

文章编号: 2095-2163(2022)09-0105-05

中图分类号: TK229.1

文献标志码: A

基于 GA-SVR 的循环流化床锅炉床温预测

盛家豪¹, 钱进^{1,2}, 王一桂³, 黄凤启⁴

(1 贵州大学 电气工程学院, 贵阳 550025; 2 重庆大学 动力工程学院, 重庆 400044;

3 中国电建集团贵州工程有限公司, 贵阳 550025; 4 国投盘江发电有限公司, 贵州 盘州 553000)

摘要: 床温是影响循环流化床(CFB)锅炉经济、安全、可靠工作的重要运行参数之一。本文从小样本数据容量出发,利用遗传算法的全局寻优能力和支持向量回归的泛化能力,对某 300 MW 现役循环流化床锅炉设计了床温预测模型。通过分析循环流化床运行机理,选择床温特性模型的输入变量,然后利用遗传算法对支持向量回归算法的惩罚系数、径向核参数和损失函数进行全局寻优,构建 CFB 锅炉床温 GA-SVR 预测模型。将 GA-SVR 模型与 BP 神经网络、BP-SVR 和卷积神经网络三种预测模型分别进行对比,分析各模型的性能指标,结果表明在小样本数据情况下,GA-SVR 模型预测误差更小,响应更迅速,对 CFB 锅炉的运行和控制具有指导作用。

关键词: 循环流化床锅炉; 床温; 遗传算法; 支持向量回归

CFB boiler bed temperature prediction based on Genetic Algorithm and Support Vector Regression

SHENG Jiahao¹, QIAN Jin^{1,2}, WANG Yigui³, HUANG Fengqi⁴

(1 College of Electrical Engineering, Guizhou University, Guiyang 550025, China; 2 College of Power Engineering, Chongqing University, Chongqing 400044, China; 3 Power China Guizhou Engineering Co., Ltd., Guiyang 550025, China; 4 Guotou Panjiang Power Generation Co., Ltd., Panzhou Guizhou 553000, China)

[Abstract] Bed temperature is one of the important parameters affecting the operating efficiency, safety and reliability of the Circulating Fluidized Bed (CFB) boilers. Based on the small capacity sample data, combined with the global optimization ability of Genetic Algorithm (GA) and the generalization ability of Support Vector Regression (SVR), the bed temperature prediction model is designed for a 300 MW circulating fluidized bed boiler in service. By analyzing the operation mechanism of this specific boiler, the input variables of the bed temperature characteristic model are selected, and then the genetic algorithm is used to optimize the penalty coefficient, radial kernel parameters and loss function of the support vector regression algorithm to construct the GA-SVR prediction model of CFB boiler bed temperature. Comparison is conducted with BP neural network, BP-SVR and Convolution Neural Network(CNN). The results show that the GA-SVR model has smaller prediction error and faster response in the case of small sample data, which has a guiding role for the operation and control of CFB boilers.

[Key words] Circulating Fluidized Bed boiler; bed temperature; Genetic Algorithm; Support Vector Regression

0 引言

中国“双碳”目标及实现能源结构转型的提出,使传统化石燃料发电行业面临严峻挑战。循环流化床(Circulating Fluidized Bed, CFB)锅炉具有燃料适应性广、可掺烧生物质、降负荷比大、碳排放强度及其他有害污染物排放量相对较低等优点^[1]。在未来新能源、可再生能源为主的能源供应结构中,必将作为深度调峰、旋转备用的主力,确保 CFB 锅炉经济可靠运行研究有重要意义。

CFB 锅炉床温是影响 CFB 锅炉经济、高效、安全运行的重要参数之一。床温过高,会导致锅炉的热传导增快,受热面温度过高,引起高温结渣腐蚀和超温爆管等问题,同时也会导致 NO_x 排放量增加;而床温过低将引起锅炉热效率降低,锅炉汽水参数不达标,底渣和飞灰中可燃物含量增高,从而影响锅炉的经济运行^[2]。因此,一些国内外学者对 CFB 锅炉的床温特性进行了大量的建模研究。基于机器学习和神经网络等技术建立的分布参数、时变、非线性、多变量紧密耦合控制对象的动态模型具有较好的自适应能力和泛化能力,在电站锅炉运行及污染

基金项目: 贵州省科技支撑计划项目([2020]2Y040)。

作者简介: 盛家豪(1998-),男,硕士研究生,主要研究方向:热能利用与节能工程; 钱进(1965-),男,副教授,硕士生导师,主要研究方向:热力过程多尺度数值模拟与仿真技术。

通讯作者: 钱进 Email: jqian@gzu.edu.cn

收稿日期: 2022-02-08

物排放控制建模方面得到了广泛的应用。尤海辉等人^[3]提出基于减法聚类的模糊神经网络模型预测循环流化床入炉垃圾热值。钱虹等人^[4]实现了基于深度循环神经网络的 SCR 烟气脱硝系统出口 NO_x 排放预测。白建云等人^[5]构建了基于 BP 神经网络的 NO_x 质量浓度在线测量模型。张妍等人^[6]利用粒子群优化粒子滤波算法实现了循环流化床床温预测。但上述方法存在所需基础数据量大、模型复杂、训练和计算时间长、精确度不高等问题。对于小样本数据,段萌等人^[7]提出基于卷积神经网络的小样本图像识别方法。王琦等人^[8]建立基于 BP-SVM 的 SO₂ 特性模型,实现对 SO₂ 浓度的快速、准确预测。

本文在分析 CFB 锅炉运行机理的基础上,从小样本容量数据预测出发,结合已有数据,选择合适的模型输入变量,通过 GA 算法对 SVR 的重要参数进行全局寻优,构建 CFB 锅炉床温预测模型,将 GA-SVR 模型与 BP 神经网络、BP-SVR 和 CNN 模型分别进行性能对比分析,以实现某 300 MW CFB 锅炉的床温预测。

1 遗传算法和支持向量回归原理

1.1 遗传算法

遗传算法(Genetic Algorithm, GA)源自对自然界物种繁衍进化的数学和计算机模拟研究。遗传算法本质上是一种高效、并行、全局搜索的方法,能在搜索过程中自动获取和积累有关搜索空间的知识,并自适应地控制搜索过程以求得最佳解^[9]。其一般步骤为^[10]:

(1) 编码及初始化。将所需求解的数据集编译成遗传空间的基因型串结构数据集,并随机生成原始种群。

(2) 适应度评估。选择一个评判基因优劣的指标函数作为遗传算法的适应度,计算种群所有个体的适应度。

(3) 个体筛选。通过轮盘赌法保留评价指标高的个体。

(4) 遗传操作。对经筛选保留的各个体基因进行交叉和变异操作,产生子代个体,形成新一代种群。

(5) 准则判断。根据设定的准则、如迭代次数和指标大小判断 GA 算法是否继续执行。若执行算法,则返回步骤(2);否则进行全局搜索,找出所有子代中适应度最高的个体作为算法的最优解。

1.2 支持向量回归

支持向量机(Support Vector Machine, SVM)是一类按监督学习方式对数据进行二元分类的广义线性分类器,而支持向量回归是在 SVM 的基础上,针对回归问题而衍生出来的算法。对于样本 (x, y) ,传统回归模型的预测误差为模型预测值与数据真实值的差值,而 SVR 能容忍预测值 $f(x)$ 与真实值 y 之间最多有 ε 的偏差,即仅当 $f(x)$ 与 y 之间的差别绝对值大于 ε 时才计算损失,于是 SVR 问题可形式化为式(1)^[11]:

$$\min_{\omega, b, \xi_i, \xi_i^*} \frac{1}{2} \|\omega^2\| + C \sum_{i=1}^m (\xi_i + \xi_i^*) \quad (1)$$

约束条件为:

$$f(x_i) - y_i \leq \varepsilon + \xi_i$$

$$y_i - f(x_i) \leq \varepsilon + \xi_i$$

$$\xi_i \geq 0, \xi_i^* \geq 0, i = 1, 2, \dots, m$$

其中,式(1)的第一项为最小化结构风险,第二项为经验风险, C 是对二者进行折中的常数,设为惩罚系数; ξ_i, ξ_i^* 为松弛变量; ε 为损失边界,可设为损失函数。

求解可用拉格朗日乘子法转为对偶问题,研究推得的数学公式可写为:

$$\max_{\alpha, \alpha^*} \sum_{i=1}^m (y_i(\alpha_i^* - \alpha_i) - \varepsilon(\alpha_i^* + \alpha_i)) - \frac{1}{2} \sum_{i=1}^m \sum_{j=1}^m (\alpha_i^* - \alpha_i)(\alpha_j^* - \alpha_j) K(x_i, x_j) \quad (2)$$

约束条件为:

$$\sum_{i=1}^m (\alpha_i^* - \alpha_i) = 0$$

$$0 \leq \alpha_i, \alpha_i^* \leq C$$

其中, α_i^*, α_i 为拉格朗日乘子, $K(x_i, x_j)$ 为核函数。

因 SVR 继承了 SVM 的稀疏性,式(2)的解中,只有少量样本的系数 $(\alpha_i^* - \alpha_i)$ 不等于0,这些样本 (x_i, x_j) 就是支持向量。设支持向量个数为 n ,由支持向量决定 SVR 回归,其数学定义式分别如下:

$$f(x) = \sum_{i=1}^n (\alpha_i^* - \alpha_i) K(x_i, x) + b \quad (3)$$

$$b = y_i + \varepsilon - \sum_{j=1}^n (\alpha_j^* - \alpha_j) K(x_j, x_i) \quad (4)$$

2 床温的 GA-SVR 预测模型

本文研究对象为某 300 MW 循环流化床锅炉,选择床温预测模型的输入变量和输出变量,进行数据预处理;通过遗传算法对 SVR 参数进行全局寻

优;采用最佳参数训练 SVR,最终将模型用于床温预测。

2.1 床温特性分析

CFB 锅炉床温的高低主要由其运行参数、如给煤量、一次风量、二次风量、循环灰量等决定^[12]。

对于同一煤种,一次风量和循环灰量不变时,床温与给煤量保持一致的变化趋势^[13]。而给煤量不变时,一次风量的增加能促进密相区燃料的燃烧热解,但也会增大烟气携带床料的固体热损失,同时由于一次风温较低,过量的一次风也会冷却床层,使床温下降;而一次风量过低,无法提供燃烧热解所需氧量,密相区过度缺氧燃烧,导致床温降低^[14]。此外一次风量过小,可能无法维持床层的正常流化状态,将严重威胁 CFB 锅炉的安全运行。二次风的作用是保证稀相区挥发分和残碳的正常燃烧,稀相区氧量充足时,因二次风温相对于稀相区整体的温度偏低,二次风量的增加,将冷却烟气和流化物料;当稀相区氧量不足时,提高二次风量能促进挥发分和残

碳的燃烧,提高烟气和流化物料温度,而少量烟气和物料经旋风分离器分离后又将返回密相区,从而可间接影响床温。而循环灰量增加时,循环灰从炉膛底部带离的热量增加,从而降低床温;循环灰量不足时,床层热量无法被及时带离,将导致床温升高^[15]。

2.2 模型数据

本文基于某 300 MW 循环流化床锅炉机组变负荷工况下的运行数据,从锅炉运行机理和床温特性分析出发,选择负荷(x_1)、锅炉总烟气流量(x_2)、烟气氧含量(x_3)、石灰石颗粒给料量(x_4)、给煤量(x_5)、热一次风温(x_6)、热二次风温 A 侧(x_7)、热二次风温 B 侧(x_8)、一次风总量(x_9)、二次风 A 侧总量(x_{10})、二次风 B 侧总量(x_{11})、粉尘排放量(x_{12})作为模型的输入参数,炉膛平均床温(x_{13})作为输出参数。先对样本进行粗大误差处理,剔除数据坏点,选取连续 9 天内锅炉机组的 216 组数据后再以 6:4 的比例,将运行数据随机划分为训练集和验证集。部分数据见表 1。

表 1 某 300 MW 循环流化床锅炉运行数据

Tab. 1 Operation data of a 300 MW CFB boiler

序号	x_1 / MW	x_2 / (Nm ³ ·h ⁻¹)	x_3 / %	x_4 / (t·h ⁻¹)	x_5 / (t·h ⁻¹)	x_6 / ℃	x_7 / ℃	x_8 / ℃	x_9 / (Nm ³ ·h ⁻¹)	x_{10} / (Nm ³ ·h ⁻¹)	x_{11} / (Nm ³ ·h ⁻¹)	x_{12} / (t·h ⁻¹)	x_{13} / ℃
1	151.09	728 050.55	8.47	1.61	28.67	259.65	250.68	264.50	441 034.21	30 919.71	45 985.53	28.34	790.48
2	150.91	802 004.94	7.81	1.61	29.31	259.19	247.73	261.59	409 087.30	42 327.89	57 245.38	5.47	805.96
3	151.22	847 940.02	6.72	1.63	30.09	257.00	250.08	262.43	412 414.41	43 795.18	57 979.18	12.83	795.33
...
214	256.92	1 204 644.27	6.73	2.86	76.98	275.17	259.96	281.49	439 275.42	177 471.80	181 660.96	37.10	921.89
215	262.37	1 198 615.08	5.19	2.93	84.73	275.02	260.31	282.12	444 136.32	174 117.89	180 344.38	21.59	932.80
216	260.02	1 202 217.05	7.39	2.90	87.03	269.24	254.04	276.82	435 886.46	181 908.25	190 604.84	44.93	931.31

2.3 模型建立

(1)数据预处理:在参数寻优和 SVR 建模前,需将所选的运行数据无量纲化,以消除各数据量纲的影响。本文选择归一化处理各输入变量,数学定义公式见如下:

$$x^* = \frac{x - x_{\min}}{x_{\max} - x_{\min}} \quad (5)$$

其中, x 为某变量空间; x_{\min} 为变量空间 x 中最小的值; x_{\max} 为变量空间 x 中最大的值。训练集和验证集的输入输出变量均经过归一化处理后映射到 $[0, 1]$ 区间内。

(2)SVR 参数寻优:建立预测模型前,需确定 SVR 的相关参数,以保证模型的准确性。本文采用遗传算法,对其 3 个主要参数:惩罚系数 C 、径向核参数 $K(x_i, x_j)$ 和损失函数 ε 全局寻优。以确定系数 R^2 (R-Square) 为遗传算法的适应度,再设置 SVR

为待优化函数,并绑定所需优化的参数 C 、 ε 和 $K(x_i, x_j)$; 设置参数的区间范围并确定初始值: $C = 2$, $\varepsilon = 0.5$, $K(x_i, x_j) = 1$ 。规定遗传为 40 代,每代 20 人,并按照步骤计算,最终得到最优化参数 $C = 7.840 8$, $\varepsilon = 0.013$, $K(x_i, x_j) = 0.299 7$ 。遗传算法每代适应度值如图 1 所示。

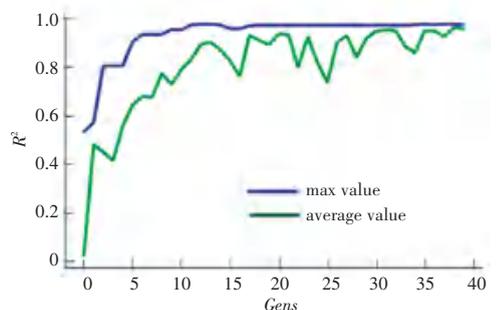


图 1 遗传算法适应度

Fig. 1 Fitness of genetic algorithm per generation

(3)SVR 建模:将经过归一化处理的数据集 x_1, x_2, \dots, x_{12} 作为 GA-SVR 模型的输入,将锅炉平均床温作为模型的输出,建立锅炉床温的 GA-SVR 预测模型。

3 预测结果与模型对比

GA-SVR 模型对验证集的预测结果如图 2 所示。

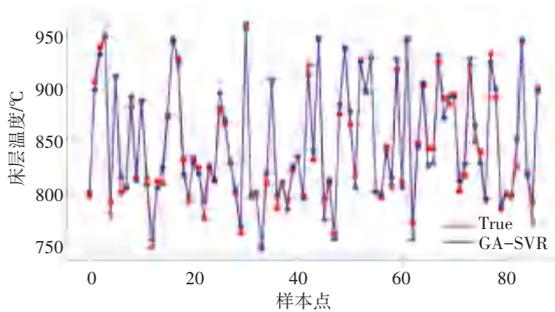


图 2 验证集的床温预测结果

Fig. 2 Bed temperature prediction results of the verification set

选取确定系数 R^2 (R-Square)、均方误差 (MSE)、模型计算时间为模型的评价指标,具体数学定义公式分别如下:

$$R^2 = \frac{\sum_{i=1}^m (y_i^* - \bar{y}_i)^2}{\sum_{i=1}^m (y_i - \bar{y}_i)^2} \quad (6)$$

$$MSE = \frac{1}{m} \sum_{i=1}^m (y_i - y_i^*)^2 \quad (7)$$

其中, y_i 为真实值; y_i^* 为预测值; \bar{y}_i 为真实值均值; m 为样本个数。

在研究运行后可得到基于 GA-SVR 的 CFB 锅炉床温预测模型的确定系数为 0.977 6, 均方误差为 0.16%, 计算时间为 0.008 9 s, 验证集的最大预测误差为 24.4 ℃。

为验证 GA-SVR 模型对于小样本数据预测的优越性, 本文将 GA-SVR 模型与 BP 神经网络、BP-SVR 和 CNN 三种常用预测模型进行了对比。对此拟展开分述如下。

3.1 GA-SVR 与 BP 神经网络预测结果对比

本文采用常见的单隐层 BP 神经网络结构, 将 12 个运行参数作为输入变量, 按经验公式计算得出隐层神经元参考个数为 10, 以 CFB 锅炉床温作为输出变量, 选取 MSE 为损失函数, sigmoid 函数为各神经元的激活函数, 网络总迭代次数为 1 000 次, 建立了基于 BP 神经网络的床温预测模型。将 BP 模型

与 GA-SVR 模型的预测误差和评价指标进行对比, 具体如图 3 所示。各指标对比结果见表 2。

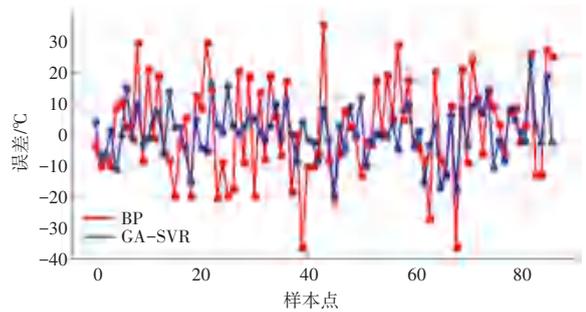


图 3 BP 模型与 GA-SVR 模型预测误差对比

Fig. 3 Comparison of prediction errors between BP model and GA-SVR model

表 2 BP 模型与 GA-SVR 模型评价指标对比

Tab. 2 Comparison of evaluation indexes between BP model and GA-SVR model

模型	确定系数	均方误差/%	计算时间/s
BP	0.932 5	0.48	15.504 0
GA-SVR	0.977 6	0.16	0.008 9

由图 3 和表 2 可知, 对于同一验证集, GA-SVR 模型预测误差分布较为均匀, 几乎全部样本点的预测误差都小于 BP 模型预测误差; GA-SVR 模型的最大误差为 24.4 ℃, 相较于 BP 模型的 37.2 ℃ 也有了明显的降低; GA-SVR 模型的各项评价指标明显优于 BP 模型。尤其是时间指标, BP 模型的计算时间为 GA-SVR 模型的 1 742 倍。

3.2 GA-SVR 与 BP-SVR 预测结果对比

将 BP 神经网络与支持向量机结合, 对输入变量进行影响力分析计算, 筛除影响力小于 5% 的输入变量, 再将降维后的变量作为 SVR 模型的输入。为保证模型间具有可比性, 对 SVR 参数也进行寻优, 建立 BP-SVR 的 CFB 锅炉床温预测模型。BP-SVR 与 GA-SVR 模型预测误差与评价指标对比, 具体如图 4 所示。各项指标对比结果见表 3。

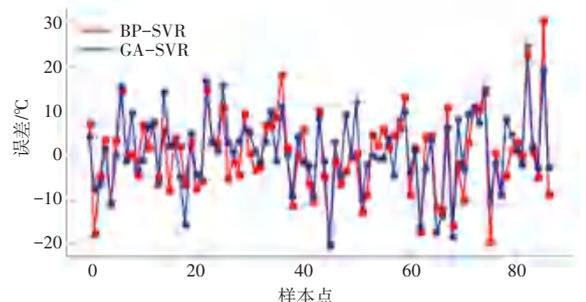


图 4 BP-SVR 模型与 GA-SVR 模型预测误差对比

Fig. 4 Comparison of prediction errors between BP-SVR model and GA-SVR model

表3 BP-SVR模型与GA-SVR模型评价指标对比

Tab. 3 Comparison of evaluation indexes between BP-SVR model and GA-SVR model

模型	确定系数	均方误差/%	计算时间/s
BP-SVR	0.974 7	0.18	0.005 9
GA-SVR	0.977 6	0.16	0.008 9

由图4和表3可知,BP-SVR预测模型与GA-SVR预测模型的误差主要集中在[-10℃,10℃]区间内,两模型预测折线图整体差异不大,均能较好反映床温真实值。GA-SVR模型的 R^2 和MSE指标略优于BP-SVM模型,但BP-SVR模型的计算时间有了一定的改善。BP-SVR模型在部分样本点的预测结果与真实值存在较大偏差,精度较低,其最大误差为30.1℃,高于GA-SVR模型的24.4℃,表明输入变量维度的减少能降低模型的复杂程度,提高模型计算速率,但也会损失部分数据信息,造成模型精度的轻微下降。

3.3 GA-SVR与CNN预测结果对比

本文选择的CNN模型采用LeNet-5结构,即2卷积层/2池化层/2全连接层,设定卷积层Conv1和Conv2的卷积核个数均为40,卷积核尺寸为3×3;全连接层Fc1和Fc2的神经元个数分别为20和10;选取池化函数为最大值函数,激活函数为ReLU函数,训练次数为40次,建立CNN的CFB锅炉床温预测模型。CNN与GA-SVR模型预测误差与评价指标对比,具体如图5所示。各指标对比结果见表4。

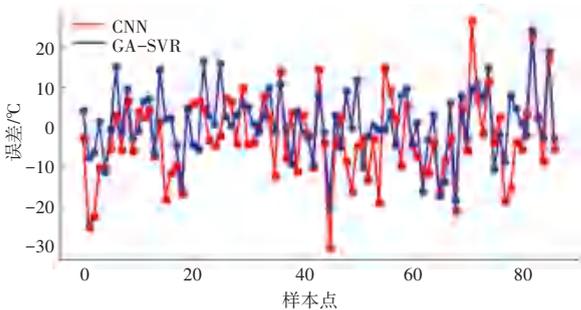


图5 CNN模型与GA-SVR模型预测误差对比

Fig. 5 Comparison of prediction errors between CNN model and GA-SVR model

表4 CNN模型与GA-SVR模型评价指标对比

Tab. 4 Comparison of evaluation indexes between CNN model and GA-SVR model

模型	确定系数	均方误差/%	计算时间/s
CNN	0.965 9	0.24	3.688 8
GA-SVR	0.977 6	0.16	0.008 9

由图5和表4可知,从整体上看CNN模型预测误差相比GA-SVR模型有负误差增大的情况,且预测误差折线图的波动较大,存在多个误差大于20℃的样本点,最大误差为30.0℃,也高于GA-SVR的

24.4℃。从评价指标上看,CNN模型的评价指标 R^2 和MSE均劣于GA-SVR,且预测所需时间更长。

4 结束语

本文通过GA遗传算法精确且快速地对SVR算法的惩罚系数C、径向核参数 $K(x_i, x_j)$ 和损失函数 ϵ 全局寻优,能有效解决SVR算法最优参数选取困难的问题。对于同一测试样本,分别建立基于BP神经网络、BP-SVR和CNN的床温预测模型,并预测同一验证集,得出GA-SVR、BP神经网络、BP-SVR和CNN四种预测模型的最大绝对误差分别为:24.4℃、37.2℃、30.1℃和30.0℃,最大相对误差分别为:2.4%、4.4%、3.8%和3.7%;通过比较分析4种模型的评价指标和误差可得出面向小样本数据容量,GA-SVR预测模型拥有更高的预测精度和更低的时间成本,因此面向小数据样本容量的GA-SVR循环流化床锅炉床温预测模型能实现对CFB锅炉床温的实时准确预测,具有较好的工程应用价值。

参考文献

- [1] 黄中,杨娟,车得福.大容量循环流化床锅炉技术发展应用现状[J]. 热力发电,2019,48(06):1-8.
- [2] 王超,宋国良,吕清刚.循环流化床超低NO_x与SO₂排放技术研究进展[J]. 洁净煤技术,2021,27(04):17-25.
- [3] 尤海辉,马增益,唐义军,等.循环流化床入炉垃圾热值软测量[J]. 浙江大学学报(工学版),2017,51(06):1163-1172.
- [4] 钱虹,柴婷婷,张超凡.基于深度循环神经网络的SCR烟气脱硝系统出口NO_x排放预测研究[J]. 热能动力工程,2020,35(08):77-84.
- [5] 白建云,朱竹军,张培华.基于BP神经网络的循环流化床锅炉生成NO_x质量浓度在线软测量[J]. 热力发电,2016,45(12):78-83.
- [6] 张妍,王东风,韩璞.粒子群优化粒子滤波算法及其在循环流化床床温辨识中的应用[J]. 华北电力大学学报(自然科学版),2015,42(02):104-110.
- [7] 段萌,王功鹏,牛常勇.基于卷积神经网络的小样本图像识别方法[J]. 计算机工程与设计,2018,39(01):224-229.
- [8] 王琦,范常浩,白建云,等.变量选择与支持向量机相结合的SO₂排放特性建模[J]. 热力发电,2018,47(03):68-75.
- [9] GOLDBERG D E, HOLLAND J H. Holland. Genetic algorithms and machine learning[J]. Machine Learning, 1988,3(2-3):95-99.
- [10] 薛同来,赵冬晖,韩菲.基于GA优化的SVR水质预测模型研究[J]. 环境工程,2020,38(03):123-127
- [11] 周志华. 机器学习[J]. 中国民商, 2016, 3(21): 93-93.
- [12] 张栾英,李建强,谷俊杰,等.100 MW循环流化床锅炉床温控制策略[J]. 华北电力大学学报,2004,31(01):44-47.
- [13] 刘爱军,刘孝国,蔡宝玲,等.循环流化床锅炉床温主要影响因素仿真研究[J]. 热力发电,2011,40(09):73-75.
- [14] 杨儒,张悦,冷辉.循环流化床床温的多变量建模[J]. 计算机仿真,2019,36(04):52-56,93.
- [15] 赵周明,孙健秋,郭涛,等.床温对CFB锅炉运行的影响及其调整措施分析[J]. 电站系统工程,2013,29(04):35-37.