

# 基于自适应遗传算法的极限学习机改进算法

丁胜夺, 谭 昆, 田 琨, 吴顺成

(中国石油天然气股份有限公司安全环保技术研究院, 北京 102206)

**摘要:**针对极限学习机结构参数过多, 初始化阶段输入层和隐层间的权值矩阵以及隐层的阈值随机生成导致该算法性能不稳定、易出现过拟合现象的缺陷, 在自适应遗传算法的基础上对极限学习机算法进行优化提升。首先, 该算法将传统遗传算法中的交叉变异概率改进为随群体适应度水平和当前遗传代数更新的自适应参数, 优化算法对局部和全局最优解的搜索能力; 然后, 将极限学习机中的待定参数作为染色体带入遗传算法中进行交叉变异操作寻找适应度最佳个体。通过仿真实验将该改进算法和 ELM、BP 神经网络模型进行比较, 实验结果表明该改进算法在分类问题以及回归问题中都得到了更加准确的预测结果, 极限学习机的泛化性能得到显著提升。

**关键词:**遗传算法; 极限学习机; 自适应; 分类算法; 回归预测

中图分类号: TP311

文献标识码: A

文章编号: 1673-629X(2022)0026-05

## Improved Algorithm of Extreme Learning Machine Based on Adaptive Genetic Algorithm

DING Sheng-duo, TAN Kun, TIAN Kun, WU Shun-cheng

(China National Petroleum Corporation Safety and Environmental Technology Research Institute Co., Ltd.,  
Beijing 102206, China)

**Abstract:** In view of the defects of the extreme learning machine algorithm, such as too many parameters, the connection weights between the input layer and the hidden layer and the random generation of the threshold of the hidden layer, which lead to the unstable performance of the algorithm and the over-fitting phenomenon, an improved extreme learning machine algorithm based on the adaptive genetic algorithm is proposed. Firstly, the crossover mutation probability of traditional genetic algorithm is changed into an adaptive parameter that changes with population fitness and genetic algebra, and the global and local searching ability of genetic algorithm is enhanced. Then, the parameters to be determined in the limit learning machine are taken as chromosomes into the genetic algorithm for cross mutation operation to find the optimal individual. The improved algorithm was compared with ELM and BP neural network models through simulation experiments, and the experimental results showed that the improved algorithm obtained more accurate prediction results in both classification problems and regression problems, and effectively improved the generalization performance of the extreme learning machine.

**Key words:** genetic algorithm; extreme learning machine (ELM); adaption; classification algorithm; regression prediction

## 0 引言

极限学习机(extreme learning machine, ELM)是一种分类、回归以及特征学习算法<sup>[1]</sup>, 针对传统反向传播(BP)神经网络<sup>[2]</sup>求解参数过多、神经网络训练时间消耗较大、易陷入局部极值等缺陷, 该算法以求解线性方程组的方法取代标准算法中梯度下降方法进行迭代求解的过程, 求取最小二乘解得到输出层权值, 简化了繁琐的迭代过程, 形成一种速度快、结构简单、无需太多人工干预的学习算法。在解决数据分类、回归预测等实际问题中的良好性能表现使其得到了广泛的应用。

遗传算法(genetic algorithm, GA)是现代智能算法中的一种经典学习算法, 该算法从达尔文的进化学说中得到启示, 以遗传进化和自然选择为理论基础, 将生物学中适者生存法则和生物种群中不同个体结合并交换组合彼此染色体信息的机制用于寻找种群中最优个体<sup>[3]</sup>。在实际应用<sup>[4-6]</sup>中, 该算法将待优化问题参数化并选择合适的编码方式转化为等长染色体码组, 然后模仿自然界生物进化过程进行循环迭代, 经过逐轮的选择、交叉、变异操作搜索空间中适应度更好的区域, 直至得到最优解对应的点<sup>[7-10]</sup>。

收稿日期: 2021-08-29

作者简介: 丁胜夺(1994-), 男, 助理工程师, 硕士, 研究方向为大数据与人工智能。

ELM算法初始化过程中随机生成输入层和隐层之间权值矩阵、隐层偏置,虽然提高了模型的学习效率,但也容易导致网络稳定性差,结构风险大,降低模型的泛化性能。为了克服该算法的以上缺陷,众多的学者提出了相应的改进算法<sup>[11-13]</sup>。Kong等<sup>[14]</sup>提出一种使用蝙蝠算法优化ELM隐层节点数的算法,从一定程度上降低了算法结构过拟合的风险,但网络的权值和阈值未得到优化;Sun等<sup>[15]</sup>利用粒子群算法对ELM算法中的权值和阈值网络进行优化,取得了一定的效果,但同时增加了模型的结构风险;彭显刚等<sup>[16]</sup>通过遗传算法对ELM算法的权值矩阵及偏置进行寻优,但空间搜索能力的局限性容易导致遗传算法陷入局部极值。

为了改善ELM的缺陷,该文提出了一种使用自适应遗传算法进行ELM参数优化的算法。经典遗传算法用于优化问题时可发挥其全局搜索能力的优势,但局部搜索能力的不足容易导致算法在进化后期收敛停止<sup>[17]</sup>。且经典的遗传算法仅凭借适应度值来衡量当前解的优劣<sup>[18]</sup>,使得在算法初期某些适应度大的解得到过度的关注而陷入次优或其他局部极值,出现“早熟”的现象。为此,为改善遗传算法的上述缺陷,结合算法的全局搜索能力和交叉操作频度相关、局部搜索能力和变异操作频度相关的特征,根据当前个体的适应度和当前的进化代数,自适应的调节算法的交叉和变异概率。

## 1 ELM和改进的遗传算法

### 1.1 ELM算法

对于包含 $N$ 个样本的数据集 $(x_i, t_i)_{i=1}^N, x_i \in R^n, t_i = [t_{i1}, \dots, t_{ik}]^T, t$ 为对应的样本类别, $j$ 为样本编号, $m$ 为样本维数, $k$ 为类别数量。假设该模型包含 $L$ 个隐层节点,激活函数为 $g(x)$ ,则该ELM模型<sup>[1]</sup>的输出表达式为:

$$f_i = \sum_{j=1}^L \beta_j g(w_j x_i + b_j), i = 1, 2, \dots, N \quad (1)$$

式中, $\beta_j$ 表示第 $j$ 个隐层节点到网络输出节点之间的连接权值; $g(x)$ 为隐层激活函数, $w_j = [w_{j1}, w_{j2}, \dots, w_{jm}]$ 表示输入节点和第 $j$ 个隐层节点之间的连接权值向量; $b_j$ 代表第 $j$ 个隐层节点的偏置。

为了求取最优的网络参数 $w$ 和 $\beta$ ,尝试最小化网络的预测输出和样本真实值之间的差值使其为零,即

$$\sum_{i=1}^N \|f_i - t_i\| = 0, \text{可转化为:} \quad (2)$$

$$\sum_{j=1}^L \beta_j g(w_j x_i + b_j) = t_i, i = 1, 2, \dots, N$$

转化为矩阵形式为:

$$H\beta = T \quad (3)$$

$$H(w_1, \dots, w_L, b_1, \dots, b_L, x_1, \dots, x_N) = \begin{pmatrix} g(w_1 x_1 + b_1) & \dots & g(w_L x_1 + b_L) \\ \vdots & \ddots & \vdots \\ g(w_1 x_N + b_1) & \dots & g(w_L x_N + b_L) \end{pmatrix}_{N \times L} \quad (4)$$

式中, $H$ 表示网络隐含层的输出向量矩阵,矩阵中的第 $i$ 列即为相应的隐层节点的拟合输出值; $\beta = [\beta_1, \beta_2, \dots, \beta_L]^T$ ;  $T = [t_1, t_2, \dots, t_N]^T$ 。

在实际应用中,通常训练样本的数量 $N$ 远大于算法隐层节点的数量,当所选取的激活函数是连续可微的,所求网络的隐含层参数 $w$ 和 $b$ 在训练过程中可以视为常数,无需调整网络中全部参数<sup>[19-21]</sup>。然后上述模型可以通过求解最小化问题: $\min_{\beta} \|H\beta - T\|$ 得到权值向量 $\beta$ ,其求解可根据广义逆原理计算得:

$$\hat{\beta} = H^+ T \quad (5)$$

式中,根据线性代数的求解法则需要求解 $H$ 的广义逆矩阵 $H^+$ , $H$ 表示隐层输出矩阵。

### 1.2 改进的遗传算法

遗传算法的主要操作流程可描述如下:

Step1:进行种群的初始化。将种群的规模设定为 $N$ ,在定义区间内随机生成 $N$ 个作为初始群体 $Y(0)$ ,将待求参数视为遗传基因进行固定长度的二进制编码,设定最大进化代数 $T$ 、交叉概率 $P_c$ 以及变异概率 $P_m$ 。

Step2:适应度评估。根据适应度公式,计算并记录当前群体中所有个体的适应度值。

Step3:进化。选择:采用轮盘赌的方式可结合群体适应度分布选择得到 $L$ ( $L \geq N$ )对染色体作为父代;交叉:在选择操作中得到的过渡群体中随机选定 $L/2$ 对染色体,然后使其中的个体以交叉概率 $P_c$ 与其结对个体进行单点交叉操作;变异:对交叉所得群体以变异概率 $P_m$ 对其中的个体进行变异操作。最后,在 $L$ 个候选个体中以适应度排序后选择 $N$ 个体作为 $t+1$ 代种群。

Step4:终止进化。当算法进行到第 $T$ 代时终止操作,以当前群体中适应度值最大的个体作为所求解<sup>[22]</sup>。

针对传统遗传算法交叉以及变异概率在进化过程中大小改变,容易导致算法后期收敛速度慢且易收敛到局部极值的缺陷对该算法进行改进,使个体的交叉变异概率随自身适应度以及种群进化水平而自适应变化。在算法运行前期,大部分个体的环境适应能力偏低,所以适应度低于群体均值的个体此时需要较大的交叉概率提升算法的在前期的全局搜索能力。相对应的,在算法的运行后期,适应能力极差的个体已经在逐

轮迭代中被剔除,此时对于适应度高于平均值的个体可以减小其交叉概率,削弱其全局搜索能力。基于以上思想改进后交叉概率为:

$$P_c = \begin{cases} P_{cmax} - \frac{(P_{cmax} - P_{cmin})(f_{max} - f)}{(f_{max} - f_{avg})} \cdot \frac{iter}{iter_{max}} & f \geq f_{avg} \\ P_{cmin} & f < f_{avg} \end{cases} \quad (6)$$

式中,  $f$  表示当前代数中个体的适应度值,  $f_{max}$  和  $f_{avg}$  为相应的最大值和平均值,  $P_{cmax}$  和  $P_{cmin}$  分别为初始化设定的最大最小交叉操作发生概率,  $iter$  为个体当前所处的进化代数,  $iter_{max}$  为初始设定的最大进化代数。

在执行变异操作时,与交叉操作类比,在算法开始运行时,个体的适应度通常偏低,此时为了将群体中优良的染色体保留下来应分配给适应度值低于平均值的个体一个较低的变异概率;在算法演化的后期,对适应度高于平均适应度的个体可适当增加其变异的概率,提高此时局部搜索能力。所以,改变后的变异概率为:

$$P_m = \begin{cases} P_{mmin} & f < f_{avg} \\ P_{mmin} + \frac{(P_{mmax} - P_{mmin})(f_{max} - f)}{f_{max} - f_{avg}} \cdot \frac{iter}{iter_{max}}, & f \geq f_{avg} \end{cases} \quad (7)$$

式中,  $P_{mmax}$  为初始设定的最大变异操作执行概率,  $P_{mmin}$  为最小概率,其他参数同式(6)。

## 2 基于改进遗传算法的 ELM 算法

为提高 ELM 的预测性能,降低模型过拟合的结构风险,提出了基于改进遗传算法进行初始参数优化的极限学习机(IGA-ELM)。在该算法中,将 ELM 中随机初始化的连接权值和阈值等参数转化为遗传算法中的染色体,将 ELM 算法在相应参数下所得的训练误差作为染色体的适应度值。执行遗传算法的寻优操作过程,逐轮进化寻优得到适应度最高的个体作为 ELM 的初始化参数。该算法结合了遗传算法突出的寻优能力和 ELM 的强学习能力。

基于改进自适应遗传算法的 ELM 算法流程如下:

Step1: 确定 ELM 和遗传算法的基本参数。包括 ELM 的输入神经元数量、隐层神经元数量以及输出层神经元数量,遗传算法中的初始化种群大小  $N$ 、进化过程最大执行代数  $T$ 、交叉、变异操作执行概率等;

Step2: 通过随机初始化得到多组 ELM 参数并对其进行等长二进制编码<sup>[23]</sup>,得到初始种群;

Step3: 将参数分别带入 ELM 网络中进行训练,得到误差平方和作为适应度函数;

Step4: 将当前染色体种群以及每个个体所对应的适应度值带入遗传算法进行选择操作,然后以自适应概率进行交叉变异操作得到新种群,转至 Step3 直到执行至第  $T$  代,求出适应度最高的个体;

Step5: 将所得最优参数带入 ELM 网络,求解隐含层输出矩阵  $H$  以及其广义逆矩阵  $H^+$ ;

Step6: 计算 ELM 输出层权值  $\beta$ ;

Step7: 将待测样本带入模型中进行预测。

算法流程如图 1 所示。

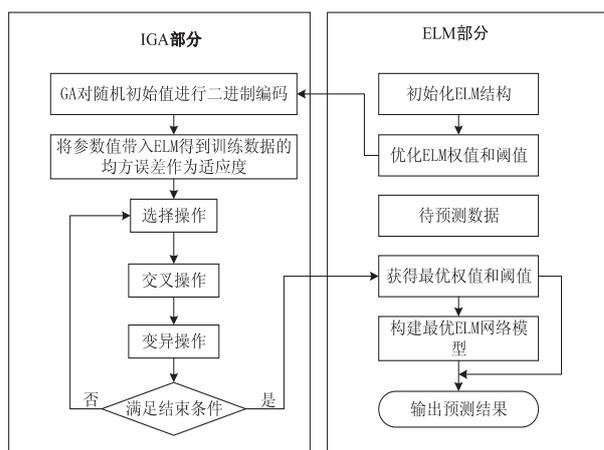


图 1 IGA-ELM 简要流程

## 3 实验结果及分析

### 3.1 实验设置

为了验证 IGA-ELM 算法的性能提升效果,将此改进算法和标准 ELM 算法以及 BP 神经网络算法在分类以及回归预测的性能上进行实验对比,实验数据采用 UCI 中的公共数据集,数据集的属性信息如表 1、表 2 所示,其中表 1 数据为分类实验采用的数据,表 2 数据为回归实验采用的数据。实验环境如下: 8 G 内存的 Inter(R) Core(TM) i7-8750H CPU @ 2.2 GHz 2.21 GHz 计算机, MATLAB R2018b 版本。

表 1 用于分类实验的 UCI 数据集及其相关特征

数据集	样本数	特征数	类别数	训练数据	测试数据
Iris	150	4	3	120	30
Abalone	4 177	8	29	3 342	835

表 2 用于回归实验的 UCI 数据集及其相关特征

数据集	样本数	特征数	训练数据	测试数据
Automobile	205	26	164	41
SML2010	4 137	24	3 310	827

在实验中,为了消除不同量纲造成的数值差别过大对实验模型的影响将所有输入数据属性值归一化到  $[-1, 1]$  之间,将分类问题中的输出映射到  $[1, n]$  ( $n$  表示类别总数)之间的整数上,回归问题的输出不做

特殊处理;ELM中设置 Sigmoid 函数作为隐层神经元的激活函数,并为对比试验的公平性将 ELM 及 BP 神经网络中的隐层神经元数量设定为等值 40;遗传算法中,设定染色体数目为 20,算法执行到 100 代时停止操作,代沟设置为 0.95,交叉概率区间为 $[0.6, 0.8]$ ,变异概率区间为 $[0.001, 0.1]$ 。实验为了消除偶然性,ELM 算法和 BP 神经网络算法中的实验结果采用 4 次 5 折交叉实验的平均结果,IGA-ELM 算法中的实

验结果为 10 次运行结果的平均值。

### 3.2 实验结果及分析

在仿真实验中,IGA-ELM 算法在 6 个 UCI 数据集上具有代表性的进化过程如图 2 所示,其中图 2(a)、图 2(b)显示出该算法在分类数据集上的分类误差的变化,图 2(c)、2(d)显示了该算法在回归数据集上预测值和真实值之间误差的变化。

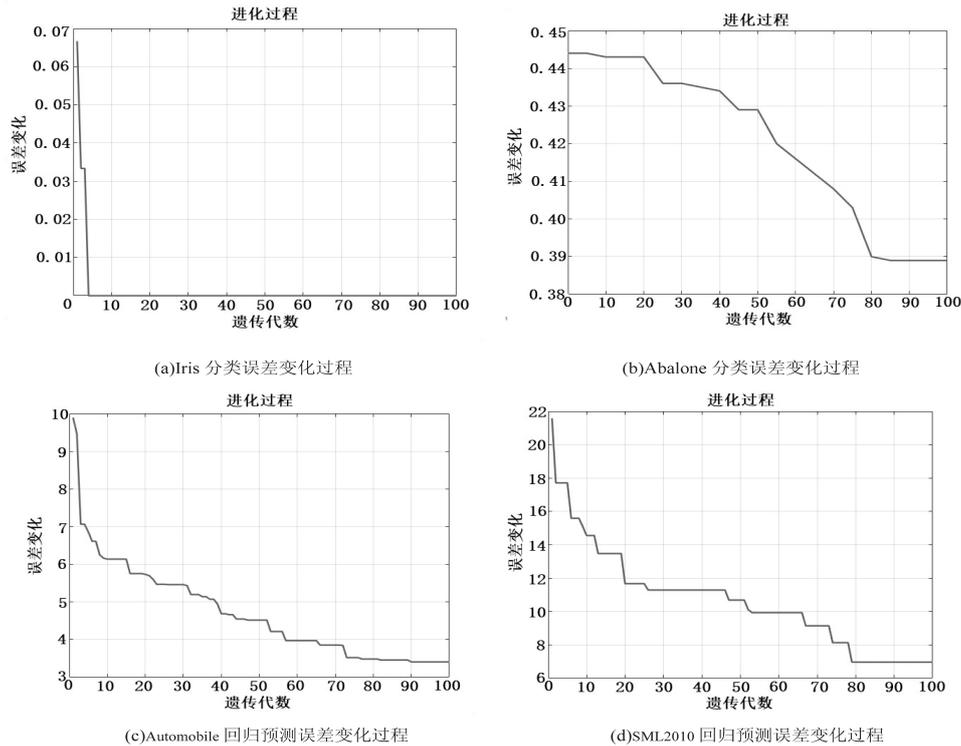


图 2 各数据集在 IGA-ELM 算法进化过程中的误差变化

在仿真实验中,Iris 数据集在 IGA-ELM 算法中分类预测误差如图 2(a)所示,由于该数据集数据量偏少且种类特征明显,ELM 算法的快速学习能力使其在遗传算法的增强下分类性能得到了快速的提高,在进化到第 5 代左右时 ELM 对测试数据的预测精度已经达到了 100%。在 Abalone、Automobile、SML2010 等数据集中,IGA 算法进化初期使得 ELM 的性能有了显著的增强,这是由于进化初期群体相对适应度不高,此时的交叉概率较大,IGA 的全局寻优能力较强,而此时的变异概率较低,保证优良的个体得以保留,此阶段群体分类精度逐步提升;在进化中期,大多数数据集上群体适应度在一定时间内处于稳定水平,在 Automobile 数据集上群体的适应度值有缓慢的破浪式的提升但相对初期放慢了收敛的速度;在进化后期,群体的适应度水平较高,此时的交叉概率随着遗传代数的增加而降低,相应的变异概率相对进化前期会有一定的提升,此阶段的全局搜索能力下降,局部搜索能力逐渐提升,故此阶段的误差有一定的降低,致使算法更快的收敛到全局最

优解。

此外,进行了 IGA-ELM 算法和 ELM 算法、BP 神经网络算法的对比实验,在分类问题中,将模型对测试数据集的分类正确率作为衡量指标,在回归问题中将模型预测值和真实值之间的误差值作为衡量指标,进行了对比。实验结果如表 3、表 4 所示。由表中数据可知,IGA-ELM 算法在分类问题以及回归问题上都取得了最优的效果,其中在 Iris、Abalone 数据集上,IGA-ELM 算法相对于 ELM 算法分类准确率分别提高了 6%、7.5%,相对于 BP 神经网络算法正确率分别提高了 6.7%、6.6%;在 Automobile、SML2010 数据集上,IGA-ELM 算法相对于 ELM 算法回归预测误差分别降低了 6.453、14.555,相对于 BP 神经网络算法误差分别降低了 6.351、19.564。

表 3 各算法在分类数据集上的预测正确率比较

数据集	ELM	BP	IGA-ELM
Iris	0.940	0.933	1.000
Abalone	0.554	0.563	0.629

表 4 各算法在回归数据集上的误差比较

数据集	ELM	BP	IGA-ELM
Automobile	9.825	9.723	3.372
SML2010	21.573	26.582	7.018

#### 4 结束语

为提高 ELM 模型稳定性以及泛化性能,提出了一种 IGA-ELM 算法,并通过在公共数据集上的实验证明该算法相对于 ELM 算法和传统的 BP 神经网络算法无论是分类性能还是回归预测性能都有显著的提高。实验结果也表明了改进自适应遗传算法的全局搜索能力在 ELM 参数寻优中的应用起到了积极的作用,两种算法的优势得到充分结合,提高了算法的适用性,对于离线操作或者精度要求较高的应用场合有一定的参考意义。此外,尽管目前针对 ELM 的一些分类算法被提出,但该算法仍具有参数相对较多,可解释性不强等缺点,将 ELM 算法扩展到半监督和无监督学习任务的研究极少,以上都是有待解决的问题。

#### 参考文献:

- [1] HUANG G B, ZHU Q Y, SIEW C K. Extreme learning machine: a new learning scheme of feedforward neural networks [C]//IEEE international joint conference on neural networks. Beijing: IEEE, 2005.
- [2] ZHANG X Y, WANG S, YUN X. Bidirectional active learning: a two-way exploration into unlabeled and labeled data set[J]. IEEE Transactions on Neural Networks & Learning Systems, 2017, 26(12): 3034-3044.
- [3] 孙晓燕, 陈姗姗, 巩敦卫, 等. 基于区间适应值交互式遗传算法的加权多输出高斯过程代理模型[J]. 自动化学报, 2014, 40(2): 172-184.
- [4] AYBAR-RUIZ A, JIMÉNEZ-FERNÁNDEZ S, CORNEJO-BUENO L, et al. A novel grouping genetic algorithm - extreme learning machine approach for global solar radiation prediction from numerical weather models inputs[J]. Solar Energy, 2016, 132: 129-142.
- [5] HAMIAN M, DARVISHAN A, HOSSEINZADEH M, et al. A framework to expedite joint energy-reserve payment cost minimization using a custom - designed method based on mixed integer genetic algorithm [J]. Engineering Applications of Artificial Intelligence, 2018, 72: 203-212.
- [6] SEGHIR F, KHABABA A. A hybrid approach using genetic and fruit fly optimization algorithms for QoS-aware cloud service composition [J]. Journal of Intelligent Manufacturing, 2017(6): 1-20.
- [7] 高萍, 刘松, 程顺, 等. 基于 BP 神经网络和遗传算法优化 S Zorb 装置汽油辛烷值损失 [J]. 石油炼制与化工, 2021, 52(7): 88-95.
- [8] 轩华, 王晶, 张慧贤, 等. 混合遗传算法求解含机器可利用约束的 HFSP [J]. 计算机应用与软件, 2021, 38(6): 176-181.
- [9] 席亮, 王瑞东. 基于自适应遗传算法的神经网络结构优化算法 [J]. 哈尔滨理工大学学报, 2021, 26(1): 39-44.
- [10] 贺毅朝, 王熙照, 李文斌, 等. 基于遗传算法求解折扣  $|0-1|$  背包问题的研究 [J]. 计算机学报, 2016, 39(12): 2614-2630.
- [11] ATEE H A, AHMAD R, NOOR N M, et al. A novel extreme learning machine-based cryptography system [J]. Security & Communication Networks, 2016, 9(5): 45-54.
- [12] BUI D T, NGO P T T, PHAM T D, et al. A novel hybrid approach based on a swarm intelligence optimized extreme learning machine for flash flood susceptibility mapping [J]. Catena, 2019, 179(3): 13-25.
- [13] 陈晓青. 改进 PSO 算法与 ELM 在基因数据分类中的应用 [D]. 杭州: 中国计量大学, 2017.
- [14] 孔令春, 孙琼琼, 杨照峰. 蝙蝠算法优化极限学习机的电力负荷预测模型 [J]. 辽宁工程技术大学学报: 自然科学版, 2016, 35(1): 89-92.
- [15] SUN W, WANG C, ZHANG C. Factor analysis and forecasting of CO<sub>2</sub> emissions in Hebei, using extreme learning machine based on particle swarm optimization [J]. Journal of Cleaner Production, 2017, 162(1): 9-15.
- [16] 彭显刚, 郑伟钦, 林利祥, 等. 考虑负荷自适应检测和修复的鲁棒极限学习机短期负荷预测方法 [J]. 中国电机工程学报, 2016, 36(23): 6409-6417.
- [17] ZHANG R, TAO J. A nonlinear fuzzy neural network modeling approach using improved genetic algorithm [J]. IEEE Transactions on Industrial Electronics, 2017(99): 1-11.
- [18] 崔光照, 李小广, 张勋才, 等. 基于改进的粒子群遗传算法的 DNA 编码序列优化 [J]. 计算机学报, 2010, 33(2): 311-316.
- [19] 单英浩, 付青, 耿炫, 等. 基于改进 BP-SVM-ELM 与粒子化 SOM-LSF 的微电网光伏发电组合预测方法 [J]. 中国电机工程学报, 2016, 36(12): 3334-3343.
- [20] ZHENG Dehua. An improved LDA-based ELM classification for intrusion detection algorithm in IoT application [J]. Sensors, 2020, 20(6): 10-21.
- [21] 张宁, 魏秀业, 徐晋宏. 基于 LMD 样本熵与 ELM 的行星齿轮箱故障诊断 [J]. 机械传动, 2020, 44(4): 152-157.
- [22] DONNELLY K, VARLOW B R. AC conductivity effects of non-linear fillers in electrical insulation [C]//Conference on electrical insulation & dielectric phenomena. Singapore: IEEE, 2000: 132-135.
- [23] 张志毅, 袁荣湘, 杨同忠, 等. 基于粗糙集和小生境遗传算法的电网故障诊断规则提取 [J]. 电工技术学报, 2009, 24(1): 158-163.