型会使预测效果变差,所以先对高炉生产数据进行处理.

小波变换可将信号分解成具有不同频率和尺度的子信号,并且能够保留信号的重要部分^[5].而且,小波变换对非平稳信号具有良好的去除噪声能力,因此本文采用小波变换对高炉现场数据进行去除噪声处理.

从现场采集7个高炉生产参数(风量、风温、风压、风速、顶压、顶温、透气性指数)的数据样本序列,表2列出了原始采样数据的相关信息.分别对这7个参数进行去除噪声处理.本文以风量参数为例,采用MATLAB小波工具箱中的WDEN()函数进行去噪处理.经过多次实验可知,选用的小波函数为sym8,阈值选择标准为固定阈值原则(sqrtwolog),采用软门限阈值处理方式,并且根据每一层小波分解的噪声水平估计来调整阈值,由于分解层次大于5层后去噪效果已不明显,故分解层次选为5层.图1为从高炉现场采集的风量数据的去噪过程.

表2 样本数据信息

Table 2 Information of sampling data

No.	参数	平均值	最小值	最大值	标准差	量纲
输入	1 风量	4180.12	4131.04	4242.13	12.35	m ³ /min
	2 风温	1179.86	1165.74	1189.54	2.34	°C
	3 风压	346.05	316.74	357.94	5.20	kPa
	4 风速	256.68	247.47	276.40	3.38	m/s
	5 顶温	200.33	126.84	326.86	39.95	°C
	6 顶压	190.20	186.13	214.58	1.19	kPa
输出	1 透气性指数	0.81	0.61	0.89	0.03	m ³ /(min · kPa)

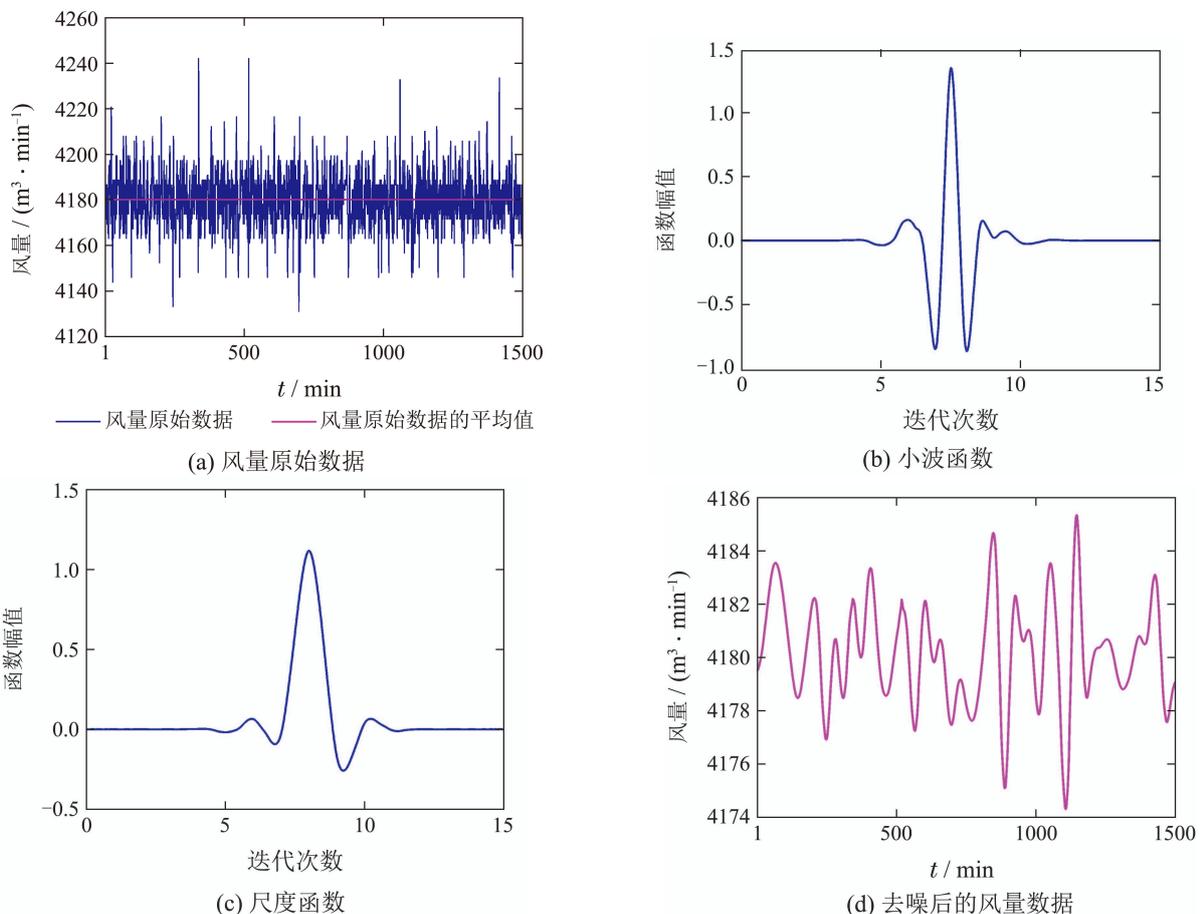


图1 风量数据的去噪过程

Fig. 1 The de-noising process of blast volume

从图 1 中的 (a) 图可知风量数据在平均值 4180.12m³/min 附近上下波动, 并且波动幅度较大; 图 1 中的 (b) 图是选用的小波函数 sym8; 图 1 中的 (c) 图表示小波函数 sym8 的尺度函数; 图 1 中的 (d) 图去噪后的风量数据。

3 基于 PLS-ML-ELM 的高炉透气性指数预测模型 (Prediction model of permeability index for blast furnace based on PLS-ML-ELM)

由于影响高炉透气性指数的因素众多, 并且和 ELM 算法相比, ML-ELM 算法具有多层学习能力、深度特征提取和更高的准确度等优点, 本文采用 ML-ELM 算法建立预测模型。但是, ML-ELM 最后一层隐藏层存在多重共线性问题, 影响高炉现场对透气性指数预测提出的高精度要求。因此, 本文将 ML-ELM 算法和 PLS 方法结合, 消除 ML-ELM 的多重共线性问题, 提高预测模型精度和泛化能力, 从而实现精确预测透气性指数的目的。本章首先简要介绍与 ML-ELM 算法相关的 ELM 算法、ELM-AE 算法, 并且介绍 ML-ELM 算法。然后详细介绍改进的 ML-ELM 算法, 最后介绍采用改进的 ML-ELM 算法所建立的高炉透气性指数预测模型。

3.1 超限学习机 (Extreme learning machine, ELM)

由黄广斌等人提出的 ELM 是一个简单、高效并且具有单隐藏层的学习算法^[6-7]。ELM 的结构图如图 2 所示。假设有 N 个独立样本 (x_i, t_i) , 其中: $x_i = [x_{i1} \ x_{i2} \ \cdots \ x_{in}]^T \in \mathbb{R}^n$ 表示第 i 个样本有 n 维特征, $t_i = [t_{i1} \ t_{i2} \ \cdots \ t_{im}]^T \in \mathbb{R}^m$ 表示目标矢量, 则 l 个隐藏层节点和激活函数 $g(x)$ 可以由下面的数学模型表示

$$\sum_{i=1}^l \beta_i g_i(x_j) = \sum_{i=1}^l \beta_i g(w_i \cdot x_j + b_i) = o_j, \quad (3)$$

$$k = 1, 2, \dots, N,$$

其中: $w_i = [w_{i1} \ w_{i2} \ \cdots \ w_{in}]^T$ 是连接第 i 个隐藏层节点和全部输入节点的权值矢量, $\beta_i = [\beta_{i1} \ \beta_{i2} \ \cdots \ \beta_{im}]^T$ 是连接第 i 个隐藏层节点和全部输出节点的权值矢量, b_i 是第 i 个隐藏层节点的偏置值, $w_i \cdot x_j$ 是 w_i 和 x_j 的内积, 并且 $g(w_i \cdot x_j + b_i)$ 是第 i 个隐藏层节点的输出。在本文中, 输出节点被选择为线性的。

上述 N 个方程可以简要地写为下面的形式:

$$H\beta = T, \quad (4)$$

其中

$$H = \begin{bmatrix} g(w_1 \cdot x_1 + b_1) & \cdots & g(w_n \cdot x_1 + b_n) \\ \vdots & & \vdots \\ g(w_1 \cdot x_N + b_1) & \cdots & g(w_n \cdot x_N + b_n) \end{bmatrix}_{N \times n}, \quad (5)$$

$$\beta = \begin{bmatrix} \beta_1^T \\ \vdots \\ \beta_n^T \end{bmatrix}_{n \times m}, \quad T = \begin{bmatrix} t_1^T \\ \vdots \\ t_N^T \end{bmatrix}_{N \times m}, \quad (6)$$

H 称为隐藏层的输出矩阵, H 的第 i 列是相对于输入 x_1, x_2, \dots, x_N 的第 i 个隐藏层节点输出矢量。

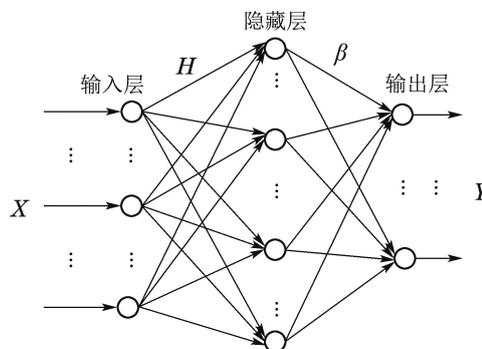


图 2 ELM 的结构图

Fig. 2 The structure of ELM

ELM 算法使用最小二乘法计算输出权值 β , 可以由下式表示

$$\beta = H^\dagger T = (H^T H)^{-1} H^T T. \quad (7)$$

3.2 超限学习机-自动编码器 (Extreme learning machine-autoencoder, ELM-AE)

广泛应用于深度学习^[8]的自动编码器是无监督的人工神经网络, 它的输出和输入相同, 并且能很好地逼近输入信号^[9]。Chamara L L 等人^[10]提出的 ELM-AE 是神经网络的新方法, 它可以像深度学习中的自动编码器那样重建输入信号。

ELM-AE 是无监督神经网络, 并且 ELM-AE 的输出和 ELM-AE 的输入相同。图 3 是 ELM-AE 的模型。ELM-AE 的输入权值和隐藏层的偏置矢量是随机生成的, 而且是正交的。对于 N 个独立的样本, $x_i \in \mathbb{R}_n \times \mathbb{R}_j$ ($i = 1, 2, \dots, N$), ELM-AE 隐藏层的输出可以表示为

$$h = g(ax + b), \quad (8)$$

其中: $a^T a = I$, $b^T b = I$ 。隐藏层输出和输出层输出之间的数学关系可以表述为

$$h(x_i)\beta = x_i^T, \quad i = 1, 2, \dots, N, \quad (9)$$

其中 β 为输出权值。这里的输出权值有 3 种求解情况:

情况 1 训练样本数目大于隐藏层节点数目

$$\beta = \left(\frac{I}{C} + H^T H\right)^{-1} H^T X. \quad (10)$$

情况 2 训练样本数目等于隐藏层节点数目

$$\beta = H^{-1} X. \quad (11)$$

情况 3 训练样本数目小于隐藏层节点数目

$$\beta = H^T \left(\frac{I}{C} + HH^T \right)^{-1} X. \quad (12)$$

3.3 多层超限学习机(Multi-layer extreme learning machine, ML-ELM)

Chamara L L等人^[10]提出的ML-ELM可以利用无监督学习去训练每层的参数. 和现在流行的深度学习算法相比, ML-ELM最突出的特点是ML-ELM不需要调整网络, 这正是ML-ELM可以用很短的时间去训练网络的原因^[11]. 图4为ML-ELM的模型.

为了训练每层参数, ML-ELM采用ELM-AE算法. ML-ELM的隐藏层激活函数可以是线性或者非线性分段. 如果第 k 层隐藏层的节点数等于第 $k-1$ 层隐藏层的节点数, 激活函数 $g(x)$ 应该是线性的, 否则 $g(x)$ 应该是非线性分段的, 所以第 k 层隐藏层的输出可以表示为

$$H_k = g((\beta^k)^T H_{k-1}), \quad (13)$$

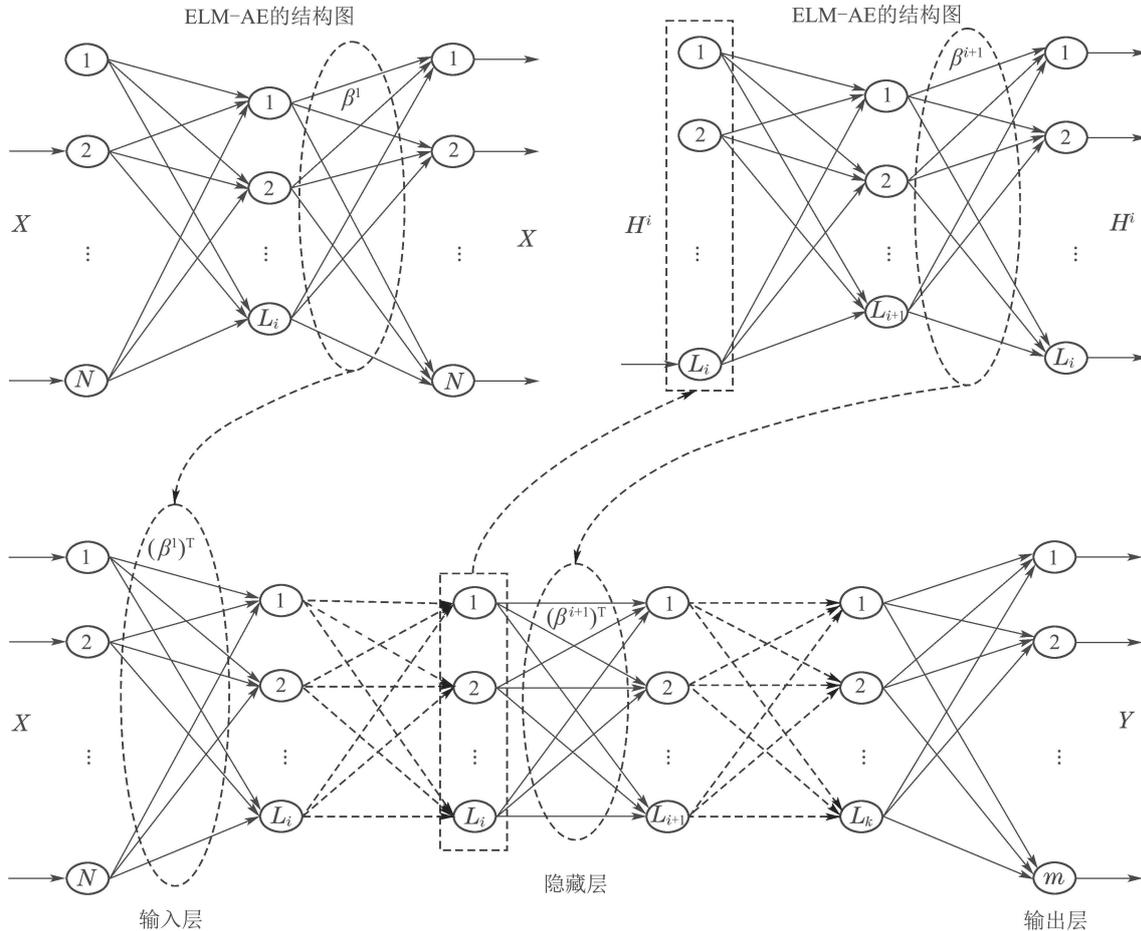


图4 ML-ELM的结构图

Fig. 4 The structure of ML-ELM

3.4 基于偏最小二乘的改进多层超限学习机(The improved multi-layer extreme learning machine based on partial least square, PLS-ML-ELM)

在ML-ELM算法中, 当最后一层隐藏层的输出矩

其中: H_k 是第 k 层隐藏层的输出, 第 $k-1$ 层隐藏层的输出矩阵 H_{k-1} 是第 k 层隐藏层的输入矩阵, $g(\cdot)$ 表示隐含层的激活函数, 注意: $k-1=0$, 本文可以解释这种情况为 H_1 是第一隐藏层, H_0 是输入层.

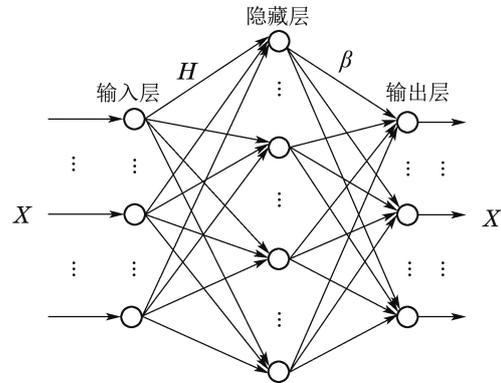


图3 ELM-AE的结构图

Fig. 3 The structure of ELM-AE

阵高度共线性时, ML-ELM算法的泛化能力会受到限制. 集合了多元回归和主成分分析的优点的偏最小二乘算法是一个有效建立多自变量和多因变量关系的方法^[12-13], 而且, 偏最小二乘方法可以解决多重共线

性问题和消除噪声. 此外, 最后一层隐藏层和输出层之间的输出权值可以通过偏最小二乘直接计算得到. 所以, 改进的ML-ELM算法可以使计算结果具有更好的泛化能力和鲁棒性.

假设最后一层隐藏层的输出矩阵为 $H_{Last} \in \mathbb{R}^{N \times m}$, 其中: N 为样本个数, m 为最后一层隐藏层的节点数, 并且输出层的输出为 $Y \in \mathbb{R}^{N \times l}$, 其中 N 为样本个数, l 为输出节点数, 本文使用偏最小二乘方法建立 H_{Last} 和 Y 之间的关系. 数学表达式为

$$Y = H_{Last}\beta_{PLS} + \xi, \quad (14)$$

其中: β_{PLS} 是由偏最小二乘方法计算得出的, ξ 是噪声.

PLS-ML-ELM的主要思路如下: 首先, 从 H_{Last} 中提取出第一成分 u_1 (u_1 是 h_1, h_2, \dots, h_m 的线性组合, 其中 $H_{Last} = [h_1 \ h_2 \ \dots \ h_m]_{N \times m}$), 并且从 Y 中提取出第一成分 v_1 (v_1 是 y_1, y_2, \dots, y_l 的线性组合), 同时, u_1 和 v_1 的相关程度必须达到最大; 然后, 建立 y_1, y_2, \dots, y_l 和 u_1 之间的回归模型. 如果回归方程达到满意的精度, 则算法终止, 否则, 从 $H_{Last} = [h_1 \ h_2 \ \dots \ h_m]_{N \times m}$ 和 $Y = [y_1 \ y_2 \ \dots \ y_l]_{N \times l}$ 中继续分别提取第 2 成分 u_2 和 v_2 , 建立回归方程, 直到回归方程达到满意的精度为止^[14]. 假设从 H_{Last} 中提取出 r 个成分 u_1, u_2, \dots, u_r , 并且建立了 y_1, y_2, \dots, y_l 和 u_1, u_2, \dots, u_r 之间的回归方程, 然后得到 y_1, y_2, \dots, y_l 和原变量 h_1, h_2, \dots, h_m 之间的回归方程, 最终得到输出权值 β_{PLS} . 具体步骤如下:

第0步 给定一个含有 N 个采样数据 (x_i, t_i) 的数据集, 其中 $x_i = [x_{i1} \ x_{i2} \ \dots \ x_{im}]^T \in \mathbb{R}^m$ 表示第 i 个采样数据有 m 维特征, $t_i = [t_{i1} \ t_{i2} \ \dots \ t_{il}]^T \in \mathbb{R}^l$ 表示目标向量, 随机生成输入层和各隐藏层的权值和偏差, 根据公式(13)计算出各隐藏层的输出矩阵, 直到得到最后一层隐藏层的输出矩阵 H_{Last} .

第1步 分别从 H_{Last} 和输出层的输出矩阵 Y 中提取出第一对成分 u_1 和 v_1 , 并且使得 u_1 和 y_1, y_2, \dots, y_l 之间的相关性最大. 而且, u_1 和 v_1 应该满足两个要求:

- 1) u_1 和 v_1 尽可能多地分别从所在变量组提取信息;
- 2) u_1 和 v_1 的相关性达到最大. u_1 和 v_1 可以由公式(15)描述

$$\begin{cases} u_1 = H_{Last}p_1 = \begin{bmatrix} h_{11} & \dots & h_{1m} \\ \vdots & \ddots & \vdots \\ h_{N1} & \dots & h_{Nm} \end{bmatrix} \begin{bmatrix} p_{11} \\ \vdots \\ p_{1m} \end{bmatrix} = \begin{bmatrix} u_{11} \\ \vdots \\ u_{N1} \end{bmatrix}, \\ v_1 = Yq_1 = \begin{bmatrix} y_{11} & \dots & y_{1l} \\ \vdots & \ddots & \vdots \\ y_{N1} & \dots & y_{Nl} \end{bmatrix} \begin{bmatrix} q_{11} \\ \vdots \\ q_{1l} \end{bmatrix} = \begin{bmatrix} v_{11} \\ \vdots \\ v_{N1} \end{bmatrix}, \end{cases} \quad (15)$$

其中: $p_1 = [p_{11} \ \dots \ p_{1m}]^T$ 和 $q_1 = [q_{11} \ \dots \ q_{1l}]^T$ 分别为 H_{Last} 和 Y 的载荷因子.

所以, 可将要求1)和2)转换为条件极值问题:

$$\begin{aligned} \max (u_1, v_1) &= (H_{Last}p_1, Yq_1) = p_1^T H_{Last}^T Y q_1, \\ \text{s.t.} \quad \begin{cases} p_1^T p_1 = \|p_1\|^2 = 1, \\ q_1^T q_1 = \|q_1\|^2 = 1. \end{cases} \end{aligned} \quad (16)$$

运用拉格朗日乘子法, 即

$$L = p_1^T H_{Last}^T Y q_1 - \lambda_1(p_1^T p_1 - 1) - \lambda_2(q_1^T q_1 - 1), \quad (17)$$

L 关于 $p_1, q_1, \lambda_1, \lambda_2$ 的偏导数分别为

$$\begin{cases} \frac{\partial L}{\partial p_1} = H_{Last}^T Y q_1 - 2\lambda_1 p_1 = 0, \\ \frac{\partial L}{\partial p_2} = Y^T H_{Last} p_1 - 2\lambda_2 q_1 = 0, \\ \frac{\partial L}{\partial \lambda_1} = -(p_1^T p_1 - 1) = 0, \\ \frac{\partial L}{\partial \lambda_2} = -(q_1^T q_1 - 1) = 0. \end{cases} \quad (18)$$

运用等式(18)可以得到方程(19)

$$2\lambda_1 = 2\lambda_2 = p_1^T H_{Last}^T Y q_1 = \langle H_{Last} p_1, Y q_1 \rangle. \quad (19)$$

注意

$$\theta_1 = 2\lambda_1 = 2\lambda_2 = p_1^T H_{Last}^T Y q_1,$$

所以 θ_1 是条件极值问题的目标函数值. 因此, 这里有

$$H_{Last}^T Y q_1 = \theta_1 p_1, \quad (20)$$

$$Y^T H_{Last} p_1 = \theta_1 q_1. \quad (21)$$

利用等式(20)-(21)得到等式(22):

$$H_{Last}^T Y Y^T H_{Last} p_1 = \theta_1^2 p_1. \quad (22)$$

观察等式(22), p_1 是矩阵 $H_{Last}^T Y Y^T H_{Last}$ 的特征矢量, θ_1^2 是对应的特征值, 并且 θ_1 是目标函数值. 根据等式(21)由 p_1 求得 q_1 , 则得分矢量可由载荷因子 p_1 和 q_1 表示

$$\begin{cases} u_1 = H_{Last} p_1, \\ v_1 = Y q_1. \end{cases} \quad (23)$$

第2步 分别建立 h_1, h_2, \dots, h_m 和 $u_1, y_1, y_2, \dots, y_l$ 和 v_1 的回归方程. 注意 $E_0 = H_{Last}$, $F_0 = Y$. 所以回归模型是

$$\begin{cases} E_0 = u_1 \alpha_1^T + E_1, \\ F_0 = v_1 \gamma_1^T + F_1, \end{cases} \quad (24)$$

其中: $\alpha_1 = [\alpha_{11} \ \alpha_{12} \ \dots \ \alpha_{1m}]^T$, $\gamma_1 = [\gamma_{11} \ \gamma_{12} \ \dots \ \gamma_{1l}]^T$ 是参数矢量, E_1, F_1 是残差矩阵. 回归参数矢量 α_1, γ_1 的最小二乘估计是

$$\begin{cases} \alpha_1 = \frac{E_0^T u_1}{\|u_1\|^2}, \\ \gamma_1 = \frac{F_0^T v_1}{\|v_1\|^2}. \end{cases} \quad (25)$$

第3步 用残差矩阵 E_1, F_1 代替 E_0, F_0 .

如果残差矩阵 F_1 中的元素趋近于0, 由第1对得分矢量建立的回归方程满足精度要求, 并且停止. 否则, 重复第1步和第2步, 并且第2对得分矢量可以表示为

$$\begin{cases} u_2 = E_1 p_2, \\ v_2 = F_1 q_2, \end{cases} \quad (26)$$

所以回归模型是

$$\begin{cases} E_0 = u_1 \alpha_1^T + u_2 \alpha_2^T + E_2, \\ F_0 = v_1 \gamma_1^T + v_2 \gamma_2^T + F_2, \end{cases} \quad (27)$$

其中: α_2 和 γ_2 是回归参数矢量, 可以表示为

$$\begin{cases} \alpha_2 = \frac{E_1^T u_2}{\|u_2\|^2}, \\ \gamma_2 = \frac{F_1^T v_2}{\|v_2\|^2}. \end{cases} \quad (28)$$

第4步 继续第2步和第3步, 直到 r 个主成分被保留. 与此同时, 余下的 $m - r$ 个成分有很小的变化, 可以被认为是噪声或者是造成最后一层隐藏层的多重共线性的原因. 而且, 偏差 E_r 和 F_r 非常小. 所以这里有

$$\begin{cases} E_0 = \sum_{i=1}^r u_i \alpha_i^T + u_2 \alpha_2^T + \dots + u_r \alpha_r^T + E_r = \\ U \alpha^T + E_r, \\ F_0 = \sum_{i=1}^r v_i \gamma_i^T + v_2 \gamma_2^T + \dots + v_r \gamma_r^T + F_r = \\ V \gamma^T + F_r, \end{cases} \quad (29)$$

而且, u_k 和 v_k 之间有着内在联系, 这个联系可以描述为^[14]

$$v_k = u_k b_k, \quad k = 1, 2, \dots, r, \quad (30)$$

所以 F_0 的等式可以重写为

$$\begin{aligned} F_0 &= V \gamma^T + F_r = \\ &\sum_{i=1}^r u_i b_i \gamma_i^T + u_2 b_2 \gamma_2^T + \dots + u_r b_r \gamma_r^T + F_r = \\ &U B \gamma^T + F_r, \end{aligned} \quad (31)$$

其中 $\hat{U} = E_0 P$, 回归方程被描述为

$$\hat{F}_0 = E_0 P B \gamma^T + F_r. \quad (32)$$

根据上面的分析, 输出层的输出权值被表述为

$$\hat{\beta}_{PLS} = P B \gamma^T, \quad (33)$$

其中: P 是成分矩阵, B 是对角矩阵, 并且 γ^T 是 F_0 的权重矩阵.

3.5 基于PLS-ML-ELM高炉透气性指数预测模型的建立(Establishing the prediction model of permeability index for blast furnace based on PLS-ML-ELM)

基于PLS-ML-ELM的高炉透气性指数预测模型的建模过程如下所示.

步骤1 确定PLS-ML-ELM的网络结构.

通过本文上述分析可知, 以风温、风量、风压、风速、顶温、顶压6个参数作为PLS-ML-ELM高炉透气性指数预测模型输入层的输入变量, 即输入层有6个输入节点. 以透气性指数作为该预测模型的输出变量, 即输出层有1个输出节点.

步骤2 建立PLS-ML-ELM网络模型.

- 1) 各隐藏层节点的激活函数为Sigmoid函数;
- 2) 各隐藏层节点个数是随机选取的, 并且输入层与隐藏层、各隐藏层之间的连接权值是任意生成的且是正交的, 各隐藏层节点的偏置是随机生成的且是正交的;
- 3) 计算各隐藏层的输出矩阵 H_i , 其中 i 表示第 i 层隐藏层;
- 4) 计算最后一层隐藏层与输出层连接权值 H_{Last} .

根据上述步骤, 可以得到基于PLS-ML-ELM高炉透气性指数预测模型结构图, 如图5所示.

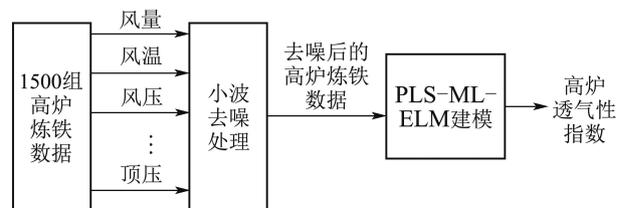


图5 高炉透气性指数预测模型

Fig. 5 The prediction model of permeability index of blast furnace

步骤3 验证PLS-ML-ELM网络模型.

如果运用该模型所得到的输出预测值与实际测量值的均方误差达到要求, 则模型建立完毕. 如果均方误差不符合要求, 则返回步骤2重新建立模型.

4 仿真实验和分析(Simulation experiment and analysis)

利用武钢某高炉的现场生产数据, 现场数据的采样间隔为1 min, 并且以1 min采样周期能很好地反映现场数据的变化. 经过上述数据去除噪声处理后, 选取1500组数据. 其中1200组数据作为模型的训练数据, 300组数据作为模型的验证数据. 由于这7个参数

的量纲各不相同且数值差异较大, 会对预测结果带来很大的波动, 进而影响预测模型的精度, 所以本文对去噪处理后的数据进行归一化处理. 然后, 分别采用 ELM, PLS-ELM^[15], ML-ELM 及 PLS-ML-ELM 模型进行仿真验证.

为了确保所建立的预测模型的性能达到最佳目标, 本文经过反复调整和分析后, ELM 隐藏层的节点数确定为 100, PLS-ELM 隐藏层的节点数为 100, ML-ELM 分别采用 2 层、3 层和 4 层隐藏层结构进行验证, ML-ELM 预测模型各层节点数在表 3 中详细给出. PLS-ML-ELM 预测模型采用 3 层隐藏层结构.

表 3 隐藏层层数不同的 ML-ELM 的各隐藏层节点数目

Table 3 Numbers of nodes of ML-ELMs which have different number of hidden layers

隐藏层 层数	第 1 隐藏层	第 2 隐藏层	第 3 隐藏层	第 4 隐藏层
2	200	100		
3	150	150	100	
3	200	200	150	
3	300	300	150	
3	300	300	200	
4	150	150	100	70
4	200	200	150	100
4	300	300	150	100
4	300	300	200	100

根据上述分析, 本文所有仿真实验是在 MATLAB R2012a 软件上运行的. 首先, 本文根据表 3 的相关信息

采用 ML-ELM 预测模型预测高炉透气性指数, 预测结果如表 4 所示. 比较表 4 给出的预测结果, 可知具有 3 层或 4 层隐藏层的 ML-ELM 预测模型的预测精度明显优于具有 2 层隐藏层的 ML-ELM 预测模型; 具有 3 层隐藏层的 ML-ELM 预测模型和具有 4 层隐藏层的 ML-ELM 预测模型在预测精度方面没有明显差别, 而且具有 3 层隐藏层的 ML-ELM 预测模型的训练时间和测试时间均小于具有 4 层隐藏层的 ML-ELM 预测模型. 本文根据上述分析, 以具有 3 层隐藏层的 ML-ELM 和提出的具有 3 层隐藏层的 PLS-ML-ELM 预测模型为代表, 与 ELM 和 PLS-ELM 预测模型进行对比. 图 6 为运用 PLS-ML-ELM 预测模型得到的高炉透气性指数预测结果与实际数据的对比图, 图 7 给出 PLS-ML-ELM 预测透气性指数的相对误差. 从图 6 和图 7 可知 PLS-ML-ELM 预测模型能够很好预测高炉透气性指数. 为了进一步比较各预测模型的预测效果, 图 8 给出根据 PLS-ML-ELM, ML-ELM, PLS-ELM 和 ELM 预测模型得到的高炉透气性指数预测结果与真实值的对比结果, 图 9 给出 PLS-ML-ELM, ML-ELM, PLS-ELM 和 ELM 预测透气性指数的相对误差的对比结果. 由图 8 可以看出, 相比于 ML-ELM, PLS-ELM, ELM, PLS-ML-ELM 预测模型具有更好的预测效果, 并且能更好的预测透气性指数的变化趋势. 从图 9 可以得知, PLS-ML-ELM 预测模型的相对误差明显小于 ML-ELM 预测模型及 ELM, PLS-ELM 预测模型的相对误差, PLS-ML-ELM 预测模型相对误差控制在 1% 之内, 而 ML-ELM, PLS-ELM 和 ELM 预测模型的相对误差分别在 2%, 4% 和 5% 之内, 由此可知 PLS-ML-ELM 具有较高的预测精度, 可以满足高炉操作现场的需求.

表 4 基于隐藏层层数不同的 ML-ELM 透气性指数预测结果比较

Table 4 Comparison of prediction results for permeability index using the ML-ELM algorithm with different numbers of hidden layers

隐藏层 层数	各隐藏层 节点数	RMSE		训练 时间/s	测试 时间/s	SD	CV
		训练	测试				
2	200, 100	0.0081	0.0067	0.3120	0.0156	0.0192	0.0240
3	150, 150, 100	0.0061	0.0049	0.4680	0.0156	0.0164	0.0202
3	200, 200, 150	0.0053	0.0044	0.6396	0.0312	0.0157	0.0197
3	300, 300, 150	0.0047	0.0035	0.9672	0.0312	0.0151	0.0190
3	300, 300, 200	0.0042	0.0030	1.0452	0.0312	0.0150	0.0189
4	150, 150, 100, 70	0.0059	0.0049	0.7156	0.0468	0.0162	0.0203
4	200, 200, 150, 100	0.0053	0.0042	1.3884	0.0624	0.0153	0.0192
4	300, 300, 150, 100	0.0046	0.0035	1.2636	0.0624	0.0151	0.0189
4	300, 300, 200, 100	0.0041	0.0030	1.5756	0.0624	0.0150	0.0188

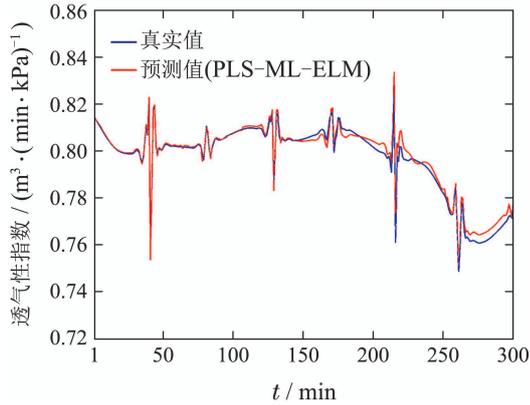


图6 PLS-ML-ELM透气性指数预测结果

Fig. 6 Prediction results of permeability index based on PLS-ML-ELM

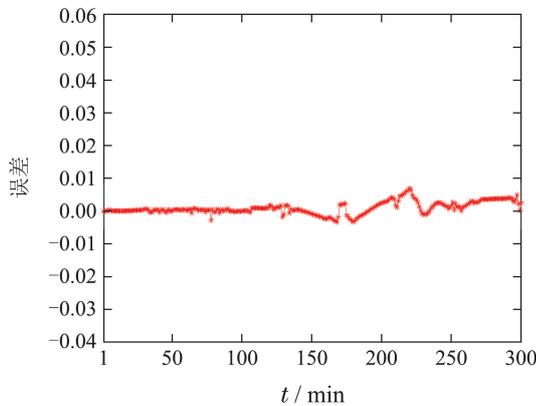


图7 PLS-ML-ELM透气性指数预测相对误差

Fig. 7 Relative prediction errors of permeability index based on PLS-ML-ELM

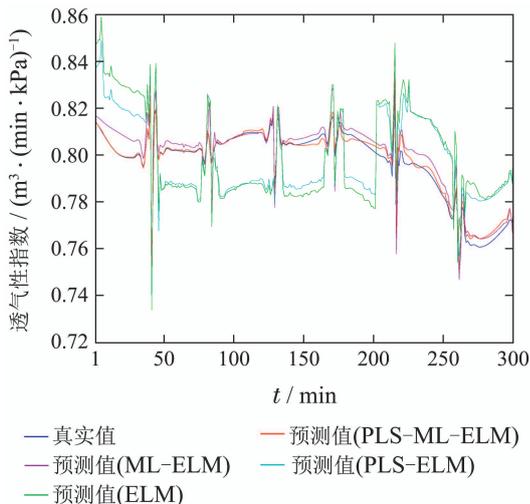


图8 不同预测模型的透气性指数预测结果

Fig. 8 Prediction results of permeability index based on different prediction models

本文采用均方根误差RMSE、标准差SD、变异系数CV及训练时间、验证时间作为各预测模型的评价标准。其中RMSE, SD和CV的表达式为

$$\left\{ \begin{aligned} \text{RMSE} &= \sqrt{\frac{\sum_{i=1}^n (X_{\text{obs},i} - X_{\text{model},i})^2}{n}}, \\ \text{SD} &= \sqrt{\frac{\sum_{i=1}^N (X_i - \bar{X})^2}{n-1}}, \\ \text{CV} &= \frac{\text{SD}}{\bar{X}} \times 100\%. \end{aligned} \right. \quad (34)$$

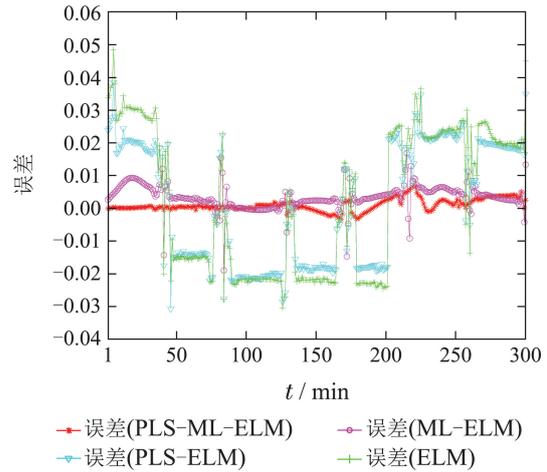


图9 不同预测模型的透气性指数相对误差

Fig. 9 Relative prediction errors of permeability index based on different prediction models

本文运用不同的预测模型,经过50次仿真,取平值得到各预测模型的预测结果,如表5所示。由表5可知,ELM预测模型的训练和测试均方误差分别为0.0265和0.0221,远大于其他3种预测模型,这表明ELM预测模型的预测精度远不及其余3种预测模型。PLS-ELM预测模型的训练和测试均方误差分别为0.0210和0.0189,ML-ELM预测模型的训练和测试均方误差分别为0.0061和0.0049,PLS-ML-ELM预测模型的训练和测试均方误差分别为0.0052和0.0039,由此可知,PLS-ML-ELM预测模型的预测精度优于ML-ELM和PLS-ELM。在训练时间和测试时间方面,如表5所示,ELM预测模型的训练和测试时间分别为0.0624 s和0.0022 s;PLS-ELM预测模型的训练和测试时间分别为0.0780 s和0.0024 s;ML-ELM预测模型的训练和测试时间分别为0.4680 s和0.0156 s;PLS-ML-ELM预测模型的训练和测试时间分别为0.5770 s和0.0312 s。从时间方面看,PLS-ML-ELM预测模型耗时最长,但是其预测精度远大于其余3种预测模型。此外,本文采用现有的BP算法建立预测模型,其训练和预测时间分别为13.9316 s和0.1275 s,相对于BP预测模型,PLS-ML-ELM预测模型的预测时间不足1 s,是十分快速的,并且是可以接受的。PLS-ML

-ELM预测模型的标准差最小,表明采用其得到的预测值最接近真实值;并且PLS-ML-ELM预测模型的变异系数最小,表明利用其得到的预测值与真实值的偏离程度最小.由此可知,PLS-ML-ELM预测模型具有更好的泛化能力和稳定性.与ELM, PLS-ELM和ML-ELM相比,本文提出的PLS-ML-ELM预测模型能够实现高炉透气性指数的精准预测,从而满足高炉现场的操作需求.并且表明在处理工业数据方面具有

更好的预测性能和鲁棒性.此外,本文还对表3中具有3层隐藏层的ML-ELM预测模型的各种情况进行改进,即采用具有3层隐藏层的PLS-ML-ELM预测模型.相应的预测结果由表6给出.表6可以进一步说明提出的PLS-ML-ELM预测模型在预测精度上优于ML-ELM预测模型,并且具有更好的泛化能力.同时,进一步说明PLS-ML-ELM预测模型能够精准预测高炉透气性指数.

表 5 透气性指数预测模型的预测结果比较

Table 5 Comparison of prediction results for permeability index prediction models

模型	各隐藏层 节点数	RMSE		训练 时间/s	测试 时间/s	SD	CV
		训练	测试				
ELM	100	0.0265	0.0221	0.0624	0.0022	0.0239	0.0298
PLS-ELM	100	0.0210	0.0189	0.0780	0.0024	0.0191	0.0240
ML-ELM	150, 150, 100	0.0061	0.0049	0.4680	0.0156	0.0164	0.0202
PLS-ML-ELM	150, 150, 100→70	0.0052	0.0039	0.5770	0.0312	0.0158	0.0198

表 6 基于PLS-ML-ELM的透气性指数预测结果比较

Table 6 Comparison of prediction results for permeability index based on PLS-ML-ELM

模型	各隐藏层 节点数	RMSE		训练 时间/s	测试 时间/s	SD	CV
		训练	测试				
PLS-ML-ELM	150, 150, 100→70	0.0052	0.0039	0.5770	0.0312	0.0158	0.0198
	150, 150, 100→60	0.0046	0.0035	0.4524	0.0312	0.0153	0.0192
	150, 150, 100→50	0.0045	0.0033	0.4992	0.0312	0.0143	0.0179
PLS-ML-ELM	200, 200, 150→110	0.0050	0.0038	0.5148	0.0312	0.0158	0.0197
	200, 200, 150→100	0.0048	0.0036	0.6396	0.0312	0.0152	0.0191
	200, 200, 150→90	0.0047	0.0035	0.6552	0.0312	0.0144	0.0180
PLS-ML-ELM	300, 300, 200→150	0.0043	0.0029	1.0920	0.0312	0.0153	0.0193
	300, 300, 200→100	0.0041	0.0027	1.1388	0.0312	0.0147	0.0184

5 结论(Conclusions)

针对高炉透气性指数传统预测模型的缺陷,本文提出了一种基于PLS-ML-ELM的透气性指数预测方法.该方法利用高炉现场生产数据,通过分析高炉机理和炼铁制度,并且考虑现场数据存在噪声,使用小波变换对数据进行去除噪声处理.最后利用ML-ELM建立预测模型,并且运用PLS优化ML-ELM最后一层隐藏层的输出矩阵,消除了最后一层隐藏层的多重共线性问题.仿真实验表明,该方法可以精准预测透气性指数,并且为高炉后续操作提供有效的决策与支持.

参考文献(References):

- [1] LIU Xiangguan, LUO Shihua, LIU Yuanhe, et al. Nonlinear mixed control for silicon contents of hot metal in BF ironmaking processes [J]. *Control Theory & Applications*, 2006, 23(3): 391 - 396. (刘祥官, 罗世华, 刘元和, 等. 高炉炼铁过程炉温的非线性混合控制 [J]. 控制理论与应用, 2006, 23(3): 391 - 396.)
- [2] LIU Hui, LI Peiran, BAO Zhejing, et al. Intelligent predictive modeling of blast furnace system [J]. *Journal of Central South University (Science and Technology)*, 2012, 43(5): 1787 - 1794. (刘慧, 李沛然, 包哲静, 等. 高炉生产过程的智能预测建模 [J]. 中南大学学报(自然科学版), 2012, 43(5): 1787 - 1794.)
- [3] ZHOU Chuandian. *Blast Furnace Ironmaking Production of Technical Manual* [M]. Beijing: Metallurgical Industry Press, 2005. (周传典. 高炉炼铁生产技术手册 [M]. 北京: 冶金工业出版社, 2005.)
- [4] GAO Chuanhou, JIAN Ling, CHEN Jiming, et al. Data-driven modeling and predictive algorithm for complex blast furnace ironmaking process [J]. *Acta Automation Sinica*, 2009, 35(6): 725 - 730. (邵传厚, 渐令, 陈积明, 等. 复杂高炉炼铁过程的数据驱动建模及预测算法 [J]. 自动化学报, 2009, 35(6): 725 - 730.)
- [5] MADADI Z, ANAND G V, PREMKUMAR A B. Signal detection in generalized gaussian noise by nonlinear wavelet denoising [J]. *IEEE Transactions on Circuits and Systems I: Regular Papers*, 2013, 60(11): 2973 - 2986.
- [6] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: theory and applications [J]. *Neurocomputing*, 2006, 70(1): 489 - 501.
- [7] LIANG N Y, HUANG G B, SARATCHANDRAN P, et al. A fast and accurate online sequential learning algorithm for feedforward

- networks [J]. *IEEE Transactions on Neural Networks*, 2006, 17(6): 1411 – 1423.
- [8] BENGIO Y. Learning deep architectures for AI [J]. *Foundations and trends in Machine Learning*, 2009, 2(1): 1 – 127.
- [9] DING S, ZHANG N, XU X, et al. Deep extreme learning machine and its application in EEG classification [J]. *Mathematical Problems in Engineering*, 2015(2015): 1 – 11.
- [10] CHAMARA L L, ZHOU H, HUANG G B. Representational learning with ELMs for big data [J]. *IEEE Intelligent Systems*, 2013, 28(6): 31 – 34.
- [11] TANG J, DENG C, HUANG G B. Extreme learning machine for multilayer perceptron [J]. *IEEE Transactions on Neural Networks & Learning Systems*, 2016, 27(4): 809 – 821.
- [12] WOLD S, TRYGG J, BERGLUND A, et al. Some recent developments in PLS modeling [J]. *Chemometrics and Intelligent Laboratory Systems*, 2001, 58(2): 131 – 150.
- [13] ABDI H. Partial least square regression (PLS regression) [J]. *Encyclopedia for Research Methods for the Social Sciences*, 2003: 792 – 795.
- [14] GELADI P, KOWALSKI B R. Partial least-squares regression: a tutorial [J]. *Analytica Chimica Acta*, 1986, 185(86): 1 – 17.
- [15] ZHAO L, WANG D, CHAI T. Estimation of effluent quality using PLS-based extreme learning machines [J]. *Neural Computing and Applications*, 2013, 22(3/4): 509 – 519.

作者简介:

苏晓莉 (1990–), 女, 博士研究生, 目前研究方向为复杂系统建模与控制、机器学习和流程工业自动化, E-mail: sxlsuccess6@163.com;

尹怡欣 (1957–), 男, 教授, 目前研究方向为复杂系统建模与控制、人工智能和流程工业自动化, E-mail: yyx@ies.ustb.edu.cn;

张森 (1971–), 女, 副教授, 目前研究方向为复杂系统建模与控制 and 现代大型高炉布料系统, E-mail: zhangsen@ustb.edu.cn.