

One-Class 分类器研究

潘志松¹, 陈 斌^{2,3}, 繆志敏¹, 倪桂强¹

(1. 解放军理工大学指挥自动化学院, 江苏南京 210007; 2. 扬州大学信息技术学院, 江苏扬州 225009;
3. 南京航空航天大学信息科学与技术学院, 江苏南京 210016)

摘 要: 传统的多类分类问题需要多类样本训练分类器, 然而由于样本获取(复杂性或代价)的原因很多情况下只能获取一类样本, 故只能利用这一类样本进行学习, 形成数据描述从而实现分类, 故称之为 One-Class 分类。鉴于目前单类分类研究存在领域相关方法和通用方法百花齐放的格局, 本文首先给出了当前 One-Class 分类器的研究综述, 然后重点针对基于核方法的单类分类器进行分析, 将该类方法分为对偶方式和核距离方式两类, 并分析各自的特点。最后论文介绍了目前的单类分类器的应用领域, 指出其在故障分析、异常检测、疾病诊断和敌我识别等现实问题中的重要作用。

关键词: 单类; 核方法; 分类器; 异常检测

中图分类号: TP181 **文献标识码:** A **文章编号:** 0372-2112 (2009) 11-2496-08

Overview of Study on One-Class Classifiers

PAN Zhi-song¹, CHEN Bin^{2,3}, MIAO Zhi-min¹, NI Gui-qiang¹

(1. PLA University of Science and Technology, Nanjing, Jiangsu 210007, China;

2. Department of Computers, Yangzhou University, Yangzhou, Jiangsu 225009, China;

3. Department of Computers, Nanjing University of Aeronautics and Astronautics, Nanjing, Jiangsu 210016, China)

Abstract: In Traditional binary or multiple classification problem of machine learning, each class of samples is necessary for classifier design; however, in some case only one class of samples can be acquired (due to the complexity or the expensive costs), so we have to learn from the only one class of samples and form the data description for classification. The classification problem is named as one-class classification. Since now there exist the domain-specific and the generic methods, this paper first presents an overview on one-class classification, then emphasizes on the analysis of the kernel-based one-class classifiers and divides this class of methods into two types, that is, dual-based and kernel-induced distance based. Hereafter, the characteristics of these two types of methods are analyzed. Finally, we summarize the implementation techniques and applications of one-class classification in fault analysis, anomaly detection, and disease diagnosis and hostility recognition.

Key words: one/single-class; kernel methods; classifier; anomaly detection

1 引言

传统模式识别方法一般需要多个类别的样本, 从而设计两分类和多分类的分类器。而在机器故障诊断、疾病分析、信用卡欺诈、入侵检测以及敌我识别等应用领域, 我们会遇到极端情况, 即只有一类的样本可用于训练分类器。例如, 有些场合中, 几乎无法获取多类的样本(如攻击样本), 或者获取样本代价极高(卫星网络管理中的故障诊断, 无法为了获取异常样本而故意破坏卫星设备), 造成各类样本数严重不平衡。由于两分类方法的分类边界需由两类实例来“支持”, 当由于某种原因某一

边不能提供很好的“支持”, 譬如上述的某一类的实例很少的不平衡类别问题, 则分类面就会严重偏离, 因而两类方法不能很好地适应这种分类问题。所以只能从已采样的目标类样本中进行学习并形成一一个对该类别的数据描述, 然后根据给定或设计的相似性度量, 并设定阈值来判别新样本的归属^[1]。因只使用一类实例, 故称为单类分类(One-class classification)^[2]。目前提高单类分类器的性能主要从降低单类分类的错误率入手, 而单类分类中分类错误有两种: 错误的接受和错误的拒绝。前者指非目标类被分为目标类; 后者指目标类被分为非目标类^[3]。因而单类分类器在最小化错误拒绝率的同时, 最

小化错误的接受率.其第一个目标通过最小化训练集上的经验风险实现,而后一个目标则通过最小化数据描述的体积来实现.

针对目前提出的各种应用领域相关的以及通用的单类分类器,本文首先就其意义和分类做了一个综述,然后重点对主流的基于核方法的单类分类器技术进行了分析,将各种基于核的 one class 算法分为两大类:对偶方式和核距离方式.最后就目前单类分类器的应用研究领域进行了回顾,并指出了单类分类器的实用价值.

2 单类分类器的分类

目前单类分类问题已有了许多解决方法,根据其原理大致可分为四类:第一类是密度估计法;第二类基于神经网络的方法;第三类是基于聚类的方法;第四类方法是基于支持域的方法.下面对这些方法进行详细讨论.

2.1 基于密度的方法(Density-based classifiers)

对一类样本识别的最简单也是最直接的方式就是通过参数化或非参数化方法来估计训练样本的密度模型,并设置一密度阈值,小于该阈值的被认为是异常(outlier)^[4,5].其中最简单的密度模型就是高斯密度估计^[6],假设目标类的条件概率 $p_G(x|^{(v)})$ 为:

$$p_G(x|^{(v)}) = \frac{1}{\sqrt{(2\pi)^N \det(\Sigma)}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right) \quad (1)$$

其中 μ 为训练样本的均值, Σ 为训练样本的方差.对于新的样本 x ,分类器定义如下:

$$h(x) = \begin{cases} 1, & \text{if } p_G(x|^{(v)}) \geq \tau \\ 0, & \text{if } p_G(x|^{(v)}) < \tau \end{cases} \quad (2)$$

其中 τ 为判断是否异常的阈值, $h(x)$ 为判别函数,1 代表是目标类(target),0 代表野值(outlier).设定 τ 的常用做法是以目标类样本上的错误率最小为依据.

由于单模态模型太简单,缺乏足够的灵活性,故 Sain 等人^[7]采用高斯混合模型对一般性的数据分布进行建模,其目标类的条件概率 $p_{MoG}(x|^{(v)})$ 为:

$$p_{MoG}(x|^{(v)}) = \frac{1}{\sqrt{(2\pi)^N}} \sum_{j=1}^J \frac{1}{\sqrt{\det(\Sigma_j)}} \exp\left(-\frac{1}{2}(x-\mu_j)^T \Sigma_j^{-1}(x-\mu_j)\right) \quad (3)$$

其中 μ_j 是混合参数, Σ_j , μ_j 和 α_j 都可以使用 EM 算法进行优化.该方法主要问题是混合分量数量是数据依赖的故难以选择,且多个高斯模型的混合需要更多的样本来克服维数灾难.

除上述参数化方法之外,Bishop 等人应用 Parzen 窗

估计在训练集上对概率密度进行建模^[8,9],然后根据经验风险设置相应的概率密度阈值,分类时将所有概率密度低于该阈值的测试样本判为异常.另一种无参数的密度估计的方法也由 Breunig 等^[10]提出,其设计了基于样本点与其局部邻域样本点分隔程度的局部异常度(LOF,Local Outlier Factor)进行异常检测,但其本质仍然是利用 KNN 估计样本点的密度^[11].

如目标训练样本结构简单维数较低,且训练样本数较多,则基于密度的单类分类器方法比较有效.但高维有限样本情况下,由于密度估计的方法不能真实地反映模式的特征,很难对目标类数据的稀疏区域做出正确的判断.统计学习理论指出,概率密度函数的估计问题是一个比分类本身更难的问题,因而,利用 Parzen 窗函数方法进行对目标类进行构建单类分类器时,违背了一个常识性的原理:不要试图将解决一个更难的问题作为媒介,去解决另外一个特定的问题^[12].

2.2 基于神经网络的方法(ANN-based classifiers)

准确的说,单分类方法的研究是从神经网络开始的,比较知名的是 Japkowicz 的自联想器:Autoassociator^[13].基于区分的多层感知器和基于识别的 Autoassociator,其用两类实例来训练 MLP,MLP 输出的类别只有两种,输出层使用一个结点即可. Autoassociator 输出层和输入层结点数相同,输出层用来重现输入层,具有联想功能. Autoassociator 中间层较“窄”,具有压缩输入的功能.仅用目标类实例时训练 Autoassociator,对于未知实例,输出层给出了它的联想,如果这两者之间差别较小,就可以认定该实例属于目标类.

一些神经网络的方法也应用到新颖性检测上,例如自动编码器(Auto-Encoders)^[3]和学习矢量量化(LVQ, Learning Vector Quantization)^[3]和自组织特征映射(SOM, Self-organizing Map)^[14].潘等人^[15]对 SOM 模型进行了改进,将目标类训练样本输入到 SOM 网络中进行训练,以有序方式映射到 Voronoi 空间上,形成了只包含目标类信息的 SOM 原型.

该类方法是在成熟的神经网络研究基础之上,继承了神经网络的自组织与自适应能力,能够在学习过程中自适应地发现蕴含在目标样本数据中的内在特性及规律性,对大规模和非线性的实际问题有很好的分类效果.但其不足是因仅有一类样本供学习,改进的算法不可避免的要面对各种阈值的选择,比如训练神经网络需要确定网络隐层的数量以及每层神经元的数量,而这些都不能预先给定,只能通过尝试法来确定合理的网络规模.

2.3 基于聚类的方法(Clustering-based classifiers)

聚类方法也可以对一类数据进行学习,从而应用

到单类分类器的设计上,例如 K 均值^[12]和 K 中心方法^[16]. 该类方法假定目标类样本满足某种聚类假设,对已有的聚类算法进行改进,以满足单类学习的特殊要求. K 均值单类分类器首先执行 K 均值聚类,然后所有训练样本到其最近聚类中心的平均距离作为阈值,新样本到最近聚类中心的距离超过该阈值的就判为异常. K 中心单类分类器和 K 均值单类分类器的型式上相同,不同的是其选取所有训练样本到其聚类中心的距离的最大值作为阈值. 然而直观上,落在两个聚类中心之间的新样本更可能是正常样本,但是如果其到两个聚类中心的距离超过该阈值,就被判为异常. 并且这两种方法严重依赖于聚类中心的初始化,并且聚类数目 K 的选取仍然是个开放问题. 正如笔者在开篇中指出的单类分类的最重要的问题是度量的选择和阈值选取. 基于聚类的单类分类器设计中最重要的问题也是对聚类的相似性度量的选择和对判别异常的阈值的定义,这些都是数据依赖的. 该方法的另一个问题就是对野值点敏感.

由于上述基于 K 均值和 K 中心方法存在簇类数 K 选择的开放问题,所以其只能实用于簇类数已知的情形. 且硬标号聚类方法对野值的不鲁棒,使得上述方法应用价值有限. 陈斌^[17]等人提出了基于单簇类聚类的数据描述方法,将其聚类数设置为 1,避免了簇类数的选择.

2.4 基于支持域的方法(Do main - based classifiers)

基于支持域方法主要是通过对目标数据的学习,形成一个围绕目标类的边界或区域,如超平面、超球等,并且最小化目标数据支撑域的体积,以达到错误接受率最小的目的. 其借鉴了 SVM(Support Vector Machines)的最大间隔(Margin)理论^[18],因而适合处理小样本、高维和存在噪声数据的单类分类问题,代表方法有 One-class SVM^[19]和 SVDD(Support Vector Data Description)^[20]. One-class SVM 训练时寻找一个由支持向量表示的超平面,并最大化超平面和原点(视为反例)间的间隔;SVDD 则寻求一个包含所有正常类训练样本的最小超球. 和这两种数据描述性状不同,Scholkopf 等^[21]提出了 Slab SVM 引入两个超平面来约束目标数据所在区域,形成一条带状数据描述.

由于提供直观的数据描述,且借助于核技巧,可以方便地在高维特征空间中进行求解,因而该类方法成为目前流行的单类分类方法.

2.5 基于两类分类的单类分类方法

除上述单类分类方法之外,不少两类分类的方法也被利用到单类分类中,其人工生成异常类样本,补全两类样本分布后利用两类分类方法例如 SVM 等进行分类.

例如 Fisher 判别分析,采用目标类样本和原点的对称化方法构建反类样本,然后利用 FDA(Fisher Discriminant Analysis)寻求一个分类面,是两类尽可能分开,同时保证其类内的紧性^[22,23]. 而 Stewart 等^[24]则根据目标类样本空间上的均匀分布产生异常类样本,然后构建一个 SVM 获得最终分类面,然而这种情况考虑的反类分布最为复杂,一般不符合异常样本的真实分布. 因而 B áhalmi 等人^[25]根据目标类样本的边界的方向信息,产生相应的反类样本,然后通过 SVM 求解分类面.

单类应用中,极少数情形下可能存在珍贵的异常类样本,但其数量远小于目标类样本数,形成样本分布极端不平衡的问题. 除了使用标准的两类方法之外,还有部分方法专门针对这种样本分布极端不平衡情形,给不同实例的错分以不同的惩罚代价^[26],使分类面不致偏向反类. 但目前 One-Class 分类器还不是代价敏感的,实际上代价敏感对于提高类似 SVDD 这样的基于区域的分类器的能力很重要. 燕继坤等人^[27]利用了代价敏感等方法实现了 CSVDD,该方法可使得代价高的样本对分类面形成起更重要的作用. 潘志松等人^[28]也利用了代价敏感方法根据具体的实例的发生频率等信息构造了 One-Class 分类器.

3 基于核的单类分类器方法

由于现实中的目标类样本的分类问题十分复杂,而且数据分布大都呈现高维且非线性. 传统的算法在解决此类问题的时候不可避免地产生维数灾难,导致计算复杂性的指数增长. 核基理论是近年来发展起来的一种新的机器学习方法,其实质是通过核诱导的隐映射将低维输入空间中的非线性问题变换至高维(甚至无穷维)特征空间中的较易解决的(可能的)线性问题,并以内积形式刻画,从而通过核代入最终在特征空间中获得原问题的解决^[29],因此核方法广泛应用于各种学习方法中,例如支持向量机(SVM)、神经网络设计、PCA 及聚类分析等.

3.1 基于核的单类分类方法的发展

20 世纪 90 年代后,随着核基理论的发展,很多基于该理论的单类分类器也应运而生. 从 1995 年 Scholkopf 等^[30]提出硬间隔的最小超球应用于单类分类后,1999 年, Tax 和 Duin 又提出一种软间隔的最小超球 SVDD 方法,如图 1(b)所示. 其主要思路是寻求一个能把大多数训练样本包围起来的最小超球. 如目标集不是“球形”,SVDD 会因包含过多的空白区域而不够紧凑,增加了错误接受率的可能性. 因此,Wei 等人^[31]利用超椭球代替 SVDD 中的超球以考虑数据的结构信息,类似的椭球模型还有最小体积包含椭球(minimum volume enclosing ellipsoid, MVEE)^[32]以及核最小体积覆盖椭

球(kernel minimum volume covering ellipsoid, KMVCE)^[33], 它们均是通过优化椭球体积来寻找最小超椭球。

2001 年, Scholkopf 等人^[19]从统计学习理论的观点深入研究了 one-class 问题, 设计出了与支持向量机完全类似的简明算法——one-class SVM, 其基本思路是寻找超平面, 使大部分目标样本位于超平面的一侧, 而大部分非目标样本位于超平面的另一侧^[16], 如图 1(a) 所示。One-class SVM 作为一种基于支持域的方法, 它通过在训练发现一个由支持向量表示的超平面, 其能最大化和原点之间的间隔。因为只有目标类的训练实例, Scholkopf 把原点虚设为非目标类唯一的实例, 在尽量正确区分更多目标类训练实例的情况下, 使原点到超平面距离最大化。然而在使用 RBF 核后, 该方法和 SVDD 等价^[19], 且二者计算上最终都归结为凸二次规划问题, 复杂性平均为 $O(n^3)$ ^[34]。通过正则化技术, 两种方法都能实现经验风险和推广能力的平衡。它们都能通过核化得到目标类样本的更为灵活的数据描述。由于他们出色的分类性能, 它们成为当前单类分类器的主流研究方向。

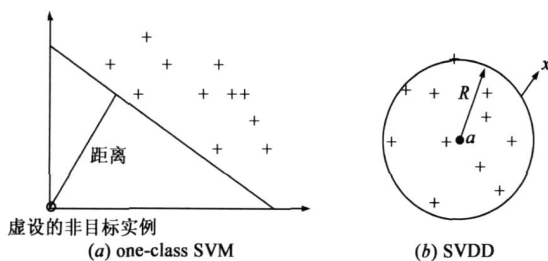


图1

由于 One-class SVM 并未考虑数据自身的紧性, 所给出的分类面可能无法紧致地覆盖目标区域, Scholkopf 等^[21]提出了 Slab SVM 引入两个超平面来约束目标数据所在区域, 这里的目标数据是对相应物体表面的测定。Tao 等^[35]将 Slab SVM 进一步引申, 将两个超平面之间的条带区域作为目标数据所在空间, 并按照两类支持向量机的设计思想, 从统计学理论出发定义了单类问题的损失函数、推广误差以及间隔, 使单类分类器如同 SVM 一样具有了完备的统计学习理论体系, 并且提出了相应的单类算法 Slab OCC。而由 Lanckriet^[36]等提出的单类最小最大概率机 (Single Class Minimax Probability Machine, SCMPM) 采用了概率约束来寻找一个尽量远离原点的超平面。根据 Chebychev 不等式, 可将目标数据的马氏距离与落入负半空间的概率密度上界建立联系, 通过最大化原点和超平面之间的马氏距离来优化目标数据所在的正半空间, 并保证数据分布在最坏情形下仍至少有指定百分比的目标数据落入超平面内部。特别有趣的是, 他们的核版本算法与计算几何中最小覆盖球算法相吻合。

笔者认为这些算法都可以归结到核框架下来。One-class SVM (SVDD) 使用核方法的一个共同技巧是对超平面法矢量(目标类中心) 采用了一个对偶形式的表示形式, 即超平面法矢量(目标类中心) 可表示为目标类的训练数据的线性组合, 因而位于高维特征空间无法在原输入空间中获得相应目标类中心的刻画。这是第一类基于核的单类分类器的工作原理。

第二类单类分类器的思路是摒弃了目标类中心的对偶形式表示, 其单类数据描述仍在输入空间而不是对偶空间中形成。其实质是核代入为原输入空间诱导出了一类不同于欧氏距离的新的距离度量^[37], 而核的灵活性可导出不同的距离度量, 从而在原空间中实现不同“核”度量的单类分类器。

3.2 对偶理论下的单类分类器

3.2.1 支持向量数据描述 (SVDD)

核基理论下的 SVDD 的基本思想是: 首先通过核映射 (\cdot) 将输入样本映射到一个高维空间, 在这个高维空间构造一个包含所有训练样本点的最小球体; 在球面上的样本点即为 SVDD 所求得的支持向量。假设模型 $f(x, w)$ 表示一类紧密的有界数据集, 因此我们可以借助一个超球体去包含并描述它。这个球体可以用中心 a 和半径 R 表示, 而且使训练集的所有样本都落在此球体内。由于训练样本中一般含有噪声或野值(也叫新颖值), 因此上述优化结果对噪声或野值敏感, 缺少鲁棒性。为提高结果的鲁棒性, 仿照 SVM 为每个样本 x_i 引入松弛变量 $i \geq 0$, $\forall i$, 软间隔的 SVDD 通过优化下面的表达式来发现包含大多数目标类训练样本的超球:

$$\text{struct } (R, a, i) = R^2 + C \sum_i i \quad (4)$$

$$\text{s.t. } \|\phi(x_i) - a\|^2 \leq R^2 + i, \quad i \geq 0, \quad \forall i$$

其中参数 C 是正则化参数。利用 Lagrange 函数求解上述约束下的最小化问题, 可得原问题的对偶问题,

$$L = \sum_i i \phi(x_i) \cdot \phi(x_i) - \sum_{i,j} i_j \phi(x_i) \cdot \phi(x_j) \quad (5)$$

$$\text{s.t. } (1) \sum_i i = 1,$$

$$(2) 0 \leq i \leq C, \quad \forall i$$

由 KKT 条件, 对应 $i = 0$ 的样本在超球内; 对应 $0 < i < C$ 的样本在超球面上, 故而称作支持向量, 其余对应于 $i = C$ 的部分样本在超球外故判为野值点 (Outlier)。

通过核代入 $\langle \phi(x_i) \cdot \phi(x_j) \rangle = K(x_i, x_j)$, 上述对偶问题变成如下形式:

$$L = \sum_i i K(x_i, x_i) - \sum_{i,j} i_j K(x_i, x_j) \quad (6)$$

$$\text{s.t. } (1) \sum_i i = 1,$$

$$(2) 0 \leq i \leq C, \quad \forall i$$

相应决策函数为:

$$\begin{aligned}
 f_{SVDD}(z; \cdot, R) &= \text{sign}(\| \phi(z) - \phi(a) \|^2 - R^2) \\
 &= \text{sign} \left(K(z, z) - 2 \sum_i \alpha_i K(z, x_i) \right. \\
 &\quad \left. + \sum_{i,j} \alpha_i \alpha_j K(x_i, x_j) - R^2 \right) \quad (7)
 \end{aligned}$$

输出为 1 则为目标类, - 1 则为异常.

3.2.2 单类支持向量机 One-class SVM

One-class SVM 首先将输入空间通过核函数映射到高维空间,在高维空间将它们与原点尽可能地分开.对于测试样本 $x, f(x)$ 表明了它在高维空间位于超平面的正负方向, $f(x)$ 为正的认为是正常类,反之异常类. One-class SVM 求解如下的二次规划问题:

$$\begin{aligned}
 \min_{\alpha, \xi} \quad & \frac{1}{2} \|w\|^2 + \frac{1}{l} \sum_i \xi_i \\
 \text{s.t.} \quad & (\alpha \cdot \phi(x_i)) - \xi_i \geq 0
 \end{aligned} \quad (8)$$

由于非 0 松弛变量 ξ_i 是对目标函数的惩罚,那么决策函数:

$$f(x) = \text{sign}(\alpha \cdot \phi(x) - \xi) \quad (9)$$

对训练样本集中的绝大部分样本 x_i 都为正,而 $\|w\|$ 仍然保持最小.它们之间的折衷由 $(0, 1/l)$ 控制.和 SVM 一样,为了求解上述公式,引入 Lagrange 乘子,可得下式:

$$\begin{aligned}
 L(w, \alpha, \xi) &= \frac{1}{2} \|w\|^2 + \frac{1}{l} \sum_i \xi_i \\
 &\quad - \sum_i \alpha_i ((w \cdot \phi(x_i)) - \xi_i + \xi_i) \\
 &\quad - \sum_i \lambda_i \xi_i
 \end{aligned} \quad (10)$$

代入核函数通过计算可得如下的对偶问题:

$$\begin{aligned}
 \min \quad & \frac{1}{2} \sum_{i,j} \alpha_i \alpha_j k(x_i, x_j) \\
 \text{s.t.} \quad & 0 \leq \alpha_i \leq \frac{1}{l}, \sum_i \alpha_i = 1
 \end{aligned} \quad (11)$$

在最优化后,可以看到原来的不等式约束(19)变成了 α_i, ξ_i 都非 0 时的等式约束.由于支持向量是位于超平面上,我们由此可以从某个支持向量 x_i 及其对应的 α_i 求出:

$$w = (\alpha \cdot \phi(x_i)) = \sum_j \alpha_j k(x_j, x_i) \quad (12)$$

3.2.3 采用线性规划算法的单类分类器方法

由于 One-class SVM 需求解二次规划问题,复杂度较高;Colin Campbell 等^[38]提出了一种采用线性规划算法的新颖性检测方法,而避免求解二次规划问题.该方法与 One-class SVM 有一定的相似之处,其基本思想也是在高维空间寻找一个最优分类面.但其不以原点作为参照最大化 Margin,而在保证所有样本均落入正半空间的条件下,最小化所有样本判别函数的均值.

假设输入空间的样本为 $x_i (i = 1, \dots, m)$,映射函数为 $\phi(x_i)$,那么在高维空间,我们可以得到权向量的形式为 $w = \sum_j \alpha_j \phi(x_j)$.因此定义 $w \cdot \phi(x_i) + b = 0$ 为高维空间的分类面.

该方法的目的是在输入空间找到一个包裹所有训练样本点的表面,所有在这个边界之外的样本都被认为是异常类别.这个表面可以定义成某种非线性函数形式 $f(z) = 0$.而在高维空间对应的就是超平面 $f(z) = \sum_i \alpha_i K(z, x_i) + b$,这个超平面与训练样本点的映射紧紧地靠在一起,并且对于所有的训练样本点,其值都大于或等于 0.通过最小化输出函数 $f(z)$ 的均值,来实现超平面与样本点映射紧凑这个目标.因此,其可以表示为如下形式:

$$W(\alpha, b) = \sum_{i=1}^m \left(\sum_{j=1}^m \alpha_j K(x_i, x_j) + b \right) \quad (13)$$

约束为:(1) $\sum_{j=1}^m \alpha_j K(x_i, x_j) + b \geq 0$;

$$(2) \alpha_i \geq 0, \sum_{i=1}^m \alpha_i = 1 \quad (14)$$

2006 年冯爱民等^[39]在上述方法中利用 K-NN 方法获得样本的局部密度信息,并强化高密度区的数据而弱化低密度区的数据对分类超平面的影响,使分类超平面自动靠近高密度区,取得一定的性能提高.

3.3 基于核距离的单类分类器

SVDD 等为代表的单类分类器算法在使用核方法的一个共同技巧是对目标类的数据描述采用一个对偶形式的表示形式,即目标类的类中心可表示为训练数据的线性组合.这些目标类的类中心属于高维特征空间,目标类的类中心在特征空间中形成,无法在原输入空间中获得相应类中心的刻画,其根本原因是在特征空间中的类中心并不一定在原输入空间中存在相应的原像(Pre-image);此外,单类分类器的分类结果不易在输入空间中获得直观的解释.

基于核距离的单类分类器,摒弃目标类中心的对偶形式表示,其单类的轮廓形成仍在原输入空间中进行,因而获得了输入空间中直观的目标类数据描述,克服了第一类核方法的不足.其实质是核代入为原输入空间诱导出了一类不同于欧氏距离的新的距离度量,而核的灵活性可导出不同的距离度量,从而在原空间中实现不同度量的聚类,这是一种核方法应用的新解释.

基于核距离的单类分类器主要是要定义核距离.定义非线性映射 $\phi: x \rightarrow F$,其中 $x \in X, X$ 为样本集, F 为特征空间.用形式化目标函数代替欧氏距离:

$$J(w_j) = \| \phi(X) - \phi(w_j) \|^2 \quad (15)$$

求其极小值. 依据 Mercer 条件的核函数的进行核化: 有

$$J(w_j) = \phi(X) - \phi(w_j) \quad (16)$$

$$= K(X, X) + K(w_j, w_j) - 2K(X, w_j)$$

当计算出的距离 d 大于该阈值 D 时, 则认为是反类, 否则认为是目标类样本.

陈斌等提出的单聚类簇类的核化版本^[17], 其使用特征空间的核诱导距离代替原空间的欧氏距离, 其优化可能性 1 - Means 算法中的目标:

$$\min J(U, V) = \sum_{j=1}^n u_j^m d^2(\phi(x_j), v) + \sum_{j=1}^n (1 - u_j)^m$$

s. t. $0 \leq u_j \leq 1$, for all i . (17)

由于这种距离仍缺乏对数据中几何结构的刻画能力, 在 One-class SVM 中 Tsang 等人^[40]引入马氏距离, 变换了单类分类器的距离度量, 提出了马氏距离的 One-class SVM, 并避免二次锥规划求解, 在一定程度上克服了原算法采用欧氏距离的不足.

基于核距离的单类分类器方法实质是核代入为原输入空间诱导出了一类不同于欧氏距离的新的距离度量后, 在输入空间中训练学习算法, 且其求解无需核矩阵的求逆过程, 故较第一类方法更快速. 此外核的灵活性可导出不同的距离度量, 从而在原空间中实现不同度量的学习, 同时所诱导出的基于径向基函数的距离度量函数是一类非线性鲁棒估计准则, 该准则使得该方法具有较强的抗噪声和野值的能力, 而诱导新距离度量的新观点可应用于其它类似问题.

4 单类分类器的应用领域

由于单类分类器的优良特性, 这些模型已经广泛应用于文本分类^[41, 42]、入侵检测^[9, 43]、手写体识别^[17, 36]、图像处理等领域^[40]. 单类分类器一般可以用于解决如下类别样本数目不平衡问题:

- (1) 比如两个训练类别中的样本数目比例高至 9:1, 如医学图像等异常样本数据极少的情况.
- (2) 异常类样本获取代价较高的问题: 比如某些卫星网络故障诊断, 要获取故障样本所付出的代价太高, 不可能为了获取故障样本而特意让卫星系统出现故障.
- (3) 异常类样本数目几乎无穷问题: 比如入侵检测中攻击数量和种类层出不穷, 而且有的攻击产生变种, 根本就不清楚该拿哪些攻击样本来训练入侵检测器.

单类分类方法同样可以用于野值与新颖值的检测. 在这种场合, 野值与新颖值被看作成异常类数据^[31]. 由于单类分类器只训练正常类的训练数据, 因此在解决方法上不同于传统的两类(或者多类)分类问题. 在单类分类问题的错误率监测上, 存在着两个标

准: 正常类错误率与异常类错误率. 这两个错误率是相关的, 正常类错误率高的往往导致异常类错误率低, 反之亦然. 因此, 在最后评价一个单类分类器的性能, 应从这两个方面综合考虑.

单类分类器由于只需要一类数据作为训练样本, 减化了数据预处理的时间, 在卫星通信系统和大型网络管理系统的故障诊断和入侵检测中都有着极大的应用前景.

5 结束语

本文对分类器设计中的热点问题——单类分类器进行了系统的研究并给出了综述. 首先根据设计原理对流行的几种单类分类方法进行了简单回顾. 然后重点阐述了基于核的单类分类方法, 并将其分类两类. 其中一类通过核空间的对偶方式获得原问题的解, 并给出了对正常和异常的判别界; 另一类方法则通过核函数诱导出的“核”距离来解决线性不可分问题, 并通过传统算法获得原空间下单类分类器的直观描述. 最后文章对单类分类器的应用研究作了一个归纳总结.

参考文献:

- [1] Moya M, Koch M, Hostetler L. One-class classifier networks for target recognition applications [A]. World Congress on Neural Networks '93 [C]. Portland: International Neural Network Society, 1993. 797 - 801.
- [2] Platt J, Scholkopf B, Shawe-Taylor J, et al. Estimating the support of high dimensional distribution [R]. Redmond: Microsoft Research, 1999.
- [3] Tax D. One-class classification concept-learning in the absence of counter-examples [D]. Netherlands: Universiteit Delft, 2001.
- [4] Tarassenko L, Hayton P, Brady M. Novelty detection for the identification of masses in mammograms [A]. Proc. of the Fourth International IEE Conference on Artificial Neural Networks [C]. London: IEE press, 1995. 442 - 447.
- [5] Chawla A B. Learning from labeled and unlabeled data using graph min-cuts [A]. Proc. 18th International Conference Machine Learning [C]. San Francisco: Morgan Kaufmann Publishers, 2001. 19 - 26.
- [6] Bishop C. Neural Networks for Pattern Recognition [M]. New York: Oxford University Press, 1995.
- [7] Sain S R, Gray H L, Woodward W A, et al. Outlier detection from a mixture distribution when training data are unlabeled [J]. Bulletin of the Seismological Society of America, 1999, 89 (1): 294 - 304.
- [8] Bishop C. Novelty detection and neural network validation [A]. IEE Proceedings Vision, Image and Signal Processing [C], London: IEE press, 1994. 217 - 222.

- [9] Yeung D, Chow C. Parzen-window network intrusion detectors [A]. Proceedings of the Sixteenth International Conference on Pattern Recognition [C], London: IEE press, 1994. 385 - 388.
- [10] Breunig M M, Kriegel H P, Ng R T, et al. LOF: identifying density-based local outliers [J]. SIGMOD, 2000, 29 (2) : 93 - 104.
- [11] Duda R, Hart P, Stork D. Pattern Classification [M]. New York: John Wiley & Sons, 2001.
- [12] Cherkassky V M. Learning from Data [M]. New York: John Wiley & Sons, 1998.
- [13] Japkowicz N. Concept-learning in the absence of counter-examples: an autoassociation-based approach to classification [D]. New Jersey: The State University of New Jersey, 1999.
- [14] Ypma E, W R P. Novelty Detection Using Self-Organizing Maps [C]. Proc. of ICONIP, Berlin: Springer, 1997. 1322 - 1325.
- [15] 潘志松,倪桂强,谭琳. 基于系统执行迹的异常检测中单类分类算法和免疫框架设计[J]. 南京理工大学学报, 2006, 30(1) : 48 - 52.
Pan Z S, Ni G Q, Tan L. One-class classification and immune framework in abnormal detection[J]. Journal of Nanjing University of Science And Technology, 2006, 30(1) : 48 - 52. (in Chinese)
- [16] Ypma A, Duin R. Support objects for domain approximation [A]. ICANN 98[C]. Sweden: Springer, 1998. 719-724.
- [17] 陈斌,冯爱民,陈松灿,李斌. 基于单簇类聚类的数据描述[J]. 计算机学报, 2007, 30(8) : 1325 - 1332.
Chen B, Feng A M, Chen S C, Li B. One-cluster clustering based data description [J]. Chinese Journal of Computers, 2007, 30(8) : 1325-1332. (in Chinese)
- [18] Vapnik V N. The Nature of Statistical Learning Theory [M]. Berlin: Springer-Verlag, 1999.
- [19] Scholkopf B, Platt J, Shawe-Taylor J, et al. Estimating the support of high-dimensional distribution[J]. Neural Computation, 2001, 13(7) : 1443 - 1471.
- [20] Tax D, Duin R. Support vector domain description[J]. Pattern Recognition Letters, 1999, 20(11 - 13) : 1191 - 1199.
- [21] Scholkopf B, Giesen J, Spalinger S. Kernel methods for implicit surface modeling [A]. Advances in Neural Information Processing Systems [C]. Cambridge USA: MIT Press, 2005. 17. 1193 - 1200.
- [22] Roth V. Kernel fisher discriminants for outlier detection[J]. Neural Computation, 2006, 18(4) : 942 - 960.
- [23] Roth V. Outlier detection with one-class kernel fisher discriminants [A]. Advances in Neural Information Processing Systems [C]. Cambridge USA: MIT Press, 2005. 17. 1169 - 1176.
- [24] Steinwart I, Hush D, Scovel C. A classification framework for anomaly detection[J]. Journal of Machine Learning Research, 2005, 6(12) : 211 - 232.
- [25] Banhalmi A, Kocsor A, Busa-fekete R. Counter-example generation-based one-class classification [A]. The 18th European Conference on Machine Learning (ECML) [C], Berlin/ Heidelberg: Springer-Verlag, 2007. 543 - 550.
- [26] Elkan C. The foundation of cost-sensitive learning [A]. Proc. of the 7th International Joint Conference on Artificial Intelligence (IJCAI2001) [C]. Seattle: Morgan Kaufmann, 2001. 973 - 978.
- [27] 燕继坤,王勇,曹春霞,郑辉. 样本错误加权的支持向量数据描述[J]. 计算机工程, 2005, 31(2) : 24 - 26.
Yan J K, Wang Y, Cao C X, Zheng H. Example error weighted support vector data description [J]. Computer Engineering, 2005, 31(2) : 24 - 26. (in Chinese)
- [28] 罗隽,丁力,潘志松,胡谷雨. 异常检测中频率敏感的单分类算法研究[J]. 计算机研究与发展, 2007, 44(z2) : 235 - 239.
Luo J, Ding L, Pan Z S, Hu G Y. Research on sequence-call-frequency-based one-class algorithm in abnormal detection [J]. Journal of Computer Research and Development, 2007, 44(z2) : 235 - 239. (in Chinese)
- [29] Chen Y, Zhou X, Huang T S. One-class SVM for learning in image retrieval. [A]. Proc. IEEE Int'l Conf. on Image Processing [C]. London: IEE press, 2001. 34 - 37.
- [30] Scholkopf B, Burges C, Vapnik V. Extracting support data for a given task [A]. First International Conference on Knowledge Discovery & Data Mining [C], Menlo Park: AAAI Press, 1995. 252 - 257.
- [31] Wei X K, Huang G B, Li Y H. Mahalanobis ellipsoidal learning machine for one class classification [A]. Proc of the 6th Int. Conf. on Machine Learning and Cybernetics [C], London: IEE press, 2007. 3528 - 3533.
- [32] Juszczak P. Learning to recognise-A study on one-class classification and active learning [D]. Netherlands: Universiteit Delft, 2006.
- [33] Dolia A N, de Bie T, Harris C J, et al. The minimum volume covering ellipsoid estimation in kernel-defined feature spaces [A]. Proc. of the 17th European Conference on Machine Learning (ECML 2006) [C]. Berlin/ Heidelberg: Springer-Verlag, 2006. 630 - 637.
- [34] Chin K K. Support vector machines applied to speech pattern classification [D]. Great Britain: University of Cambridge, 1998.
- [35] Tao Q, Wu G, Wang J. A new maximum margin algorithm for one-class problems and its boosting implementation [J]. Pattern Recognition, 2005, 38(10) : 1071 - 1077.
- [36] Lanckriet G R G, Ghaoui L E, Jordan M I. Robust novelty detection with single-class MPM [A]. Advances in Neural Information Processing Systems [C]. Cambridge: MIT Press, 2003. 15: 905 - 912.

- [37] 张道强. 基于核的联想记忆及聚类算法的研究与应用 [D]. 南京:南京航空航天大学,2005.
Zhang D Q. Kernel-based associative memories ,clustering algorithms and their applications [D]. Nanjing :Nanjing University of Aeronautics & Astronautics ,2005. (in Chinese)
- [38] Campbell C ,Bennett K P. A linear programming approach to novelty detection [A]. Advances on Neural Information Processing Systems [C]. Cambridge :MIT Press ,2000. 11 ,395 - 401.
- [39] 冯爱民,陈斌. 基于局部密度的单类分类器 LP 改进算法 [J]. 南京航空航天大学学报,2006,38(6):727 - 731.
Feng A M,Chen B. Improved LP algorithms of one-class classifier based on local density factor [J]. Journal of Nanjing University of Aeronautics & Astronautics , 2006, 38 (6) : 727 - 731. (in Chinese)
- [40] Tsang I W, Kwok J T, Li S. Learning the kernel in mahalanobis one-class support vector machines [A]. International Joint Conference on Neural Networks [C]. Vancouver : IEEE press ,2006. 1169 - 1175.
- [41] Christodoulou C I, Pattichis C S. Unsupervised pattern recognition for the classification of EMG signals [J]. IEEE Transactions on Biomedical Engineering,1999,46(2):169 - 178.
- [42] Kang W S, Jin Y C. SVDD-based method for face recognition

system [A]. Proceedings of the SCIS & ISIS [C]. Tokyo :Fuji Technology ,2006. 1302 - 1306.

- [43] Markos M ,Sameer S. Novelty detection :a review-part I :statistical approaches [J]. Signal Processing ,2003,83(12):2481 - 2497.

作者简介:



潘志松 男,1973年生,江苏南京人.2003年毕业于南京航空航天大学,博士,副教授,主要研究方向为模式识别、机器学习和在信息安全、海量数据处理以及敌我识别方面的应用.

E-mail :hotpzs@hotmail.com



陈斌 男,1974年生,江苏泰州人,博士生,讲师,主要研究方向为模式识别、数据分析与异常检测.

E-mail :chb@yzu.edu.cn