

文章编号: 0253-374X(2014)08-1256-05

DOI: 10.3969/j.issn.0253-374x.2014.08.018

基于支持向量机的烧结能耗及性能指标预测模型

王俊凯, 乔 非, 祝 军, 倪嘉呈

(同济大学 电子与信息工程学院, 上海 201804)

摘要: 针对烧结过程中能耗和性能指标预测方法精度不高、训练时间长的问题,首先,在总结当前预测建模方法的基础上,将回归型支持向量机(support vector machine for regression, SVR)引入烧结生产系统,分析了2种建模模式;然后,给出基于SVR预测建模一般流程;最后,以某大型钢铁企业为例进行验证,并与传统的多元线性回归、反向传播(back propagation, BP)神经网络、径向基函数(radial basis function, RBF)网络和极限学习机(extreme learning machine, ELM)等预测方法在相同模式内和不同模式间进行比较。结果表明,SVR方法可快速获得理想的预测结果,在预测精度和时间效率上具有优势。

关键词: 烧结; 能耗; 性能指标; 预测模型; 回归型支持向量机

中图分类号: TP274

文献标志码: A

SVR-based Predictive Models of Energy Consumption and Performance Criteria for Sintering

WANG Junkai, QIAO Fei, ZHU Jun, NI Jiacheng

(College of Electronics and Information Engineering, Tongji University, Shanghai 201804, China)

Abstract: As to the unsatisfactory accuracy and long training time in current predictive methods of energy consumption and other performance criteria for sintering process, firstly, based on summary of existing predictive methods, support vector machine for regression (SVR) was introduced into sintering production system, and two modeling modes were proposed. Then, the general procedures of predictive modeling based on SVR were given. After that, the proposed method was verified in a scenario derived from a large-scale iron and steel enterprise, compared with other predictive methods like traditional multiple linear regression, BP neural network, RBF network and extreme learning machine within the same mode and between different modes. The result

shows that SVR method can achieve satisfied predictive results rapidly, which have advantages in prediction accuracy and time efficiency over other methods.

Key words: sintering; energy consumption; performance criteria; predictive model; support vector machine for regression

烧结是钢铁企业中耗能最大的工序之一,研究烧结工序节能降耗意义重大。同时,烧结为高炉炼铁提供主要原料,与后续生产的产量、质量以及能源消耗有着密切关系,因而研究和预测烧结矿的物理化学指标具有重要意义。

从能耗预测的角度来看,烧结工序中固体燃料消耗占到75%~80%,是烧结降耗的一个主要方向^[1]。然而目前相关研究还非常匮乏,仅有利用BPNN^[2-3],RBFN^[4]为代表的方法通过历史数据挖掘输入参数与能耗之间的关系,仍存在建模方法精度不高、模型训练时间长的问题。从性能指标预测角度来看,这些指标可以分为生产指标、质量指标和化学指标三类。这方面的研究相对较多,大致也可分为机理建模、专家系统和数据挖掘三类。机理模型往往存在较多前提假设,导致与实际脱节严重,并且在描述复杂过程和对象时无能为力。专家系统也存在模型精度有限、推广性不强的缺点。针对烧结过程具有非线性、大滞后、参数众多且耦合严重的特点,应用基于数据的方法进行相关问题的建模和分析正被广泛接受。目前,BP(back propagation)神经网络^[5]、模糊神经网络^[6]、灰色神经网络^[7]等方法已有应用。神经网络的参数确定是影响模型精度的关键问题,故又有自适应遗传神经网络^[8]等方法用于网络参数调整。然而,已有的这些研究存在以下几个问题:①预测指标偏少,未考虑固体能耗、成品率等重要指标;

收稿日期: 2013-09-09

基金项目: 国家自然科学基金(61273046, 61034004); 安徽省钢铁产业技术创新规划研究(09020203014)

第一作者: 王俊凯(1989—), 男, 博士生, 主要研究方向为可持续制造。E-mail: 2010wangjk@tongji.edu.cn

通讯作者: 乔 非(1967—), 女, 教授, 博士生导师, 主要研究方向为系统工程、控制工程、管理工程。E-mail: fqiao@tongji.edu.cn

②输入层参数考虑不足,未考虑主要的操作参数和设备参数;③训练速度慢,容易陷入局部极小,对学习参数的选择敏感.

针对以上问题,本文基于回归型支持向量机方法研究了烧结矿能耗及性能指标的预测模型,探究了2种建模模式,在获取典型钢铁企业烧结生产历史数据的基础上,从预测精度和时间效率上比较了其与多种建模方法效果的优劣,对模型预测结果进行分析探讨.

1 SVR 预测建模原理

SVR 算法的主要思想是:在传统支持向量机分类的基础上引入了不敏感损失函数(ϵ),寻找一个最优分类面使得所有训练样本离该最优分类面的误差最小^[9].

设含有 l 个训练样本的训练集样本对为 $\{(x_i, y_i), i=1, 2, \dots, l\}$, 其中, x_i 是第 i 个训练样本的输入列向量, $x_i = [x_i^1, x_i^2, \dots, x_i^d]^T$, $x_i \in \mathbf{R}^d$, \mathbf{R}^d 表示 d 维实数; $y_i \in \mathbf{R}$ 为相应的输出值.

设在高维特征空间中建立的线性回归函数为

$$f(\mathbf{x}) = \omega \Phi(\mathbf{x}) + b \quad (1)$$

式中: $f(\mathbf{x})$ 为回归函数返回的预测值; $\Phi(\mathbf{x})$ 为非线性映射函数; ω, b 为系数.

定义 ϵ 线性不敏感损失函数

$$L(f(\mathbf{x}), y, \epsilon) = \begin{cases} 0, & |y - f(\mathbf{x})| \leq \epsilon \\ |y - f(\mathbf{x})| - \epsilon, & |y - f(\mathbf{x})| > \epsilon \end{cases} \quad (2)$$

其中, $y = \{y_i | i=1, 2, \dots, l\}$ 为样本的真实值.

类似于 SVM 分类情况,引入松弛变量 ξ_i 和 ξ_i^* , 并将上述寻找 ω 和 b 的问题用数学语言描述出来, 即

$$\min \frac{1}{2} \|\omega\|^2 + C \sum_{i=1}^l (\xi_i + \xi_i^*) \quad (3)$$

$$\text{s. t. } y_i - \omega \Phi(x_i) - b \leq \epsilon + \xi_i, \quad i = 1, 2, \dots, l \\ -y_i + \omega \Phi(x_i) + b \leq \epsilon + \xi_i^*, \quad \xi_i \geq 0, \xi_i^* \geq 0$$

式中: C 为惩罚因子, C 越大表示对训练误差大于 ϵ 的样本惩罚越大, ϵ 规定了回归函数的误差要求, ϵ 越小表示回归函数的误差越小.

引入拉格朗日函数,该问题即可转换为对偶形式求解,即

$$\max_{\alpha, \alpha^*} \left[-\frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l (\alpha_i - \alpha_i^*)(\alpha_j - \alpha_j^*) K(x_i, x_j) - \sum_{i=1}^l (\alpha_i + \alpha_i^*) \epsilon + \sum_{i=1}^l (\alpha_i - \alpha_i^*) y_i \right] \quad (4)$$

s. t. $\sum_{i=1}^l (\alpha_i - \alpha_i^*) = 0, 0 \leq \alpha_i \leq C, 0 \leq \alpha_i^* \leq C$
式中, α_i 和 α_i^* 是针对式(3)2个约束条件构造的拉格朗日乘子,核函数 $K(x_i, x_j) = \Phi(x_i)\Phi(x_j)$.

SVR 模型只需确定核函数类型及其参数和惩罚因子即可,故相比 BP 神经网络而言参数个数少,无需反复凑试,容易确定.

2 基于数据预测建模模式

在烧结能耗及性能指标的预测中,常用的预测方法包括:多元线性回归、BP 神经网络、RBF(radical basis function) 网络、ELM(extreme learning machine)^[10] 等. 这些方法大致可分为两类,对于线性回归、SVR(support vector machine for regression) 等回归方法,每个回归模型只能对一个指标进行预测,即只能建立多输入单输出模型,称为 MISO 模式,如图 1a 所示. 对 MISO 模式,若要建立多个输出的预测模型,则要根据输出个数建立相应个数的 MISO 模型. 而对于 BP, RBF, ELM 等神经网络,则可以直接建立多输入多输出模型,称为 MIMO 模式,如图 1b 所示. 本文针对这 2 类建模模式分别建立相应模型并测试验证,比较模式内和模式间不同方法的预测精度和时间效率.

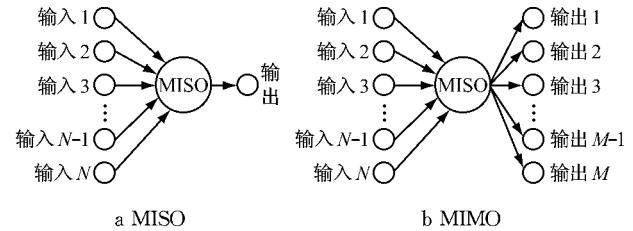


图 1 预测建模模式

Fig. 1 Modes of predictive modeling

3 基于 SVR 的预测模型建立

建模过程一般分为以下 5 步,如图 2 所示.

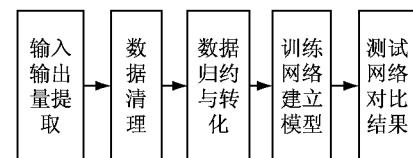


图 2 基于 SVR 预测建模过程

Fig. 2 Predictive modeling process based on SVR

步骤 1: 输入输出量提取. 影响烧结工序能耗及性能指标的因素多种多样且关系复杂. 不仅混合料

配比值、溶剂添加量等与烧结能耗和性能指标有密切联系,一些操作参数如料层厚度、点火温度和状态参数如风箱负压等也与烧结矿的产量、质量及能耗密不可分。

步骤 2: 数据清理。历史数据中异常值对模型的精度影响很大,因而离群点检测是很重要的一步。这里采用基于密度的方法进行离群点检测。

步骤 3: 数据归约与转化。将原料加入量归一化,转换成 0~1 之间的数据,所有原料配比之和为 1。将生产参数数据转化为 0~1 之间的归一化数据,以消除单位和数量级对于模型精度的影响。

步骤 4: 训练网络建立模型。将数据分为训练集与测试集,确定核函数类型及其参数和惩罚因子 C 的大小,并用训练集数据训练 SVR 模型。

步骤 5: 测试网络对比结果。用测试集测试模型精度,得到测试结果后先反归一化处理,再比较预测值与真实值之间的误差。采用 4 项指标对模型进行评价,分别是平均相对误差 e_{mean} 、最大相对误差 e_{max} 、均方差 E_{mse} 、决定系数 R^2 。

$$e_{\text{mean}} = \frac{1}{m} \sum_{k=1}^m |\tilde{y}_k - y_k| / y_k \quad (5)$$

$$e_{\text{max}} = \max\{|\tilde{y}_k - y_k| / y_k\} \quad (6)$$

$$E_{\text{mse}} = \frac{1}{k} \sum_{k=1}^m (\tilde{y}_k - y_k)^2 \quad (7)$$

$$R^2 =$$

$$\frac{\left(m \sum_{k=1}^m \tilde{y}_k y_k - \sum_{k=1}^m \tilde{y}_k \sum_{k=1}^m y_k\right)^2}{\left(m \sum_{k=1}^m \tilde{y}_k^2 - \left(\sum_{k=1}^m \tilde{y}_k\right)^2\right) \left(m \sum_{k=1}^m y_k^2 - \left(\sum_{k=1}^m y_k\right)^2\right)} \quad (8)$$

式中: m 为测试集样本个数; y_k ($k=1, 2, \dots, m$) 为第 k 个样本的真实值; \tilde{y}_k ($k=1, 2, \dots, m$) 为第 k 个样本的预测值。

4 实例研究

4.1 基于 SVR 的实例模型建立

以某年产 650 万 t 钢规模的钢铁联合企业为例,选取 2010 年 1~5 月的 122 组烧结生产历史数据进行分析建模,仿真试验在 MATLAB 2013a 平台上进行。

(1) 变量选择。输入变量分为 3 类:原料参数、状态参数和操作参数。本例中共有 26 个输入,其中原料参数包含 8 种原料的配比值,状态参数包括南北两侧各 7 个风箱的负压值,操作参数包括 2 个点火嘴的温度以及料层厚度和台车速度。输出变量也有 3

类:生产指标、质量指标和化学指标。本例中共 11 个,其中生产指标包括固体能耗(solid energy consumption, SEC) 和成品率(rate of finished products, RFP);质量指标主要指转股指数(drum index, DI);化学指标则是 7 种化学成分的质量分数(全铁 TFe、氧化亚铁 FeO、氧化硅 SiO₂、氧化钙 CaO、氧化镁 MgO、氧化铝 Al₂O₃、硫分 S 对应的质量分数分别为 w_{TFe} , w_{FeO} , w_{SiO_2} , w_{CaO} , w_{MgO} , $w_{\text{Al}_2\text{O}_3}$ 和 w_{S})和碱度(R)。因而,需建立 11 个 SVR 子模型。

(2) 数据清理。为了清除数据中的噪声,采用基于密度的方法检测离群点,获得较为干净的 113 条样本。为使结果与文献[3]具有可比性,再从 113 条样本中随机选择 105 条,其中 90 条作为训练集,15 条作为测试集。

(3) 数据归约。对原始数据进行归一化处理,规范到 [0.2, 0.8] 区间。归一化公式如下:

$$\tilde{x}_i = \frac{Q_i - Q_{\min}}{Q_{\max} - Q_{\min}} a + c \quad (9)$$

式中: Q_i 为各输入变量的第 i 个值; Q_{\max} , Q_{\min} 分别为该输入变量中的最大值和最小值; \tilde{x}_i 为归一化后该输入变量的第 i 个量; a, c 为参数,这里 $a=0.6, c=(1-a)/2$ 。

(4) 创建模型。采用 RBF 核函数,惩罚因子 C 和 RBF 核函数中的方差 g 通过交叉验证法获得。为了减少计算时间,当模型性能相同时优先选择 C 比较小的参数组合。

(5) 测试网络。SVR 模型训练完成后,利用 15 个测试样本对模型进行测试验证,反归一化,记录每个输出变量的平均相对误差、最大相对误差、均方差、决定系数等性能指标。

4.2 结果比较及分析

研究不同的模型模式下多种预测方法的性能,对于 MISO 模式,采用多元线性回归(LR)、多输入单输出 BP 神经网络(BP_MISO)作对比;对于 MIMO 模式,采用了多输入多输出 BP 神经网络(BP_MIMO)和 RBF 以及 ELM 等方法进行比较。需要说明的是,为使各试验结果具有说服力,对于其他对照模型,已通过多次凑试将模型参数调至最佳状态。例如,对于 BP_MISO 模型,每个子模型隐层节点个数为 10,隐层采用 Sigmoid 传输函数,输出层采用线性传输函数,学习速率为 0.05,训练精度为 0.001,最大训练次数为 5 000 次,采用 LM 训练函数,动量因子取 0.9。ELM 模型隐层节点数为 15,采用 Sigmoid 传输函数。RBF 网络模型中径向基函数

扩展速度取 0.3。同时,每种模型都运行 10 次,取每个指标的平均值作为最终结果。

(1) MISO 模式模型结果。如表 1 所示,从相对误差来看,SVR 预测平均相对误差均在 5% 以内,除了硫分质量分数预测误差达到 4.78% 外,其余指标均在 2% 以内,最低仅为 0.18%,较线性回归和 BP 网络预测精度高。从均方差来看,线性回归各指标存

在较大波动,原因在于某些指标与输入因素间线性关系不大;BP 和 SVR 模型中各指标均方差则相对平稳,这与两者出色的非线性拟合能力有关,其中 SVR 各指标均方差均维持在较低水平,表现出更优的性能。从决定系数也可以看出,SVR 中绝大多数指标优于其他 2 种方法。

表 1 MISO 模式模型测试结果比较

Tab.1 Comparison of testing results between MISO models

指标	$e_{\text{mean}} / \%$			$e_{\text{max}} / \%$			E_{mse}			$R^2 / \%$		
	LR	BP_MISO	SVR	LR	BP_MISO	SVR	LR	BP_MISO	SVR	LR	BP_MISO	SVR
SEC	3.09	2.88	1.54	10.07	7.10	4.68	0.0047	0.0029	0.0010	88.30	69.05	86.86
RFP	1.34	2.29	1.47	5.09	7.75	2.89	0.0152	0.0125	0.0031	48.05	28.66	57.22
DI	0.30	0.51	0.36	0.65	1.10	1.40	0.0373	0.0172	0.0148	24.46	12.01	22.23
w_{TFe}	0.38	0.27	0.18	0.94	0.72	0.39	0.0231	0.0161	0.0054	1.41	7.56	5.26
w_{FeO}	2.00	3.17	1.88	5.71	7.06	7.07	0.0422	0.0230	0.0153	0.45	0.01	27.47
w_{SiO_2}	1.42	1.87	0.80	4.49	5.28	1.58	0.0089	0.0357	0.0046	31.84	5.49	65.66
w_{CaO}	1.42	2.44	0.98	2.56	5.06	2.51	0.0219	0.0203	0.0043	48.22	4.50	48.00
w_{MgO}	2.81	2.50	1.75	7.42	4.88	4.91	0.0058	0.0167	0.0103	19.47	8.58	12.01
$w_{\text{Al}_2\text{O}_3}$	1.37	1.84	1.66	6.27	4.87	4.56	0.0016	0.0116	0.0114	44.06	35.14	61.91
w_{S}	6.97	8.04	4.78	17.27	25.10	18.75	0.0010	0.0140	0.0090	23.22	14.84	58.07
R	1.06	1.37	0.75	2.60	6.84	4.37	0.0006	0.0073	0.0033	61.64	17.07	71.95

(2) MIMO 模式模型结果比较。如表 2 所示, BP_MIMO 是文献[3] 中提出的方法, 从相对误差来看, 三者平均相对误差基本都在 3% 以内, 只有硫分质量分数相对误差偏高, 这是因其实际数值较小的

缘故。三者的误差相差不大, 这也说明 3 种神经网络都对解决该类问题具有较好的效果。从均方差和决定系数来看, ELM 略优于其他两者。

表 2 MIMO 模式模型测试结果比较

Tab.2 Comparison of testing results between MIMO models

指标	$e_{\text{mean}} / \%$			$e_{\text{max}} / \%$			E_{mse}			$R^2 / \%$		
	BP_MIMO	RBF	ELM	BP_MIMO	RBF	ELM	BP_MIMO	RBF	ELM	BP_MIMO	RBF	ELM
SEC	2.99	2.69	2.14	4.99	5.90	4.36	0.0023	0.0022	0.0016	76.40	66.69	77.27
RFP	1.56	1.51	1.17	3.62	3.58	3.30	0.0056	0.0040	0.0024	59.86	32.45	67.85
DI	0.28	0.59	0.28	1.07	1.77	0.86	0.0090	0.0340	0.0168	5.11	1.61	5.47
w_{TFe}	0.24	0.29	0.24	0.66	0.70	0.71	0.0121	0.0178	0.0126	2.51	28.33	9.11
w_{FeO}	2.14	1.49	2.77	6.31	3.82	5.24	0.0127	0.0055	0.0168	10.84	35.96	7.65
w_{SiO_2}	1.52	2.23	1.08	4.26	5.05	2.15	0.0201	0.0395	0.0079	0.50	24.44	41.38
w_{CaO}	1.14	1.57	1.04	2.69	3.49	3.15	0.0046	0.0085	0.0044	45.48	45.94	48.47
w_{MgO}	1.95	2.63	2.63	3.24	7.34	5.02	0.0093	0.0204	0.0172	41.27	6.15	26.14
$w_{\text{Al}_2\text{O}_3}$	2.12	1.74	1.43	7.16	3.57	4.92	0.0199	0.0098	0.0095	37.42	71.94	71.18
w_{S}	6.76	8.77	7.48	12.24	22.35	17.58	0.0095	0.0193	0.0117	18.61	15.96	24.39
R	1.96	1.14	1.53	5.36	2.52	3.98	0.0092	0.0029	0.0064	15.87	23.37	17.73

(3) 模式间对比分析。比较 MISO 和 MIMO 模式模型中性能表现最好的 SVR 和极限学习机模型, 如图 3 所示, 从这 2 项主要指标来看, SVR 和 ELM 都取得了较好的性能, SVR 有 8 个指标仍略优于 ELM, 例如 SEC 提升了 0.60%, TFe 质量分数提升了 0.06%, FeO 质量分数提升了 0.89%。值得一提的是, S 质量分数的预测在 6 种方法中仅 SVR 方法的平均相对误差在 5% 内。

(4) 时间效率分析。如图 4 所示, 线性回归方法训练时间和测试时间最短, 但模型效果不佳; 2 种 BP 神经网络都存在训练时间长的问题; RBF 和 ELM 时间效率都较好, 其中 ELM 网络因不用调整神经元的连接权值和神经元阈值而学习速度更快。SVR 模型虽然需要建立 11 个子模型, 但因其拓扑结构由支持向量决定, 避免了传统神经网络需要反复试凑确定网络结构的问题, 其训练时间略逊于线性回归和

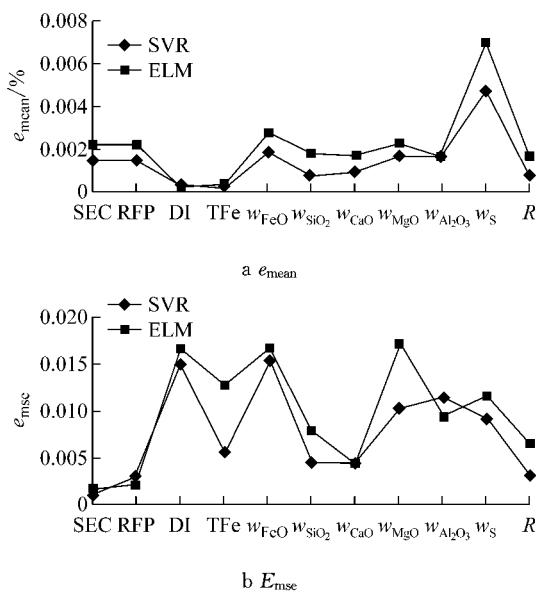


图3 SVR 和 ELM 试验结果对比

Fig.3 Comparison of results between SVR and ELM

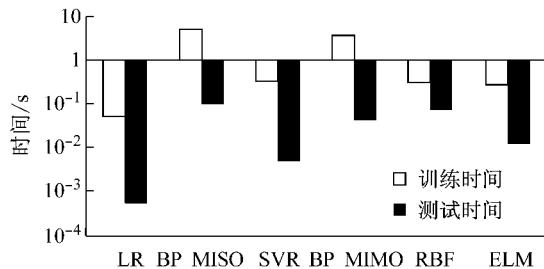


图4 6种预测模型训练时间与测试时间比较

Fig.4 Comparison of training and test time between six predictive models

ELM,而测试时间仅次于线性回归法,比 ELM 优.

5 结语

将回归型支持向量机应用于钢铁企业烧结能耗及性能指标预测模型研究中,该模型较好地实现了基于实际生产数据预测烧结能耗及其性能指标的功能,通过多种模型的比较研究表明,SVR 在该领域的应用取得了较好的模型精度和时间效率,这一模型可以与烧结配料优化模型结合,通过优化配料方案预测未来能耗值,如果能耗未能达到预期目标,再调整优化模型中的约束,生成新的优化解集,如此迭代,直至找到满意的配料比,这也是今后的一个研究方向.

参考文献:

- [1] 王社斌,许并社. 钢铁生产节能减排技术[M]. 北京:化学工业出版社,2008.
WANG Shebin, XU Bingshe. Technology of energy saving and emission reduction for iron and steel production[M]. Beijing: Chemical Industry Press, 2008.
- [2] 韩庆虹,金永龙. 人工神经网络在烧结固体燃料预测中的应用[J]. 冶金能源,2005, 24(2): 9.
HAN Qinghong, JIN Yonglong, ZHANG Junhong. Application of neural networks in the prediction of solid fuel consumption in sintering process [J]. Energy for Metallurgical Industry, 2005, 24(2): 9.
- [3] 孟辉,乔非,李莉. 基于BP神经网络的烧结能耗预测模型[J]. 机械工程师,2012(2): 45.
MENG Hui, QIAO Fei, LI Li. Predictive model of energy consumption in sintering process based on BP neural network [J]. Mechanical Engineer, 2012(2): 45.
- [4] 冯朝辉,张华,王艳红. 烧结工序能耗预测与优化研究[J]. 烧结球团,2012, 37(6): 13.
FENG Zhaoxian, ZHANG Hua, WANG Yanhong. Study on prediction and optimization of sintering process energy consumption[J]. Sintering and Pelletizing, 2012, 37(6): 13.
- [5] 王炜,陈畏林,贾斌,等. 基于线性规划和神经网络的优化烧结配料系统开发[J]. 烧结球团,2006, 31(1): 27.
WANG Wei, CHEN Weilin, JIA Bin, et al. A sinter feed proportion optimum system based on linear programming and neural network[J]. Sintering and Pelletizing, 2006, 31 (1): 27.
- [6] ER Mengjoo, LIAO Jun, LIN Jianya. Fuzzy neural networks-based quality prediction system for sintering process[J]. IEEE Transactions on Fuzzy Systems, 2000, 8(3): 314.
- [7] 鲍雅萍,马金元,宋强. 基于灰色神经网络的烧结矿碱度组合预测[J]. 控制理论与应用,2008, 25(4): 791.
BAO Yaping, MA Jinyuan, SONG Qiang. Combination forecasting of sintered ore alkalinity based on grey neural network[J]. Control Theory and Application, 2008, 25(4): 791.
- [8] 刘勇,程武山,孙鑫. 基于自适应遗传神经网络的烧结矿化学成分超前预报模型[J]. 烧结球团,2005, 6(12): 16.
LIU Yong, CHENG Wushan, SUN Xin. A prediction model of sinter chemistry based on adaptive genetic-neural network[J]. Sintering and Pelletizing, 2005, 6(12): 16.
- [9] Shevad S K, Keerthi S S, Bhattacharyya C, et al. Improvements to the SMO Algorithm for SVM Regression[J]. IEEE Transactions on Neural Networks, 2000, 11(5): 1188.
- [10] Sun Z L, Choi T M, Au K F, et al. Sales forecasting using extreme learning machine with applications in fashion retailing [J]. Decision Support Systems, 2008, 46: 411.