

基于粒子群优化算法的支持向量机参数选择^①



贺心皓, 罗旭

(成都信息工程大学 通信工程学院, 成都 610225)

通讯作者: 贺心皓, E-mail: episodecrazy02@163.com

摘要: 由于支持向量机的主要参数的选择能够在很大程度上影响分类性能和效果, 并且目前参数优化缺乏理论指导, 提出一种粒子群优化算法以优化支持向量机参数的方法. 该方法通过引入非线性递减惯性权值和异步线性变化的学习因子策略来改善标准粒子群算法的后期收敛速度慢、易陷入局部最优的缺陷. 实验结果表明, 相对于标准粒子群算法, 本方法在参数优化方面具有良好的鲁棒性、快速收敛和全局搜索能力, 具有更高的分类精确度和效率.

关键词: 支持向量机; 粒子群优化算法; SVM 参数优化; 惯性权值非线性递减; 异步变化学习因子

引用格式: 贺心皓, 罗旭. 基于粒子群优化算法的支持向量机参数选择. 计算机系统应用, 2019, 28(8): 241-245. <http://www.c-s-a.org.cn/1003-3254/7011.html>

Support Vector Machine Parameter Selection Based on Particle Swarm Optimization Algorithm

HE Xin-Hao, LUO Xu

(School of Communication Engineering, Chengdu University of Information Technology, Chengdu 610225, China)

Abstract: Since the selection of the main parameters of the support vector machine can affect the classification performance and effect to a large extent, and the current parameter optimization lacks theoretical guidance, a particle swarm optimization algorithm is proposed to optimize the parameters of the support vector machine. This method improves the shortcomings of the standard particle swarm optimization algorithm with slow convergence rate and easy to fall into local optimum by introducing nonlinear decreasing inertia weight and asynchronous linear variation learning factor strategy. The experimental results show that compared with the standard particle swarm optimization algorithm, the proposed method has good robustness, fast convergence and global search ability in parameter optimization, and has higher classification accuracy and efficiency.

Key words: support vector machine; particle swarm optimization algorithm; SVM parameter optimization; inertia weight nonlinear decrement; asynchronous change learning factor

引言

支持向量机 (Support Vector Machines, SVM) 是一种应用广泛的机器学习方法, 具有理论知识清晰完备, 适应性和泛化能力良好的优点, 核心思想是在特征空间中寻找到一个最优超平面将两类样本尽可能大的分开, 能够较好的处理小样本、非线性和克服“维数灾难”问题, 并且表现出优秀的分类能力和泛化能力而被

广泛应用于分类和回归等领域. 但是 SVM 对核函数的参数选取对分类效果影响很大, 不合适的参数可能使得分类器性能大大降低.

针对 SVM 核参数的选取问题, 目前尚没有统一有效的方法. 传统的参数选择方法如实验法、网格搜索法等由于耗时过长和不必要的验证流程等缺点, 更常用的方法是群智能算法如蚁群算法、遗传算法和粒子

① 收稿时间: 2019-01-24; 修改时间: 2019-02-26, 2019-03-04; 采用时间: 2019-03-04; csa 在线出版时间: 2019-08-08

群算法等优化支持向量机核参数。

粒子群算法由于算法结构简单、寻优能力相对较好,近年来选择粒子群算法优化 SVM 参数成为研究热点之一.如文献[1]用标准粒子群算法对 SVM 参数进行优化,并将该模型应用于聚丙烯腈生产的软测量模型中,通过与标准 SVM 比较,证明了所提方案有效性;随着研究深入,标准 PSO 算法在应用过程中暴露出不少的缺点,如易陷入局部最优值、后期收敛的速度较慢等缺点,因此研究者也对其提出不少改进方法,如文献[2-4]等通过改进惯性权值,或者改变学习机制等方法以提高粒子群算法性能。

本文通过引入非线性惯性权值递减和异步线性变化的学习因子策略改进粒子群算法,并且通过机器学习公开数据集 UCI 数据集的实验数据证明了其有效性。

1 支持向量机基本思想和参数

支持向量机是由 Vapnik 等人在 1995 年提出的基于统计学理论和结构风险最小化原理的机器学习方法,其基本思想是找到一个分类面将两类样本分开.当线性可分时,其最优的分类面要求正确地将样本分开且分类间隔最大;当线性不可分时,寻找一个多维超平面从而将样本分开。

对于训练样本集: $(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N), x \in R^n, y \in (-1, 1)$ 最优分类超平面的二次规划问题如下:

$$\begin{cases} \min \frac{1}{2} w^T \times w + C \sum_{i=1}^N \xi_i \\ \text{s.t. } y_i (w^T x_i + b) \geq 1 - \xi_i, \xi_i \geq 0, i = 1, 2, \dots, N \end{cases} \quad (1)$$

其 Lagrange 函数为:

$$L(w, b, l) = \frac{1}{2} w^T \times w + C \sum_{i=1}^N \xi_i - \sum_{i=1}^N l_i [y_i (w^T x_i + b) - 1 + \xi_i] - \sum_{i=1}^N a_i x_i \quad (2)$$

式中, $\lambda_i \geq 0, i = 1, 2, \dots, N$. 式 (1) 的对偶问题可表示为:

$$\begin{cases} \max Q(l) = \sum_{i=1}^N l_i - \frac{1}{2} \sum_{i,j=1}^N y_i y_j \lambda_i \lambda_j (x_i \cdot x_j) \\ \text{s.t. } \sum_{i=1}^N y_i \lambda_i = 0, 0 \leq \lambda_i \leq C \end{cases} \quad (3)$$

可得决策函数为:

$$f(x, l) = \text{sgn} \left(\sum_{SV} y_i \lambda_i (x_i \cdot x_j) + b \right) \quad (4)$$

对线性不可分的对偶函数为:

$$\begin{cases} \max Q(l) = \sum_{i=1}^N l_i - \frac{1}{2} \sum_{i,j=1}^N \lambda_i y_j \lambda_j (\phi_{x_i} \times \phi_{x_j}) \\ \text{s.t. } \sum_{i=1}^N y_i \lambda_i = 0, 0 \leq \lambda_i \leq C \end{cases} \quad (5)$$

式中, $\phi_{x_i} \times \phi_{x_j} = K(x_i, x_j)$ 即为核函数, 其决策函数为:

$$f(x, l) = \text{sgn} \left(\sum_{SV} y_i \lambda_i K(x_i, x_j) + b \right) \quad (6)$$

不同的核函数可以生成不同的支持向量机分类器, SVM 中常用的核函数有线性核、多项式核、径向基核和 Sigmoid 核函数等. 本文选择径向基核函数作为 SVM 的核函数, 即:

$$K(x_i, x) = \exp(\gamma \|x_i - x\|^2) \quad (7)$$

其中, $\gamma = \frac{1}{2\sigma^2}$.

影响核函数的参数有惩罚因子 C 和核函数 σ , 惩罚因子 C 是对错分的样本的惩罚程度的控制, 越大表示惩罚越重, 但其泛化能力也会同时降低. σ 是核函数的宽度参数, 表示对径向范围的控制. 惩罚因子 C 和核函数 σ 的选取对 SVM 分类器影响至关重要, 因此要获得优秀的支持向量机的分类性能, 需要选取合适的 C 和 σ [5].

2 改进粒子群算法优化 SVM

2.1 粒子群算法概述

粒子群算法是基于对鸟类的捕食过程进行模拟的一种新型智能仿生优化算法. 该算法把群体中的个体视为多维空间中的一个没有质量和体积的点, 且以一定的速度飞行, 在此迭代过程中根据自身的飞行经验和同伴的飞行经验对自身的飞行速度进行动态调整以修正前进方向和速度大小. 粒子群算法就是通过每个粒子由适应度函数调整粒子至较优的区域以搜寻到问题的最优解。

PSO 初始化一群随机粒子 (粒子不考虑质量和体积), 粒子在 N 维空间位置表示为矢量 $x_i = (x_1, x_2, \dots, x_n)$, 飞行速度表示为矢量 $v_i = (v_1, v_2, \dots, v_n)$, 每个粒子都有一个由目标函数确定的适应值 (fitness value),

其个体最好位置 p_{best} 值和整个群体中发现的最好的位置 g_{best} 值以及当前位置, 通过迭代找到最优解, 在每次迭代中粒子通过 p_{best} 值和 g_{best} 值来更新自己.

$$\begin{cases} v_i = wv_i + c_1r_1(p_{besti} - x_i) + c_2r_2(g_{besti} - x_i) \\ x_i = x_i + v_i \end{cases} \quad (8)$$

式中, p_{best} 表示粒子 i 是个体最优值, g_{best} 表示粒子群的全局最优值; C_1 和 C_2 是两个常数, 称为学习因子; r_1, r_2 是 0 到 1 的随机数; w 被称为惯性因子, w 的值越大, 全局寻优能力越强, 收敛越慢, w 值越小, 局部寻优能力越强, 收敛较快.

2.2 粒子群优化算法和支持向量机参数选择

惯性权重是粒子群算法中一个重要的参数之一, 它控制了全局搜索能力和局部搜索能力的平衡. 传统的粒子群算法有一些缺点, 如容易陷入局部最优值和后期震荡和收敛速度慢等, 国内外学者对算法的改进展开了研究. 文献[6]提出惯性权重进行了研究, 提出惯性权重 w 从 0.9 至 0.4 的线性递减的策略能够保证有更好的全局搜索能力和后期的局部搜索能力; 文献[7]在递减的惯性权值的基础上, 提出了抛物线和指数曲线的非线性的权值递减策略; 文献[8]提出一种惯性权重线性递增的粒子群算法; 文献[9]提出一种惯性权重先增后减的粒子群算法; 文献[10]提出一种非线性递减的惯性权重粒子群算法; 文献[11,12]提出了惯性权重指数递减和余弦递减的粒子群算法; 文献[13]提出了一种惯性权值随粒子的位置和目标函数性质变化的动态改变惯性权重的粒子群算法等.

1) 具有非线性递减惯性权重粒子群算法

本文使用非线性递减权值策略将惯性权重 w 设置为变量, 通过非线性递减权值实现粒子的调整, (当 $n=1$ 时, 即 Shi 提出的线性递减惯性权重策略), 其公式如下:

$$w(k) = w_{end} + \left(\frac{i_{max} - i}{i_{max}} \right)^n \cdot (w_{ini} - w_{end}) \quad (9)$$

式中, w_{ini} 、 w_{end} 分别为 w 的最大值和最小值, i 表示当前迭代次数, i_{max} 表示最大迭代次数, n 是指数值.

2) 异步线性改变学习因子策略

由标准粒子群算法公式 (8) 可知, C_1 控制了“自我认知”部分, 即粒子自身之前的飞行经验对之后飞行方向的影响, C_2 控制了“社会认知”部分, 即种群中所有粒子的飞行经验对每个粒子之后飞行方向的影响. 在标准粒子群算法中, Shi 等建议 C_1 和 C_2 都取值为 2, 在

其他文献中有学者也提出 C_1 和 C_2 可以取不同的值, 其取值也基本在 0 至 4 之间.

学习因子在过程中采用不同的变化策略称之为异步变化. A.Ratnawecra 等提出了一种异步线性改变学习因子的策略^[14], 该策略是在算法早期阶段 C_1 取值较大而 C_2 取值较小, 使粒子能够更多地向自我的学习以此加强全局搜寻能力; 在搜寻的后期阶段 C_2 取值较大, 使粒子能够快速收敛到全局最优解. 学习因子的更新公式如下:

$$\begin{cases} c_1 = c_{1ini} - (c_{1ini} - c_{1fin})(t/T_{max}) \\ c_2 = c_{2ini} + (c_{2fin} - c_{2ini})(t/T_{max}) \end{cases} \quad (10)$$

式中, c_{1ini} 、 c_{2ini} 分别表示学习因子的初始值; c_{1fin} 、 c_{2fin} 分别表示学习因子的迭代最终值; t 和 T_{max} 分别表示当前迭代次数和最大迭代次数.

综上所述, 本文引入非线性递减惯性权重和异步线性改变学习因子的粒子群优化算法, 用以寻找 SVM 模型中的惩罚因子 C 和核函数参数 σ 最优解, 每个粒子的位置代表为 (C, σ) .

其算法流程如下,

- ① 初始化参数设置, 包括设置粒子种群规模、粒子维数、最大迭代次数、惯性权重最大值和最小值和指数值、加速因子初始值和迭代最终值、随机产生粒子初始位置和速度、个体极值和全局极值.
- ② 根据式 (8)~式 (10) 更新每个粒子的速度和位置.
- ③ 根据适应度函数计算粒子的适应度值.
- ④ 将粒子的当前适应度值和粒子个体极值进行比较, 若优于粒子个体极值, 则进行替换.
- ⑤ 若当前粒子的适应度值优于全局极值的适应度值, 则用当前粒子适应度值替换全局极值.
- ⑥ 若达到最大迭代次数, 则输出迭代出的最优解 (C, σ) , 输出; 否则返回至步骤②.
- ⑦ 将优化后的 C 和 σ 利用 SVM 分类器进行训练, 并用文本测试集进行测试.

本算法优化过程如图 1 所示.

3 实验与分析

实验所用硬软件环境信息如下, AMD A6-3420M APU with Radeon (tm), RAM 4 GB (含一 2 GB samsung 内存条), Windows 8 64 位, 平台为 Matlab R2014a.

为了验证本文提出的粒子群优化的支持向量机分类模型的有效性, 本文选取了机器学习公开数据集 UCI 的 4 组数据进行测试, 下载地址如下: <http://archive.ics.uci.edu/ml/datasets.html>.

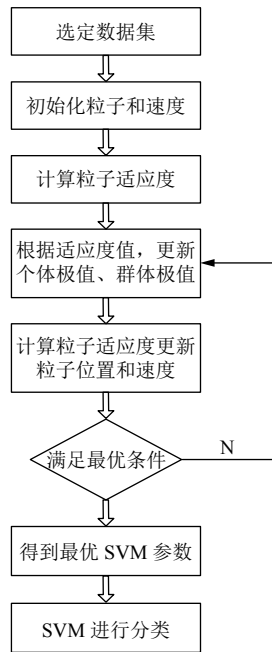


图1 粒子群优化算法的过程

3.1 实验参数设定

在参考大量文献之后^[1,5-7,15], 本文对粒子群参数作出如下合理设置. 种群规模 $N=20$, 最大迭代次数 $T_{max} = 100$, 惩罚因子 C 最小值 $C_{min} = 2^{-5}$, 最大值 $C_{max} = 2^{15}$, 误差阈值为 10^{-3} , 核参数 σ 的最小值 $\sigma_{min} = 2^{-15}$, 最大值 $\sigma_{max} = 2^3$, 惯性因子最小值 $w_{end} = 0.4$, 惯性因子最大值 $w_{ini} = 0.9$, 非线性递减惯性权重粒子群算法中的指数 $n = 1.2$, $C_{1ini} = 2$, $C_{1fin} = 0.5$, $C_{2ini} = 0.5$, $C_{2fin} = 2$; 标准粒子群算法的 $w = 0.9$, $C_1 = C_2 = 2$;

3.2 实验结果

在表1的数据集通过采用 K 折交叉验证法(K -fold cross validation), 即把表1中的每个样本数据随机地分为 K 个子样本集, 每次都去一个子样本集作为测试集, 剩余 $K-1$ 个样本集作为训练集, 以此进行验证.

表1 实验数据信息

数据集	类别个数	属性个数	样本个数
Sonar	2	60	208
Seeds	3	7	210
Ionosphere	2	34	351
AustralianCredit	2	14	690

如下表2所示, 通过标准粒子群优化支持向量机算法^[1]和本文提出粒子群优化算法在UCI数据集上的实验结果, 表中的最大值、最小值和平均值表示经过了10-fold CV交叉验证后的分类准确率的最大值、最小值和平均值. 图2为两种算法在数据集上的准确率平均值对比图.

表2 两种算法在数据集上的交叉验证结果对比

数据集	分类器	最大值 (%)	最小值 (%)	平均值 (%)
Sonar	PSO-SVM	88.7953	84.3872	86.0151
	IPSO-SVM	91.173	88.6843	89.9312
Seeds	PSO-SVM	94.1143	91.7955	93.0931
	IPSO-SVM	95.3612	92.4135	94.711
Ionosphere	PSO-SVM	94.1482	91.7243	92.881
	IPSO-SVM	95.3756	92.9913	94.3752
AustralianCredit	PSO-SVM	87.364	82.6631	84.6397
	IPSO-SVM	89.3335	85.5787	86.1989

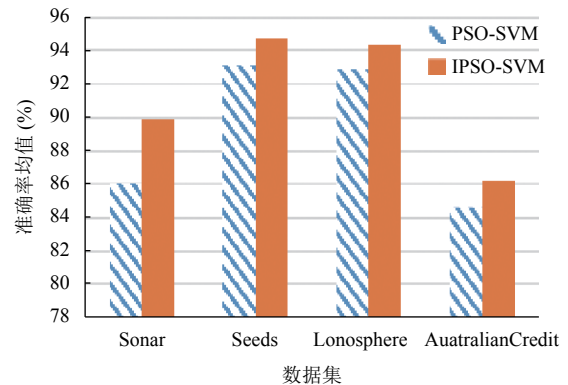


图2 两种算法的准确率均值对比图

4 结论

支持向量机模型的分类精确率很大程度上由关键参数的选取决定, 本文引入非线性递减惯性权重和异步线性变化的学习因子策略. 随着迭代次数增加, 学习因子 C_1 的值减小而 C_2 值增加, 保证了前期个体最优质影响和后期全局最优解的影响, 加快后期的收敛; 而非线性递减惯性权重在保证粒子在迭代初期对整个搜索空间能进行快速搜索, 随着惯性权重非线性减小能使大部分搜索范围集中于最优值领域中, 使搜索速度和精度都有提高. 实验结果表明, 非线性递减惯性权重和异步线性改变学习因子粒子群算法优化支持向量机比标准粒子群优化的支持向量机有更高的分类精确度和效率.

参考文献

- 1 邵信光, 杨慧中, 陈刚. 基于粒子群优化算法的支持向量机

- 参数选择及其应用. 控制理论与应用, 2006, 23(5): 740-743, 748. [doi: 10.3969/j.issn.1000-8152.2006.05.014]
- 2 路志英, 李艳英, 陆洁, 等. 粒子群算法优化 RBF-SVM 沙尘暴预报模型参数. 天津大学学报, 2008, 41(4): 413-418. [doi: 10.3969/j.issn.0493-2137.2008.04.006]
 - 3 李杰. 改进粒子群算法优化支持向量机的工程造价预测. 计算机系统应用, 2016, 25(6): 202-206. [doi: 10.15888/j.cnki.csa.005210]
 - 4 胡旺, 李志蜀. 一种更简化而高效的粒子群优化算法. 软件学报, 2007, 18(4): 861-868.
 - 5 刘伟丽. 基于粒子群算法和支持向量机的中文文本分类研究[硕士学位论文]. 郑州: 河南工业大学, 2010.
 - 6 Shi YH, Eberhart R. A modified particle swarm optimizer. 1998 IEEE International Conference on Evolutionary Computation Proceeding. IEEE World Congress on Computational Intelligence. Anchorage, AK, USA. 1998. 69-73.
 - 7 陈贵敏, 贾建援, 韩琪. 粒子群优化算法的惯性权值递减策略研究. 西安交通大学学报, 2006, 40(1): 53-56. [doi: 10.3321/j.issn:0253-987X.2006.01.013]
 - 8 Zheng YL, Ma LH, Zhang LY, *et al.* On the convergence analysis and parameter selection in particle swarm optimization. Proceedings of the 2003 International Conference on Machine Learning and Cybernetics. Xi'an, China. 2003. 1802-1807.
 - 9 崔红梅, 朱庆保. 微粒群算法的参数选择及收敛性分析. 计算机工程与应用, 2007, 43(23): 89-91, 131. [doi: 10.3321/j.issn:1002-8331.2007.23.028]
 - 10 Chatterjee A, Siarry P. Nonlinear inertia weight variation for dynamic adaptation in particle swarm optimization. Computers & Operations Research, 2006, 33(3): 859-871.
 - 11 刘伟. 一种改进惯性权重的 PSO 算法. 计算机工程与应用, 2006, 45(7): 46-48.
 - 12 陈国初, 俞金寿. 增强型微粒群优化算法及其在软测量中的应用. 控制与决策, 2005, 20(4): 377-381. [doi: 10.3321/j.issn:1001-0920.2005.04.004]
 - 13 王启付, 王战江, 王书亭. 一种动态改变惯性权重的粒子群优化算法. 中国机械工程, 2005, 16(11): 945-948. [doi: 10.3321/j.issn:1004-132X.2005.11.002]
 - 14 Ratnaweera A, Halgamuge SK, Watson HC. Self-organizing hierarchical particle swarm optimizer with time-varying acceleration coefficients. IEEE Transactions on Evolutionary Computation, 2004, 8(3): 240-255. [doi: 10.1109/TEVC.2004.826071]
 - 15 赵文芳, 王京丽, 尚敏, 等. 基于粒子群优化和支持向量机的花粉浓度预测模型. 计算机应用, 2019, 3. [doi: 10.11772/j.issn.1001-9081.2018122608]