文章编号:1671-6833(2026)04-0074-06

基于 SMC 策略的线性不可分数据集的分类算法

张光晨[1], 李展菲[1], 何舒平[2], 夏元清[3]

(1.北方民族大学 数学与信息科学学院,宁夏 银川 750021;2.安徽大学 电气工程与自动化学院,安徽 合肥 230601;3.北京理工大学 自动化学院,北京 100081)

摘 要:针对线性不可分数据集的分类问题,通过将滑模控制(SMC)思想用于支持向量(SVM)核函数参数优化过程,提出一种基于 SMC 策略的 SVM 核函数参数优化算法。设计误差方程和滑模面,建立 SVM 分类目标函数与 SMC 的关联,并推导出核函数参数迭代更新规则及代价函数。通过动态调整 SVM 核函数参数,在减少支持向量数量的同时提升分类性能。实验部分采用 Iris、Heart disease 等 6 个 UCI 数据集验证算法有效性。结果表明:与传统 SVM 相比,所提算法在 Iris 数据集上支持向量减少 56.25%,测试准确率保持 100%;在 Heart-disease 数据集上测试准确率提升 13.58 百分点。所提算法与现有优化算法对比,在多个数据集上表现出更高的分类精度。

关键词:支持向量机; 滑模控制策略; 分类算法; 核函数; 参数

中图分类号: TP181;O224 **文献标志码**:A **doi**:10.13705/j.issn.1671-6833.2026.04.004

二分类算法逻辑简洁、执行效率高,在机器学习与数据分析领域有着广泛应用[1-2]。常见的二分类算法包括逻辑回归模型、决策树以及支持向量机(support vector machine,SVM)等。作为一种监督式机器学习算法,SVM 核心思想是通过支持向量构建分类超平面以实现对数据的划分。对于线性可分的数据集,SVM 通过寻求最优分类边界实现唯一性分类,这一特性显著区别于其他二分类算法,因此在诸多领域得到广泛应用[3-5]。然而线性分类方法存在对噪声敏感、特征处理依赖性高及泛化能力受限等缺陷,限制了其实际应用范围。线性不可分的数据集在实际问题中普遍存在,使得线性分类方法面临更大的挑战。为此,研究者们引入核函数技术,通过将低维不可分数据映射至高维特征空间,实现非线性数据解耦与分类边界的优化。例如 Adeyanju 等[6]研究分析了不同核函数在同一支持向量机模型上的分类性能,实现了 7 种人脸面部表情的识别分类,且平均准确率达到 99.33%;Zhou 等[7]提出了基于贝叶斯优化支持向量机的方法,并将该方法应用于识别城市污水泵中的多种气蚀状态,仿真实验证明,在不同流量工况下,该算法的识别准确率可达到 99.3%及以上。

近年来,随着对 SVM 研究的不断深入,许多学者使用智能优化算法优化 SVM 核函数参数,以达到缩短分类时间并提高分类精度的目标。Zhang 等[8]提出一种改进的麻雀搜索的 SVM(ISSA-SVM)算法来优化 SVM 核函数参数,与普通 SVM 相比,该分类算法精度明显提高,并缩短了分类时间。Abbaszadeh 等[9]提出了基于粒子群优化改进支持向量分类器(PSO-SVM)模型参数的算法,以克服传统网格搜索方法的不足,实验证明该方法在朱斑岩铜矿的地质建模中准确率可达到 97%以上。Li 等[10]提出了一种基于灰狼社会等级的支持向量回归模型,实验结果表明所提算法具备更强的全局搜索能力、更快的收敛速度和更好的稳定性。

目前使用的智能优化算法大多属于开环算法,无法实时反馈分类数据误差并更新参数值,导致分类失准。另一方面,受外界干扰和网络限制的双重影响,数据集之间的耦合非线性现象显著增强,这些特征在数据集中普遍存在。针对上述问题,已有研究者尝试借鉴滑模控制思想中的反馈机制与机器学习相结合以增强自适应性。滑模控制(sliding mode

收稿日期:2026-01-10;**修订日期**:2026-02-12

基金项目:国家自然科学基金资助项目(62563002);2024 年度宁夏回族自治区青年拔尖人才培养项目;北方民族大学创新项目(YCX24288)

作者简介:张光晨(1985—),男,山西晋中人,北方民族大学副教授,博士,主要从事复杂网络智能控制、数据分析与决策、大模型理论及应用研究,E-mail:guangchen_123@163.com。

引用本文:张光晨,李展菲,何舒平,等. 基于 SMC 策略的线性不可分数据集的分类算法[J]. 郑州大学学报(工学版),2026,47(4):74-79.[Zhang Guangchen, Li Zhanfei, He Shuping, et al. Classification algorithm for linearly inseparable datasets based on SMC strategy[J]. Journal of Zhengzhou University (Engineering Science),2026,47(4):74-79.]

control,SMC)是处理非线性问题和不确定性系统的闭环方法,对克服模型误差、参数变化和外部干扰具有良好的鲁棒性[11]。近年来有学者将 SVM 与 SMC 策略进行了融合。刘凌等[12]针对永磁同步电机的调速性能受电机系统状态变量影响的问题,提出了一种采用 SVM 智能预测永磁同步电机可变参数非奇异快速终端滑模控制策略。在基于 SMC 理论的 SVM 核函数参数优化算法方面,Yalsavar 等[13]提出了一种基于滑模控制算法的 SVM 核函数参数优化算法,与传统 SVM 算法相比,该算法收敛速度更快,鲁棒性更强,然而其未明确定义分类目标的函数形式,降低了模型的透明度和可解释性,进而缺乏分类算法收敛性的理论支撑。

鉴于 SMC 的良好性能,本文提出一种基于 SMC 策略和非线性 SVM 核函数参数优化相结合的分类算法。该方法旨在提高分类精度并寻找更小的支持向量集。本文的主要工作如下。

(1)定义了分类误差的具体形式,通过引入适定的代价函数,为整个学习过程提供了系统的框架和范式。

(2)将 SMC 策略应用于 SVM 核函数参数的寻优中,基于 SMC 策略和平方损失函数定义了相应的算法代价函数,证明了代价函数最小值的存在性,并给出了算法的收敛性和稳定性的证明。

1　模型建立

在本文中,将借助核函数对线性不可分数据集进行分类算法设计。

定义 1　SVM 分类函数[14]为

$$y = \mathrm{sign}\left[\sum_{i=1}^{N}\alpha_i y_i K(x,x_i) + b\right]。\tag{1}$$

式中:$\sum_{i=1}^{N}\alpha_i y_i K(x,x_i) + b$ 为不同类别数据点的分类超平面;$K(x,x_i)$ 为核函数;x 为输入的待分类数据点;y 为 x 的输出;x_i 为训练样本,用于训练 SVM 模型,y_i 为 x_i 的输出,则训练样本可表示为 $D = \{(x_1, y_1),(x_2,y_2),\cdots,(x_N,y_N)\}$;$\alpha_i$ 为权重变量;b 为位移项,表示超平面与坐标原点之间的距离;N 为样本总数。

在 SVM 分类结果中,当 $y_d = y$ 时实现了正确分类;当 $y_d \neq y$ 时存在两种错误分类情况,即

$$\begin{cases} y_d = 1, y = -1; \\ y_d = -1, y = 1。 \end{cases}\tag{2}$$

式中:y_d 为待分类数据点 x 的真实标签值。

本文将使用加权的高斯型函数[15]作为核函数,即 $K(x,x_i) = \exp(-\gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x})$,其中,$\boldsymbol{A} \geqslant 0$ 为半正定矩阵;列向量 $\boldsymbol{x} = (x - x_1, x - x_2, \cdots, x - x_N)^{\mathrm{T}}$;$\gamma$ 为待优化的参数。

引理 1　当 $\boldsymbol{A} \geqslant 0, \gamma > 0$ 时,$K(x,x_i) = \exp(-\gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x})$ 是核函数。

证明　首先,因 $\boldsymbol{A} \geqslant 0$ 则存在矩阵 $\boldsymbol{C}$ 满足 $\boldsymbol{A} = \boldsymbol{C}^{\mathrm{T}}\boldsymbol{C}$,对于 $\widetilde{K}(x,x_i)$,有

$$\widetilde{K}(x,x_i) = \gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x} = (\sqrt{\gamma}\boldsymbol{C}\boldsymbol{x})^{\mathrm{T}}(\sqrt{\gamma}\boldsymbol{C}\boldsymbol{x}) = \phi(\boldsymbol{x})^{\mathrm{T}}\phi(\boldsymbol{x})。\tag{3}$$

式中:$\phi(\boldsymbol{x}) = \sqrt{\gamma}\boldsymbol{C}\boldsymbol{x}$。

根据核函数的定义可知 $\widetilde{K}(x,x_i)$ 是核函数。之后根据核函数的性质[16-17],可知 $K(x,x_i) = \exp(-\gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x})$ 是核函数。

至此,引理 1 得证。

将引理 1 的核函数代入式(1)可得加权高斯核分类函数 y 的表达式为

$$y = \mathrm{sign}\left(\sum_{i=1}^{N}\alpha_i y_i \exp(-\gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}) + b\right)。\tag{4}$$

在使用 SMC 优化 SVM 的核函数参数 γ 时,算法迭代过程中需要解决求导问题。然而符号函数求导之后不可逆,因此将符号函数替换为 sigmoid 函数以求解和表述分类问题。具体而言,本文使用 $\mathrm{sigmoid}(u) = \dfrac{1}{1 + \mathrm{e}^{-u}}$ 和 $-\mathrm{sigmoid}(-u) = \dfrac{-1}{1 + \mathrm{e}^{u}}$ 两个表达式分别处理当 $y_d \neq y$ 时的两种错误分类数据点,其中 $u = \sum_{i=1}^{N}\alpha_i y_i \exp(-\gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}) + b$。

通过误差 e 建立起 SMC 与 SVM 之间的联系。

定义 2　误差 $e = |y(\gamma) - y_d|$,其中

$$y(\gamma) = \begin{cases} \mathrm{sigmoid}\left(\sum_{i=1}^{N}\alpha_i y_i \exp(-\gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}) + b\right), \\ y_d = 1, y = -1; \\ -\mathrm{sigmoid}\left(-\sum_{i=1}^{N}\alpha_i y_i \exp(-\gamma \boldsymbol{x}^{\mathrm{T}}\boldsymbol{A}\boldsymbol{x}) - b\right), \\ y_d = -1, y = 1。 \end{cases}\tag{5}$$

基于误差 e,进一步设计滑模面和趋近律方程分别为

$$\dot{e} + \lambda e = S \equiv 0;\tag{6}$$

$$\ddot{e} + \lambda \dot{e} = -f(S)。\tag{7}$$

式中:λ 为严格正常数;S 为所设计的滑模面;$f(S)$ 为任意非递减的奇函数。所设计的滑模面与趋近律方程在保证算法渐进稳定性的同时可以提高算法的

分类精度。

2 参数迭代的主要结果

2.1 基于 SMC 策略的最优核函数参数 γ^* 的求解

针对两种错误分类数据点,设计合理的 SMC 策略,以选取 γ 最优解。

定理 1 存在 γ 值,使得滑模面方程式(6)与趋近律方程式(7)成立。

证明 满足定理 1 的 γ 最优解是存在的。

当 $y_d = 1, y = -1$ 时,由定义 2 中的式(5)可得

$$y(\gamma) = \text{sigmoid}\Big(\sum_{i=1}^{N} \alpha_i y_i \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}) + b)\Big)。\quad (8)$$

根据滑模面与趋近律方程,可求得误差 e 关于核函数参数 γ 的一阶导数 $\dot{e}$ 与二阶导数 $\ddot{e}$ 分别为

$$\dot{e} = \Big(\sum_{i=1}^{N} (-\alpha_i y_i \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}) \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})\Big) \cdot \text{sign}(y(\gamma) - y_d) y(\gamma)(1 - y(\gamma)); \quad (9)$$

$$\ddot{e} = \Big(\sum_{i=1}^{N} \alpha_i y_i \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x} \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})\Big)^2 \cdot \text{sign}(y(\gamma) - y_d)(1 - 2y(\gamma)) y(\gamma)(1 - y(\gamma)) + \Big(\sum_{i=1}^{N} \alpha_i y_i (\boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})^2 \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})\Big) \cdot \text{sign}(y(\gamma) - y_d) y(\gamma)(1 - y(\gamma))。\quad (10)$$

进一步由滑模面公式 $S = \dot{e} + \lambda \ddot{e}$, 可得

$$S = \Big(\sum_{i=1}^{N} (-\alpha_i y_i \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}) \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})\Big) \cdot \text{sign}(y(\gamma) - y_d) y(\gamma)(1 - y(\gamma)) + \lambda |y(\gamma) - y_d|; \quad (11)$$

$$\dot{S} = \text{sign}(y(\gamma) - y_d)(1 - 2y(\gamma)) y(\gamma)(1 - y(\gamma)) \cdot \Big(\sum_{i=1}^{N} m_i q_i(\gamma)\Big)^2 + \text{sign}(y(\gamma) - y_d)(1 - y(\gamma)) y(\gamma) \cdot \sum_{i=1}^{N} n_i q_i(\gamma) + \lambda \text{sign}(y(\gamma) - y_d)(1 - y(\gamma)) y(\gamma) \cdot \sum_{i=1}^{N} m_i q_i(\gamma) = -f(S)。\quad (12)$$

式中: $m_i = \alpha_i y_i \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$; $q_i(\gamma) = \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})$; $n_i = m_i \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x}$。

仅对 $y_d = 1, y = -1$ 的错误分类数据点进行讨论, $y_d = -1, y = 1$ 的情况类似可得。用 $y_d = 1, y = -1$ 替换式(12)中的 y_d 和 $y(\gamma)$, 可得

$$6\Big(\sum_{i=1}^{N} m_i q_i(\gamma)\Big)^2 + 2\lambda\Big(\sum_{i=1}^{N} m_i q_i(\gamma)\Big) + 2\sum_{i=1}^{N} n_i q_i(\gamma) = -f(S)。\quad (13)$$

令 $\boldsymbol{M} = (m_1, m_2, \cdots, m_i, \cdots, m_N)^{\mathrm{T}}$, $\boldsymbol{P} = (n_1, n_2, \cdots, n_i, \cdots, n_N)^{\mathrm{T}}$, $\boldsymbol{q} = (q_1(\gamma), q_2(\gamma), \cdots, q_i(\gamma), \cdots, q_N(\gamma))^{\mathrm{T}}, i = 1, 2, \cdots, N$。通过式(13)求解 γ, 将其表示为矩阵的形式:

$$6\boldsymbol{q}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{M}^{\mathrm{T}} \boldsymbol{q} + 2(\boldsymbol{P}^{\mathrm{T}} + \lambda \boldsymbol{M}^{\mathrm{T}}) \boldsymbol{q} + f(S) = 0。\quad (14)$$

$\boldsymbol{q}$ 的最优解通过对式(14)关于 $\boldsymbol{q}$ 求偏导来确定:

$$\frac{\partial(6\boldsymbol{q}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{M}^{\mathrm{T}} \boldsymbol{q} + 2(\boldsymbol{P}^{\mathrm{T}} + \lambda \boldsymbol{M}^{\mathrm{T}}) \boldsymbol{q} + f(S))}{\partial \boldsymbol{q}} = 6\boldsymbol{q}^{\mathrm{T}} \boldsymbol{M} \boldsymbol{M}^{\mathrm{T}} + 6\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}} \boldsymbol{q} + 2(\boldsymbol{P}^{\mathrm{T}} + \lambda \boldsymbol{M}^{\mathrm{T}}) = 0。\quad (15)$$

可得 $\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}} \boldsymbol{q} = -\frac{1}{6}(\boldsymbol{P} + \lambda \boldsymbol{M})$。

因此, $\boldsymbol{q}$ 的一个估计值可以由 $\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}}$ 的伪逆矩阵 $(\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}})^{\dagger}$ 计算得到:

$$\widetilde{\boldsymbol{q}} = -\frac{1}{6}(\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}})^{\dagger}(\boldsymbol{P} + \lambda \boldsymbol{M})。\quad (16)$$

向量 $\widetilde{\boldsymbol{q}}$ 中的元素 $\widetilde{q}_i(\gamma)$ 为

$$\widetilde{q}_i(\gamma) = -\frac{1}{6}\sum_{j=1}^{N} (\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}})^{\dagger}_{i,j}(n_j + \lambda m_j)。\quad (17)$$

式中: $(\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}})^{\dagger}_{i,j}$ 为 $(\boldsymbol{M} \boldsymbol{M}^{\mathrm{T}})^{\dagger}$ 的第 i 行第 j 列元素。

然而, $q_i(\gamma) = \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})$ 恒为正值,因此只有正元素 $\widetilde{q}_i(\gamma)$ 才能满足 $q_i(\gamma) = \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})$。因此,利用正元素 $\widetilde{q}_i(\gamma)$ 可以计算出 γ 的最优解:

$$\gamma^* = \frac{1}{k}\sum_{i=1}^{k} \gamma_i。\quad (18)$$

式中: k 为正元素 $\widetilde{q}_i(\gamma)$ 的个数。

至此,定理 1 得证。

在保证误差的同时,定义代价函数为

$$J = (y(\gamma) - y_d)^2 + (\boldsymbol{q}^{\mathrm{T}}(6\boldsymbol{M}\boldsymbol{M}^{\mathrm{T}})\boldsymbol{q} + 2(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{q} + \lambda \boldsymbol{M}^{\mathrm{T}})\boldsymbol{q})。\quad (19)$$

式中:等号右侧第 1 项为平方损失函数,促使算法关注减少较大的误差,从而使模型的预测值 $y(\gamma)$ 更加接近实际值 y_d; 第 2 项为引入的先验知识和约束,通过调整优化方向保证算法泛化能力的同时使其满足特定的条件。

定理 2 存在 λ 使得代价函数式(19)达到最小值。

证明 代价函数 J 存在最小值。

要使得 J 达到最小,分别对 J 关于 m_i 和 n_i 求偏导有

$$\frac{\partial J}{\partial m_i} = \frac{\partial(6(\boldsymbol{M}^{\mathrm{T}}\boldsymbol{q})^2)}{\partial m_i} = 2(6\boldsymbol{M}^{\mathrm{T}}\boldsymbol{q} + \lambda) q_i; \quad (20)$$

$$\frac{\partial J}{\partial n_i} = \frac{\partial(2(\boldsymbol{P}^{\mathrm{T}}\boldsymbol{q} + \lambda \boldsymbol{M}^{\mathrm{T}}\boldsymbol{q}))}{\partial n_i} = 2q_i。\quad (21)$$

令式(20)、(21)都等于 0,即 $2(6\boldsymbol{M}^{\mathrm{T}}\boldsymbol{q} + \lambda) q_i = 0$, $2q_i = 0$, 而 $q_i(\gamma) = \exp(-\gamma \boldsymbol{x}^{\mathrm{T}} \boldsymbol{A} \boldsymbol{x})$ 恒为正值,不等于

0,则只有当 $2(6\boldsymbol{M}^{\mathrm{T}}\boldsymbol{q}+\lambda)q_i=0$ 时,J 达到最小值,可得

$$\lambda=-6\boldsymbol{M}^{\mathrm{T}}\boldsymbol{q}。\tag{22}$$

至此,定理 2 得证。

对于两种错误分类数据点,根据其预测的 y,使用以上方法来计算其 γ 值,并将得到的 γ 在下一次迭代中用作核函数参数。参数优化过程将持续进行,直到达到总迭代次数或验证误差在预定义的迭代次数内没有变化,同时在代价函数达到最小时,可求得 λ 的最优解。

2.2　基于 SMC 策略的核函数参数算法

算法 1　基于 SMC 策略的核函数参数算法。

输入:待分类数据集,3 个迭代计数器 r_1、r_2、r_3,核函数参数初始 γ,最大迭代数 $R_{\max}$;

输出:数据集分类结果、参数 γ^*、支持向量数量集合 SV^* 以及分类准确率。

步骤 1　预处理:空值/异常值检验→归一化→按 7:2:1 比例划分训练集 D_{tr}、测试集 D_{te}、验证集 D_{val}。

步骤 2　闭环 SVM 训练:用参数 γ 训练 SVM;用 D_{tr} 计算训练误差 $TE=\dfrac{1}{|D_{\mathrm{tr}}|}(\sum e_i)$,用 D_{val} 计算验证误差 VE。

步骤 3　滑模反馈更新 γ:从 D_{tr} 中筛选出所有误分类样本,计算向量 $\widetilde{\boldsymbol{q}}=-\dfrac{1}{6}(\boldsymbol{M}\boldsymbol{M}^{\mathrm{T}})^{\dagger}(\boldsymbol{P}+\lambda\cdot\boldsymbol{M})$;$\gamma_{\mathrm{new}}=\dfrac{1}{k}\sum\gamma_i$;$\lambda=-6\boldsymbol{M}^{\mathrm{T}}\boldsymbol{q}$。

步骤 4　收敛检测:当训练过程达到总迭代次数或验证误差在预定义的迭代次数内没有变化。

步骤 5　使用 γ_{new} 对数据集进行分类,并输出分类准确率和参数 γ^*。

3　实验结果分析

本文使用 UCI(University of California Irvine)数据库中的 6 个数据集进行实验,主要情况如下:Iris 数据集[18]中的每个实物都是一株鸢尾花,共 150 个案例,4 个特征,这些特征用于对鸢尾花品种进行区分;Heart disease[19]数据集包含 303 个案例、13 个特征,用于判断患者是否患有心脏病;Haberman's survival[20]数据集是对接受过乳腺癌手术的患者的研究数据,包含 306 个案例、3 个特征,用于研究患者的术后生存情况;Letter recognition[21]数据集一共包含 20 000 个案例、16 个特征,本文只选取首字母为"A"和"N"的案例进行区分实验,共 1 536 个案例;Connectionist bench(sonar,mines vs. rocks)[22]数据集包含 208 个案例、60 个特征,用以识别海洋环境中的水雷和岩石;使用 Monk's problems[23]数据集中的 Monks-1 版本,包含 432 个案例、6 个特征。以上数据集用于测试 SMC-SVM 分类算法在非线性规则下的分类性能和鲁棒性,以评估所提算法是否优于其他 SVM 分类算法。

对所选的 6 个数据集进行缺失值以及异常值检验后,进行归一化处理,并将每个数据集按 7∶2∶1 比例生成训练集、测试集以及验证集。训练集用于更新核函数,测试集用于评估、比较算法的性能,验证集用于终止算法。

以 Iris 数据集为例,先通过 SMC-SVM 算法寻优得到适配该数据集的最优核函数参数,再将此参数应用于"Setosa 类鸢尾花与其他类鸢尾花"的二分类任务中(即区分 Setosa 品种与 Versicolor、Virginica 品种的混合类别)。图 1 展示了 Iris 数据集通过主成分分析法(PCA)降至二维特征空间后,叠加 SMC-SVM 算法的分类结果。

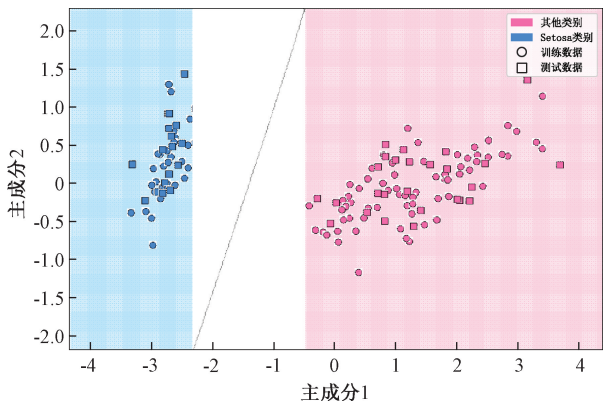


图 1　SMC-SVM 算法下鸢尾花数据集的分类结果

Figure 1　Classification result graph of the Iris dataset based on the SMC-SVM algorithm

将本文提出的 SMC-SVM 算法与传统网格搜索法 SVM 展开对比实验,具体实验结果见表 1。结果表明,在上述数据中,SMC-SVM 算法在训练集和测试集上的分类准确率普遍优于网格搜索法 SVM,且其分类所需的支持向量个数更少。

为进一步验证所提算法的性能,本文将 SMC-SVM 算法分别与 PSO-SVM 算法[9]、GWO-SVM 算法[10]展开对比实验,具体对比结果如表 2 所示。由表 2 可知,相较于 PSO-SVM 与 GWO-SVM 两种智能优,SMC-SVM 算法在大部分测试分类任务中展现出更高的准确率,充分体现了其在分类性能优化方面的优势。

表1 SMC-SVM算法与网格搜索法SVM的结果对比

Table 1 Comparison of the results between SMC-SVM algorithm and grid search method SVM

数据集	支持向量数量		训练准确率/%		测试准确率/%		支持向量的减少率/%
	SVM	SMC-SVM	SVM	SMC-SVM	SVM	SMC-SVM	
Iris[18]	48	21	100	100	100	100	56.25
Heart disease[19]	121	80	81.46	87.07	69.23	82.81	33.88
Haberman's survival[20]	102	93	80.11	80.11	73.75	73.75	8.82
Letter recognition[21]	124	84	100	100	100	100	32.20
sonar[22]	68	65	100	100	92.68	92.68	4.41
Monks-1[23]	226	134	99.40	99.49	80.84	97.01	40.71

表2 SMC-SVM算法与PSO-SVM、GWO-SVM算法的结果对比

Table 2 Comparison of results between SMC-SVM and PSO-SVM, GWO-SVM algorithm

数据集	初始γ	测试准确率/%			最优γ		
		PSO-SVM	GWO-SVM	SMC-SVM	PSO-SVM	GWO-SVM	SMC-SVM
Iris[18]	0.01	100	100	100	0.065 5	0.141 8	0.198 2
Heart disease[19]	0.01	85.8	86.8	87.07	0.000 1	0.009 8	0.011 0
Haberman's survival[20]	0.10	76.45	77.1	80.11	0.170 3	0.086 4	0.370 9
Letter recognition[21]	0.01	99.8	99.86	100	0.226 5	0.063 9	0.560 9
sonar[22]	0.01	87.3	83.51	92.68	0.028 8	0.176 1	0.100 0

4 结论

本文提出的基于SMC策略的SVM核函数参数优化算法有效解决了线性不可分数据集的分类问题。通过引入SMC策略,算法结合闭环控制思想,用来动态调整SVM的核函数参数,提高了模型对数据分布的适应能力。通过误差建立了SVM算法与SMC之间的关联性,并证明了满足滑模面方程、趋近律方程以及保证代价函数达到最小的参数值的存在性,以确保算法的稳定性和有效性。实验结果表明,与传统SVM相比,所提出的算法在支持向量数量和分类精度上均有改善。同时,与智能优化算法相比,该算法在大部分数据集上具有更高的分类准确率,验证了所提出的基于SMC策略的SVM核函数参数优化算法具有较高分类准确率的同时,还增强了模型的鲁棒性。该算法为未来在机器学习领域中处理非线性和不确定性问题时应用SMC策略提供了新的思路。

参考文献:

[1] Bishop C M. Pattern recognition and machine learning [M]. New York: Springer, 2006.

[2] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.

[3] Li Xiaotong, Cheng Fan, Tian Jing, et al. Application of SMOTE technique and optimization algorithm support vector machine in predicting adverse outcome of chronic heart failure[J]. Chinese Journal of Health Statistics, 2024, 41(6): 802-806. [李晓桐, 程璠, 田晶, 等. 结合SMOTE技术与优化算法的支持向量机在慢性心衰不良结局预测中的应用[J]. 中国卫生统计, 2024, 41(6): 802-806.]

[4] Zheng Xiang, Zhang Honghu, Wei Yanyan. Research on infrared image fault diagnosis method for substation equipment based on improved SVM[J]. Electrical Measurement & Instrumentation, 2024, 61(12): 49-55. [郑祥, 张鸿鹄, 魏岩岩. 基于改进SVM的变电站设备红外图像故障诊断方法研究[J]. 电测与仪表, 2024, 61(12): 49-55.]

[5] Zhao Xuyang, Zhang Yanbin, Wang Zhongyong, et al. Design of acoustic magnetic label detection system based on SVM and FPGA implementation [J]. Journal of Zhengzhou University (Engineering Science), 2021, 42(3): 13-18. [赵旭阳, 张延彬, 王忠勇, 等. 基于SVM的声磁标签检测系统设计及其FPGA实现[J]. 郑州大学学报(工学版), 2021, 42(3): 13-18.]

[6] Adeyanju I A, Omidiora E O, Oyedokun O F. Performance evaluation of different support vector machine kernels for face emotion recognition[C]//Proceedings of the 2015 SAI Intelligent Systems Conference (IntelliSys). Piscataway: IEEE, 2015: 804-806.

[7] Zhou Peijian, Zeng Weitao, Zhang Wenwu, et al. Multi-cavitation states identification of a sewage pump using CEEMDAN and BOA-SVM[J]. Journal of Water Process Engineering, 2024, 61: 105299.

[8] Zhang Hong, Zhang Yifan. An improved sparrow search algorithm for optimizing support vector machines [J]. IEEE Access, 2023, 11: 8199-8206.

[9] Abbaszadeh M, Soltani-Mohammadi S, Ahmed A N. Optimization of support vector machine parameters in modeling of Iju deposit mineralization and alteration zones using particle swarm optimization algorithm and grid search method [J]. Computers & Geosciences, 2022, 165: 105140.

[10] Li Changpeng, Peng Tianhao, Zhu Yanmin. A cutting pattern recognition method for shearers based on ICEEMDAN and improved grey wolf optimizer algorithm-optimized SVM [J]. Applied Sciences, 2021, 11(19): 9081.

[11] Zhang Guangchen, Xia Yuanqing, Li Xufei, et al. Multievent-triggered sliding-mode control for a class of complex dynamic network[J]. IEEE Transactions on Control of Network Systems, 2022, 9(2): 835-844.

[12] Liu Ling, Si Jiewen, Lin Qilian. A fast terminal sliding mode control of permanent magnet synchronous motor with variable parameters predicted by SVM[J]. Journal of Xi'an Jiaotong University, 2021, 55(6): 53-60. [刘凌, 司杰文, 林起联. 支持向量机预测可变参数的永磁同步电机快速终端滑模控制[J]. 西安交通大学学报, 2021, 55(6): 53-60.]

[13] Yalsavar M, Karimaghaee P, Sheikh-Akbari A, et al. Sliding mode control based support vector machine RBF kernel parameter optimization[C]//Proceedings of the 2019 IEEE International Conference on Imaging Systems and Techniques (IST). Piscataway: IEEE, 2019: 1-6.

[14] 宋慧玲, 张仲广, 夏冰. 基于支持向量机的分类问题研究[M]. 哈尔滨: 黑龙江大学出版社, 2022.

[15] 蒋刚. 核函数理论与信号处理[M]. 北京: 科学出版社, 2013.

[16] Cristianini N, Shawe-Taylor J. An introduction to support vector machines and other kernel-based learning methods [M]. Cambridge, UK: Cambridge University Press, 2000.

[17] Rasmussen C E, Williams C K I. Gaussian processes for machine learning[M]. Cambridge, MA, USA: The MIT Press, 2005.

[18] Fisher R A. Iris[DS/OL]. (1988-06-30)[2025-12-11]. https://doi. org/ 10. 24432/ C56C76.

[19] Janosi A, Steinbrunn W, Pfisterer M, et al. Heart disease[DS/OL]. (1988-06-30)[2025-12-11]. https://doi. org/ 10. 24432/ C52P4X.

[20] Haberman S. Haberman's survival[DS/OL]. (1999-03-03)[2025-12-11]. https://doi. org/ 10. 24432/ C5XK51.

[21] Slate D. Letter recognition[DS/OL]. (1990-12-31)[2025-12-11]. https://doi. org/ 10. 24432/ C5ZP40.

[22] Sejnowski T, Gorman R. Connectionist bench (sonar, mines vs. rocks)[DS/OL]. [2025-12-11]. https://doi. org/ 10. 24432/ C5T01Q.

[23] Wnek J. MONK's problems[DS/OL]. (1992-09-30)[2025-12-11]. https://doi. org/ 10. 24432/ C5R30R.

Classification Algorithm for Linearly Inseparable Datasets Based on SMC Strategy

ZHANG Guangchen[1], LI Zhanfei[1], HE Shuping[2], XIA Yuanqing[3]

(1. School of Mathematics and Information Science, North Minzu University, Yinchuan 750021, China; 2. School of Electrical Engineering and Automation, Anhui University, Hefei 230601, China; 3. School of Automation, Beijing Institute of Technology, Beijing 100081, China)

Abstract: For the classification problem of linearly inseparable datasets, a support vector machine(SVM) kernel function parameter optimization algorithm was proposed based on the sliding mode control (SMC) strategy by applying the SMC idea to the SVM kernel function parameter optimization process. By designing the error equation and sliding surface, the association between SVM classification objective function and SMC was established, and then the iteration update rules of kernel function parameters and cost function was derived. The algorithm improved classification performance while reducing the number of support vectors by dynamically adjusting the kernel parameters of SVM. In the experiment part, six UCI datasets, such as Iris and Heart disease, were used to verify the validity of the algorithm. The results showed that compared with traditional SVM, the proposed algorithm reduced the number of support vectors by 56.25% on the Iris dataset, and the test accuracy remained at 100%. Test accuracy increased by 13.58 percentage points on the Heart disease dataset. Furthermore, the proposed algorithm, compared with existing optimization algorithms, showed a higher classification accuracy on some datasets.

Keywords: support vector machines; sliding mode control strategy; classification algorithm; nuclear function; parameter