

文章编号: 1001-0920(2013)03-0413-05

基于粒子群优化算法的最大相关最小冗余 混合式特征选择方法

姚旭, 王晓丹, 张玉玺, 权文

(空军工程大学 防空反导学院, 西安 710051)

摘要: 在分析粒子群优化(PSO)算法和简化PSO算法的基础上, 提出一种基于PSO的最大相关最小冗余的Filter-Wrapper混合式特征选择方法. Filter模型是基于互信息和特征的相关冗余综合测度, Wrapper模型是基于改进的简化粒子群算法. 在PSO搜索过程中, 引入相关冗余度量标准来选择特征子集, 将Filter融合在Wrapper中, 利用Filter的高效率和Wrapper的高精度提高搜索的速度和性能. 最后以支持向量机(SVM)为分类器, 在公共数据集UCI上进行实验, 实验结果表明了所提出算法的可行性和有效性.

关键词: 特征选择; 粒子群优化; Filter; Wrapper; 互信息

中图分类号: TP391

文献标志码: A

A maximum relevance minimum redundancy hybrid feature selection algorithm based on particle swarm optimization

YAO Xu, WANG Xiao-dan, ZHANG Yu-xi, QUAN Wen

(School of Air and Missile Defense, Air Force Engineering University, Xi'an 710051, China. Correspondent: YAO Xu, E-mail: ffly132@163.com)

Abstract: A Filter-Wrapper hybrid feature selection approach with maximum relevance and minimum redundancy based on particle swarm optimization(PSO) algorithm is proposed on the analysis of PSO algorithm and simplified PSO algorithm. The Filter is based on mutual information and the composite measure of feature relevance and redundancy, while the Wrapper is based on a simply modified PSO algorithm. The relevance and redundancy criterion is introduced to select features in the PSO's searching procedure. Meantime, the Filter is fused into the Wrapper. The speed and performance of the search are improved with the higher efficiency of the Filter and the greater accuracy of the Wrapper. The experiment results based on UCI data sets with support vector machine(SVM) as the classifier show the effectiveness and feasibility of the algorithm.

Key words: feature selection; particle swarm optimization; Filter; Wrapper; mutual information

0 引言

在现实世界的很多分类问题中, 并不是数据集中所有的特征都对分类有贡献. 从理论上讲, 增加特征向量的维数可以提供更多的判别能力. 但实际上, 过多的特征不仅可能降低学习性能而且还可能造成过拟合, 导致降低学习模型的泛化能力^[1], 因此特征选择是分类问题中的一个关键步骤.

从广义上讲, 特征选择算法可以分为Filter, Wrapper和Hybrid^[2]. Filter算法与后续学习算法无关, 速度快, 但评估与后续学习算法的性能偏差较大; Wrapper算法利用后续学习算法的训练准确率

评估特征子集, 偏差小、计算量大, 不适合大数据集; Hybrid算法则融合了两者的优点, 试图在效率和正确率之间寻找一个平衡. 很多关于Hybrid的研究基本都采用顺序方法^[3-7], 即第1阶段利用某种评价准则约减特征维数, 第2阶段在此基础上构建Wrapper模型来得到期望的特征子集. 文献[3]首先利用互信息作为评价准则对特征进行排序, 选择前 k 个特征组成一个候选特征子集; 然后以Shepley值为评价准则建立Wrapper模型来选择最终的特征子集. 文献[4]首先利用F-score对特征进行排序, 然后以最小聚类错误率为准则建立Wrapper模型. 文献[5]首先利用基于ReliefF的特征加权算法NReliefF和降维算

收稿日期: 2011-11-23; 修回日期: 2012-02-14.

基金项目: 国家自然科学基金项目(60975026, 61273275).

作者简介: 姚旭(1982—), 女, 博士生, 从事智能信息处理和机器学习的研究; 王晓丹(1966—), 女, 教授, 博士生导师, 从事智能信息处理和机器学习等研究.

法 DR 对特征进行约减;然后将 NReliefF 和 KNN 结合起来,利用分类正确率来选择特征等.这些方法都在一定程度上提高了效率和正确率,但也存在一定的问题,如在特征排序中评价准则的选取是关键,如何针对特定的问题选择合适的准则是一个难点.在第 2 阶段,对于高维数据 Wrapper 方法耗时较高.最重要的是,这些方法只是将 Filter 和 Wrapper 简单地串联在一起,并没有真正将两种方法融合为一个搜索过程.本文提出了一个基于 PSO 算法的 Filter-Wrapper 混合特征选择方法,将基于互信息的 Filter 模型融合在基于 PSO 算法的 Wrapper 模型中,因此它集成了 Filter 的高效率和 Wrapper 的高正确率的优点.

1 粒子群优化算法

粒子群优化算法模拟的是鸟群觅食行为,是一种基于集群智能的随机寻优算法.设搜索空间为 D 维,群体中第 i 个粒子的位置为 $X_i = (x_{i1}, x_{i2}, \dots, x_{iD})$,速度为 $V_i = (v_{i1}, v_{i2}, \dots, v_{iD})$,它经历过的最佳位置为 $P_i = (p_{i1}, p_{i2}, \dots, p_{iD})$,群体经历过的最佳位置为 $P_{gbest} = (p_{g1}, p_{g2}, \dots, p_{gD})$.在标准粒子群算法中,粒子分别按下式更新位置和速度^[8]:

$$v_{id}^{k+1} = \omega v_{id}^k + c_1 r_1 (p_{id}^k - x_{id}^k) + c_2 r_2 (p_{gd}^k - x_{id}^k), \quad (1)$$

$$x_{id}^{k+1} = x_{id}^k + v_{id}^k. \quad (2)$$

其中: ω 为惯性权重; k 为迭代次数; r_1, r_2 为 $(0,1)$ 之间的随机数; c_1, c_2 为加速因子,通常都取 2.每一维的位置和速度都被限制在一定的范围内,粒子“飞行”过程中如果位置和速度超过边界范围则取边界值.

文献 [9] 在证明了 PSO 进化过程与粒子速度无关的基础上提出了简化粒子群优化 (SPSO) 算法. SPSO 算法的进化公式为^[9]

$$x_{ij}^{k+1} = \omega x_{ij}^k + c_1 r_1 (p_{ij} - x_{ij}^k) + c_2 r_2 (p_{gj} - x_{ij}^k). \quad (3)$$

其中: x_{ij}^k 表示第 k 代第 i 个粒子的第 j 维分量; r_1, r_2 为 $(0,1)$ 之间的随机数; c_1, c_2 为学习因子, ω 为惯性因子; p_{ij} 为第 i 个粒子当前搜索到的最优位置的第 j 维分量; p_{gj} 为整个粒子群当前的最优位置的第 j 维分量.

2 基于互信息的相关冗余分析

特征按其其与类别的关系可以分为强相关特征、弱相关特征和无关特征^[10].强相关特征影响着类别的分布,如果缺少了强相关特征,则必然改变类别的分布情况,因此它是最优子集的一部分;弱相关特征在一定条件下影响类分布,但不一定是必须的;无关特征对类别分布没有影响,应首先删除.因此特征选择的任务就是去除无关特征,以及相对于其他特

征冗余的弱相关特征.为了度量特征空间中的相关性, Bonnländer^[11]首次引入了互信息的概念.互信息 (MI) 概念的引入是为了衡量两个变量间相互依赖的强弱程度,它表示两个变量间共同拥有信息的含量,也可以理解为在已知一个变量的情况下,另外一个变量不确定性的减少程度.给定两个随机变量 X 和 Y ,若它们的边缘概率分布分别为 $p(x)$ 和 $p(y)$,联合概率分布为 $p(x, y)$,则它们之间的互信息 $I(X; Y)$ 定义为

$$\begin{cases} I(X; Y) = \int_y \int_x p(x, y) \log \frac{p(x, y)}{p(x)p(y)} dx dy, \\ I(X; Y) = \sum_y \sum_x p(x, y) \log \frac{p(x, y)}{p(x)p(y)}. \end{cases} \quad (4)$$

很多文献都利用互信息来度量特征的相关和冗余,比较典型的有文献 [12] 提出的 MIFS-U (MIFS-Uncertainty) 方法,文献 [13] 提出的 mRMR (max-relevance and min-redundancy) 方法以及对它们的一些改进方法^[14-15].文献 [15] 分析了文献 [12-14] 中方法的缺陷,提出了一种新的相关冗余分析标准

$$J(f_j) = NI(C; f_j) - \frac{1}{|S_{i-1}|} \sum_{f_i \in S_{i-1}} NI(f_i; f_j). \quad (5)$$

其中

$$NI(f_i; f_j) = \frac{I(f_i; f_j)}{\log_2(N)}, \quad NI(C; f_j) = \frac{I(C; f_j)}{\log_2(|\Omega_C|)},$$

N 为特征的量化水平,详细说明参见文献 [15]; $|\Omega_C|$ 为类别空间的维数.本文将相关冗余度量标准 (5) 作为 PSO 算法中的一个中间步骤来提高算法的搜索速度和性能.

3 基于粒子群优化算法的最大相关最小冗余混合式特征选择方法

特征选择问题是一个典型的 0-1 组合优化问题.利用 PSO 算法进行特征选择已在多个文献中出现,如文献 [16-18],但他们的研究一般建立在离散二进制 PSO 算法的基础上.由于 PSO 算法与生俱来的连续性,导致离散 PSO 算法丧失了某些连续 PSO 算法的特性.为了保持连续 PSO 算法的特性,在特征选择过程中,本文进行如下处理:用 1 表示特征被选中,用 0 表示特征未被选中.将粒子的位置向量中的每一维表示成该维为 1 的概率,始终限定每维的搜索区域在 $[0, 1]$ 区间内,那么就可以通过设定相应的阈值,根据每维概率的大小来确定该维特征是否被选中.具体做法如下:将特征选择问题表示成一个由 0 和 1 组成的向量 $F_i = (f_{i1}^k, f_{i2}^k, \dots, f_{iD}^k)$, $f_{ij}^k = 1$ 表示第 j 维特征被选中,否则被淘汰.定义位置向量向特征向量的转化公式为

$$f_{ij}^k = \begin{cases} 1, & x_{ij}^k \geq r_0; \\ 0, & \text{else.} \end{cases} \quad (6)$$

其中 r_0 为区间 $(0.2, 0.8)$ 上的随机数.这样就解决了

由连续空间域到离散问题域的转化. 对于每个粒子的评价也就相应地转化成对每个特征向量的评价.

由PSO算法的进化公式可以看出, 粒子通过跟踪个体极值 P_i 和群体极值 P_{gbest} 来更新自己. 算法运行过程中, 如果某个粒子发现一个当前最优位置, 则其他粒子将迅速向其靠拢. 如果该最优位置为局部最优点, 则粒子群无法在解空间内重新搜索, 算法将陷入局部最优, 出现所谓的早熟收敛现象. 在本文算法中, 当PSO算法陷入局部极值时, 通过对群体极值 P_{gbest} 进行自适应 t 分布变异, 使其跳出局部极值. 如果停滞进化周期超过给定阈值, 则对 P_{gbest} 进行 t 分布变异, 变异公式如下:

$$P'_{\text{gbest}} = P_{\text{gbest}} + \sigma_j * t_n, \quad (7)$$

其中

$$\sigma_j(k) = \left[\alpha^k \frac{f_{\max}(k) - f_{\min}(k)}{f_{\max}(k)} + (1 - \alpha^k) \alpha^k \right] \cdot (P_{\text{gbest}}^{\max} - P_{\text{gbest}}^{\min}).$$

$f_{\max}(k)$ 为第 k 代最好粒子的适应度值, 即 $f(P_{\text{gbest}})$. $f_{\min}(k)$ 为第 k 代中最差粒子的适应度值; P_{gbest}^{\max} 和 P_{gbest}^{\min} 为群体极值 P_{gbest} 的最大值和最小值; α 是一个比1略小的常数, 本文取 $\alpha = 0.97^{[19]}$.

第2节讨论了特征的相关冗余度量标准, 在这里定义一个相关冗余指标 Ψ_{ij} . 在算法中, 首先选出与类别互信息最大的特征加入到已选特征集合, 余下的特征则根据 Ψ_{ij} 来选择, 直到达到预设的特征个数. 相关冗余指标定义如下:

$$\Psi_{ij} = J(f_j) * x_{ij}. \quad (8)$$

其中: x_{ij} 为第 i 个粒子的第 j 维的位置分量; $f_j \in F \setminus S$, F 为原始特征集合, S 为已选特征集合. 本文将式(8)作为相关冗余的度量标准融入到PSO算法的寻优过程中, 提出一种Filter和Wrapper混合的特征选择方法(MRPSO). 输入训练集 $X = \{x'_{11}, x'_{12}, \dots, x'_{1p_1}, x'_{21}, \dots, x'_{Cp_C}\}$, C 为类别数, $p_i (i \in \{1, 2, \dots, C\})$ 表示第 i 类的样本数, 原始特征集为 $F = \{f_i | i = 1, 2, \dots, D\}$, 最大迭代次数为 T , 变异触发阈值为 T' , 种群大小为 N , 所选特征个数为 d , 变异次数为 n . 算法的具体步骤如下:

Step 1: 初始化. 特征子集 $S_i^k = \varphi$, 生成 N 个初始粒子.

Step 2: 将每个粒子按式(6)转化成特征向量, 计算每个粒子的适应度值 $f_i(1)$, 将每个粒子的位置作为该粒子经历的最佳位置, 即 $P_i = x_i(1)$; 种群经历的最佳位置 $P_{\text{gbest}} = x_i(1)$, 其中 $i = \arg \max_j (f_j(1))$.

Step 3: 如果满足收敛条件(群体最大适应度满足条件或达到最大迭代次数), 转向Step 7, 否则, 转

向Step 4.

Step 4: 更新粒子位置, 利用互信息求出所有特征与类别的相关性 $I(f_i, C)$, 选出使互信息最大的特征 f , 将其加入到 S_i^k 中, 即 $f = \arg \max I(f_i, C)$, $S_i^k = S_i^k + \{f\}$.

Step 5: 当 $k \leq T (k = k + 1)$, 粒子的进化停滞周期小于 T' 时, 进行以下步骤:

Step 5.1: 当 $|S| < d$ 时, 重复执行如下操作: 选择满足 $f_j = \arg \max \Psi_{ij}$ 的 f_j 加入到 S_i^k 中, 即 $S_i^k = S_i^k + \{f_j\}$;

Step 5.2: 计算特征子集的适应度函数值 $f(S_i^k)$, 如果粒子的个体极值和群体极值优于历史位置, 则更新 P_i 和 P_{gbest} , 并更新粒子位置向量, 令 $S_i^k = \varphi$, 转向Step 5.1;

Step 6: 如果粒子的进化停滞周期超过 T' , 则利用式(7)对 P_{gbest} 进行 n 次变异操作, 从中选择适应度值最好的 P' . 如果 $f(P') > f(P_{\text{gbest}})$, 则 $P_{\text{gbest}} = P'$, 转向Step 5, 否则转向Step 7.

Step 7: 输出 P_{gbest} 对应的特征子集 S , 算法结束.

4 实验结果及分析

4.1 实验数据及参数设置

实验中的数据均来自UCI^[20]数据库中的数据集, 实验选择了其中8组数据(特征维数范围为8~60, 样本范围为208~6435), 关于实验数据的详细描述如表1所示. 在进行实验前, 对数据集中的数据进行了归一化处理.

表1 UCI数据集各数据描述

Problem	Train	Attributes	Classes
Sonar	208	60	2
Soybean	307	35	19
Ecoli	336	8	8
Ionosphere	351	34	2
Breast-c-w	699	10	2
Segment	2310	19	7
Waveform	5000	21	3
Satimage	6435	36	6

考虑到惯性权重 ω 对整个搜索过程的影响, 一般采用线性递减的方法改变 ω , 但是粒子搜索过程并不是一个线性过程, 本文采用文献[21]中的策略, 利用非线性方法自适应调整惯性权重 ω , 即

$$\omega_0 = \omega_{\min} + (\omega_{\max} - \omega_{\min}) \exp \left(- \frac{1}{1 + (1 + t/t_{\max})} \right). \quad (9)$$

其中: ω_{\max} 和 ω_{\min} 为惯性权重的最大值和最小值, $\omega_{\max} = 0.9$, $\omega_{\min} = 0.4$; t 为当前迭代次数; t_{\max} 为最大迭代次数.

变异触发阈值 T' 是算法中的一个关键参数. 如果设置得过大, 则会导致实际上粒子并没有收敛

到群体极值 $P_{g_{best}}$ 却被误判为收敛到 $P_{g_{best}}$, 从而对 $P_{g_{best}}$ 进行变异, 这样就延长了搜索时间; 如果设置得过小, 则可能会把实际上已经收敛到 $P_{g_{best}}$ 判断为未收敛到 $P_{g_{best}}$, 这样就无法及时跳出局部收敛. 经过多次实验, 本文设定 $T' = 5$.

t 算子的变异能力与自由度 n 和变异尺度 σ 有关. 本文已经设计了动态变化的变异尺度, 下面分析自由度 n 的选取. 为了更好地实现变异操作, 采用自适应的 t 变异. 当算法陷入局部收敛时, 利用不同的 t 算子对群体极值进行变异, 选择其中最好的结果. 选用的算子有 $t_{1.5}$, t_2 , $t_{2.5}$, t_3 和 t_{10+k} , k 为进化代数.

实验中, 种群大小 $N = 30$, 最大迭代次数 $T = 200$, 变异次数 $n = 100$, 学习因子 $c_1 = c_2 = 2$; 各个算法的适应度函数均为分类正确率; 以 SVM 为分类器, 来自 PRTTool 工具箱, 实验机器配置为 2G 内存, 2.80G CPU, 算法基于 Matlab 7.10 (R2010a) 实现.

4.2 实验结果和分析

为了验证本文算法 MRPSO 的性能, 在公共数据集 UCI 上进行实验, 并与基于 BPSO 算法和 CBPSO 算法^[18]的特征选择算法以及典型的 Filter 式特征选择方法 ReliefF 进行比较. 实验采用 10 重交叉验证, 从分类正确率、所选择的特征子集规模和算法收敛的迭代次数 3 个角度对本文算法进行验证.

4.2.1 8 个数据集上的分类正确率

表 2 和表 3 给出了 8 个数据集上利用 5 种不同特征选择算法得到的分类正确率以及最好情况下最优特征子集的规模. 其中: CBPSO(1) 表示基于 Logistic 映射的 CBPSO 算法, CBPSO(2) 表示基于 tent 映射的 CBPSO 算法.

分析表 2 和表 3 可以看出, 所有的特征选择算法都约减了特征的维数. 除了在 Sonar 数据集上以外, 本文方法都获得了最小的特征子集规模. 在数据集 Sonar, Soybean, Ecoli, Ionosphere, Breast-c-w, Segment, Waveform 和 Satimage 上, 与原始特征维数相比, 维数缩减比例分别为 58.3%, 54.3%, 50%, 61.8%, 60%, 57.9%, 57.1% 和 55.6%, 这表明了本文算法可以有效地缩减特征维数. 从分类正确率的角度来看, 在数据集 Sonar, Soybean, Ecoli, Ionosphere, Breast-c-w, Segment, Waveform 和 Satimage 上, 本文算法取得的分类正确率分别为 91.85%, 93.99%, 88.22%, 93.85%, 97.56%, 94.96%, 90.31% 和 85.24%. 除了在数据集 Soybean 和 Satimage 上以外, 本文算法均得到了最高的分类正确率, 但所选择的特征个数却是最少的. 这表明本文算法是有效的, 它并不以牺牲正确率来获得最小的特征子集. 与 ReliefF 方法相比, 本文方法和

基于 BPSO 的方法在 8 个数据集上的分类正确率均有所提高. 这充分说明了与 Filter 式特征选择方法相比, Wrapper 式特征选择方法能获得更高的分类正确率; 而本文提出的将 Filter 融合在 Wrapper 中的混合式特征选择方法取得了更优的分类结果.

表 2 不同特征选择算法得到的分类正确率 %

Dataset	MRPSO	BPSO	CBPSO(1)	CBPSO(2)	ReliefF
Sonar	91.85	90.56	91.56	91.85	86.34
Soybean	93.99	94.03	93.48	94.11	91.69
Ecoli	88.22	87.84	87.25	87.52	80.65
Ionosphere	93.85	92.89	93.72	93.44	91.25
Breast-c-w	97.56	97.43	97.47	97.53	96.67
Segment	94.96	93.68	94.15	94.84	92.25
Waveform	90.31	88.63	89.24	89.65	85.27
Satimage	85.24	83.15	83.54	85.35	82.03

表 3 不同特征选择算法得到的最优特征子集的规模

Dataset	MRPSO	BPSO	CBPSO(1)	CBPSO(2)	ReliefF
Sonar	25	43	22	25	52
Soybean	16	34	22	33	34
Ecoli	4	4	7	7	7
Ionosphere	13	14	30	31	32
Breast-c-w	4	9	8	9	9
Segment	8	12	15	13	16
Waveform	9	13	16	15	19
Satimage	16	22	25	24	34

4.2.2 迭代次数与分类正确率比较

由于本文算法是针对陷入局部收敛的群体极值进行变异, 是从本质上解决局部收敛的问题. 相对于仅仅自适应的调整算法中一些参数的做法, 从理论上讲本文算法能得到较好的结果. 而且本文算法是基于连续 PSO 算法, 相对于离散二进制 PSO 算法更具优势. 因此为了说明本文算法的优点, 将其与基于 BPSO 算法和基于 CBPSO 算法的特征选择算法^[18]进行比较, 在数据集 Sonar, Ionosphere, Breast-c-w 和 Segment 上进行实验. 实验结果如图 1~图 4 所示. 图中: CBPSO(1) 表示基于 Logistic 映射的 CBPSO, CBPSO(2) 表示基于 tent 映射的 CBPSO.

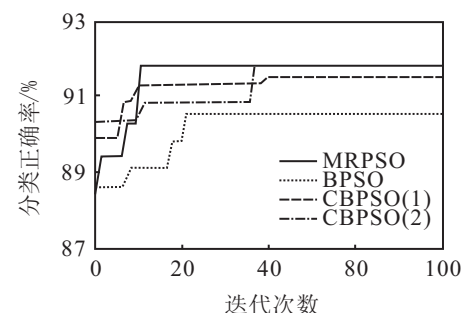


图 1 Sonar 数据集上迭代次数与正确率比较

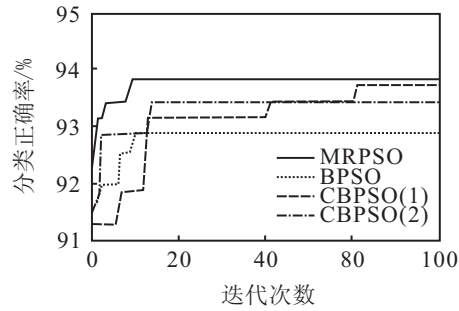


图2 Ionosphere数据集上迭代次数与正确率比较

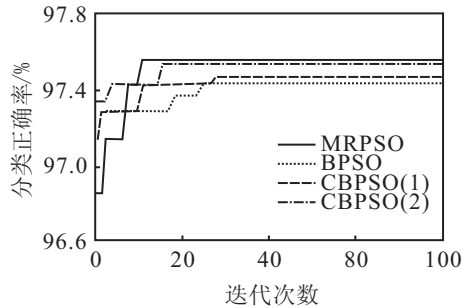


图3 Breast-c-w数据集上迭代次数与正确率比较

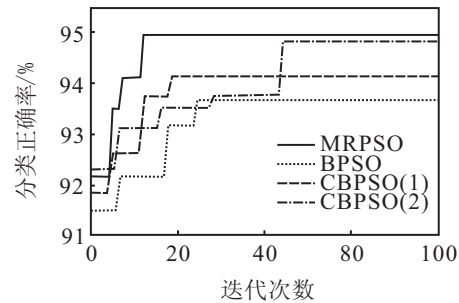


图4 Segment数据集上迭代次数与正确率比较

从图1~图4可以看出, 不管搜索的起点值大小如何, 本文方法均能以最快的速度收敛获得较高的分类正确率. 由于本文方法是对群体极值进行变异, 从本质上解决了搜索过程中的局部收敛问题, 而CBPSO算法只是自适应地调整惯性权重, 因此本文方法更容易跳出局部收敛, 以最少的迭代次数得到较高的分类正确率. 相对于BPSO算法, 本文算法基本都获得了最高的分类正确率, 且得到的特征子集规模都远小于BPSO算法得到的子集规模, 因此体现了连续PSO算法的优点.

5 结论

本文提出了一种集成Filter和Wrapper优点的特征选择方法. 与以往采用顺序法的混合Filter-Wrapper不同, 本文方法将Filter融合于基于PSO算法的Wrapper模型中. Filter模型基于互信息, 利用相关冗余综合测度来权衡特征被选择的概率, 因此该方法提高了PSO算法的收敛速度, 优化了特征子集. 同时, 为了避免PSO算法在搜索过程中陷入局部收敛, 采用自适应的t分布对群体极值进行变异, 从本质上

解决了局部收敛问题. 实验结果表明, 该算法能够在保证SVM分类精度的情况下, 有效地减少特征维数, 加速了SVM训练, 具有较好的性能.

参考文献(References)

- [1] Yu L, Liu H. Efficient feature selection via analysis of relevance and redundancy[J]. J of Machine Learning Research, 2004, 5(1): 1205-1224.
- [2] Dash M, Liu H. Feature selection for classification[J]. Intelligent Data Analysis, 1997, 1(2): 131-156.
- [3] Liu Jihong, Wang Guoxiong. A hybrid Feature Selection Method for Data Sets of thousands of variables[C]. The 2nd Int Conf on Advanced Computer Control. Shenyang, 2010, 2: 288-291.
- [4] Sarojini Llango B, Ramaraj N. A hybrid prediction model with F-score feature selection for type II diabetes databases[C]. Proc of the 1st Amrita ACM-W Celebration on Women in Computing. India: ACM, 2010: 1-4.
- [5] 刘杰, 金弟, 杜慧君, 等. 一种新的混合特征选择方法RRK[J]. 吉林大学学报, 2009, 39(2): 419-423.
(Liu J, Jin D, Du H J, et al. New hybrid feature selection method[J]. J of Jilin University, 2009, 39(2): 419-423.)
- [6] 毛俐旻, 姚淑萍, 胡昌振. 一种新型混合特征选择方法及其在入侵检测中的应用[J]. 北京理工大学学报, 2008, 28(3): 218-221.
(Mao L M, Yao S P, Hu C Z. A new hybrid attribute selection method and its application in intrusion detection[J]. Trans of Beijing Institute of Technology, 2008, 28(3): 218-221.)
- [7] 陈鑫, 梁海洁, 廖腾峰. 基于TSVM分类器和混合型特征选择方法的入侵检测研究[J]. 微电子学与计算机, 2010, 27(8): 242-244.
(Chen X, Liang H J, Liao T F. Intrusion detection based on TSVM and feature selection[J]. Microelectronics and Computer, 2010, 27(8): 242-244.)
- [8] Kennedy J, Eberhart R. Particle swarm optimization[C]. Proc of IEEE Int Conf on Neural Networks. Perth: IEEE Press, 1995: 1942-1948.
- [9] HU Wang, LI Zhi-shu. A simpler and more effective particle swarm optimization algorithm[J]. J of Software, 2007, 18(4): 861-868.
- [10] John G H, Kohavi R, Pfleger K. Irrelevant feature and the subset selection problem[C]. Proc of the 11th Int Conf on Machine Learning. New Jersey: Morgan Kaufmann Publishers, 1994: 121-129.
- [11] Bonnländer B V, Weigend A S. Selecting input variables using mutual information and nonparametric density evaluation[C]. Proc of the 1994 Int Symposium on Artificial Neural Networks. Tainan, 1994: 42-50.