

生物氧化提金中基于 PSO-LSSVM 的氧化还原电位建模研究

李 伟, 南新元*, 吴 琼
(新疆大学 电气工程学院, 乌鲁木齐 830047)

摘 要: 生物氧化预处理过程是一个复杂非线性的耦合过程, 该过程关键参数氧化还原电位通常难以准确检测。为了预估该参数, 将 PSO 算法与 LSSVM 相结合构建生物氧化预处理过程氧化还原电位预估模型。该模型采用改进的 PSO 算法优化 LSSVM 模型参数, 克服了参数选择的盲目性和耗时, 具有学习速度迅速、预测精度较高以及泛化能力强的优点。以新疆某金矿的实际数据进行仿真研究, 结果表明: 改进的 PSO-LSSVM 方法建立的模型的预测值与实测值拟合较好, 对于生物氧化预处理过程的关键参数氧化还原电位的预估有一定的指导意义。

关键词: 生物氧化预处理; 粒子群参数优化; 最小二乘支持向量机; 氧化还原电位; 建模
中图分类号: TP273 **文献标识码:** A **文章编号:** 1004-0676(2014)04-0060-05

Oxidation-reduction Potential Modeling Based on PSO and LSSVM for Gold Extraction by Bio-oxidation

LI Wei, NAN Xinyuan*, WU Qiong
(School of Electrical Engineering, Xinjiang University, Urumqi 830047, China)

Abstract: The biological oxidation pretreatment process is a complex nonlinear and coupling process. It is difficult to measure the oxidation reduction potential by using conventional physical sensors. A soft sensor model based on the combination of PSO and LSSVM was established to estimate the parameters. To overcome the time consuming and blindness in parameter selection of traditional LSSVM model, PSO algorithm was employed to optimize the parameters of LSSVM soft sensor model. The model had a fast learning speed, high precision and well generalization ability. The results obtained from the real data of a gold mine in Xinjiang for simulation showed that predicted values of PSO-LSSVM model fitted measured values well, suggesting that the model established in our study is useful to measure the key parameters for biological oxidation pretreatment process of oxidation-reduction potential.

Key words: bio-oxidation pretreatment; particle swarm parameter optimization (PSO); least squares support vector machine (LSSVM); oxidation reduction potential (ORP); modeling

生物氧化预处理过程是黄金生产中的一种重要工艺过程, 作为该过程控制的关键参数(氧化还原电位)难以准确检测。氧化还原电位反映着矿浆的氧化程度, 也在一定程度上反映了资源利用率和提金率。生物氧化预处理过程是多个参数相互耦合、非线性的过程, 工作机理十分复杂, 这些复杂性和不确定

性给各种工艺流程的控制和氧化还原电位的测定带来极大挑战。因此, 需要建立氧化还原电位的预测模型, 对于生物氧化预处理过程操作和控制优化具有重要的指导意义^[1-3]。

LSSVM(最小二乘支持向量机)是 Suykens J A K 于 1999 年提出的一种新型支持向量机, 它是 SVM 的

收稿日期: 2014-04-02

基金项目: 新疆维吾尔自治区自然科学基金项目(2012211A004)。

第一作者: 李 伟, 男, 硕士研究生, 研究方向: 软测量建模与工业控制优化。E-mail: liwei0306@hotmail.com

*通讯作者: 南新元, 男, 硕士, 副教授, 研究方向: 计算机控制技术和现场总线。E-mail: nxyxd@sina.com

一种改进, 通过求解线性方程组得出模型参数, 避免传统 SVM 中求解二次规划问题的复杂性^[4]。LSSVM 具有泛化能力优异、拟合精度高等优势, 已成功用于模式识别、回归分析等领域^[5-7]。LSSVM 模型中的参数直接影响了拟合能力和泛化性能。高丙朋等^[8]通过将迭代算法和 LSSVM 相结合, 对生物氧化提金过程中的工艺参数进行优化, 避开了外部环境中的干扰因素对复杂生物氧化预处理过程数学建模的影响, 且计算量小, 易于掌握, 但在该方法中是以固定的提金率作为目标值来进行参数优化。

由于生物氧化预处理过程反应机理复杂, 导致机理建模极其困难。随着生产企业自动化水平的大幅度提升, 生产运行中积累了大量的运行数据, 为生物氧化预处理过程中 ORP(氧化还原电位)预测模型的建立提供了有利条件。本文在研究 PSO(粒子群参数优化)和 LSSVM 的基础上提出了一种改进的 PSO-LSSVM 的建模方法, 参数优化采用动态调整策略。并以实际生产数据对 ORP 进行建模预测, 结果表明, 方法可行有效。

1 最小二乘支持向量机

最小二乘支持向量机算法是支持向量机算法的一种改进, 通过构造损失函数求解线性方程组, 进而取代了支持向量机算法的二次优化求解问题, 该方法的基本原理表述如下:

设样本为 n 维向量, 给定了 l 个样本数据 $(x_1, y_1), \dots, (x_k, y_k), \dots, (x_l, y_l) \in R^n \times R$ 。其中, x_k 为样本输入, y_k 为样本输出, LSSVM 利用非线性映射 $\varphi(x)$ 将样本从原空间 R^n 映射到一个高维特征空间 $\varphi(x)=(\varphi(x_1), \varphi(x_2), \dots, \varphi(x_l))$, 在高维特征空间中采用结构风险最小化原则构造最优决策函数, 并利用原空间的核函数取代高维特征空间的点积运算以避免复杂运算, 从而将非线性函数估计问题转化为高维特征空间中的线性函数问题。设构造的最优决策函数具有下式的形式:

$$y(x)=\omega^T \cdot \varphi(x)+b \quad (1)$$

求解目的就是利用结构风险最小化原则, 寻找参数 ω 和 b , 使得对于样本外的输入 x , 有 $|y - \omega^T \varphi(x_k) - b| \leq \varepsilon$ 。等价于求解下式的优化问题:

$$\min J = \|\omega\|^2 / 2 + c \cdot R_{cpm} \quad (2)$$

其中, $\|\omega\|^2$ 是置信区间, 控制模型的复杂程度; $c > 0$ 是误差惩罚参数, 表示函数的平滑度和允许误差大于 ε 的数值之间的折衷; R_{cpm} 为经验风险, 即 ε 不敏感损失函数。选择不同的损失函数便可以构造

不同的支持向量机。常用的损失函数有线性 ε 损失函数、二次 ε 损失函数、Huber 损失函数等。

标准支持向量机优化目标中的损失函数为误差 ξ_i (允许错分的松弛变量), 而最小二乘支持向量机则选择了误差 ξ_i 的二次项, 其优化问题为:

$$\min_{\omega, b} J = \frac{1}{2} \omega^T \omega + \frac{1}{2} c \sum_{i=1}^l \xi_i^2 \quad (3)$$

$$s.t. y_i = \omega^T \cdot \varphi(x_i) + b + \xi_i, \quad i=1, 2, \dots, l \quad (4)$$

式(4)为优化函数的约束条件, 来源于式(1), 式(3)是优化函数, 为了求解优化函数的最小值, 式(3)和式(4)联立, 进而用拉格朗日法求解上述优化问题。根据目标函数和约束条件建立 Lagrangian 函数, 用拉格朗日法求解上述优化问题:

$$L(\omega, b, \xi, \alpha) = \frac{1}{2} \omega^T \cdot \omega + \frac{1}{2} c \sum_{i=1}^l \xi_i^2 - \sum_{i=1}^l \alpha_i (\omega^T \cdot \varphi(x_i) + b + \xi_i - y_i) \quad (5)$$

根据优化条件:

$$\frac{\partial L}{\partial \omega} = 0, \frac{\partial L}{\partial b} = 0, \frac{\partial L}{\partial \xi} = 0, \frac{\partial L}{\partial \alpha} = 0 \quad (6)$$

可得:

$$\omega = \sum_{i=1}^l \alpha_i \varphi(x_i), \sum_{i=1}^l \alpha_i = 0, \alpha_i = c \xi_i \quad (7)$$

$$\omega^T \cdot \varphi(x_i) + b + \xi_i - y_i = 0$$

定义核函数 $K(x, x_i) = \Psi(x) \cdot \Psi(x_i)$ 代替非线性映射, 其中核函数是满足 Mercer 条件的任意对称函数。根据式(7), 求解的优化问题转化为求解线性方程:

$$\begin{bmatrix} 0 & 1 & \dots & 1 \\ 1 & K(x_1, x_1) + \frac{1}{c} & \dots & K(x_1, x_l) \\ \vdots & \vdots & \ddots & \vdots \\ 1 & K(x_l, x_1) & \dots & K(x_l, x_l) + \frac{1}{c} \end{bmatrix} \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_l \end{bmatrix} = \begin{bmatrix} 1 \\ y_1 \\ \vdots \\ y_l \end{bmatrix} \quad (8)$$

最后得最小二乘支持向量机的估计函数为:

$$f(x) = \sum_{i=1}^l \alpha_i K(x, x_i) + b \quad (9)$$

选择不同的核函数, 即可构造不同的支持向量机。LSSVM 常用的核函数有径向基核函数、多项式核函数、感知器核函数。径向基函数具有较强的泛化能力, 本文选用径向基核函数:

$$K(x, x_i) = \exp \left[-\frac{\|x - x_i\|^2}{2\delta^2} \right]$$

此时 LSSVM 参数选择问题变为核函数参数 δ 和正则化参数 γ 的优化组合, 为了得到具有较好预

测效果的模型,有必要对这些参数进行优化调整。

2 粒子群算法

粒子群优化算法是计算智能领域的一种群体智能的优化算法,它源于对鸟群捕食行为的研究,最早是由Kenney与Eberhart提出。在PSO中,每个优化问题的解都看作搜索空间其中的一个粒子,所有的粒子都有一个由适应度函数决定的适应值,每个粒子都有一个速度决定它们飞翔的方向和距离,粒子们就追随当前的最优粒子在解空间中搜索最优解。

PSO 算法首先初始化一群随机粒子,通过多次迭代搜索最优解。在每一次的迭代中,粒子通过跟踪 2 个极值来更新自己:一个是每个粒子在历代搜索中自身所达到的最优解,即个体极值 P_{best} ;另一个是粒子群所有粒子在历代搜索中所达到的最优解,即全局极值 g_{best} 。群体中第 i 个粒子在 n 维空间的位置表示为 $x_i=(x_{i1},x_{i2},\dots,x_{in})$,其速度 $v_i=(v_{i1},v_{i2},\dots,v_{in})$,第 i 个粒子的个体极值为 $P_i=(P_{i1},P_{i2},\dots,P_{in})$,整个粒子群的的全局极值 $g_i=(g_1,g_2,\dots,g_n)$ 。在找到这 2 个极值时,粒子利用下式来更新自己的速度和位置。

$$v_i(t+1)=\omega v_i(t)+c_1 rand_1(P_{best}-x_i(t)) + c_2 rand_2(g_{best}-x_i(t)) \quad (10)$$

$$x_i(t+1)=x_i(t)+v_i(t+1) \quad (11)$$

式中, c_1 和 c_2 为学习因子,通常在(0,2)之间取值; $rand_1$ 和 $rand_2$ 为(0,1)之间的随机数; ω 为惯性权重。

粒子群优化算法中,惯性权重 ω 的选择很关键,较大的 ω 有较好的全局搜索能力,较小的 ω 则局部搜索能力较强。因此,惯性权重应该随着迭代次数的增加而不断减小,从而使得粒子群算法在初期全局收敛能力较强,而后期则局部收敛能力较强^[9-10]。

3 改进的 PSO 优化 LSSVM 算法

由标准 PSO 算法的模型可以得知,粒子容易陷入未成熟收敛。而粒子的更新速度主要由惯性权重 ω 和自我学习因子 c_1 和社会学习因子 c_2 来决定的。通过调整 ω 的大小,使搜索性能在两者间达到一个平衡。Eberhart 和 Shi 提出了一种线性权重递减策略^[11],可以显著改善算法的收敛性能,但不能很好地适应具有复杂非线性变化的优化问题。在基本 PSO 算法中, c_1 和 c_2 是固定的,这将导致粒子群过早收敛于局部极值。我们期望的是在算法前期, c_1

较大, c_2 较小,自我学习能力较大,社会学习能力较小,可以使粒子在整个搜索空间飞行;在算法后期, c_1 较小, c_2 较大,自我学习能力较小,社会学习能力较大,寻找到全局最优解。

因此,可以通过调整 ω 和 c_1 、 c_2 的值来提高算法的收敛性能,获得全局最优解。本文采用一种自适应粒子群优化 LSSVM 算法^[12-14],惯性权重由适应值来动态调整;学习因子可以构造简洁的线性函数自适应改变,从而对最小二乘向量机中的核函数参数 δ 和正则化参数 γ 进行自主调整、寻优以达到最佳组合。

由个体适应值动态调整惯性权重的算法如下:

当 $f(x_i) < f(g_{best})$ 时:

$$\omega = \omega - (\omega - \omega_{min}) \left| \frac{f(x_i) - f(p_i)}{f(g_{best}) - f(p_i)} \right| \quad (12)$$

当 $f(g_{best}) < f(x_i) < f(p_i)$ 时:

$$\omega = \omega_{min} + (\omega_{max} - \omega_{min}) \frac{1}{1 + \exp(-t_{max}/t)} \quad (13)$$

其中, ω_{max} 为最大惯性权重,一般为 0.9; ω_{min} 为最小惯性权重,一般为 0.4; t_{max} 为最大迭代次数, t 为迭代次数。

自适应学习因子 c_1 和 c_2 的调整公式如下所示:

$$\begin{cases} c_1 = -2 \times t / t_{max} + 2.5 \\ c_2 = 2 \times t / t_{max} + 0.5 \end{cases} \quad (14)$$

这种自适应 PSO 算法在保证了解法的收敛性的前提下,收敛速度得以加快,求解精度也得以提高。因此,基于自适应 PSO 的 LSSVM 算法流程为:

(1) 初始化粒子群,设置 PSO 参数,包括:种群规模,惯性权重的最大值与最小值,最大迭代次数 t_{max} ,初始化惯性权重,粒子的初始速度和位置。

(2) 利用训练样本进行 LSSVM 回归训练,并用适应度函数 $f(x)$ 计算粒子的适应度值:

$$f(x_i) = \frac{1}{n} \sum_{i=1}^n (f_i - y_i)^2 \quad (15)$$

(3) 针对每个粒子,在当前适应值 $f(x_i)$ 和历史最优位置适应值 $f(p_i)$ 之间比较,若 $f(x_i) < f(p_i)$ 则令 $p_i = x_i$,并调整惯性权重;在群体所有粒子当前适应值 $f(x_i)$ 与群体最优位置适应值 $f(g_{best})$ 之间比较,若 $f(x_i) < f(g_{best})$,则全局最优解 $g_{best} = x_i$,并调整惯性权重。

(4) 根据改进粒子群优化算法的进化方程将粒子的速度和位置进行更新。

(5) 判断是否满足结束条件,即是否达到最大迭代次数或者粒子的适应度值达到了设定的要求,若满足条件,则停止迭代,否则返回步骤 2。

(6) 利用寻优得到的全局最优值对 LSSVM 模型运算。整个流程示意图如图 1 所示。

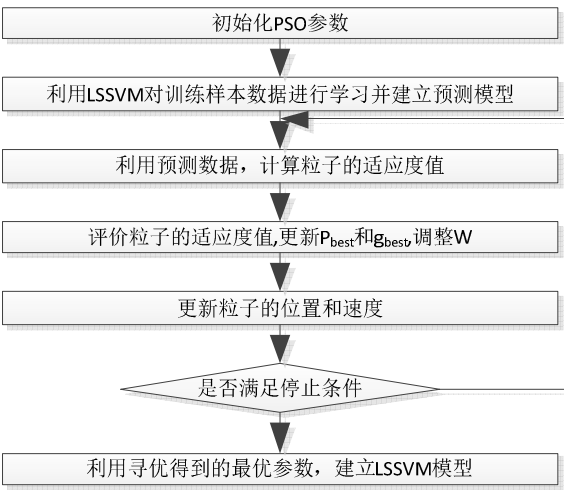


图 1 PSO-LSSVM 算法流程图
Fig.1 Flowchart of PSO-LSSVM algorithm

4 基于 PSO-LSSVM 的氧化还原电位建模与仿真

4.1 实验背景

生物氧化预处理过程是细菌等微生物与包裹在金矿石外面的含砷、含硫等金矿石氧化反应，使金矿石暴露，进而方便黄金提取出来的一种工艺流程。生物氧化预处理过程是多个参数相互耦合的过程，工作机理十分复杂，这些复杂性和不确定性给氧化还原电位的测定带来极大困难。因此针对其氧化预处理过程建立模型对关键变量进行有效预测对于实际生产有非常重要的意义。

在分析工艺流程的基础上，经综合分析比较后，

选择温度、pH、充气量、矿浆浓度作为辅助变量输入，来估算所要求的氧化还原电位值。氧化还原电位仿真实验采用的数据都可通过生产现场的生物氧化槽内的温度传感器、pH 值传感器、充气量传感器、密度传感器获得。

4.2 仿真结果与分析

本次实验共采集实际生产中的 110 组样本数据，样本数据会存在随机误差。对数据进行预处理后，剩余 96 组数据，其中 64 组用于模型训练，其余 32 组用于模型测试。

在 Matlab2010a 仿真环境下，粒子群规模取 30；解空间为 2 维，分别表示正规化参数 γ 和核函数参数 δ ，核函数参数 δ 的取值范围为[0.01,10]，正规化参数 γ 的取值范围为[0,1000]，最大迭代次数 $t_{\max}$ 取 1000。几种不同建模方法性能如表 1 所示。

表 1 建模方法性能比

Tab.1 Performance comparison of modeling method

算法	训练误差	预测误差
SVM	0.0098	0.2587
LSSVM	1.3761×10^{-3}	0.1094
PSO-LSSVM	5.0761×10^{-4}	0.0425

通过上文中所述的 PSO 优化 LSSVM 算法步骤的模型训练后，得到最终优化的正规化参数 γ 为 153，核函数参数 δ 为 0.2。图 2 为 PSO-LSSVM 模型训练结果，图 3 为实际值和 PSO-LSSVM 模型预测值的拟合曲线，图 4 为用 LSSVM 模型预测输出。从图 3 可看出，预测值和实际值拟合的较好，训练均方根误差 0.00050761，测试均方根误差 0.0425。

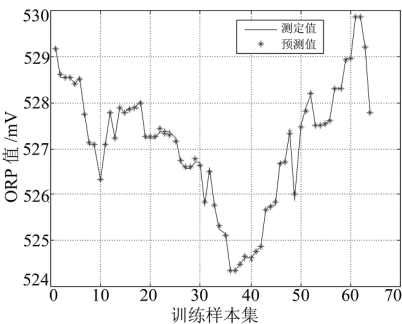


图 2 PSO-LSSVM 模型训练图
Fig.2 Training result of
PSO-LSSVM model

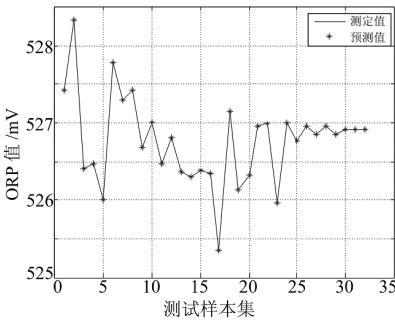


图 3 PSO-LSSVM 模型预测图
Fig.3 Prediction result of
PSO-LSSVM model

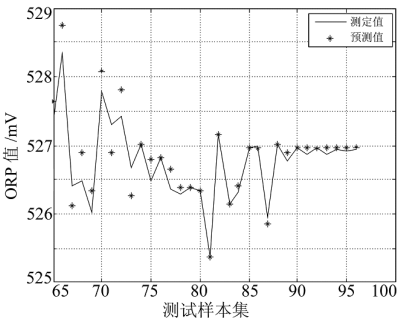


图 4 LSSVM 模型预测图
Fig.4 Prediction result of
LSSVM model

对比图 3 和图 4 两种模型预测的结果, 图 3 预测值和真实值的拟合程度优于图 4。因此, 可得知 PSO-LSSVM 模型预测能力较强, 能够较准确地反映氧化还原电位的变化情况, 从而进一步说明了利用改进 PSO 优化 LSSVM 模型参数的优越性。

5 结论

LSSVM 是一种比较常用的机器学习算法, 具有较强的泛化能力和回归精度, 但是实现的前提是要选取到最优的参数, 现多采用的是建立在经验和试验基础之上的参数选择法, 缺乏理论指导。本文提出一种自适应粒子群优化算法对 LSSVM 的参数进行寻优, 并运用于生物氧化预处理过程中氧化还原电位的预测。结果表明, 基于改进 PSO 的 LSSVM 优化参数的模型预测精度较高, 具有较强的泛化能力, 模型的性能优于 LSSVM 模型。该建模方法为生物氧化预处理过程关键参数的测量提供了新途径, 具有较高的工程实际应用意义。

参考文献:

- [1] 孙理鑫. 难处理金矿生物氧化过程及其反应器的工程基础研究[D]. 上海: 华东理工大学, 2012.
- [2] 韩晓光, 郭普金, 具滋范. 生物氧化提金技术工业生产实践[J]. 黄金, 2006, 27(11): 38-41.
Han X, Guo P, Ju Z. Industrial practice of bio-oxidization gold recovery technique[J]. Gold, 2006, 27(11): 38-41.
- [3] 瞿建威, 杨洪英, 巩恩普, 等. 细菌氧化预处理含砷难处理金矿的研究进展[J]. 贵金属, 2005, 26(1): 66-70.
Zi J, Yang H, Gong E, et al. A review on biooxidation pretreatment of arsenic-bearing refractory gold ores[J]. Precious Metals, 2005, 26(1): 66-70.
- [4] Suykens J A K, Vandewalle J. Least squares support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.
- [5] Li H, Guo S, Zhao H, et al. Annual electric load forecasting by a least squares support vector machine with a fruit fly optimization algorithm[J]. Energies, 2012, 5(11): 4430-4445.
- [6] 李轩, 李建宏, 何玉珠, 等. 基于遗传优化最小二乘支持向量机的导引头故障诊断[J]. 计算机应用研究, 2011, 28(3): 899-902.
Li X, Li J, He Y, et al. Intelligent fault diagnosis for seeker based on least squares support vector machine with genetic algorithm[J]. Application Research of Computers, 2011, 28(3): 899-902.
- [7] Zhu B, Wei Y. Carbon price forecasting with a novel hybrid ARIMA and least squares support vector machines methodology[J]. Omega, 2013, 41(3): 517-524.
- [8] 高丙朋, 南新元, 魏霞. 基于迭代 LS-SVM 生物氧化提金预处理工艺参数优化算法的研究[J]. 贵金属, 2012, 33(2): 40-43.
Gao B, Nan X, Wei X. Study on optimization algorithm of process parameters for gold extraction pretreatment based on iterative LS-SVM[J]. Precious Metals, 2012, 33(2): 40-43.
- [9] Kennedy J, Eberhart R. Particle swarm optimization[C]//Proceedings of IEEE international Conference on Neural Networks. Perth: Piscataway, 1995, 4(2): 1942-1948.
- [10] Van den Bergh F, Engelbrecht A P. A study of particle swarm optimization particle trajectories[J]. Information Sciences, 2006, 176(8): 937-971.
- [11] Shi Y, Eberhart R C. Parameter selection in particle swarm optimization[C]//Evolutionary Programming VII, Springer Berlin Heidelberg, 1998: 591-600.
- [12] 方科, 黄元亮, 刘新东. 基于自适应 PSO 算法的 LS-SVM 牵引变压器绝缘故障诊断模型[J]. 电力自动化设备, 2011(3): 85-89.
Fang K, Huang Y, Liu X. Insulation fault diagnosis model based on adaptive PSO and LS-SVM for traction transformer[J]. Electric Power Automation Equipment, 2011(3): 85-89.
- [13] 李威霖, 傅攀, 张尔卿, 等. 基于粒子群优化 LS-SVM 的车刀磨损量识别技术研究[J]. 计算机应用研究, 2014, 31(4): 1094-1097.
Li W, Fu P, Zhang E, et al. Application of particle swarm optimization-least square support vector machine in tool wear monitoring[J]. Application Research of Computers, 2014, 31(4): 1094-1097.
- [14] 刘伟, 周育人. 一种改进惯性权重的 PSO 算法[J]. 计算机工程与应用, 2009, 45(7): 46-48.
Liu W, Zhou Y. Modified inertia weight particle swarm optimizer[J]. Computer Engineering and Applications, 2009, 45(7): 46-48.