

## 基于 DYCORS 算法的 OVA-SVM 参数优化与应用研究

余晨曦, 尹彦力

重庆师范大学 数学科学学院, 重庆 401331

**摘要:**目的 现有的参数优化方法普遍存在时间成本较大、内存占用较大、难以解决高维数据情况、难以找到全局最优解等问题, DYCORS 算法可以在节约时间成本和内存的前提下, 对高维数据问题也能找到全局最优解, 故针对现有参数优化方法存在的问题, 提出了针对 OVA-SVM 模型参数分块优化的 YDYCORS 算法。方法 OVA-SVM 的参数中对模型影响较大的有惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代终止参数  $t$ , 由于同时调节 5 个参数计算量较大, 难以找到最优解, 而 DYCORS 算法可以减少迭代次数, 对于高维数据问题也同样适用, 在 DYCORS 算法的基础上进行参数分块调节: 先调节影响最大的参数  $C$ 、 $k$ 、 $\gamma$ , 再固定最优参数  $C$ 、 $k$ 、 $\gamma$ , 调节剩余参数中影响较大的参数  $d$  和  $t$ , 最后同时调节已获得的 5 个最优参数, 如此对参数进行分块调节, 提升参数优化的效果。结果 通过 MNIST 和 IRIS 两个数据集上的实验结果对比可以发现: 运用 YDYCORS 算法对 OVA-SVM 参数进行分块调节后, 能得到与手动调参和直接用 DYCORS 同时调节 5 个参数更高的模型准确率, 从而也能进一步提升模型性能。结论 最终实验结果表明: DYCORS 算法能有效解决 OVA-SVM 参数优化中时间成本较大、内存占用较大、难以解决高维数据、难以找到全局最优解等问题, 尤其是改进后的 YDYCORS 算法能进一步提升 OVA-SVM 的模型准确率, 获得较佳的模型效果。

**关键词:**超参数优化; 支持向量机; DYCORS 算法

中图分类号: O224 文献标识码: A doi: 10.16055/j.issn.1672-058X.2024.0001.005

### Optimization and Application of OVA-SVM Parameters Based on DYCORS Algorithm

YU Chenxi, YIN Yanli

School of Mathematical Sciences, Chongqing Normal University, Chongqing 401331, China

**Abstract: Objective** The existing parameter optimization methods generally have problems such as large time cost, large memory occupation, difficulty in solving high-dimensional data, and difficulty in finding global optimal solutions. The DYCORS algorithm can find the global optimal solution for high-dimensional data problems even with the saving of time cost and memory. Therefore, in view of the problems existing in the existing parameter optimization methods, the YDYCORS algorithm for block optimization of OVA-SVM model parameters was proposed. **Methods** Among the parameters of OVA-SVM, the penalty parameter  $C$ , the kernel type  $k$ , the RBF kernel function parameter  $\gamma$ , the ploy kernel function parameter  $d$ , and the iteration termination parameter  $t$  have a greater impact on the model. Due to the large computational effort of adjusting five parameters simultaneously, it is difficult to find the optimal solution. The DYCORS algorithm can also be applied to high-dimensional data problems by reducing the number of iterations, and then the parameters were adjusted in blocks based on the DYCORS algorithm. The most influential parameters  $C$ ,  $k$ , and  $\gamma$  were adjusted first, then the optimal parameters  $C$ ,  $k$ , and  $\gamma$  were fixed, the more influential parameters  $d$  and  $t$  among the remaining parameters were adjusted, and finally the five optimal parameters that had been obtained were adjusted

收稿日期: 2022-12-05 修回日期: 2023-03-02 文章编号: 1672-058X(2024)01-0038-07

作者简介: 余晨曦(1999—), 女, 四川西昌人, 硕士研究生, 从事机器学习研究。

通讯作者: 尹彦力(1999—), 女, 重庆人, 硕士研究生, 从事系统理论研究. Email: 18223108804@126.com.

引用格式: 余晨曦, 尹彦力. 基于 DYCORS 算法的 OVA-SVM 参数优化与应用研究[J]. 重庆工商大学学报(自然科学版), 2024, 41(1): 38—44.

YU Chenxi, YIN Yanli. Optimization and application of OVA-SVM parameters based on DYCORS algorithm[J]. Journal of Chongqing Technology and Business University (Natural Science Edition), 2024, 41(1): 38—44.

simultaneously, so that the parameters were adjusted in blocks to improve the effect of parameter optimization. **Results** Through the comparison of the experimental results on MNIST and IRIS data sets, it can be found that after using the YDYCORDS algorithm to adjust the parameters of OVA-SVM in blocks, the model accuracy can be higher than the accuracies of manual parameter adjustment and directly using DYCORDS to adjust the five parameters at the same time, which can also further improve its model performance. **Conclusion** The final experimental results show that DYCORDS algorithm can effectively solve the problems of OVA-SVM parameter optimization, such as high time cost, large memory occupation, difficulty in solving high-dimensional data problems, and difficulty in finding the global optimal solution. In particular, the improved YDYCORDS algorithm can further improve the accuracy of the OVA-SVM model and obtain a better model effect.

**Keywords:** hyperparametric optimization; support vector machine; DYCORDS algorithm

## 1 引言

支持向量机(SVM)是一种常见的二元分类模型,它分为硬间隔SVM、软间隔SVM、核化SVM等,主要用于模式识别、数据分析、分类问题等领域。由于SVM通常能解决二元分类问题,因此它在解决多元分类问题时需要先将其转化为二元分类问题,再用SVM进行分类。对于二元分类问题,通常将能使数据完美分成两部分的分割面称为超平面,二维空间中的超平面是一条直线,三维空间中的超平面是一个平面。SVM的目标就是寻找一个最优超平面,使得所有数据点都尽可能地远离此分割平面,从而将一个空间中不同类别的数据完美地分为两类。

一对多法(One Versus All,简称OVA)<sup>[1]</sup>是最早的基于SVM对多类别进行分类的方法,其思想是将多类别问题转化为两类别问题来实现。

OVA-SVM是把支持向量机(SVM)和One Versus All(OVA)方法结合后提出的一个多类别分类判别模型,其性能依赖于模型参数的选择,其参数包括惩罚参数 $C$ 、核函数类型 $k(k, \text{kernel})$ 、RBF核函数参数 $\gamma(\gamma, \text{gamma})$ 、poly核函数的参数 $d(d, \text{degree})$ ,以及迭代终止参数 $t$ 等。如何选取最优的参数设置,进一步提升模型性能至关重要。OVA-SVM是目前应用比较多的一种机器学习模型,具有良好的泛化能力,其性能依赖于超参数的选择。当前较为常见的超参数优化方法有手动调参、网格寻优、随机寻优、贝叶斯优化方法、基于梯度的优化方法、进化寻优等。

手动调参,即研究者们通过一步步地手动调节超参数来实现对其的优化。研究者们可以判断如何设置超参数可以获得更好的模型效果,但这一方法依赖大量的经验,且人工调节较为耗时。

网格寻优,王兴玲等<sup>[2]</sup>将网格搜索用于向量机核函数的参数确定问题。使用网格寻优,首先需要为每一个超参数构建其各自的模型,再对各个模型的性能进行评析,从而获得性能最佳的模型和最优超参数。其优点在于该方法中的所有样本均能得到验证;其缺

点在于当超参数的个数较多时,计算量会直线上升,比较耗时,并且这一方法无法保证能获得最优的超参数值,也难以保证能获得全局最优解。

随机寻优,Droste等<sup>[3]</sup>对黑盒优化中随机搜索启发式算法进行了研究。该方法可以基于时间和计算资源情况,选择合理的计算次数,并控制组合的数量,在超参数网格的基础上随机选择组合来训练模型。通常情况下,随机寻优方法可以分辨哪些超参数对模型性能影响最大,从而优先获得这些超参数的最优值。与网格寻优相比,该方法能更快地找到最优参数,也能通过较少的迭代次数获得正确的参数设置;其缺点在于它只适用于低维数据情况,并且只能够针对单独的超参数组合模型进行训练,并评估各自的性能。

贝叶斯优化,Shukla等<sup>[4]</sup>提出通过贝叶斯优化算法自动设置超参数,走向自主强化学习。该方法通过构造一个函数的后验分布来工作,它可以在前一模型的基础上对后一模型的超参数选择产生影响,在一定程度上解决了模型独立时,各模型之间失去联系的问题。随着观测次数的增加,后验分布得到改善,算法更加确定参数空间中哪些是值得探索的区域,哪些是不值得探索的区域。其优点在于该方法通过不断采样来推测函数的最大值,但采样点不多;其缺点在于容易陷入局部最优值的区域内重复采样,难以找到全局最优解,并且需要进行多次迭代,时间长且成本高。

基于梯度的优化方法,Lu等<sup>[5]</sup>提出基于梯度下降法的机器学习增强优化方案。该方法应用场景并不广泛,主要通过梯度下降算法对超参数进行优化。其优点在于梯度下降算法并不高的计算量,使其可以应用于大部分高规模数据集中;其缺点在于由于超参数优化通常是非凸的,并不是一个完全平稳的过程,所以这一方法的应用场景并不广泛,难以在各场景中找到全局最优解。

进化寻优,严盛雄<sup>[6]</sup>提出了基于进化算法的LCL型滤波器参数优化设计。该算法是一个动态过程,类似于自然进化中不断变化的环境,它经常被用来寻找

其他技术不易求解的近似解。由于优化问题可能太耗时并且计算资源占用很大,因此通常难以找到精确的解决方案,在这种情况下,进化算法通常可以找到一个局部最优解。其优点在于可以产生出不受人类偏见影响的解决方案;其缺点在于不能保证找到全局最优解,没有可靠的算法终止条件和足够的理论基础,并且算法复杂度较高,容易造成资源的消耗。

综上所述,现有的参数优化方法多样,但普遍还存在以下问题:时间成本较大,内存占用较大,难以解决高维数据情况,难以找到全局最优解。现有工作表明:OVA-SVM 的性能高度依赖其模型参数的选择,如何选取最优的参数设置,进一步提升模型性能至关重要,而DYCORS 算法可以在节约时间成本和内存的前提下,对高维数据问题也能找到全局最优解。故本文针对现有参数优化方法存在的问题,提出针对 OVA-SVM 模型参数分块优化的 YDYCORS 算法,该算法在参数优化过程中,与现有方法相比更高效,使用范围更广,更易找到全局最优解。

## 2 DYCORS 算法与参数优化过程设计

### 2.1 DYCORS 算法

DYCORS 算法由 Regis 等<sup>[7]</sup>提出,算法结合了动态坐标扰动与算法重启策略,是目前最有效的昂贵“黑箱”函数优化算法之一。对于大部分测试函数,DYCORS 算法最终都能求得目标函数值。大部分高维问题最有效的算法还是 DYCORS,它不仅能解决上文中提到的时间成本与内存占用较大的问题,而且也能适用于高维数据的情况,并得到全局最优解。该算法通过动态坐标扰动,每次按照扰动概率(一个单调递减的函数)对当前最优点的每一维坐标进行扰动,以此获得一个新的采样点。本文将该算法应用于 OVA-SVM 的参数优化问题,以提升 OVA-SVM 的性能。其中 DYCORS 算法如表 1 所示。

表 1 DYCORS 算法框架

Table 1 DYCORS algorithm framework

DYCORS 算法框架	
输入:	
	初始参数 $n_0 = 2(D+1)$ ;
	候选点个数 $m = 100D$ ;
	最大迭代步数 $N_{\max}$ ;
	初始迭代步 $n = n_0$ ;
	目标函数 $f$ ;
输出:	
	优化后的参数配置 $x_{\text{best}}$ ;
	$x_{\text{best}}$ 参数配置下的概念表示;

续表(表 1)

Step1:	使用拉丁超立方体设计初始的 $n_0$ 个采样点,并设置集合 $I = \{x_i\}_{i=1}^{n_0}$ ;
Step2:	计算初始采样点的函数值 $f(x_i)$ ,得到集合 $A_{n_0} = \{(x_i, f(x_i))\}_{i=1}^{n_0}$ ;
Step3:	使用集合 $A_{n_0}$ 中的点去构建或更新响应面模型 $S_n(x)$ ;
Step4:	找出目前最优的点 $x_{\text{best}} = \text{argmin}\{f(x_i): i=1, \dots, n\}$ ;
Step5:	计算坐标扰动概率 $\varphi_n$ ;
Step6:	构建候选点集 $\Omega_n$ , 具有 $m$ 个候选点 $t_{n,1:m}$ , 其中, 对于每一个候选点 $y_j \in t_{n,1:m}$ ;
	(6a) 设置 $y_j = x_{\text{best}}$ ;
	(6b) 按照扰动概率 $\varphi_n$ , 选择 $y_j$ 的扰动坐标;
	(6c) 给选择的扰动坐标加一扰动 $\delta_i$ , $\delta_i$ 服从正态分布 $N(0, \sigma_n^2)$ ;
Step7:	对候选点计算函数值度量 $V_n^{\text{ev}}(t_{n,1:m})$ 、距离度量 $V_n^{\text{dm}}(t_{n,1:m})$ 和加权分数 $W_n(t_{n,1:m})$ ;
Step8:	根据当前 $W_n(t_{n,1:m})$ , 计算当前最优点 $x^* = \text{argmin}\{W_n(t_{n,1:m})\}$ ;
Step9:	根据当前最优参数配置 $x^*$ , 训练模型, 计算当前函数值 $f(x^*)$ ;
Step10:	调整方差 $\sigma_n^2$ ;
Step11:	更新集合 $A_{n+1} = \{A_n U(x^*, f(x^*))\}$ ;
Step12:	若 $n < N_{\max}$ , 则设置 $n = n+1$ , 然后返回 Step3, 否则程序终止, 返回 $x^*$ 。

扰动概率  $\varphi_n$  (对应 Step5) 是关于迭代次数  $n_0 \leq n \leq N_{\max} - 1$  的严格单调递减函数, 其计算表达式为

$$\varphi_n = \varphi_0 \left[ 1 - \frac{\ln(n - n_0 + 1)}{\ln(N_{\max} - n_0)} \right], n_0 \leq n < N_{\max}$$

由于  $\varphi_n$  表示概率, 所以其函数值在  $[0, 1]$ , 并且  $\varphi_0 = 1$ 。Step7 和 Step8 是从候选点中选出最有潜力的点作为新的采样点, 通过计算每一个候选点的函数值度量  $V^{\text{ev}}(t)$  与距离度量  $V^{\text{dm}}(t)$  作为其度量标准, 函数定义如下:

$$V^{\text{ev}}(t) = \begin{cases} \frac{s(t) - s^{\min}}{s^{\max} - s^{\min}}, & \text{if } s^{\max} \neq s^{\min} \\ 1, & \text{otherwise} \end{cases}$$

$$V^{\text{dm}}(t) = \begin{cases} \frac{\Delta^{\max} - \Delta(t)}{\Delta^{\max} - \Delta^{\min}}, & \text{if } \Delta^{\max} \neq \Delta^{\min} \\ 1, & \text{otherwise} \end{cases}$$

其中  $\Delta(t) = \min_{1 \leq i \leq n} \|t - t_i\|$ , 有了函数值度量和距离度量以后可以算出候选点的最终加权分数:  $W(t) = \omega V^{\text{ev}}(t) + (1 - \omega) V^{\text{dm}}(t)$ , 以最终加权分数最低的候选点作为新的采样点。

DYCORS 算法可以在减少迭代次数的前提下, 对于高维数据问题也能找到全局最优解, 故本文拟采用

DYCORS 算法调节 OVA-SVM 的惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代终止参数  $t$  等,获得较现有方法更优的参数配置,进一步提升模型性能。

### 2.2 参数优化过程设计

针对 OVA-SVM 参数优化时出现的问题,提出优化后的参数调节方法,提升模型性能,本文的主要贡献包含以下两个方面:

(1) OVA-SVM 模型的性能十分依赖于模型参数的选择,而该模型需要调节的参数有惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代终止参数  $t$ ,这些参数中既包含整数又包含连续实数,并且模型本身的训练时间成本也较高。而运用 DYCORS 算法进行参数调节可以减少迭代次数,并且对整数和连续实数超参进行优化,因此 DYCORS 算法是适用于 OVA-SVM 模型参数优化的方法。

(2) 由于同时调节 5 个参数计算量较大,可能难以找到全局最优解,因此本文针对 OVA-SVM 参数优化时出现的问题,在 DYCORS 算法的基础上进行参数分块调节,提出了改进后的 YDYCORS 算法。该算法先用 DYCORS 调节对 OVA-SVM 的性能影响最大的参数:惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ ;再固定最优惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ ,用 DYCORS 调节剩余参数中影响较大的参数:ploy 核函数的参数  $d$  与迭代的终止参数  $t$ ;最后再用 DYCORS 一起调节已获得的 5 个最优参数。先对参数进行分析,再对参数进行分块调节,可以在避免出现计算量过大的前提下,提升参数优化的效果。

本文中运用 DYCORS 算法优化参数的主要过程如图 1 所示。

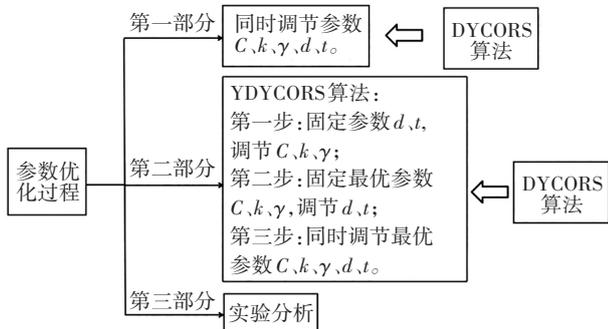


图 1 运用 DYCORS 算法优化参数的主要过程

Fig. 1 The main process of using DYCORS algorithm to optimize parameters

### 2.3 OVA-SVM 模型的参数影响分析

影响 OVA-SVM 性能的参数主要有惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代的终止参数  $t$  等。

惩罚参数  $C$ <sup>[11]</sup>,即对误差的宽容度,表示对离群点的重视程度。 $C$  越大,表示重视程度越高,允许它们被分类错误率高; $C$  越小,表示重视程度越低,允许它们被分类错误率低。

核函数<sup>[12]</sup>的引入是为了将高维空间中的内积运算转化至低维空间,从而解决线性不可分问题。核函数的选取,使得线性不可分的样本通过映射至更高维的空间中而变得线性可分,并且能解决映射过程中由于样本维度过高而造成的“维数灾难”,在操作过程中也可以忽略低维向高维映射函数的计算,只需计算低维空间中的核函数即可。常见的核函数类型( $\mathbf{x}, \mathbf{y}$  表示输入空间的向量)如表 2 所示。

表 2 常见核函数类型

Table 2 Common types of kernel functions

名称	表达式
线性核 (Linear Kernel)	$K(\mathbf{x}, \mathbf{y}) = \mathbf{x}^T \mathbf{y} + c$
多项式核 (Polynomial Kernel)	$K(\mathbf{x}, \mathbf{y}) = (\alpha \mathbf{x}^T \mathbf{y} + c)^d$
径向基核函数 (Radial Basis Function)	$K(\mathbf{x}, \mathbf{y}) = \exp\left(-\frac{\ \mathbf{x} - \mathbf{y}\ ^2}{2\gamma^2}\right)$
Sigmoid 核 (Sigmoid Kernel)	$K(\mathbf{x}, \mathbf{y}) = \tanh(\alpha \mathbf{x}^T \mathbf{y} + c)$

RBF 核函数表达式  $\exp\left(-\frac{\|\mathbf{x} - \mathbf{y}\|^2}{2\gamma^2}\right)$  中的  $\gamma$ ,即为 RBF 核函数参数;ploy 核函数表达式  $(\alpha \mathbf{x}^T \mathbf{y} + c)^d$  中的  $d$ ,即为 ploy 核函数参数;迭代终止参数  $t$  是为了避免使一个问题永远迭代下去而设置的参数。

### 3 基于 DYCORS 算法的 OVA-SVM 参数优化

本文涉及的参数主要有惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$ 、迭代终止参数  $t$ 。

针对现有超参数优化方法在调节 OVA-SVM 参数过程中存在的时间成本较大、内存占用较大、难以解决高维数据,且难以找到全局最优解等问题,提出运用 DYCORS 算法调节 OVA-SVM 参数的方法。以下将从 DYCORS 算法对参数  $C, k, \gamma, d, t$  同时调节,并对其参数优化过程进行分析。

#### 3.1 DYCORS 同时调节参数 $C, k, \gamma, d, t$

DYCORS 算法与现有的手动调参、网格寻优、随机寻优、贝叶斯优化方法、基于梯度的优化方法、进化寻优等超参数优化方法相比,可以更高效地完成参数优化,减小内存占用空间,同时能适用于高维数据问题,得到全局最优解,故 DYCORS 是一种十分适用于 OVA-SVM 参数优化的方法。

为实现 DYCORS 算法对 OVA-SVM 参数的优化,进一步提升 OVA-SVM 的性能,本节将运用 DYCORS 算法同时调节惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代的终止参数  $t$ 。

运用 DYCORS 算法进行参数优化的大致步骤如下:第一步,设置 OVA-SVM 的参数,初始采样点的个数  $n_0$ ,最大迭代步数  $N_{\max}$ ,确定模型各个参数的可行域  $D$ ;第二步,使用拉丁超立方体采样获得初始点列  $\{x_i\}_{i=1}^{n_0}$ ,  $x_i$  表示一组参数;第三步,采用每一组参数  $x_i$ ,计算初始采样点的函数值  $f(x_i)$ ,得到集合  $A_{n_0} = \{(x_i, f(x_i))\}_{i=1}^{n_0}$ ;第四步,使用集合  $A_{n_0}$  中的点去构建或更新响应面模型  $S_n(x)$ ;第五步,通过当前响应面模型得出最优解  $x^*$ ,再采用参数  $x^*$  训练模型,得到当前函数值  $f(x^*)$ ,更新集合  $A_{n+1} = \{A_n U(x^*, f(x^*))\}$ ;第六步,在达到最大迭代步数  $N_{\max}$  前,重复步骤四和步骤五,最终得到优化后的参数配置。

### 3.2 YDYCORS 算法的参数优化过程

由于同时调节 5 个参数计算量较大,可能难以获得全局最优解,因此本节对原有的 DYCORS 算法进行优化,提出改进后的 YDYCORS 算法。YDYCORS 算法对 5 个参数进行了分块,将调参过程分为了以下 3 个步骤:第一步,固定 ploy 核函数参数  $d$  与迭代终止参数  $t$ ,用 DYCORS 同时调节惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ ;第二步,在固定最优惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$  的前提下,用 DYCORS 同时调节 ploy 核函数参数  $d$  与迭代终止参数  $t$ ;第三步,用 DYCORS 同时调节最优惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代终止参数  $t$ 。

#### 3.2.1 固定参数 $d, t$ , DYCORS 同时调节参数 $C, k, \gamma$

据研究表明<sup>[8]</sup>,惩罚参数  $C$ 、核函数类型  $k$  以及 RBF 核函数参数  $\gamma$  很大程度上影响了 OVA-SVM 的性能,惩罚参数  $C$  通过影响学习机器置信范围和经验风险的比例,进而影响 OVA-SVM 的泛化能力,核函数类型与核函数的参数通过影响样本被映射到特征空间的复杂程度,进而影响 OVA-SVM 的性能。因此在提升 OVA-SVM 模型性能的过程中,惩罚参数  $C$ 、核函数类型  $k$  以及 RBF 核函数参数  $\gamma$  较其他参数更为重要。故第一步先固定参数  $d, t$ ,此时  $d, t$  的参数值是任意设置的,再运用 DYCORS 算法先对参数  $C, k, \gamma$  进行调节,具体步骤同上一节。

#### 3.2.2 固定最优参数 $C, k, \gamma$ , DYCORS 同时调节参数 $d, t$

第二步选择多项式核函数作为 OVA-SVM 的核,因此除惩罚参数  $C$ 、核函数类型  $k$  以及 RBF 核函数参数  $\gamma$  以外,ploy 核函数参数  $d$  和迭代终止参数  $t$  对 OVA-SVM 性能的影响较大。因此,固定上一小节得到的最优惩罚参数  $C$ 、核函数类型  $k$  以及 RBF 核函数参数  $\gamma$ ,运用 DYCORS 算法同时调节 ploy 核函数参数  $d$  和迭代终止参数  $t$ ,具体步骤同上一节。

#### 3.2.3 DYCORS 同时调节最优参数 $C, k, \gamma, d, t$

在经过前两步的参数调节后,OVA-SVM 参数优化的效果得到了进一步提升。为了使效果达到最佳,本小节利用参数之间的相互依赖性,在第三个步骤中用前两步获得的最优参数为初始参数值,同时调节最优惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代终止参数  $t$ ,具体步骤同上一节。

## 4 实验验证与结果分析

### 4.1 数据集

MNIST 数据集<sup>[9]</sup>来自美国国家标准与技术研究所, National Institute of Standards and Technology (NIST),是 NIST 数据库的一个子集,也是机器学习中评估模型性能时常用的数据集。MNIST 数据集分为训练集与测试集,训练图片一共有 60 000 张,可采用学习方法训练出相应的模型,测试图片一共有 10 000 张,可用于评估训练模型的性能。

IRIS 数据集<sup>[10]</sup>,也称鸢尾花卉数据集,其中共包含 150 个样本数据,该数据集是机器学习中常见的数据集,共分为以下 3 类:山鸢尾 (Iris Setosa)、杂色鸢尾 (Iris Versicolour)、维吉尼亚鸢尾 (Iris Virginica)。该数据集以鸢尾花的特征作为数据来源,包含 4 个特征:花瓣长度、花瓣宽度、花萼长度、花萼宽度。IRIS 数据集是一类多重变量分析的数据集,也是在数据分类中十分常用的测试集和训练集。

### 4.2 评估指标

本文实验的评估指标<sup>[9]</sup>主要是准确率,为阐述准确率 ( $A_{cc}$ ) 的含义,先介绍以下两个概念:

TP (True Positives): 表示模型将正实例预测为正实例的数据个数,用  $N_{TP}$  表示;TN (True Negatives): 表示模型将负实例预测为负实例的数据个数,用  $N_{TN}$  表示;准确率指标的公式:

$$A_{cc} = \frac{N_{TP} + N_{TN}}{N_{total}} \quad (N_{total} \text{ 表示待预测数据的总数})$$

### 4.3 实验环境

本文实验在具有 64 GB 内存和 NVIDIA GeForce GTX 1050 Ti GPU 和 Intel (R) Core (TM) i7-8750H CPU 的个人工作站上进行。

### 4.4 实验结果

针对现有方法存在的 4 个问题,设计如下实验:第一部分,手动调节 OVA-SVM 参数,获得在该参数设置下的模型准确率;第二部分,运用 DYCORS 算法同时调节参数  $C, k, \gamma, d, t$ ,获得在该参数设置下的模型准确率;第三部分,首先固定参数  $d, t$ ,运用 DYCORS 算法调节参数  $C, k, \gamma$ ,然后固定最优参数  $C, k, \gamma$ ,调节参数  $d, t$ ,最后同时调节最优参数  $C, k, \gamma, d, t$ ,获得在该参数设置下的模型

准确率。本实验针对 MNIST 和 IRIS 两个数据集,将 DYCORS 算法运用于 OVA-SVM 的参数优化中,其中先确定各参数的取值范围: $C$  的取值范围为 $[0, 10]$ , $k$  的取

值范围为 $[\text{linear}, \text{poly}, \text{rbf}, \text{sigmoid}]$ , $\gamma$  的取值范围为 $[0, 10]$ , $d$  的取值范围为 $[0, 5]$ , $t$  的取值范围为 $[0.000 1, 0.001]$ ,经过以上实验后可得如表 3、表 4 中的结果。

表 3 在 MNIST 数据集上的实验结果 (60%训练,40%测试)

Table 3 Experimental results on MNIST dataset (60% training, 40% testing)

调节方法	OVA-SVM 参数设置	模型准确率
手动调节	$C=3, k='linear', \gamma=8, d=3, t=0.001\ 000$	0.983 3
DYCORS 同时调节 $C, k, \gamma, d, t$	$C=5.416\ 700, k='poly', \gamma=6.250\ 000, d=3.541\ 700, t=0.000\ 888$	0.981 9
YDYCORS 第一步:固定 $d, t$ ,DYCORS 调节 $C, k, \gamma$	$C=0.848\ 400, k='rbf', \gamma=0.000\ 804$	0.986 1
YDYCORS 第二步:固定最优 $C, k, \gamma$ ,DYCORS 调节 $d, t$	$d=0.416\ 700, t=0.000\ 625$	0.986 1
YDYCORS 第三步:DYCORS 同时调节最优 $C, k, \gamma, d, t$	$C=0.843\ 200, k='rbf', \gamma=0.000\ 832, d=0.512\ 200, t=0.000\ 532$	0.986 1

表 4 在 IRIS 数据集上的实验结果 (60%训练,40%测试)

Table 4 Experimental results on IRIS dataset (60% training, 40% testing)

调节方法	OVA-SVM 参数设置	模型准确率
手动调节	$C=3, k='linear', \gamma=10, d=3, t=0.001\ 000$	0.966 7
DYCORS 同时调节 $C, k, \gamma, d, t$	$C=5.416\ 700, k='rbf', \gamma=0.416\ 700, d=1.458\ 300, t=0.000\ 663$	1
YDYCORS 第一步:固定 $d, t$ ,DYCORS 调节 $C, k, \gamma$	$C=0.820\ 500, k='linear', \gamma=5.654\ 800$	1
YDYCORS 第二步:固定最优 $C, k, \gamma$ ,DYCORS 调节 $d, t$	$d=0.416\ 667, t=0.000\ 475$	1
YDYCORS 第三步:DYCORS 同时调节最优 $C, k, \gamma, d, t$	$C=0.803\ 300, k='linear', \gamma=5.332\ 500, d=0.404\ 100, t=0.000\ 131$	1

如表 3 在 MNIST 数据集上的实验结果所示:运用 DYCORS 算法调节 OVA-SVM 模型的参数时,由于同时调节 5 个参数工作量较大,得到的模型准确率为 98.19%,但在运用 YDYCORS 算法调节 OVA-SVM 模型的参数后,模型准确率达到 98.61%,能得到更好的效果。

由表 4 在 IRIS 数据集上的实验结果可知:在运用 DYCORS 算法和 YDYCORS 算法调节 OVA-SVM 的参数时,都能得到较高的准确率与良好的模型效果。为了更清晰地反映模型准确率在不同数据集与不同调参方法下的具体情况,本文再将上述两个数据集的实验结果进行对比,如表 5 所示。

通过 MNIST 和 IRIS 两个数据集上的实验结果对比可以发现:由于数据集的选择不同,呈现的模型准确率也不同,明显在 IRIS 数据集上的模型准确率更高,但在 MNIST 数据集上,YDYCORS 算法的优势体现得更明显。运用 DYCORS 算法直接同时优化 5 个参数,计算量较大,可能无法断定是否能得到较高的模型准确率,但在运用 YDYCORS 算法对 OVA-SVM 的参数进行分块调节后,能得到与手动调参和直接用 DYCORS 同时调节 5 个参数相比更高的模型准确率,从而也能进一

步提升其模型性能。综上实验,可以进一步说明本文提出的 YDYCORS 算法可以在一定程度上解决现有超参数优化方法还存在的问题,因此本文对 OVA-SVM 参数的优化也取得了比较理想的效果。

表 5 两个数据集上的实验结果对比

Table 5 Comparison of experimental results on two data sets

调节方法	MNIST 数据集上的准确率	IRIS 数据集上的准确率
手动调节	0.983 3	0.966 7
DYCORS 同时调节 $C, k, \gamma, d, t$	0.981 9	1
YDYCORS 第 1 步:固定 $d, t$ ,DYCORS 调节 $C, k, \gamma$	0.986 1	1
YDYCORS 第 2 步:固定最优 $C, k, \gamma$ ,DYCORS 调节 $d, t$	0.986 1	1
YDYCORS 第 3 步:DYCORS 同时调节最优 $C, k, \gamma, d, t$	0.986 1	1

## 5 结论及展望

### 5.1 结论

本文针对现有的超参数优化方法存在的时间成本较大,内存占用较大,难以解决高维数据,难以找到全

局最优解等问题,运用 DYCORDS 算法对模型影响较大的惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  以及迭代终止参数  $t$  这 5 个参数进行调节,提出将参数进行分块调节的 YDYCORDS 算法:先用 DYCORDS 调节对 OVA-SVM 的性能影响最大的参数:  $C$ 、 $k$ 、 $\gamma$ ;再固定最优参数  $C$ 、 $k$ 、 $\gamma$ ,用 DYCORDS 调节剩余参数中影响较大的参数  $d$  和  $t$ ;最后用 DYCORDS 一起调节已获得的 5 个最优参数。先对参数进行分析,再对参数进行分块调节,可以在避免出现计算量过大的前提下,提升参数优化的效果。

通过 MNIST 和 IRIS 两个数据集上的实验结果对比可以发现:运用 DYCORDS 算法直接同时优化 5 个参数,由于计算量较大,可能无法断定是否能得到较高的模型准确率,但在运用 YDYCORDS 算法对 OVA-SVM 的参数进行分块调节后,能得到与手动调参和直接用 DYCORDS 同时调节 5 个参数相比更高的模型准确率,从而也能进一步提升模型性能,取得了较佳的模型效果。

## 5.2 展望

虽然本文提出的方法可以在一定程度上解决现有超参数优化方法存在的问题:如时间成本较大、内存占用较大、难以解决高维数据问题、难以找到全局最优解,但该方法能否解决超参数优化中可能存在的其他问题还有待研究。另外,本文只对惩罚参数  $C$ 、核函数类型  $k$ 、RBF 核函数参数  $\gamma$ 、ploy 核函数参数  $d$  和迭代终止参数  $t$  进行优化,而如何更全面地优化 OVA-SVM 的参数还值得未来进一步探索。同时,本文提出的 YDYCORDS 算法只对 5 个参数进行了一种形式的分块,其余的分块形式能否取得更佳的模型效果还值得研究。

## 参考文献(References):

- [1] 于彬. 基于 SVM-OVA 方法的多类别基因表达数据分类[J]. 青岛科技大学学报(自然科学版), 2008(4): 318—320.  
YU Bin. Classification of multi-class gene expression data based on SVM-OVA method [J]. Journal of Qingdao University of Science and Technology (Natural Science Edition), 2008(4): 318—320.
- [2] 王兴玲, 李占斌. 基于网格搜索的支持向量机核函数参数的确定[J]. 中国海洋大学学报(自然科学版), 2005(5): 859—862.  
WANG Xing-ling, LI Zhan-bin. Parameter determination of support vector machine kernel function based on grid search[J]. Journal of Ocean University of China (Natural Science Edition), 2005(5): 859—862.
- [3] DROSTE S, JANSEN T, WEGENER I. Upper and lower bounds for randomized search heuristics in black-box optimization [C]//Electron Colloquium Comput. Complex, 2003: 48—48.
- [4] SHUKLA P K. On gradient based local search methods in unconstrained evolutionary multi-objective optimization [C]. EMO, 2006: 96—110.
- [5] LU W F, CAI B Y, GU R. Improved particle swarm optimization based on gradient descent method [C]. CSAE, 2020: 121—126.
- [6] 严盛雄. 基于进化算法的 LCL 型滤波器参数优化设计[D]. 武汉: 华中科技大学, 2019.  
YAN Sheng-xiong. Parameter optimization design of LCL filter based on evolutionary algorithm [D]. Wuhan: Huazhong University of Science and Technology, 2019.
- [7] REGIS R G, SHOEMAKER C A. Combining radial basis function surrogates and dynamic coordinate search in high-dimensional expensive black-box optimization[J]. Engineering Optimization, 2013, 45(5): 529—555.
- [8] WANG Y F, LIU D, PEI Z F, et al. Parameter optimization and application of SVM based on improved gravitational search algorithm[J]. Computer Application Research, 2020, 37(S1): 152—154.
- [9] 肖驰. 四种机器学习算法在 MNIST 数据集上的对比研究[J]. 智能计算机与应用, 2020, 10(12): 185—188.  
XIAO Chi. A comparative study of four machine learning algorithms on MNIST dataset [J]. Intelligent Computer and Application, 2020, 10(12): 185—188.
- [10] 邹心遥, 陈敬伟, 姚若河. 采用粒子群优化的 SVM 算法在数据分类中的应用[J]. 华侨大学学报(自然科学版), 2016, 37(2): 171—174.  
ZOU Xin-yao, CHEN Jing-wei, YAO Ruo-he. Application of SVM algorithm based on particle swarm optimization in data classification [J]. Journal of Huaqiao University (Natural Science Edition), 2016, 37(2): 171—174.
- [11] LIU Q, XU Y Q, ZHOU Y. A class of exact penalty functions and penalty algorithms for nonsmooth constrained optimization problems[J]. J Glob Optim, 2020, 76(4): 745—768.
- [12] WANG H, XU D X, ALEXANDER M. Parameter selection method for support vector machine based on adaptive fusion of multiple kernel functions and its application in fault diagnosis[J]. Neural Comput Appl, 2020, 32(1): 183—193.

责任编辑:李翠薇