

改进的自适应遗传算法及应用^{*}

王倩, 李风军

(宁夏大学 数学统计学院, 银川 750021)

摘要:【目的】研究改善自适应遗传算法(Adaptive genetic algorithm, AGA)的早熟问题,并提升该算法优化的神经网络模型的预测性能。【方法】引入四分位间距替换自适应交叉和变异概率中的固定参数,首先给出了一种改进的自适应遗传算法(Modified adaptive genetic algorithm, MAGA),并将它的自适应交叉和变异概率映射到固定区间,使得优良个体在MAGA作用下仍具有良好的进化能力。其次利用MAGA对单隐层前馈神经网络(Single hidden layer feedforward neural networks, SLFNs)的初始权值和阈值进行优化,并给出了具体的算法实施步骤。【结果】6个常用测试函数的仿真实验表明:MAGA能够迅速稳定地逼近函数最优解。【结论】实证分析表明:与SLFNs及AGA优化的SLFNs相比,MAGA优化的SLFNs具有更高的分类正确率。

关键词: 自适应遗传算法;单隐层前馈神经网络;动态选择;动态变异;优化

中图分类号: O221.6

文献标志码: A

文章编号: 1672-6693(2021)02-0014-06

遗传算法(Genetic algorithm, GA)可通过模拟生物个体染色体的选择、交叉、变异实现全局寻优^[1]。由于标准遗传算法采用固定的交叉概率和变异概率,不利于寻优, Srinivas 等人^[2]提出自适应遗传算法(Adaptive genetic algorithm, AGA),但在较优个体中几乎不起作用。单隐层前馈神经网络(Single hidden layer feedforward neural networks, SLFNs)具有强大的非线性映射能力、适应能力和学习能力,已广泛地应用于各个领域^[3]。但是SLFNs存在易陷入局部极小的缺点^[4],而AGA可以进行全局寻优^[5],因此可利用AGA对SLFNs的初始权值和阈值进行优化。

针对上述问题,文献[6-12]从不同角度改进了自适应交叉和变异概率,均取得了较好的效果,但本质问题并未解决。鉴于此,本文引入四分位数的定义,给出一种改进的自适应遗传算法(Modified adaptive genetic algorithm, MAGA),仿真实验和实证分析均表明MAGA能够更加迅速地逼近全局最优解,改善了经过AGA优化的SLFNs模型容易陷入局部极值的缺点。

1 AGA

AGA的自适应性主要体现在种群中个体可以根据不同适应度调整交叉和变异概率,具体公式分别为:

$$P_c = \begin{cases} \frac{k_1(f_{\max} - f')}{f_{\max} - f_{\text{avg}}}, & f' \geq f_{\text{avg}}, \\ k_2, & f' < f_{\text{avg}} \end{cases}, \quad (1)$$

$$P_m = \begin{cases} \frac{k_3(f_{\max} - f)}{f_{\max} - f_{\text{avg}}}, & f \geq f_{\text{avg}}, \\ k_4, & f < f_{\text{avg}} \end{cases}. \quad (2)$$

其中: f' 为待交叉的两个个体中较大的适应度值; $f_{\max}$ 为种群的最大适应度; f_{avg} 表示种群的平均适应度; f 表示变异个体的适应度; k_1, k_2, k_3, k_4 均为自适应控制参数。

* 收稿日期:2020-09-30 修回日期:2021-03-20 网络出版时间:2021-04-07 15:16

资助项目:宁夏自然科学基金(No. 2020AAC03030);国家自然科学基金(No. 12061055; No. 61662060);宁夏大学研究生创新项目(No. GIP2020-31)

第一作者简介:王倩,女,研究方向为人工智能的数学基础, E-mail: 2678118992@qq.com; 通信作者:李风军,男,教授,博士生导师, E-mail: fjli@nxu.edu.cn

网络出版地址: <https://kns.cnki.net/kcms/detail/50.1165.N.20210407.1457.016.html>

2 MAGA

2.1 编码和适应度函数

由于 SLFNs 的权值和阈值数量庞大,使用二进制编码将造成染色体过长,不利于遗传操作,因此本文采用实数编码,每个基因位的值代表一个权值或阈值。在分类模型中,平均绝对误差可以更好地表达预测数据与原始数据的差异。这里将预测数据反归一化后,大于 0.5 的数据作为一类,小于 0.5 的数据作为另一类,使得预测输出与期望输出均为 0-1 型变量,采用平均绝对误差(Mean absolute deviation, MAE)的倒数作为适应度函数

$$f = \frac{1}{S_{MAE}} = \frac{m}{\sum_{i=1}^m |\hat{y}_i - y_i|}, \text{ 其中: } y_i \text{ 为期望输出, } \hat{y}_i \text{ 为实际输出, } m \text{ 为训练样本数。}$$

2.2 选择操作的改进

本文引入四分位数方法。首先将个体按适应度进行升序,然后用 3 个点将个体适应度分为 4 等份,与这 3 个点位置上相对应的适应度值称为个体适应度的四分位数。在统计上认为,第 1 四分位数到第 3 四分位数体现了整个数据的分布情况。基于此,本文提出改进的选择算子:

1) 确定一个初始种群,计算种群内每个个体的适应度值;

2) 将种群内的个体按照适应度升序排列,并计算对应的索引值,得到升序后个体的染色体;

3) 将第 3 四分位数到最大适应度的个体替换最小适应度到第 1 四分位数的个体,即只选择第 1 四分位数到第 3 四分位数位置之间的个体以及第 3 四分位数到最大适应度的个体遗传到下一代,淘汰最小适应度到第 1 四分位数的个体,从而既具有良好的种群适应度,又维持种群多样性(适应度越大越好)。

2.3 自适应交叉和变异概率的改进

从(1)式和(2)式可以看出,AGA 更倾向于要改变适应度低于平均适应度的个体。对于较优良个体几乎不发生作用,使得收敛于局部极值的可能增大。本文引入四分位数间距(Interquartile range, IQR),IQR 的值 d_{IQR} 越小表明种群的适应度越集中,当选择交叉的两个个体的适应度大于平均适应度时,说明此时种群的整体适应度较好,应给予一个较小的权重,使得良好性状得以保留。因此 IQR 是一个合适的权重指标,并且能够根据种群适应度进行动态选择和动态变异。改进后的自适应交叉和变异概率的具体形式为:

$$P_c = \begin{cases} d_{IQR} \times \frac{f_{\max} - f'}{f_{\max} - f_{\text{avg}}}, & f' \geq f_{\text{avg}} \\ k_1, & f' < f_{\text{avg}} \end{cases}, P_m = \begin{cases} d_{IQR} \times \frac{(f_{\max} - f)}{f_{\max} - f_{\text{avg}}}, & f \geq f_{\text{avg}} \\ k_2, & f < f_{\text{avg}} \end{cases}。$$

为解决改进的自适应交叉和变异概率可能存在“早熟”问题,根据经验将上述公式做适当的线性变换,将交叉和变异个体适应度大于种群平均适应度的概率分别映射到固定区间 $[0.2, 0.5]$ 和 $[0.001, 0.004]$,使得在较优良个体中,MAGA 仍具有进化能力,从而可以摆脱局部收敛,找到全局最优值。

3 仿真实验

本文在 Matlab R2018b 的环境下,选取 6 个常用的 Benchmark 测试函数^[8,10],采用文献[6]中的交叉和变异操作,设置种群规模为 12 个,最大迭代次数为 160 次,AGA 中的自适应交叉和变异概率参数 $k_1=1, k_2=1, k_3=0.5, k_4=0.5$ 。MAGA 中的自适应交叉和变异概率参数 $k_1=1, k_2=1$ 。6 个测试函数的具体信息如表 1 所示,计算公式依次为:

$$\begin{aligned} f_1(x_1, x_2) &= (1 + (x_1 + x_2 + 1)^2 (19 - 14(x_1 + x_2) + 3(x_1^2 + x_2^2) + 6x_1x_2)) \times \\ &\quad (30 + (2x_1 - 3x_2)^2 (18 - 32x_1 + 12x_1^2 + 48x_2 - 36x_1x_2 + 27x_2^2)), \\ f_2(x_1, x_2) &= x_1^2 + x_2^2 - \cos(18x_1) - \cos(18x_2), \\ f_3(x) &= \sum_{j=1}^4 [100(x_j^2 - x_{j+1})^2 + (x_j - 1)^2], \\ f_4(x) &= \sum_{j=1}^5 x_j^2 + \left(\sum_{j=1}^n 0.5j x_j \right)^2 + \left(\sum_{j=1}^n 0.5j x_j \right)^4, \\ f_5(x) &= \sum_{i=1}^{10} |x_i|^{(i+1)}, \end{aligned}$$

$$f_6(x) = \sum_{i=1}^{20} x_i^2。$$

分别对 6 个测试函数运用 AGA 和 MAGA 模拟 100 次,两种算法对应的各函数最优值的收敛情况具体如图 1~图 3 所示。计算函数的最小值(最优值)、最大值(最差值)以及平均值,具体如表 2 所示。

从图 1 可以看出:函数 f_1, f_2 应用 MAGA 进行寻优的收敛曲线均在应用 AGA 的收敛曲线下方,表明 MAGA 具有更强的寻优能力,在同步进化中更能够靠近函数最小值。函数 f_1, f_2 应用 MAGA 进行寻优相比 AGA 具有更快的收敛速度。函数 f_2 应用 AGA 进行寻优在迭代结束后仍未达到全局最小值。

从图 2 可以看出:函数 f_3 应用 MAGA 进行寻优的收敛曲线前期在应用 AGA 的收敛曲线上方,但是随着迭代进行,能够迅速靠近函数的全局最优值。而函数 f_3 应用 AGA 进行寻优时,在迭代结束后仍未达到全局最小值。函数 f_4 应用 MAGA 进行寻优的收敛曲线前期略优于应用 AGA 的收敛曲线,在迭代 50 次后,均能很好逼近全局最小值。

从图 3 可以看出:函数 f_5 和 f_6 应用 MAGA 进行寻优的收敛曲线均在应用 AGA 的收敛曲线下方。函数 f_5 应用 MAGA 和 AGA 相继接近函数的全局最优值。而函数 f_6 应用 AGA 进行寻优时,在迭代结束后仍未达到全局最小值。

6 个测试函数的收敛曲线均表明,MAGA 相比 AGA 达到函数最优值的速度更快,能够摆脱局部最优解的束缚,即在较优良个体中仍具有良好的进化能力,改善了 AGA 容易陷入局部最优值的缺点。

表 1 6 个测试函数的基本信息

Tab. 1 Basic information of the 6 test functions

函数	自变量变量个数	定义域	全局最小值
$f_1(x)$	2	$[-2, 2]$	3
$f_2(x)$	2	$[-1, 1]$	-2
$f_3(x)$	5	$[-5, 10]$	0
$f_4(x)$	5	$[-5, 10]$	0
$f_5(x)$	10	$[-2, 2]$	0
$f_6(x)$	20	$[-2, 2]$	0

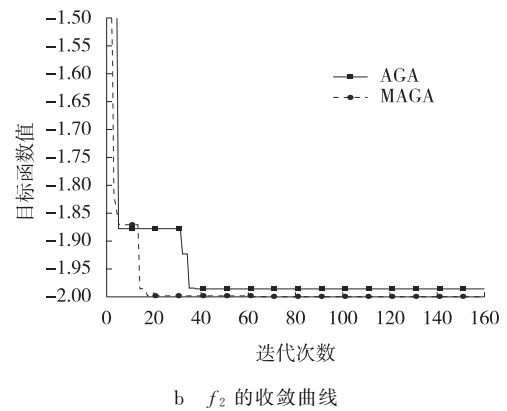
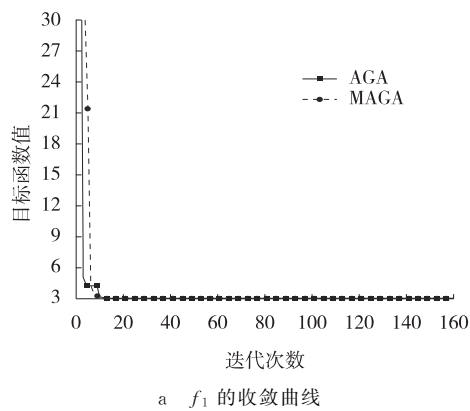


图 1 函数 f_1 和 f_2 的收敛曲线

Fig. 1 Convergence curves of functions f_1 and f_2

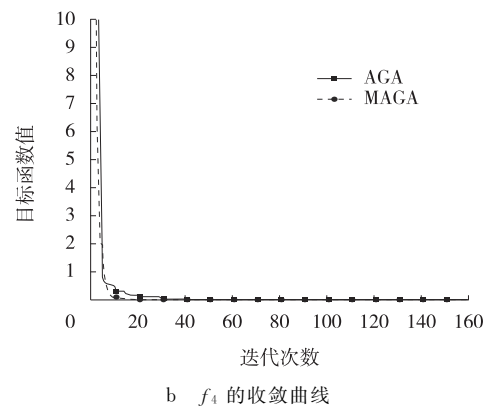
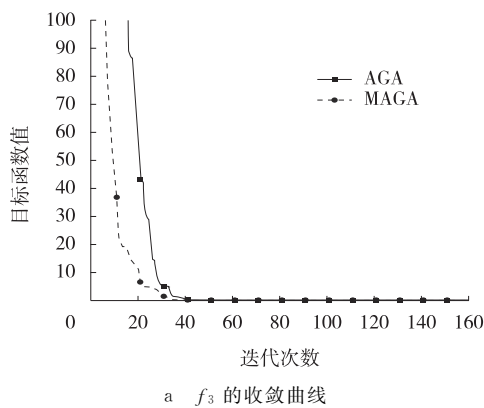


图 2 函数 f_3 和 f_4 的收敛曲线

Fig. 2 Convergence curves of functions f_3 and f_4

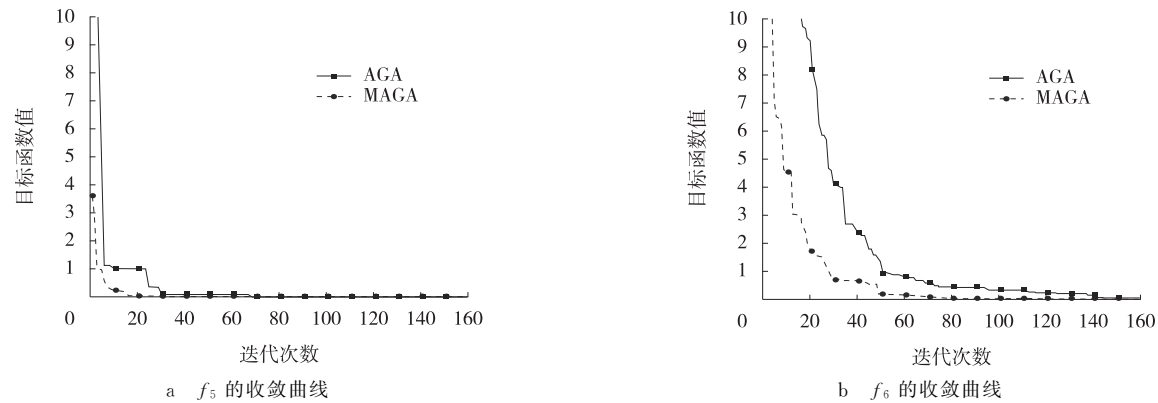


图 3 函数 f_5 和 f_6 的收敛曲线

Fig. 3 Convergence curves of functions f_5 and f_6

表 2 AGA、MAGA 计算结果

Tab. 2 Calculation results of AGA and MAGA

函数	全局最优值	算法	最优值	最差值	平均值
$f_1(x)$	3	AGA	3.000 0	84.003 6	5.663 5
		MAGA	3.000 0	3.051 5	3.006 2
$f_2(x)$	−2	AGA	−1.991 5	−0.839 5	−1.465 5
		MAGA	−2.000 0	−1.989 2	−1.991 5
$f_3(x)$	0	AGA	0.001 4	14.906 2	1.434 3
		MAGA	1.692 7e−07	13.182 6	0.869 7
$f_4(x)$	0	AGA	0.002 2	0.152 3	0.039 1
		MAGA	6.097 6e−05	0.032 3	0.008 5
$f_5(x)$	0	AGA	2.939 2e−04	0.081 7	0.022 2
		MAGA	2.366 3e−08	1.514 8e−04	1.719 2e−05
$f_6(x)$	0	AGA	0.075 3	0.424 4	0.202 6
		MAGA	0.015 7	0.222 9	0.080 7

从最优值角度看,两种算法均达到了函数 $f_1(x)$ 的全局最优值;MAGA 达到了函数 $f_2(x)$ 的全局最优值,AGA 可以近似达到函数 $f_2(x)$ 的全局最优值;对于函数 $f_3(x),f_4(x),f_5(x)$,两种算法均可以近似达到函数全局最优值,但 MAGA 的近似程度均显著高于 AGA;对于函数 $f_6(x)$,两种算法得到的最优值与函数全局最优值均具有一定差距,但 MAGA 的近似程度仍高于 AGA。

从最差值角度看,对于函数 $f_1(x)$,AGA 出现异常大的函数值,而 MAGA 在 100 次仿真模拟中最差值与最优值的差距极小;对于函数 $f_2(x),f_4(x),f_5(x),f_6(x)$,MAGA 在 100 次仿真模拟中最差值与最优值的差距均显著小于 AGA;对于函数 $f_3(x)$,两种算法均出现了与最优值相比差距较大的函数值。

从平均值角度看,MAGA 在 100 次仿真模拟中平均值与最优值的差距均显著小于 AGA。其中函数 $f_1(x),f_2(x),f_4(x),f_5(x)$ 的平均值与函数最优值十分接近,说明 MAGA 能够较稳定地对函数进行寻优。

4 MAGA 对 SLFNs 的优化算法

4.1 SLFNs

SLFNs 包括输入层、隐藏层和输出层。具有 d 个输入神经元, m 个隐藏神经元和一个输出神经元的 SLFNs 的数学表示如下^[3]:

$$N(\mathbf{x}) = \sum_{i=1}^m c_i \sigma \left(\sum_{j=1}^d w_{ij} x_j + \theta_i \right), \mathbf{x} = (x_1, x_2, \cdots, x_d) \in \mathbf{R}^d, d \geqslant 1.$$

其中: $\mathbf{w}_i = (w_{i1}, w_{i2}, \cdots, w_{id})^T \in \mathbf{R}^d$ 是隐藏层中第 i 神经元与输入神经元的连接权重, $c_i \in \mathbf{R}$ 是隐藏层中第 i 神经元与输出神经元的连接权重, $\theta_i \in \mathbf{R}$ 是阈值, σ 是激活函数。

4.2 MAGA-SLFNs 实施步骤

步骤 1:对于分类数据和回归数据,分别按类别标签和按大小比例,将数据同分布地划分为训练集和测试集,并进行归一化处理。

步骤 2:设置 SLFNs 的网络参数(输入神经元个数、隐藏层神经元个数、输出神经元个数、迭代次数、学习率、动量系数、目标精度)和 MAGA 参数(进化代数、种群规模、自适应交叉和变异概率参数)。

步骤 3:生成初始种群,并进行实数编码、计算初始适应度;

步骤 4:迭代求解最佳初始权值和阈值,首先根据改进的选择操作进行选择,然后根据改进的自适应交叉和变异概率进行交叉和变异,最后计算当代的适应度,并记录每代中最高的适应度。

步骤 5:若达到最大进化代数,则输出最优个体;否则,进入步骤 4。

步骤 6:将解码后的最优个体对 SLFNs 的初始权值和阈值进行赋值,建立 MAGA-SLFNs 模型。

5 实证分析

下面对一份肝病患者的实际数据来进行实证分析。这一实际数据来自 UCI 数据库的 ILPD,该数据集共收集了 416 份肝病患者和 167 份非肝病患者的资料。每份数据包括 8 个自变量:年龄(Age)、性别(Gender)、总胆红素(TB)、直接胆红素(DB)、碱性磷酸酶(AAP)、谷丙转氨酶(AL)、天冬氨酸转氨酶(AS)及总蛋白(TP);一个因变量:是否为肝病患者(是/否)。

首先将数据集按类别随机抽取 75%(438 份)为训练集,其中包括 309 个肝病患者和 129 个非肝病患者;余下的 25%(145 份)为测试集,其中包括 107 个肝病患者和 38 个非肝病患者。接着设置隐藏层神经元个数为 10,迭代次数为 1 000,学习率为 0.1,动量系数为 0.2,目标精度为 0.001,进化代数为 100。种群规模、自适应交叉和变异概率参数与仿真实验中一致。最后引入混淆矩阵对模型的预测能力进行评价,结果以 F1 度量呈现,具体如表 3 所示。

由表 3 可以看出 SLFNs 模型预测测试集中共 131 个肝病患者,其中 98 名预测正确,将 33 名非肝病患者预测为肝病患者,将 9 名肝病患者预测为非肝病患者,F1 度量值为 0.823 5。AGA-SLFNs 模型预测测试集中共 131 个肝病患者,其中 102 名预测正确,将 29 名非肝病患者预测为肝病患者,将 5 名肝病患者预测为非肝病患者,F1 度量值为 0.857 1。MAGA-SLFNs 模型预测测试集中共 126 个肝病患者,其中 103 名预测正确,将 23 名非肝病患者预测为肝病患者,将 4 名肝病患者预测为非肝病患者,F1 度量值为 0.884 1。相比 SLFNs 和 AGA 优化的 SLFNs,MAGA 优化的 SLFNs 的预测能力高于其他两种方法。

如图 4 所示,MAGA 的适应度曲线在迭代 60 次前均位于 AGA 上方,MAGA 在迭代不到 20 次达到最佳适应度,AGA 在迭代 60 次后达到相同适应度。适应度曲线表明 MAGA 能够迅速摆脱局部最优值的束缚,达到全局最优值。

表 3 3 种 SLFNs 模型预测结果的混淆矩阵

Tab. 3 Confusion matrix of the prediction results of the three SLFNs models

	预测值	真实值	
		是	否
SLFNs 预测值	是	98	33
	否	9	5
AGA-SLFNs 预测值	是	102	29
	否	5	9
MAGA-SLFNs 预测值	是	103	23
	否	4	15

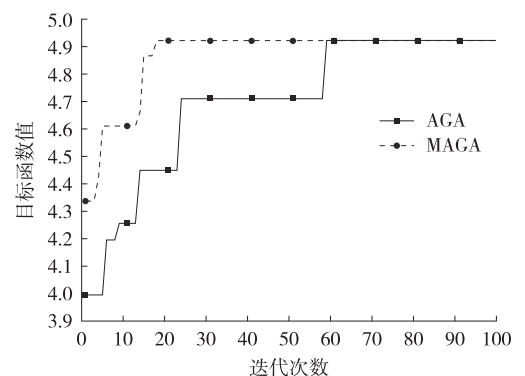


图 4 AGA 和 MAGA 的适应度曲线

Fig. 4 Fitness curve of AGA and MAGA

6 结束语

为解决 AGA 优化 SLFNs 存在未成熟收敛问题,提出了 MAGA 这一优化算法:首先通过引入四分位数改进了遗传操作中选择操作;其次以四分位数间距代替自适应遗传算法中固定的交叉和变异概率参数,并将改进后的交叉和变异概率投影到固定区间,使之在较优良个体中仍具有进化能力。最后利用 MAGA 对 SLFNs 进行优

化。仿真实验与实证分析均表明:MAGA 能够摆脱局部最优值的束缚,进一步寻找全局最优值,且在寻优的稳定性、收敛速度方面均有较大提升。

参考文献:

- [1] HOLLAND J H. Adaptation in natural and artificial systems[M]. Cambridge: MIT Press, 1975: 211-247.
- [2] SRINIVAS M, PATNAIK L M. Adaptive probabilities of crossover and mutation in genetic algorithms[J]. IEEE Trans on Systems Man and Cybernetics, 2002, 24(4): 656-667.
- [3] LI F J. Constructive function approximation by neural networks with optimized activation functions and fixed weights[J]. Neural Computing and Applications, 2019, 31: 4613-4628.
- [4] 周志华, 曹存根. 神经网络及其应用[M]. 北京: 清华大学出版社, 2004.
ZHOU Z H, CAO C G. Neural network and its application[M]. Beijing: Tsinghua University Press, 2004.
- [5] HOU W H, LIU X, CHEN D Q, et al. Finance-based scheduling for construction project in China with improved genetic algorithms[J]. Journal of Convergence Information Technology, 2012, 7(18): 29-38.
- [6] 陈闯, RYAD C, 邢尹. 改进遗传算法优化 BP 神经网络的语音情感识别[J]. 计算机应用研究, 2019, 36(2): 344-346.
CHEN C, RYAD C, XING Y. Improved genetic algorithm to optimize speech emotion recognition of BP neural network[J]. Application Research of Computers, 2019, 36(2): 344-346.
- [7] 杨从锐, 钱谦, 王锋, 等. 改进的自适应遗传算法在函数优化中的应用[J]. 计算机应用研究, 2018, 35(4): 1042-1045.
YANG C R, QIAN Q, WANG F, et al. Application of improved adaptive genetic algorithm in function optimization[J]. Application Research of Computers, 2018, 35(4): 1042-1045.
- [8] 刘浩然, 赵翠香, 李轩, 等. 一种基于改进遗传算法的神经网络优化算法研究[J]. 仪器仪表学报, 2016, 37(7): 1573-1580.
LIU H R, ZHAO C X, LI X, et al. A neural network optimization algorithm based on improved genetic algorithm[J]. Chinese Journal of Scientific Instrument, 2016, 37(7): 1573-1580.
- [9] 于光帅, 于宪伟. 一种改进的自适应遗传算法[J]. 数学的实践与认识, 2015, 45(19): 259-264.
YU G S, YU X W. An improved adaptive genetic algorithm[J]. Mathematics in Practice and Knowledge, 2015, 45(19): 259-264.
- [10] 张京钊, 江涛. 改进的自适应遗传算法[J]. 计算机工程与应用, 2010, 46(11): 53-55.
ZHANG J Z, JIANG T. Improved adaptive genetic algorithm[J]. Computer Engineering and Applications, 2010, 46(11): 53-55.
- [11] SHELOKAR P S, SIARRY P, JAYARAMAN V K, et al. Particle swarm and ant colony algorithms hybridized for improved continuous optimization[J]. Applied Mathematics and Computation, 2007, 188(1): 129-142.
- [12] 孙建永, 申建中, 徐宗本. 一类自适应遗传算法[J]. 西安交通大学学报, 2000(10): 84-88.
SUN J Y, SHEN J Y Z, XU Z B. New adaptive genetic algorithm[J]. Journal of Xi'an Jiaotong University, 2000(10): 84-88.

Operations Research and Cybernetics

Improved Adaptive Genetic Algorithm and Its Application

WANG Qian, LI Fengjun

(School of Mathematical Statistics, Ningxia University, Yinchuan 750021, China)

Abstract: [Purposes] To improve the premature problem of adaptive genetic algorithm (Adaptive genetic algorithm, AGA), and improve the prediction performance of the neural network model optimized by the AGA algorithm. [Methods] The fixed parameters of adaptive crossover and mutation probability are replaced by interquartile distance. First, a modified adaptive genetic algorithm (Modified adaptive genetic algorithm, MAGA) is proposed, and its adaptive crossover and mutation probability is mapped to the fixed interval, so that the excellent individuals still have good evolutionary ability under the action of MAGA. Secondly, the initial weights and thresholds of single hidden layer feedforward neural networks (Single hidden layer feedforward neural networks, SLFNs) are optimized by using MAGA, and the implementation steps of the algorithm are given. [Findings] The simulation experiments of six commonly used test functions show that MAGA can quickly and stably approximate the optimal solution of the function. [Conclusions] Empirical analysis shows that compared with SLFNs and AGA-optimized SLFNs, MAGA-optimized SLFNs have a higher classification accuracy rate.

Keywords: adaptive genetic algorithm; single hidden layer feedforward neural networks; dynamic selection; dynamic mutation; optimization

(责任编辑 许 甲)