

面向新异检测的启发式约减支持向量数据描述

曲建岭, 孙文柱, 邸亚洲, 高峰, 周玉平

(海军航空工程学院 青岛校区, 山东 青岛 266041)

摘要: 针对支持向量数据描述(SVDD)单类分类方法运算复杂度高的缺点, 提出一种启发式约减支持向量数据描述(HR-SVDD)方法. 以启发的方式从原有训练集中筛选出部分样本构成约减训练集, 对约减训练集进行二次规划解算, 得到支持向量和决策边界. 通过不同宽度系数高斯核SVDD特征的讨论, 证明了HR-SVDD的有效性. 人工数据集和真实数据集上的实验结果表明, HR-SVDD分类精度与传统支持向量数据描述相当, 但具有更快的运算速度和更小的内存占用.

关键词: 支持向量数据描述; 启发式约减支持向量数据描述; 新异检测

中图分类号: TP181

文献标志码: A

Heuristic reduction support vector data description for novelty detection

QU Jian-ling, SUN Wen-zhu, DI Ya-zhou, GAO Feng, ZHOU Yu-ping

(Qingdao Branch, Naval Aeronautical Engineering Institute, Qingdao 266041, China. Correspondent: SUN Wen-zhu, E-mail: sunwenzhulm@gmail.com)

Abstract: A method of heuristic reduction support vector data description(HR-SVDD) is proposed for speeding up the support vector data description(SVDD) one-class classification method. The HR-SVDD first builds a reduced training set by selecting a portion of samples from training set in heuristic, and then completes the quadratic programming using the reduced training set rather than the original training set. The efficiency of proposed method is demonstrated by discussing characteristic of Gaussian kernel SVDD with different width parameters. For demonstration, experiments on artificial and real-world datasets are conducted, and the results show that the classification accuracy of HR-SVDD is nearly identical to that of conventional SVDD, but with faster running speed and less memory usage.

Key words: support vector data description; heuristic reduction support vector data description; novelty detection

0 引言

在设备状态监控、疾病监测和网络入侵检测等领域中, 经常需要对采集的数据进行分析和判读. 这些数据的分析和判读过程都可以归结为判断数据是否正常的问题. 在这些问题中, 正常状态数据数量较多且具有较完善的描述, 异常状态数据稀缺且没有准确的描述和定义. 使用传统的两类分类方法解决上述问题时, 常会因为两类样本数量的不平衡导致决策面偏移. 为了解决此类问题, Moya等^[1]提出了单类分类方法. 该方法在训练过程中描述出目标样本的分布区域, 在测试过程中接受目标样本描述内的测试样本, 拒绝目标样本描述外的测试样本. 单类分类方法可以仅使用目标样本完成分类器训练, 因此适用于异类样本稀缺的分类问题. 使用单类分类方法对采样信号进行判别的过程称为新异检测.

支持向量数据描述(SVDD)是在支持向量机(SVM)的基础上发展而来的一种较为先进的新异检测方法. Tax等^[2-3]提出SVDD理论后, SVDD迅速成为研究热点, 并在故障检测与故障诊断^[4-6]、目标识别^[7]、噪声抑制^[8]和样本聚类^[9-10]等领域得到了广泛应用. 然而, 与SVM一样, SVDD具有较高的运算复杂度 $O(N^3)$, 其中 N 为训练样本个数, 这便导致了其运行时间随训练样本数量的增大而急剧增大^[5,11]. 目前, 许多学者致力于减少SVDD运算复杂度的研究, 如: Chu等^[12]使用核心集近似最小包络球的方法将SVDD时间复杂度降低为 $O(N)$; Choi^[13]在最小二乘支持向量机的基础上提出最小二乘单类支持向量机(LS-OCSVM), 将SVDD运算复杂度降低为 $O(N^2)$. 然而, 这两种方法虽然可以避免传统SVDD需解算二次规划的问题, 但无法得到全局最优解. Peng等^[14]

收稿日期: 2013-07-30; 修回日期: 2014-02-25.

作者简介: 曲建岭(1968—), 男, 教授, 博士生导师, 从事飞参数据应用等研究; 孙文柱(1984—), 男, 博士生, 从事单类分类的研究.

提出了一种快速 SVDD (F-SVDD) 方法, 可以将 SVDD 时间复杂度降低为常数, 但该方法不能有效处理多峰分布样本. 另外, Kim 等^[15]基于 K -means 方法, 利用“分解与结合”的策略加快了 SVDD 运行速度, 但仍然无法用于解决样本数量特别大的新异检测问题.

本文通过深入研究 SVDD 原理, 提出一种改进的快速新异检测方法, 称为启发式约减支持向量数据描述 (HR-SVDD). 该方法可以在保证分类精度的前提下, 减少训练过程中参与二次规划求解的样本数量, 极大程度地提升传统 SVDD 的运行速度.

1 SVDD

给定训练样本集 $X_{\text{tr}} = \{\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N\}$, SVDD 的目标是确定一个能够包含所有训练样本的体积最小化的超球. 设超球中心为 \mathbf{c} , 半径为 r , 那么最小化超球的目标可以通过最小化 r^2 的途径获得, 即

$$\min r^2; \quad (1)$$

$$\text{s.t. } \|\mathbf{x}_i - \mathbf{c}\|^2 \leq r^2, \forall i. \quad (2)$$

其中 r^2 可以看作结构风险. 以式 (1) 和 (2) 构成的约束不允许任何训练样本落于超球外, 为了增加 SVDD 对潜在奇异样本的容忍程度, 引入平衡参数 C 和松弛变量 ξ_i . 此时, 约束问题表示为

$$\min r^2 + C \sum_i \xi_i; \quad (3)$$

$$\text{s.t. } \|\mathbf{x}_i - \mathbf{c}\|^2 \leq r^2 + \xi_i, \xi_i \geq 0, \forall i. \quad (4)$$

其中: C 起到平衡学习能力和复杂程度的作用, ξ_i 为第 i 个样本与超球间的距离. 这是一个约束优化问题, 可通过引入 Lagrange 乘子解决. 拉格朗日函数定义为

$$L(r, \mathbf{c}, \xi, \gamma) = r^2 + C \sum_i \xi_i - \sum_i \alpha_i (r^2 + \xi_i - \|\mathbf{x}_i - \mathbf{c}\|^2) - \sum_i \gamma_i \xi_i, \quad (5)$$

其中 γ_i 和 α_i 为拉格朗日乘子, $\gamma_i \geq 0$, $\alpha_i \geq 0$, 且 γ_i 和 α_i 分别与 \mathbf{x}_i 一一对应. 此后, 由 r 、 \mathbf{c} 和 ξ 对 L 求极小值. 令式 (5) 分别对 r 、 \mathbf{c} 和 ξ 求偏导并令偏导为零, 得到约束条件

$$\mathbf{c} = \sum_i \alpha_i \mathbf{x}_i; \quad (6)$$

$$0 \leq \alpha_i \leq C, \forall i; \quad (7)$$

$$\sum_i \alpha_i = 1. \quad (8)$$

将式 (6)~(8) 代入 (5), Lagrange 函数可以重新表示为

$$L(\alpha) = \sum_i \alpha_i (\mathbf{x}_i \mathbf{x}_i) - \sum_{i,j} \alpha_i \alpha_j (\mathbf{x}_i \mathbf{x}_j). \quad (9)$$

在约束 (6)~(8) 条件下求 L 极小值是典型的二次规划问题, 目前存在成熟的计算机求解方法. 如果与样本 \mathbf{x}_i 相对应的 α_i 大于零, 则该样本为支持向量.

支持向量确定后, 由式 (6) 可求得超球中心 \mathbf{c} , 超球半径 r 为任意支持向量到超球中心的距离.

使用上述方法得到的决策边界为刚性超球, 只适用于描述球形分布样本. 实际上, 在很多分类问题中样本分布不为球形, 因此, 需要引入“核”概念使得 SVDD 的决策边界更为灵活. 式 (9) 中, 样本的运算全部为内积运算, 使用函数替代内积运算的方法称为核方法. 核方法的种类较多, 其中高斯核由于具有良好的学习能力和泛化能力^[3], 得到了广泛应用. 因此, 本文仅对高斯核进行研究. 高斯核定义为

$$K(\mathbf{x}_i, \mathbf{x}_j) = \varphi(\mathbf{x}_i) \varphi(\mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \sigma^2), \forall i, j, \quad (10)$$

其中 σ 为宽度系数. 对于高斯函数, 恒有

$$K(\mathbf{x}_i, \mathbf{x}_i) = 1, \forall i. \quad (11)$$

因此, 式 (9) 中拉格朗日函数可以重新表述为

$$L(\alpha) = - \sum_{i,j} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j). \quad (12)$$

用来评估测试样本 \mathbf{x} 是否被决策边界接受的函数称为决策函数, 表述为

$$f_{\text{SVDD}} = \sum_{i=1}^{N_{\text{SV}}} \alpha_i K(\mathbf{x}, \mathbf{x}_i) - \sum_{i,j=1}^{N_{\text{SV}}} \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j), \quad (13)$$

其中 N_{SV} 为支持向量的个数. 如果 $f_{\text{SVDD}} \geq 0$, 则 \mathbf{x} 被接受, 否则 \mathbf{x} 被拒绝.

2 启发式约减支持向量数据描述

启发式约减支持向量数据描述的步骤如下.

Step 1: 归一化训练样本集 X_{tr} .

Step 2: 使用 K -means 方法对 X_{tr} 进行聚类, 可得到 X_{tr} 的聚类中心 \mathbf{c}_j ($j = 1, 2, \dots, k$).

Step 3: 为每个样本 \mathbf{x}_i 定义距离变量

$$d_i = \min(\|\mathbf{x}_i - \mathbf{c}_j\|), \forall j, i = 1, 2, \dots, N, \quad (14)$$

其中 d_i 为第 i 个样本与最近的聚类中心 \mathbf{c}_j 的距离.

Step 4: 根据 d_i 按从大到小的顺序对 X_{tr} 中的所有样本进行排序, 得到排序后的训练集 $X_{\text{tr}}^{\text{order}}$.

Step 5: 设置约减训练集 \tilde{X}_{tr} 为空.

Step 6: 设置 $X_{\text{tr}}^{\text{order}}$ 中第 1 个样本为突出样本 \mathbf{x}_H , 并将该样本加入 \tilde{X}_{tr} .

Step 7: 在 $X_{\text{tr}}^{\text{order}}$ 中删除满足条件 $\|\mathbf{x}_H, \mathbf{x}_i\| < \delta$ ($\forall i$) 的样本 (包括 \mathbf{x}_H 本身). 其中: $\|\cdot\|$ 为欧氏距离, δ 为距离阈值.

Step 8: 重复 Step 6 和 Step 7, 直至 $X_{\text{tr}}^{\text{order}}$ 为空, 同时 \tilde{X}_{tr} 构建完成. 图 1 为训练样本集及其约减训练集示例, 参数设置为 $k = 5$, $\delta = 0.1$. X_{tr} 样本数为 200, 经启发式约减后, \tilde{X}_{tr} 样本数为 39.

Step 9: 将 \tilde{X}_{tr} 代入式 (10)~(13), 计算决策函数.

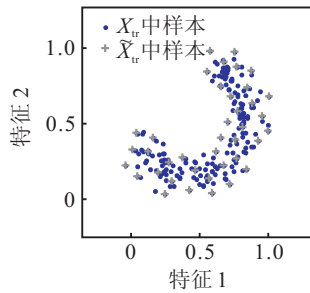


图 1 二维 Banana 训练样本集及其约减训练集示例

在 HR-SVDD 方法中, 如何寻找 Step 2 中最优的 K -means 聚类簇个数是一个重要问题. 目前, 许多成熟方法可以较好地解决该问题^[14-15]. 其中, Galluccio 等^[16]提出的基于最小生成树的方法运行速度较快且性能较好, 为此本文使用该方法确定 K -means 聚类簇个数.

3 HR-SVDD 有效性分析

由图 1 可见, \tilde{X}_{tr} 由 X_{tr} 中的部分样本组成, 且 \tilde{X}_{tr} 具有如下特征: 1) 任意两个样本间距大于等于 δ ; 2) \tilde{X}_{tr} 的样本分布形态与 X_{tr} 基本相同. \tilde{X}_{tr} 的特征决定了 HR-SVDD 能够获得与传统 SVDD(C-SVDD) 相近的决策边界. 为了方便讨论, 将高斯核 SVDD 按照不同的 σ 分为 3 类: 较小的 σ 、中等程度的 σ 和较大的 σ . 这 3 类 SVDD 决策边界示例如图 2 所示. 图 2 中: + 为训练样本, o 为支持向量, 虚线为决策边界.

对于较小的 σ , 有

$$-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \sigma^2 \ll 0. \quad (15)$$

因此, 对于不同的样本 \mathbf{x}_i 和 \mathbf{x}_j 而言, 核函数 $K(\cdot)$ 差别很小, 即

$$\sum_{p=1}^N \exp(-\|\mathbf{x}_i - \mathbf{x}_p\|^2 / \sigma^2) \approx \sum_{p=1}^N \exp(-\|\mathbf{x}_j - \mathbf{x}_p\|^2 / \sigma^2), \quad \forall i, j. \quad (16)$$

在约束条件 $0 \leq \alpha_i \leq C(\forall i)$ 和 $\sum_i \alpha_i = 1$ 下, 式(12) 取得最小值的条件为所有样本均成为支持向量且 $\alpha_i = 1/N(\forall i)$ (图 2(a)). 这种类型的 SVDD 等价于高斯窗 Parzen 单类分类器. 因为 \tilde{X}_{tr} 的样本分布形态与 X_{tr}

基本相同, 所以在这种情况下由 \tilde{X}_{tr} 和 X_{tr} 得到的决策边界几乎相同.

对于中等的 σ , 如果 \mathbf{x}_i 位于样本分布的中心区, 则 $\sum_{j=1}^N K(\mathbf{x}_i, \mathbf{x}_j)$ 项较大; 如果 \mathbf{x}_i 位于样本分布的边缘区, 则 $\sum_{j=1}^N K(\mathbf{x}_i, \mathbf{x}_j)$ 项较小. 在拉格朗日函数最小化的过程中, 如果 $\sum_{j=1}^N K(\mathbf{x}_i, \mathbf{x}_j)$ 项较大, 则 α_i 趋于 0.

$\sum_{j=1}^N K(\mathbf{x}_i, \mathbf{x}_j)$ 项最小的那部分样本所对应的 α_i 大于 0 且成为支持向量, 因此这种情况下支持向量总是位于样本分布的边缘 (图 2(b)). 由式 (13) 可见, 此类 SVDD 等同于由所有支持向量组成的加权高斯窗 Parzen 单类分类器. 在 \tilde{X}_{tr} 构建过程中, 算法通过赋予位于分布边缘样本较高的优先级使得 X_{tr} 中的潜在支持向量全部包含在 \tilde{X}_{tr} 中. 这部分潜在支持向量在二次规划解算中仍可以成为支持向量, 因此在这种情况下, 由 \tilde{X}_{tr} 和 X_{tr} 得到的决策边界几乎相同.

对于较大的 σ , $K(\mathbf{x}_i, \mathbf{x}_j)$ 可以展开为泰勒级数

$$K(\mathbf{x}_i, \mathbf{x}_j) = \exp(-\|\mathbf{x}_i - \mathbf{x}_j\|^2 / \sigma^2) = 1 - \frac{\mathbf{x}_i - \mathbf{x}_j}{\sigma^2} + \frac{2(\mathbf{x}_i \mathbf{x}_j)}{\sigma^2} + o\left(\frac{1}{\sigma^2}\right), \quad (17)$$

其中 $o(1/\sigma^2)$ 为 $1/\sigma^2$ 的高阶无穷小. 将式 (17) 代入 (12), 得到

$$L = -1 + \frac{2}{\sigma^2} \left(\sum_i \alpha_i (\mathbf{x}_i \mathbf{x}_i) - \sum_i \alpha_i \alpha_j (\mathbf{x}_i \mathbf{x}_j) \right) + o\left(\frac{1}{\sigma^2}\right). \quad (18)$$

注意到, 式 (18) 与 (9) 形式相似, 两者区别在于式 (18) 比 (9) 多了偏移量 -1 、比例项 $2/\sigma^2$ 和一个无穷小项 $o(1/\sigma^2)$. 这 3 项在二次规划中对解均不产生影响, 因此此类 SVDD 等价于刚性超球 SVDD (如图 2(c)). 这种情况下, 只有分布在最外围的少数几个样本成为支持向量. 与中等的 σ 情况类似, 由 \tilde{X}_{tr} 和 X_{tr} 解算出的支持向量几乎相同, 使得两者得到的决策边界仍然几乎相同.

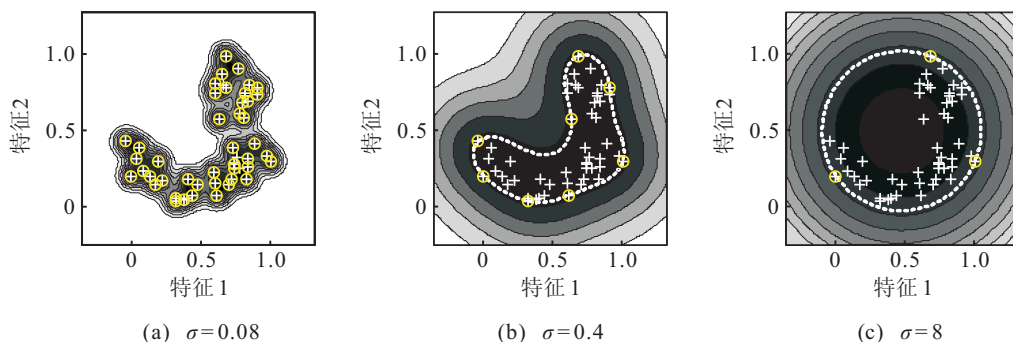


图 2 不同 σ 的高斯核 SVDD 决策边界示例

4 实验分析

本节使用人工数据集和 UCI 真实数据集分别对 HR-SVDD 性能进行测试, 并将分类结果与 C-SVDD 结果进行比较. 实验前, 给出用于度量单类分类方法分类精度的指标: 1) 漏警率(又称假阳性率 FPR), 表示异类样本被识别为目标样本的概率; 2) 检出率(又称真阳性率 TPR), 表示目标样本被正确识别的概率; 3) 接受者操作特征曲线(ROC curve), 以 FPR 为横坐标, TPR 为纵坐标的曲线称为 ROC 曲线, ROC 曲线是重要的描述分类器分类性能的指标, ROC 曲线越接近左上角表明新异检测方法的分类精度越高; 4) ROC 曲线下面积(AUC), 也是重要的描述分类器分类精度的指标, 它表示 ROC 曲线下方面积的积分, AUC 取值范围为 $0 \sim 1$, 其值越接近 1 表明分类器的分类精度越高. 用于运行实验的计算机软硬件配置为: 奔腾 Dual E2140 1.6 GHz CPU、2 GB 内存和 Matlab V7.6.0.

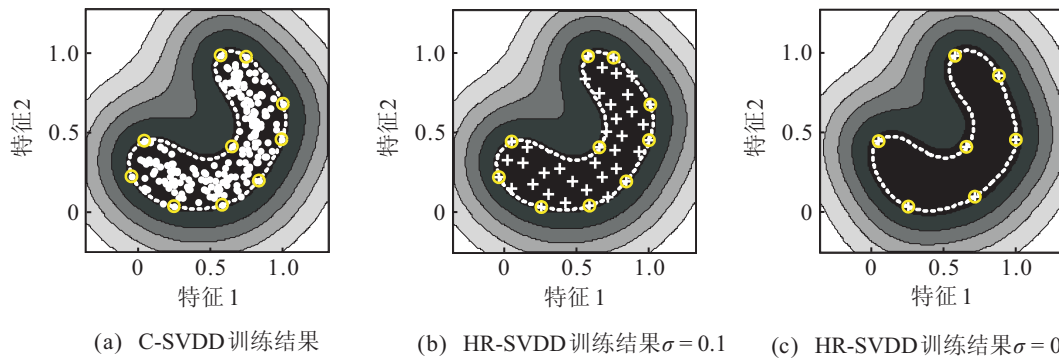


图 3 Banana 数据集训练样本集及其约减训练集示例

C-SVDD 和不同 δ 值 HR-SVDD 的 ROC 曲线如图 4 所示. 图 4 中: $\sigma = 0.4, C = 1$. 当 $\delta = 0.1$ 时, HR-SVDD 的 ROC 曲线几乎与 C-SVDD 的 ROC 曲线重合, 这表明当 $\delta = 0.1$ 时, HR-SVDD 的分类精度与 C-SVDD 几乎相同. 当 $\delta = 0.2, 0.3$ 和 0.4 时, HR-SVDD 的分类精度逐次降低.

运行在 Banana 数据集、Lithuanian 数据集和 Difficult 数据集上的 HR-SVDD 和 C-SVDD 的综合对比实验结果如表 1 所示. 实验中, 参数 σ 和 C 均设置为 0.4 和 1 . 当 $\delta = 0.05$ 时, C-SVDD 和 HR-SVDD 的 AUC 值基本相同, 但是 HR-SVDD 的运行时间远小于 C-SVDD. 如, Difficult 数据集上, 当 $\delta = 0.05$ 时, HR-SVDD 的运行时间为 36.81 s, 对应的 C-SVDD 运行

4.1 人工数据集对比实验

实验中的人工数据集由 Prtools 工具箱生成, 包括 Banana、Difficult 和 Lithuanian 数据集. 每个数据集均包含 200 个目标样本和 200 个异类样本, 实验采用留一交叉验证方法确保结果的无偏性.

C-SVDD 和 HR-SVDD 的训练结果如图 3 所示, 图 3(a) 为 C-SVDD 训练结果, C 和 σ 分别设置为 1 和 0.4 . 图 3(b) 为 HR-SVDD 训练结果, δ, C 和 σ 分别设置为 $0.1, 1$ 和 0.4 . 对比图 3(a) 和图 3(b) 可知: 当 $\delta = 0.1$ 时, \tilde{X}_{tr} 中包含了所有 X_{tr} 中的潜在支持向量, 并且由 \tilde{X}_{tr} 解算出的支持向量与 X_{tr} 解算出的支持向量完全相同, 所形成的决策边界也相同. 当 δ 增大时, \tilde{X}_{tr} 中包含的样本数量减少. 参数设置为 $\delta = 0.3, C = 1$ 和 $\sigma = 0.4$ 的 HR-SVDD 训练结果如图 3(c) 所示, 此时 \tilde{X}_{tr} 中仅包含 7 个样本, 但使用这 7 个样本形成的决策边界仍能较好地描述训练样本分布.

间为 1251.84 s. 当 $\delta = 0.1$ 时, HR-SVDD 的 AUC 值仍然接近于 C-SVDD 的 AUC 值, 运行时间比 $\delta = 0.05$ 时的 HR-SVDD 更短. 当 $\delta = 0.2$ 和 $\delta = 0.3$ 时, 部分 HR-SVDD 的 AUC 值下降较快(如 Difficult 数据集). HR-SVDD 的运行时间随 δ 的增加而缩短.

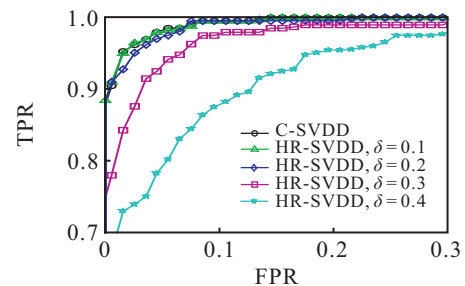


图 4 C-SVDD 和不同 δ 值 HR-SVDD 的 ROC 曲线

表 1 人工数据集上 C-SVDD 和 HR-SVDD 综合实验结果对比

方法	距离阈值	Banana 数据集		Lithuanian 数据集		Difficult 数据集	
		AUC	运行时间/s	AUC	运行时间/s	AUC	运行时间/s
C-SVDD	N/A	0.9949	1316.65	0.9825	1228.07	0.9289	1251.84
HR-SVDD	$\delta = 0.05$	0.9949	74.61	0.9824	87.70	0.9289	36.81
	$\delta = 0.1$	0.9949	26.06	0.9797	26.14	0.9244	15.69
	$\delta = 0.2$	0.9941	12.77	0.9792	12.96	0.9074	10.60
	$\delta = 0.3$	0.9872	9.28	0.9727	10.11	0.8342	10.28

表 2 真实数据集上 C-SVDD 和 HR-SVDD 综合实验结果对比

方法	距离阈值	Letter 数据集		Pendigits 数据集		Optdigits 数据集	
		AUC	运行时间/s	AUC	运行时间/s	AUC	运行时间/s
C-SVDD	N/A	0.767 8	4 022.18	0.951 4	3 838.62	0.857 1	2 140.53
	$\delta = 0.05$	0.767 8	20.07	0.951 6	13.11	0.856 8	12.05
HR-SVDD	$\delta = 0.1$	0.762 3	11.81	0.951 4	9.24	0.854 3	9.41
	$\delta = 0.2$	0.733 6	6.19	0.945 5	5.01	0.844 1	5.16
	$\delta = 0.3$	0.687 5	5.60	0.944 6	4.36	0.821 6	4.33

4.2 真实数据集对比实验

本节采用真实的 UCI 机器学习数据集对 HR-SVDD 的分类能力进行测试, 包括 Letter 数据集、Pendigits 数据集和 Optdigits 数据集. 由于真实数据集样本数较多, 采用 10 折交叉验证法提高实验结果的可信度. 实验中, 参数 σ 和 C 均设置为 0.4 和 1. HR-SVDD 和 C-SVDD 的对比结果如表 2 所示. 实验中, HR-SVDD 的 AUC 值有时甚至高于 C-SVDD. 比如在 Pendigits 数据集上, 当 $\delta = 0.05$ 时, C-SVDD 的 AUC 为 0.951 4, HR-SVDD 的 AUC 为 0.951 6. 理论上, HR-SVDD 并不能增加分类精度, 此处的 AUC 增加是由真实数据集中的奇异样本和 AUC 测量误差引起的. 其余实验结果均与人工数据集结果相似, 因此, HR-SVDD 同样适合于真实数据集单类分类.

5 结 论

本文以加速 SVDD 训练为目标, 提出了一种改进的 SVDD 方法 (HR-SVDD). 该方法在保证训练精度的前提下, 通过从原有训练集中提取约减训练集参与二次规划运算的方法, 将 SVDD 的算法复杂度从 $O(N^3)$ 降为 $O(\tilde{N}^3)$, 其中 \tilde{N} 为约减训练集中样本个数. 此外, 通过对较大 σ 、中等 σ 和较小 σ 三类 SVDD 特征的分析, 证明了 HR-SVDD 能够获得与 C-SVDD 相近的决策边界. 为验证 HR-SVDD 的性能, 本文在人工数据集和真实数据集上设计了一系列仿真实验, 并将实验结果与 C-SVDD 对比. 实验结果表明: HR-SVDD 与 C-SVDD 的分类能力相当, 但 HR-SVDD 的运行速度明显快于 C-SVDD. 因此, HR-SVDD 是一种基于 SVDD 改进的适用于大样本单类分类的有效方法.

参考文献(References)

- [1] Moya M, Hush D. Network constraints and multi-objective optimization for one-class classification[J]. Neural Networks, 1996, 9(3): 463-474.
- [2] Tax D M J, Duin R P. Support vector domain description[J]. Pattern Recognition Letters, 1999, 20(11): 1191-1199.
- [3] Tax D M J. One class classification: Concept-learning in the absence of counter-examples[D]. Netherlands: University of Delft, 2001.
- [4] Wang S, Yu J, Lapira E, et al. A modified support vector data description based novelty detection approach for machinery components[J]. Applied Soft Computing, 2012, 13(2): 1193-1205.
- [5] Liu Y H, Liu Y C, Chen Y J. Fast support vector data descriptions for novelty detection[J]. IEEE Trans on Neural Networks, 2010, 21(8): 1296-1313.
- [6] 唐明珠, 王岳斌, 阳春华. 一种改进的支持向量数据描述故障诊断方法[J]. 控制与决策, 2011, 26(7): 967-982. (Tang M Z, Wang Y B, Yang C H. Modified support vector data description for fault diagnosis[J]. Control and decision, 2011, 26(7): 967-982.)
- [7] 顾鑫, 王士同, 许敏. 领域自适应的最小包含球设计方法[J]. 控制与决策, 2013, 28(2): 177-183. (Gu X, Wang S T, Xu M. Minimum enclosing ball for domain adaptation[J]. Control and Decision, 2013, 28(2): 177-183.)
- [8] Park J, Kang D, Kim J, et al. SVDD-based pattern denoising[J]. Neural Computer, 2007, 19(7): 1919-1938.
- [9] Camastra F, Verri A. A novel kernel method for clustering[J]. IEEE Trans on Pattern Analysis and Machine Intelligence, 2005, 27(5): 801-805.
- [10] Ben-Hur A, Horn D, Siegelmann H T, et al. Support vector clustering[J]. J of Machine Learning Research, 2002, 2: 125-137.
- [11] Ortiz-García E G, Salcedo-Sanz S, Pérez-Bellido M, et al. Improving the training time of support vector regression algorithms through novel hyper-parameters search space reductions[J]. Neurocomputing, 2009, 72(16): 3683-3691.
- [12] Chu C S, Tsang I W, Kwok J T. Scaling up support vector data description by using core-sets[C]. Proc of IEEE Int Joint Conf on Neural Networks. Budapest: IEEE, 2004: 425-430.
- [13] Choi Y S. Least squares one-class support vector machine[J]. Pattern Recognition Letters, 2009, 30(13): 1236-1240.
- [14] Peng X, Xu D. Efficient support vector data descriptions for novelty detection[J]. Neural Computing and Applications, 2012, 21(8): 2023-2032.
- [15] Kim P J, Chang H J, Song D S, et al. Fast support vector data description using k-means clustering[C]. Advances in Neural Networks. Berlin: Springer, 2007: 506-514.
- [16] Galluccio L, Michel O, Comon P, et al. Graph based k-means clustering[J]. Signal Processing, 2012, 92(9): 1970-1984.