

基于追逐粒子与自适应变异的 PSO-SVM 算法 与养殖水质评价研究

魏超¹, 张刚^{1*}, 韩祥兰², 蔡永胜¹, 冯志敏¹

(1. 宁波大学 海运学院, 浙江 宁波 315211; 2. 浙江大学宁波理工学院 信息与计算机科学分院, 浙江 宁波 315110)

摘要: 为提高标准 PSO 算法对 SVM 参数优化选取的精度和速度, 提出以粒子群适应度均方差为判别依据, 用追逐粒子位置替代陷入局部优化时的群体最优位置, 对粒子的速度进行自适应调整; 利用两个距离最远的粒子位置构成的圆形区域作为最优位置的吸引域, 引入位置变异因子, 使粒子随着迭代次数有规律地跳出局部最优位置吸引域, 从而对粒子位置进行自适应变异. 通过对 4 种不同适应度函数迭代 100 次的收敛测试, 并与另外 2 类不同粒子群算法比较, 结果表明改进算法收敛精度提高了 20 倍以上. 将改进的 PSO-SVM 算法与另外 3 种分类算法对采集的 5 439 组养殖水质数据按照从优到差 5 个级别进行水质评价分类试验, 发现改进的 PSO-SVM 算法收敛速度和收敛精度最高, 对水质评价的错误率为 1.54%, 能有效地满足实际水质分类需求.

关键词: 水质评价; 支持向量机; 粒子群算法; 自适应变异

中图分类号: X824

文献标志码: A

文章编号: 1001-5132 (2018) 03-0020-07

近年来, 我国水产养殖量已占全世界水产养殖总产量约 70%^[1]. 如此大规模的水产养殖生产, 如不能对养殖水质进行实时监测分析, 极易造成难以挽回的巨大经济损失^[2-3]. 因此, 对水质的实时监测分类是当下水产养殖行业的研究热点之一.

支持向量机(Support Vector Machine, SVM)是一种解决分类、回归等机器学习问题的新型工具, 在养殖水质分类中具有较高的应用价值. 管军等^[4]利用决策二叉树分类思想构建了基于支持向量机的长江水质评价模型; 陈耀宁等^[5]构建了白云湖水质的支持向量机模型, 建立了五级的水质评价体系, 得到较好的分类效果. 但传统的 SVM 算法不能准确地选取最优参数, 在水质评价这样的非线性分类问题中极易影响分类的精度与效率. 刘帮等^[6]通过粒子群算法对 SVM 参数选取进行了优化, 并将其应用到污水水质建模中, 提高了模型泛化能力; 王道明等^[7]将粒子群算法(Particle Swarm Optimization, PSO)代入决策树型 SVM 网络

进行训练, 使得分类准确率得到明显改善. 但加入粒子群算法的 SVM 模型对参数的搜索效率有限; 另外, PSO-SVM 算法由于搜索机制的限制, 粒子在迭代过程中极易聚集在某一点造成算法早熟和陷入局部最优, 从而使得参数选取的结果不能使适应度值达到最优. 高立群等^[8]将遗传算法中的变异操作引入粒子群算法中, 通过引入一个简单的变异算子, 一定程度上解决了粒子逃离局部最优的问题, 但不可控的变异也使得整个算法后期收敛速度很慢. 因此, 寻求一种新的变异机制使粒子有目的地选取全局最优和有规则地跳出局部最优, 并且使后期收敛速度不受变异机制影响, 这需要作进一步研究.

本文针对养殖水质分类提出一种新的变异机制, 自适应地调整粒子速度与位置, 以 SVM 网络训练为适应度函数取得最优参数的优化算法; 用粒子适应度值均方差表示粒子当前的集散状态, 在粒子聚集时采用一个在搜索空间内随机运动的

收稿日期: 2017-12-05.

宁波大学学报(理工版)网址: <http://journalg.nbu.edu.cn/>

基金项目: 国家自然科学基金(51675286); 宁波市自然科学基金(2017A610085); 浙江省公益技术项目(2017C32014).

第一作者: 魏超(1993-), 男, 安徽六安人, 在读硕士研究生, 主要研究方向: 轮机工程. E-mail: 1608317146@qq.com

*通信作者: 张刚(1976-), 男, 四川遂宁人, 博士/高级工程师, 主要研究方向: 船舶自动化与机电控制. E-mail: zhanggang@nbu.edu.cn

追逐粒子位置替代全局最优位置, 更新粒子在解空间中的速度; 定义一个自适应调整的变异因子, 使粒子通过有规律地跳出局部最优吸引域进行位置变异, 并将其应用在水质分类试验中. 对试验提取的各类水质样本利用改进算法进行训练, 通过对水质参数的融合处理对养殖水质进行等级评判.

1 算法原理

1.1 支持向量机

支持向量机^[9]虽是一种二类分类算法, 但改进的 SVM 也可用来解决多分类问题, 并且分成线性和非线性两类情况^[10].

对于线性分类问题, 设满足线性可分的样本集为 (x_i, y_i) , $i=1, 2, \dots, n$, $x_i \in R$, $y_i \in \{-1, 1\}$, 线性函数为 $g(x) = w \cdot x + b$, 超平面 $w \cdot x + b = 0$ (b 是间隔向量, w 是法向量)将数据空间分成两个区域, 将问题转化为求解优化问题:

$$\min \|w\|^2 / 2 + C \sum_{i=1}^l \xi_i, \quad \text{s.t. } y_i(w \cdot x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i=1, 2, \dots, l, \quad (1)$$

其中: C 为惩罚参数, 表示对分类错误的惩罚程度; ξ_i 为非负松弛变量, 当出现分类面划分错误时, $\xi_i > 0$. 求解最优超平面转化为凸规划问题, 一般引入 Lagrange 函数进行求解, 最终所求的决策函数为:

$$f(x) = \text{sgn}(w \cdot x + b). \quad (2)$$

对于非线性分类问题, 首先通过映射函数 $\varphi(x)$, 核函数 $k(x_i, x_j) = \varphi(x_i) \cdot \varphi(x_j)$, 将训练样本从低维空间映射到高维特征空间, 使其在高维空间中线性可分, 并构造最优分类超平面. 非线性分类支持向量机算法为:

$$\min \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j k(x_i, x_j) - \sum_{j=1}^l \alpha_j, \quad \text{s.t. } \sum_{i=1}^l y_i \alpha_i = 0, 0 \leq \alpha_i \leq C, i=1, 2, \dots, l, \quad (3)$$

这是典型的二次规划问题, 令

$$f_i = f(x_i) = \text{sgn}(\sum_{i=1}^l \alpha_i^* y_i k(x_i \cdot x_j) + b^*),$$

根据最优化问题 KKT(Karush Kuhn Tucker)条件, 得 Lagrange 乘子的特解 α^* , 最终构造决策函数为:

$$f(x) = \text{sgn}(\sum_{i=1}^l \alpha_i^* y_i k(x_i \cdot x_j) + b^*). \quad (4)$$

1.2 粒子群算法

粒子群优化算法是除蚁群算法、鱼群算法外又一种群体智能优化算法^[11-12]. 该算法在解空间内初始化生成一群粒子, 用位置信息、速度信息和适应度值三项指标来表示该粒子的特征. 当粒子在解空间内搜索时, 不断迭代更新位置信息与适应度值, 跟踪、比较、更新并朝着自身最优和全局最优位置不断运动调整, 实现在可解空间内的寻优.

记由 N 个粒子组成的种群为 $x = (x_1, x_2, \dots, x_n)$, 第 j 个粒子在 t 维参数搜索区域内的位置信息为 $x_j = (x_{j1}, x_{j2}, \dots, x_{jt})^T$. 将其带入适应度函数 $f_i = f(x)$ 进行训练, 得粒子适应度值. 记第 j 个粒子的速度为 $v_j = (v_{j1}, v_{j2}, \dots, v_{jt})^T$, 个体极值为 $p_j = (p_{j1}, p_{j2}, \dots, p_{jt})^T$, 群体极值为 $G_j = (G_{j1}, G_{j2}, \dots, G_{jt})^T$, 则在每一次迭代过程中, 粒子通过个体极值 p_j 和群体极值 G_j 来更新粒子的速度和位置, 即

$$V_{jt}^{i+1} = wv_{jt}^i + c_1 r_1 (p_{jt}^i - x_{jt}^i) + c_2 r_2 (G_{jt}^i - x_{jt}^i), \quad (5)$$

$$x_{jt}^{i+1} = x_{jt}^i + v_{jt}^{i+1}, \quad (6)$$

其中: w 为惯性权重, 一般取 0.5; $j=1, 2, \dots, n$; c_1, c_2 为加速度因子; r_1, r_2 为分布于 $[0, 1]$ 之间的随机数; i 为当前迭代次数; v_{jt}^i 为粒子速度; p_{jt}^i 为第 i 次迭代粒子群个体最优位置; G_{jt}^i 为第 i 次迭代粒子群群体最优位置; x_{jt}^i 为粒子位置.

2 改进算法

2.1 追逐粒子的引入

对于群体最优位置 G_t , 其表示整个粒子群对最优目标信息的获取, 是对群体本身的认知, 一旦粒子在早期运动方式不合理, 有可能导致 G_t 在前期趋向稳定, 而粒子朝着最优目标不断搜索这一算法机制又会进一步加深 G_t 陷入局部最优, 进而出现早熟现象. 对 G_t 变异流程如下:

(1) 建立一个能够监测粒子聚散状态的评价方法, 以每代粒子适应度均方差 S 反应粒子间聚集程度. 设迭代总次数为 K , 在第 i 次迭代中, 每个粒子的适应度值是 $f_{ij} = (f_{i1}, f_{i2}, \dots, f_{in})$, 粒子群适应度平均值为 $\bar{F}_i = (\bar{F}_{i1}, \bar{F}_{i2}, \dots, \bar{F}_{ik})$, 则第 i 次迭代中粒子群 X 的适应度均方差为 S_i :

$$S_i = \frac{1}{n} \sqrt{\sum_{j=1}^n (f_{ij} - \bar{F}_i)^2}, \quad (7)$$

当粒子发生聚集时,其适应度值均方差接近于 0,因此利用粒子适应度值的均方差可以有效地反应粒子之间的聚集程度。

(2)加入一个在搜索区域内不受粒子群条件限制的粒子 Z 进行随机运动,每一次迭代过程中,更新一次粒子 Z 的随机位置信息,该粒子称之为追逐粒子;记录追逐粒子 Z 在第 i 次迭代时位置信息为:

$$Z_i = (Z_{ci}, Z_{gi}). \quad (8)$$

(3)设 S_0 为改进算法粒子适应度均方差的最小阈值,在正常情况下粒子不可能完全聚集在一点,因此 $S_i > 0$,经过多次对水质数据的实验证明, S_0 一般取 2.5 ± 0.5 . 当 $S_i > S_0$ 时,表明此时粒子在解空间范围内是正常搜索状态,由此 G_i 将不发生变异,此时粒子飞行速度表示为:

$$v_{jt}^{i+1} = wv_{jt}^i c_1 r_1 (p_{jt}^i - x_{jt}^i) + c_2 r_2 (G_{jt}^i - x_{jt}^i). \quad (9)$$

当 $S_i \leq S_0$ 时,粒子在搜索区域内出现聚集状态,表明算法寻优陷入了局部最优值,此时追逐粒子 Z 的位置 Z_i 代替 G_i 成为粒子群当代最优目标位置,更新粒子的飞行速度 v_{jt} 与方向:

$$v_{jt}^{i+1} = wv_{jt}^i c_1 r_1 (p_{jt}^i - x_{jt}^i) + c_2 r_2 (Z_{pt}^i - x_{jt}^i), \quad (10)$$

其中: Z_{pt} 为追逐粒子的位置。

2.2 自适应变异因子

以局部最优位置与最远粒子位置构成的圆形区域确定为该收敛区域的吸引域,引入一个变异因子 h 和函数 $F(i)$,根据两者的大小决定是否让粒子跳出该吸引域,加入自适应变异因子过程为:

(1)设变异因子 $h(i)$ 为一个随迭代每次产生(0,1)分布的随机变量.将每一代产生的 $h(i)$ 值与函数 $F(i)$ 在该代产生的函数值进行比较, $F(i)$ 是一个随着迭代次数增加而线性递增的函数,即

$$F(i) = i(f_2 - f_1) / k + f_1, \quad (11)$$

其中: f_1, f_2 为函数两个端值; K 为迭代总次数; i 为当前迭代次数.经实验证明 $f_1=0.5, f_2=1$ 较为合理。

(2)定义一个以适应度值最优的粒子 X_m 的位置为中心点,该粒子与适应度值最差的粒子 X_n 的距离为半径的圆形区域为该局部收敛区域的吸引域 $Q(x)$,即

$$Q(x) = \pi((x_{m1} - x_{n1})^2 + (x_{m2} - x_{n2})^2). \quad (12)$$

(3)变异因子 h 与函数 $F(i)$ 比较的结果决定粒子位置是否发生变异.当随机产生的变异因子 h 出

现 $h(i) < F(i)$ 时,说明此时外界环境不再允许粒子位置发生变异,位置表示为:

$$x_{jt}^{i+1} = x_{jt}^i + v_{jt}^{i+1}, \quad (13)$$

此时粒子将按照上一代更新得来的速度与位置进行适应度计算:当随机产生的变异因子 h 出现 $h(i) \geq F(i)$ 时,通过计算最优位置吸引域来使聚集的粒子跳出到吸引域之外进行空间搜索,对粒子位置进行变异,即

$$x_i(j,1)' = \sqrt{(x_{m1} - x_{n1})^2 + (x_{m2} - x_{n2})^2} (1 + R_{\text{rand}}) \cdot \cos(360R_{\text{rand}} \cdot \pi/180) + x_i(j,1) + v_i(j,1), \quad (14)$$

$$x_i(j,2)' = \sqrt{(x_{m1} - x_{n1})^2 + (x_{m2} - x_{n2})^2} (1 + R_{\text{rand}}) \cdot \sin(360R_{\text{rand}} \cdot \pi/180) + x_i(j,2) + v_i(j,2), \quad (15)$$

其中: $(x_i(j,1), x_i(j,2))$ 为第 i 次迭代中第 j 个粒子的位置坐标; $v_i(j,1), v_i(j,2)$ 为粒子在第 i 次迭代中在两个方向上的分速度; R_{rand} 为(0,1)随机数生成函数。

这种新的变异机制在前期可以有效地避免算法出现早熟收敛,并且这种粒子位置变异的可能性由于函数 $F(i)$ 的递增效应会随着迭代次数的增加而不断减小,保证了改进算法在后期的收敛速度.算法详细步骤为:

Step 1 初始化粒子位置和速度,设置参数 c_1 和 c_2 ,生成随机数 r_1 和 r_2 .

Step 2 计算粒子初始适应度值 f_1 .

Step 3 寻找个体极值 p_i 和群体极值 G_i .

Step 4 进行算法迭代,定义追逐粒子 Z_i ,更新当代粒子速度 v_{mt} 与位置 x_{mt} ,计算上一代群体最优位置吸引域 $Q(x)$,用粒子适应度均方差 S_i 和 S_0 进行比较: (1)若 $S_i \leq S_0$,则用式(10)计算粒子速度,然后比较 $h(i)$ 与 $F(i)$;若 $h(i) \geq F(i)$,则粒子的位置根据式(14)与式(15)发生变异,根据新的粒子速度与位置计算粒子当代适应度,转到 Step 5;若 $h(i) < F(i)$,则粒子的位置用式(13)计算,不发生位置变异,根据得到的粒子速度与位置计算粒子当代适应度,转到 Step 5. (2)若 $S_i > S_0$,则根据式(5)和式(6)计算,根据得到的粒子速度与位置计算粒子当代适应度,转到 Step 5.

Step 5 更新 p_i 和 G_i .

Step 6 判断是否满足算法终止条件;若满足,转 Step 7;否则转 Step 4.

Step 7 结束算法.

改进算法流程图如下:

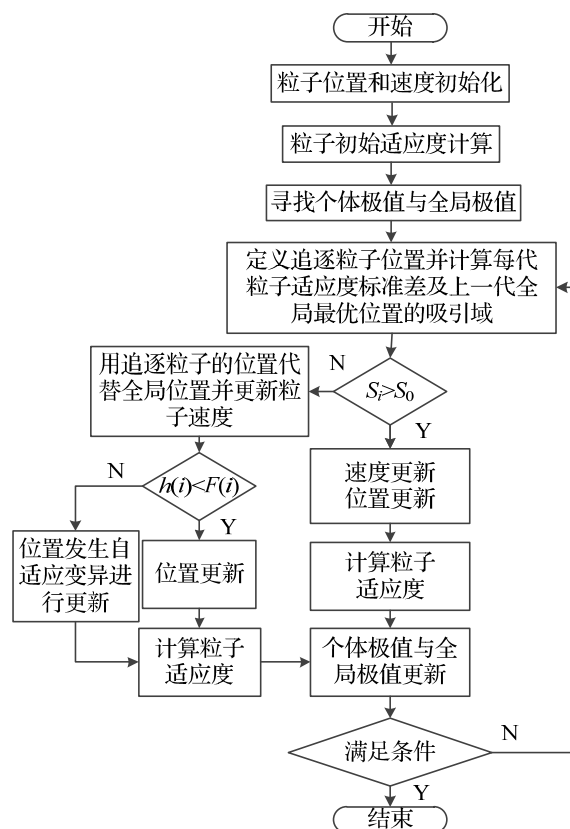


图1 改进算法流程图

3 算法测试与实验研究

3.1 算法分析测试

为了测试改进算法(简称 WPSO, 其中 W 表示追逐粒子)的收敛性能和逃出局部最优位置的能力, 将其与标准 PSO 算法和加入简单变异算子的 PSO 算法(简称 byPSO)一起进行仿真对比实验. 实验参数设置为: 粒子种群规模 $N=20$, 最大迭代次数 $K=100$, 加速度因子 $c_1=1.5$, $c_2=1.7$, 位置搜索区间为 $[-8, 8]$, 速度限制在 $[-1, 1]$.

选取 Rosenbrock、Rastrigin、 $f(x) = \sum_{i=1}^n x_i^2$ 以及 Ackley 函数^[13]分别对标准 PSO 算法、改进算法和加入简单变异 PSO 算法进行对比测试, 测试结果见图 2~5 和表 1.

由图 2~5 可知, 当粒子迭代时, 标准 PSO 算法更容易陷入局部最优值, 且难以跳出, 最大寻优步数平均为 88 代. 变异 PSO 算法在陷入局部最优时有较好的逃离能力, 但后期算法收敛速度较慢, 且

精度不高, 在 80 代左右达到最优适应度. WPSO 算法在早期得益于追逐粒子的代入, 使得算法在前期不易陷入局部最优而产生早熟收敛. 随着迭代次数的增加, 变异可能性随线性递增函数数值的增加而降低, 因此提升了后期的收敛能力, 达到最优精度时的迭代数平均为 55 代.

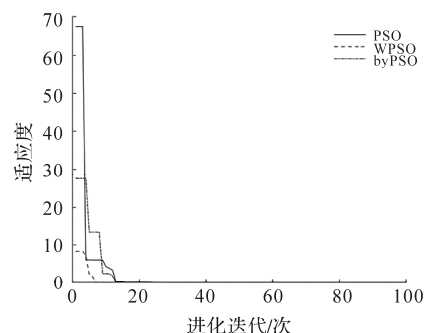


图2 Rosenbrock 函数收敛实验

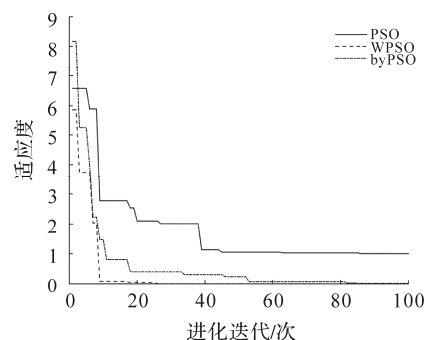


图3 Rastrigin 函数收敛实验

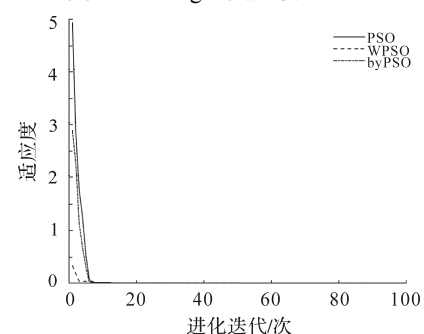


图4 $f(x)$ 函数收敛实验

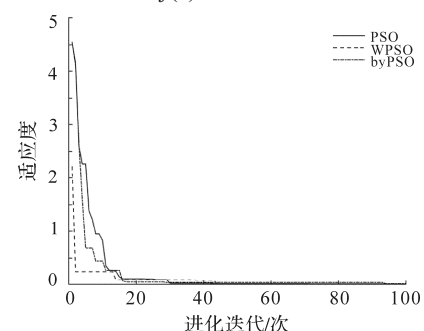


图5 Ackley 函数收敛实验

表1 3种算法在不同适应度函数下的收敛精度

算法	函数			
	Rosenbrock	Rastrigin	$f(x)$	Ackley
PSO	0.003 3	0.996 2	$2.417 7 \times 10^{-5}$	0.004 1
WPSO	$2.776 8 \times 10^{-6}$	$8.490 3 \times 10^{-4}$	0	$8.881 8 \times 10^{-16}$
byPSO	$8.387 1 \times 10^{-4}$	0.016 1	$2.501 9 \times 10^{-4}$	0.011 8

从表1可看出,在寻优步数上WPSO算法比标准PSO与byPSO算法平均减少了37.5%与31.3%。WPSO算法在寻优步数和收敛精度方面优于另两种PSO算法。

3.2 水质实验研究

根据《渔业养殖水质标准》^[14]选用温度、pH、氨氮、溶解氧和含盐量作为水质监测5类参数,通常这5类参数的融合效果可有效地反应整体水质,其融合结果按照水质由好到差依次分为I类、II类、III类、IV类和V类。

3.2.1 实验数据采集

在水产养殖现场搭建基于水质传感器-PLC-上位机的水质监测平台,将采集到的水质数据自动存储于Access数据库中^[15],共取得5439组数据,按时间均匀提取1128组为测试样本,剩下4311组为训练样本。实验现场如图6和图7所示。



图6 数据采集硬件设备



图7 工作现场图

3.2.2 构建ZPSO-SVM网络模型

首先对训练集4311组样本进行[0,1]归一化处理,设定SVM核函数为径向基函数(Radial Basis Function, RBF),引入追逐粒子并使算法进行200次迭代更新。分别用标准粒子群算法和改进的粒子群算法WPSO观察每代粒子的适应度均方差,以此判别粒子的运动形式与聚集状态,得到粒子每代均方差如图8和图9所示。

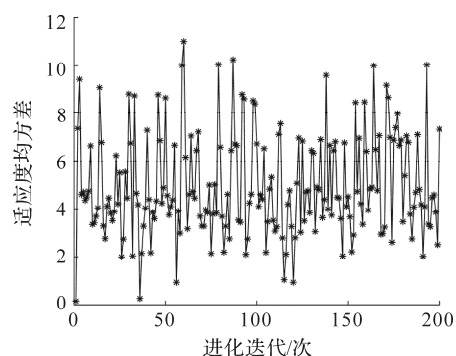


图8 PSO算法每代均方差

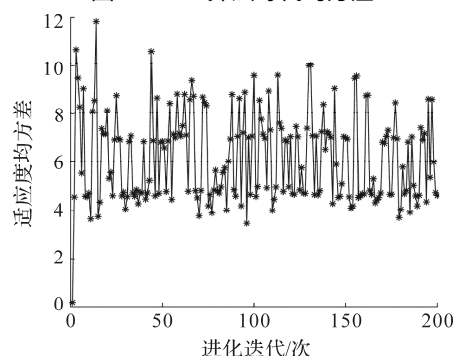


图9 WPSO算法每代均方差

从图8可看出,在标准PSO下整体显示出粒子在没有变异的环境下常会陷入局部最优。而图9显示,在WPSO算法下,由于新变异机制的引入,使得粒子呈分散搜索状态,极少陷入局部最优值。在WPSO-SVM对训练集样本分类完成时,得到训练集分类准确率为99.1696%, $c=86.4612$, $g=96.2636$,模型耗时 $t=3627.231s$;标准PSO-SVM得到训练集分类准确率为99.0773%,耗时为 $t=4601.632s$ 。准确率得到了提升,算法收敛时间缩短了21.2%。

3.2.3 实验结果

分别用WPSO-SVM模型、传统SVM模型、BP神经网络模型、标准PSO-SVM模型对1128组测试样本集进行养殖水质级别分类实验,比较水质分类准确率,结果见图10~13和表2。

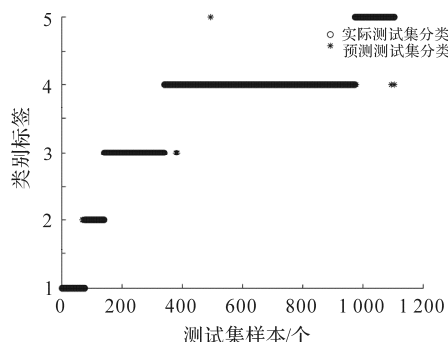


图10 WPSO-SVM 模型预测分类结果

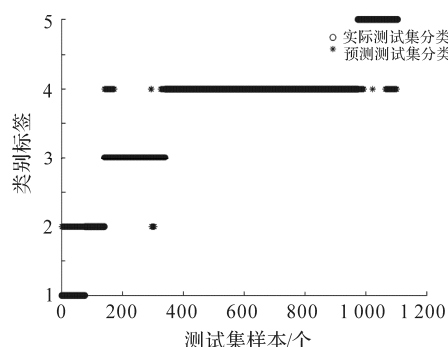


图11 传统 SVM 模型预测分类结果

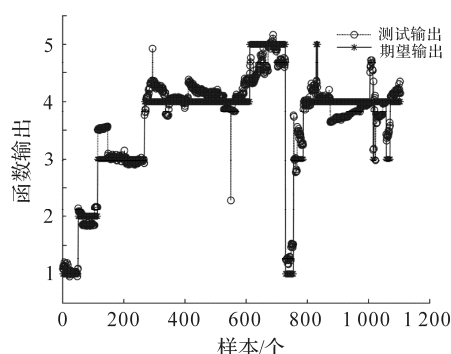


图12 BP 神经网络模型预测分类结果

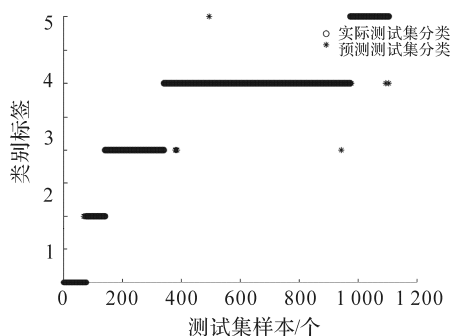


图13 PSO-SVM 模型预测分类结果

对比图10与图13发现, WPSO 算法相比于标准 PSO, 充分利用了算法本身更强的寻优能力, 得到了更加精确的模型参数, 可获得更好的分类性能。从图11和图12可知, 基于 BP 神经网络建立

的水质分类模型得到的分类效果处于传统 SVM 与引入 PSO 的模型分类效果之间, 采用 BP 方法可获得比较客观的评价结果, 但需要建立在大量数据样本的基础上。从实验结果看, WPSO-SVM 算法相比于 BP 神经网络有着更稳定的模型结构和更优的分类准确率。

表2 4种养殖水质分类模型对比

参数	BP 神经网络	SVM	PSO-SVM	WPSO-SVM
所用时间/s	0.967	0.738	4 601.632	3 627.231
分类准确率/%	95.20	83.06	98.18	98.46

从表2可看出, 加入粒子群算法的 PSO-SVM 和 WPSO-SVM 模型分类准确率都提高到了 98% 以上, WPSO-SVM 算法比 BP 神经网络算法的精度提升了 3.2%, 比传统 SVM 分类算法得到的分类准确率提升了 15.4%。改进的 WPSO-SVM 模型不论在训练时间还是分类精度上, 均要优于标准 PSO-SVM 模型。

4 结语

本文将一种改进的粒子群算法与支持向量机相结合, 利用这种新的变异机制对 SVM 中的参数进行优化选取, 得到以下结论:

(1) 引入粒子群算法建立的 PSO-SVM 水质分类模型, 预测值与实际值偏差在 2% 以内, 表明 PSO 算法引入的有效性。但是算法时间较长, 在生产过程中应提前完成建模。

(2) 用改进 PSO 算法、标准 PSO 和一般变异 PSO 算法对 4 种代表性适应度函数进行分析, 改进 PSO 算法收敛精度相比于另外 2 种算法有显著提升, 算法收敛时间缩短了 21.2%。表明追逐粒子和自适应变异引入的有效性和可靠性。

(3) 通过与 BP 神经网络、传统 SVM 算法和标准 PSO-SVM 算法对水质数据进行评价分类, 改进 PSO-SVM 模型对测试集样本分类的准确率最高可达 98.460 1%, 能有效地满足水产养殖等实际工程分类要求, 具有重要的工程应用价值。

参考文献:

- [1] 董楠楠. 世界水产品贸易竞争力与产业内贸易分析[J]. 宁波大学学报(理工版), 2007, 20(2):262-267.
- [2] 胡金有, 王靖杰, 张小栓, 等. 水产养殖信息化关键技术

- 术研究现状与趋势[J]. 农业机械学报, 2015, 46(7): 251-263.
- [3] Costea R I. ISO/TC 234-fisheries and aquaculture[J]. Aquaculture Aquarium Conservation & Legislation, 2009, 2(1):51-56.
- [4] 管军, 徐立中, 石爱业. 基于支持向量机的水质监测数据处理及状况识别与评价方法[J]. 计算机应用研究, 2006, 22(9):36-38.
- [5] 陈耀宁, 朱刚刚, 唐清华, 等. 基于支持向量机的白云湖水质评价[J]. 环境工程学报, 2014, 8(12):5535-5540.
- [6] 刘帮, 秦斌, 王欣, 等. PSO-LIBSVM 在污水水质建模中的应用[J]. 湖南工业大学学报, 2015, 29(2):89-93.
- [7] 王道明, 鲁昌华, 蒋薇薇, 等. 基于粒子群算法的决策树 SVM 多分类方法研究[J]. 电子测量与仪器学报, 2015, 29(4):611-615.
- [8] 高立群, 吴沛峰, 邹德旋. 基于变异策略的粒子群算法[J]. 东北大学学报(自然科学版), 2010, 31(11):1530-1533.
- [9] Liu W, Wang G Y, Fu J Y, et al. Water quality prediction based on improved wavelet transformation and support vector machine[J]. Advanced Materials Research, 2013, 2480(726):3547-3553.
- [10] 吴一全, 殷骏, 戴一冕, 等. 基于蜂群优化多核支持向量机的淡水鱼种类识别[J]. 农业工程学报, 2014, 30(16):312-319.
- [11] Ayed S, Mahamed G H, Omran M C, et al. Improving the performance of comprehensive learning particle swarm optimizer[J]. Journal of Intelligent & Fuzzy Systems, 2015, 30(2):735-746.
- [12] 王东风, 孟丽, 赵文杰. 基于自适应搜索中心的骨干粒子群算法[J]. 计算机学报, 2016, 39(12):2652-2667.
- [13] 纪震, 廖惠连, 吴青华. 粒子群算法及应用[M]. 北京: 科学出版社, 2009:74-78.
- [14] GB11607-89. 渔业水质标准[S].
- [15] 黄衍标, 罗广岳, 何铭金. BP 神经网络在巡逻机器人多传感器数据融合中的应用[J]. 传感技术学报, 2016, 29(12):1936-1940.

PSO-SVM algorithm based on chasing particle and adaptive mutation and evaluation of aquaculture water quality

WEI Chao¹, ZHANG Gang^{1*}, HAN Xiang-lan², CAI Yong-sheng¹, FENG Zhi-min¹

(1. Faculty of Maritime and Transportation, Ningbo University, Ningbo 315211, China; 2. Branch of Information and Computer Science, Ningbo Institute of Technology, Zhejiang University, Ningbo 315110, China)

Abstract: Aim to improve the accuracy and speed of SVM classification in water quality evaluation, the PSO optimization method is applied to optimize the key parameters of SVM algorithm in this paper. A new adaptive adjustment of particle velocity method is proposed, which is based on the consideration of taking particle swarm fitness variance as evaluation criterion and using chasing particle position to replace local optimization particle position. First, an attracting domain is defined by the distance between farthest two particles. Then, by position variation factor making the particle jump out of the local optimal attracting domain regularly with the constantly iterations. Thus, the particle location can be adaptively mutated. Convergence tests on four different functions shows that the proposed improved algorithm is increased by more than 20 times in convergence accuracy comparing to another two kinds of particle swarm algorithm. Last, the improved PSO-SVM is applied in five-level water quality classification test based on the collected 5 439 group water quality data. Contrast with the other three classification methods, the improved PSO-SVM algorithm has the highest convergence rate and convergence accuracy. The error rate of water quality evaluation is 1.54%, which can effectively meet the actual water quality classification requirements.

Key words: water quality evaluation; support vector machine; particle swarm algorithm; adaptive mutation

(责任编辑 史小丽)