

具备隐私信息保护能力的学习器研究

蒋亦樟, 王士同

(江南大学数字媒体学院, 江苏 无锡 214122)

摘要: 随着数据挖掘技术在现实生活的日益发展, 对信息数据的共享提出了更高的要求, 隐私泄露问题变得日趋严重。设计具有隐私信息保护能力的学习器成为数据挖掘中的一个重要的问题。在已有研究成果的基础上, 简要回顾了隐私保护学习器的发展现状, 对隐私保护关键技术及学习器模型进行了分析概述, 针对近年来国内外关于隐私保护学习器的研究成果进行了归纳总结。

关键词: 隐私保护; 数据挖掘; 学习器

中图分类号: TP181

文献标识码: A

doi: 10.11959/j.issn.2096-109x.2016.00062

Research on privacy-preserving learning machines

JIANG Yi-zhang, WANG Shi-tong

(School of Digital Media, Jiangnan University, Wuxi 214122, China)

Abstract: With the increasing development of data mining, more and more data mining methods need to be carried out on the privacy data, which leads to the serious problem of privacy disclosure. Therefore, Study of privacy-preserving learning machines is becoming a key data mining topic. Current privacy-preserving learning machines were surveyed and analyzed, including their basic ideas and their learning models.

Key words: privacy preservation, data mining, learning machine

1 引言

随着信息时代大数据的飞速发展, 数据挖掘技术应运而生。当今社会, 信息数据的高度共享为各领域的合作交流、学术研究提供了有利的条件, 但同时容易导致用户隐私数据或者敏感信息的泄露。因此, 如何在进行数据挖掘的同时确保用户的敏感信息不泄露, 具有非常重要的研究价值。在 1995 年的第一届 KDD 大会上, 基于隐私保护技术的数据挖掘第一次被正式提为模式识别

领域专门研究的课题, 引起了全世界的广泛关注和重视。随后, Agrawal 等又在 1999 年召开的 KDD 会议上进一步将隐私保护数据挖掘这一主题描述为未来几年的重点研究课题。此后, 隐私保护数据挖掘迅速成为新兴的研究热点。目前, 虽然国内的科研院校、组织机构, 如复旦大学、中国农业大学以及江南大学等关于这一课题都开展了大量的研究工作并取得了一些进展。但总的来说, 关于隐私保护的数据挖掘在国内仍然处于起步阶段, 还存在非常大的研究空间。

收稿日期: 2016-05-07; **修回日期:** 2016-06-02。通信作者: 蒋亦樟, jyz0512@163.com

基金项目: 国家自然科学基金资助项目 (No.61300151, No.61572236); 江苏省自然科学基金资助项目 (No.BK20130155); 江苏省产学研前瞻性联合研究基金资助项目 (No.BY2013015-02); 中央高校基本科研业务费专项基金资助项目 (No.JUSRP51614A)

Foundation Items: The National Natural Science Foundation of China (No.61300151, No.61572236), The Natural Science Foundation of Jiangsu Province (No.BK20130155), The R&D Frontier Grant of Jiangsu Province (No.BY2013015-02), The Fundamental Research Funds for the Central Universities (No.JUSRP51614A)

从数据挖掘算法的角度,常用的隐私保护数据挖掘算法包括贝叶斯、决策树、支持向量机、 k -近邻以及 Boosting 等分类法。近年来, SVM 凭借其在分类问题上较好的顽健性和准确性在数据挖掘领域得到广泛的应用。SVM 是基于统计学习 VC 维理论和结构风险最小原理的线性分类器的一种推广^[1]。同时,为了解决高维非线性数据集的分类问题,通过将核函数引入支持向量机,实现线性分类器到非线性的扩展。随着支持向量机在数据挖掘领域的飞速发展及隐私数据泄露问题的日益严重,具有隐私信息保护功能的支持向量机越来越受到重视,在学术界引起了广泛的研究和深入的探讨。

当前,基于隐私保护的学习机研究成果层出不穷,现有的隐私保护技术大致可以分为数据失真技术^[2~14]、密码学技术^[15~25]及限制发布技术^[17~36]等。通常情况下,从数据的分布方式来看,具有隐私信息保护能力的支持向量机又可分为集中式和分布式数据 2 种。然而,随着当今时代信息技术的发展,具有隐私保护能力的支持向量机研究主要集中在分布式的情况下^[37]。目前,从针对分布式数据的隐私保护支持向量机已取得的研究成果来看,大都集中在数据水平分布和垂直分布 2 个方

向。两者的不同在于,前者是根据数据的记录形式将数据分别分布到不同的站点,后者是根据数据的属性将数据分布到不同的站点。因此,在水平分布的情况下,所有站点只存储了部分记录,在垂直分布的情况下,所有站点都只存储了所有记录的部分属性^[38]。

本文在对隐私保护关键技术及典型支持向量机模型进行简要概括的基础上,对基于隐私保护的支持向量机算法中存在的问题及具有代表性的解决思路进行归纳总结,这对具有隐私保护能力的学习器研究具有重要意义。

2 隐私保护中的关键技术

隐私保护技术主要是指在各种数据库应用中保护隐私数据或敏感信息不泄露所采用的具体技术。现实生活中,不同的实际需求决定了不同的隐私保护技术,每类隐私保护技术都有各自不同的特点,因此,只局限于一种技术的隐私保护无法满足所有的应用需求。目前,关于隐私保护中的关键技术研究大致集中在数据失真技术、加密技术以及限制发布技术 3 个方面。针对不同的实际问题,几种关键技术的保护能力、使用范围也不完全相同,表 1 和表 2 分别针对不同情况给出

表 1 隐私保护技术的性能评估

隐私保护技术	隐私保护度	计算开销	数据缺失	数据依赖性	通信开销
基于数据失真的技术	中	低	高	高	低
基于密码学的技术	高	高	低	低	高
基于限制发布的技术	高	中	中	低	低

表 2 隐私保护技术的对比分析

隐私保护技术	优点	缺点	代表技术	主要应用
基于数据失真的技术	计算开销小; 实现简单	数据失真; 严重依赖数据,不同数据需设计不同的算法	随机扰动 随机应答 数据交换 阻塞 凝聚	各种数据挖掘: 关联规则挖掘; 关联规则隐藏; 决策树分类器构建等
基于密码学的技术	数据真实、无缺失; 高隐私保护度	计算开销、通信开销大; 部署复杂、实际应用难度较高	SMC	分布式下的各种数据挖掘和发布操作: 分布式关联规则挖掘; 分布式数据匿名发布; 分布式聚类; 分布式安全计算等
基于限制发布的技术	适用范围广; 通用性高; 保证真实性; 实现简单	一定程度的数据缺失; 一定程度的隐私泄露; 最优化的数据匿名开销较大	k -匿名 l -diversity t -closeness	发布匿名化数据基于: 关联规则挖掘; 决策树分类器构建等; 聚类等

了比较全面的比较分析。

2.1 基于数据失真的技术

数据失真技术是一种重要的具有隐私保护能力的技术,主要是利用一定的隐私策略对原始数据进行干扰处理,从而实现既隐藏了敏感信息又保持原始数据中的一些信息或属性不变。具体方法包括随机化^[6,12~14]、数据交换^[7]、凝聚^[3]、阻塞^[4,5,10]等。通过数据失真技术获得的数据必须符合以下几个条件:1) 不良参与者无法利用获取到的失真数据推算出其他参与方的真实信息;2) 原始数据中的某些属性在失真数据中仍然保持不变,从后者获取到的信息等价于从前者获取的信息;3) 通过失真数据可以尽可能准确地推算出隐藏在原始数据中的知识或规则^[39]。

随机化技术包括随机扰动(random perturbation)和随机化应答(randomized response)2类。随机扰动技术^[6,13,14,39]主要是在随机化过程中,在原始数据中引入遵循一定分布的噪声来实现对敏感信息的隐藏,对外不发布原始数据,只发布扰动后的数据,从而达到保护真实数据隐私的目的。与随机扰动直接发布失真数据不同,随机化应答^[6,13,14,39]是将敏感数据间接地发布给外界,并借助一种应答特定问题的方式来实现。这样,利用随机化应答技术可以在一定程度上降低不良参与方推算出原始数据是否隐藏某些敏感信息或假信息的概率。虽然通过随机应答技术发布的并不是真实的原始数据,但在发布数据量比较大的情况下,攻击者仍然可以比较准确地推算出原始数据的统计信息和汇聚信息。

应用需求的不同导致随机化技术必须设计特定的算法对转换后的数据重建数据分布。凝聚技术^[3,39]的出现较好地弥补了随机化技术中的缺点。它将原始数据划分成不同的组,存储每个属性的均值、协方差等统计信息,然后用通用的重构算法进行处理。由于每组内存储的记录相互之间是不可区分的,通过凝聚技术重构后的数据可以较好地隐藏原始数据中的隐私或敏感信息。当面对某些需要针对真实数据进行研究的应用时,随机化和凝聚技术都无法实现这个功能,而阻塞技术^[4,5,10]则可较好地解决这个问题。阻塞技术对外不发布某些特定的数据,而是将这些特定的数

据用一个不确定的符号代替,因此,在一定程度上保护了布尔关联和分类规则的不泄露,达到保护隐私的目的。

通常情况下,当进行数据发掘时,如果数据拥有者不想共享真实数据,可以预先将真实数据利用数据失真技术进行处理后再进行发布,以保护敏感信息不泄露。基于数据失真的隐私保护技术虽然有较高的效率,但由于通过处理后获得的数据是失真数据,容易造成部分信息的丢失。

2.2 基于密码学的技术

当前,数据通信的安全性是分布式环境下实现隐私保护面临的一个重要问题,基于此,加密技术应运而生,可以较好地缓解这个问题。在分布式环境下,基于密码学的隐私保护技术的实际应用大都依赖于数据的存储方式、站点的可信度及其行为。

采用密码学技术进行数据加密的方法大都用在分布式环境中,主要以安全多方计算(SMC)^[17,24]为代表。为了解决互不信任的多用户之间进行协作计算的问题,Yao于1982年通过构造多方安全协议提出多方安全计算^[17]。

以多方安全计算为代表的隐私保护技术不仅能保证原始数据的安全性,而且能确保用户之间获取到数据的准确性。然而,由于该技术复杂度比较高,在数据参与方较多的条件下,容易导致较大的计算开销,大大降低其性能。因此,当前关于密码学的隐私保护技术研究主要集中在如何降低开销、优化协议等方面^[19,39]。

2.3 基于限制发布的技术

与以上提到的2种技术不同,基于限制发布的技术可以根据实际需求自主决定发布或不发布原始数据,或者发布范化的敏感数据,从而达到保护原始数据隐私的目的。目前,基于限制发布技术的隐私保护主要以数据匿名化为代表,即在隐私泄露和数据精度之间做了一个折中处理,确保把隐私泄露的风险把握在可控范围。该技术主要是根据实际情况对外界选择性地发布原始数据或可能暴露敏感数据的信息。数据匿名化的经典算法包括: k -匿名^[26,27,29,33~44]、 l -diversity^[45,46]、 t -closeness^[47]等。

k -匿名原则是一项只针对非敏感属性项的隐

私保护技术。它要求所发布的数据表中的每一条记录都存在其他 $k-1$ 条记录不能相互区分，这些不能相互区分的 k 条记录被称为一个等价类 (equivalence class)。通常情况下，隐私保护效果受 k 值大小的影响， k 值越大，保护效果就越好，但造成信息丢失的也越多^[26,39]。 k -匿名原则如表 3 和表 4 所示。表 3 中的“姓名”和“HIV+”为隐私数据，即敏感属性，表 4 中标灰的数据为个体的非敏感属性，满足 2-匿名的原则。即等价类中的任意一条数据都无法和另一条数据相互区分。由于 k -匿名原则对隐私数据缺乏约束，恶意参与方可以通过一致性攻击和背景知识攻击来确认隐私数据与用户之间的关系^[45]，从而容易导致用户敏感信息的泄露。

表 3 原始数据

姓名	年龄	性别	邮编	HIV+
Alex	42	男	2140	Yes
John	27	男	3180	Yes
Lily	36	女	2160	No
Tom	23	男	2220	Yes
Alan	47	女	3150	No
Ben	19	男	2180	No
...	...	...	...	...

表 4 匿名化后的结果

年龄	性别	邮编	HIV+
[40~50]	男	2k~3k	Yes
[25~30]	男	3k~4k	Yes
[35~40]	女	2k~3k	No
[20~30]	男	2k~3k	Yes
[45~50]	女	3k~4k	No
[15~20]	男	2k~3k	No
...	...	...	...

l -diversity 的保护原则是确保每一个等价类内的敏感属性至少拥有 l 个不同的值^[37]。这样就可以将攻击者确认的某个体敏感信息的概率提高至 $\frac{1}{l}$ ^[39,45,46]。同样参考表 3 和表 4, HIV+

是个体的敏感特征，记录中任意一个等价类中至少包含 2 个不同的敏感特征值，同样满足 2-diversity。

t -closeness 与 l -diversity 原则两者的不同之处在于，后者在前者的基础上，进一步研究了原始数据中敏感属性的分布情况。在 t -closeness 原则中，需要确保等价类中任意一条记录的敏感特征值的分布尽可能地与它在全部数据中的分布情况相似^[39,47]。

为了更形象地说明数据匿名化技术，图 1 给出了实际应用中的数据匿名化场景。从图 1 可以看到，数据匿名化是一个复杂的过程，其在对外发布数据且确保用户敏感信息不泄露的同时，要综合考虑原始数据、背景知识、匿名化技术以及攻击状况等多种因素。数据匿名化技术虽然具有处理各种类型数据的能力，但由于其采用的是泛化技术，因此，在提高数据真实性的同时，降低了数据的精度和利用率。

3 隐私保护学习器模型研究

目前，随着支持向量机技术在大数据应用上的快速发展，隐私泄露问题也变得日趋严重。因此，具有隐私信息保护能力的支持向量机具有重要的研究价值，得到了学术界的广泛关注和重视，并取得了一定的研究进展^[48~63]。近年来，具有隐私保护能力的支持向量机研究大都集中在分布式数据，主要包括垂直分布数据、水平分布数据以及任意分布数据。本文在已取得的研究成果基础上，对这些算法研究进行归纳概括，结果如图 2 所示。

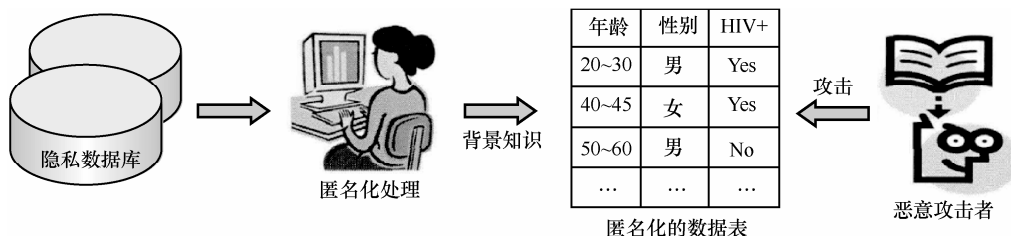


图 1 数据匿名化场景

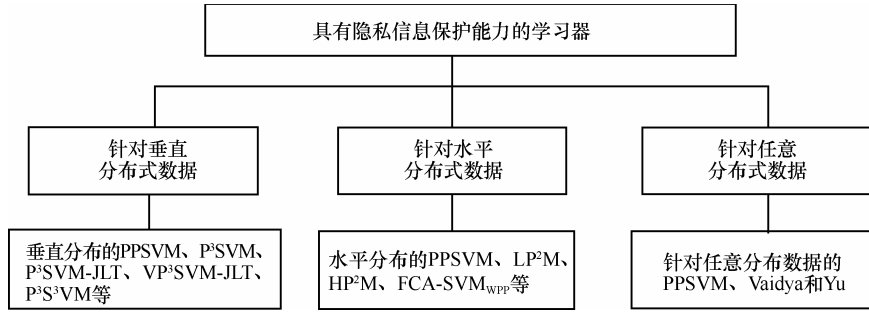


图2 隐私保护学习器模型

3.1 支持向量机模型

支持向量机的基本原理是寻找一个最优分化超平面，使不同类别的训练样本点由低维空间映射到高维特征空间后达到线性可分。支持向量机最初主要被用来处理线性分类问题，之后随着实际需求的发展，逐渐向非线性等问题推广。

3.1.1 线性支持向量机

为了进一步说明支持向量机的主要原理，图3给出了在二维线性可分条件下的最优分化超平面的实例。图3中的方形和原点标记分别表示正负2类样本点，中间黑色实线H表示分化超平面，虚线H₁和虚线H₂表示支撑超平面，前者为经过离分化超平面H最近的正类样本点的支撑超平面，后者为经过离分化超平面最近的负类样本点的支撑超平面。2个支撑超平面H₁和H₂间的距离 $\frac{2}{\|w\|}$ 定义为分类间隔（margin），它们上面的训练样本点为支持向量。从数学角度出发，最大化 $\frac{2}{\|w\|}$ 等价于最小化 $\frac{\|w\|}{2}$ 。因此，在训练样本满足线性可分或近似线性可分的情况下，最优分化超平面即使 $\frac{\|w\|}{2}$ 最小的分化超平面。

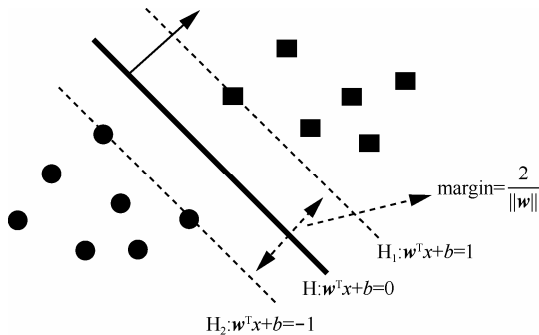


图3 二维线性可分情况下的最优分化超平面

同样，还可以将基于SVM的分类问题用数

学语言描述如下^[39]。

假设 $S = \{(x_1, y_1), \dots, (x_l, y_l)\} \in (\mathbf{R}^n \cdot y)^l$ 表示训练样本集，其中， $x_i \in \mathbf{R}^n$ ， $y_i \in y = \{+1, -1\}$ ， $i = 1, 2, \dots, l$ 。寻找 $\mathbf{R}^n$ 上的一个实值函数 $g(x)$ ，满足针对每一个输入的 x ，都可以通过决策函数 $f(x) = \text{sgn}(g(x))$ 推测出与之对应的输出 y 。

因此，在线性可分情况下，构建最优分化超平面，可以转化为以下二次规划问题

$$\begin{aligned} \min_{w, b} & \frac{1}{2} \|w\|^2 \\ \text{s.t.} & y_i (w^T x_i + b) \geq 1, i = 1, \dots, l \end{aligned} \quad (1)$$

进而，引入Lagrangian函数，引用Wolfe对偶定理，式(1)可转化为其对偶问题

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j x_i^T x_j - \sum_{i=1}^l \alpha_i \\ \text{s.t.} & \sum_{i=1}^l y_i \alpha_i = 0 \\ & \alpha_i \geq 0, i = 1, 2, \dots, l \end{aligned} \quad (2)$$

其中， $\alpha = (\alpha_1, \dots, \alpha_l)^T$ 表示Lagrangian乘子向量， α_i 不为零的样本点表示支持向量。假设满足式(2)中对偶问题的最佳Lagrangian乘子向量为 α^* ，则式(1)中的最优解为 $w^* = \sum y_i \alpha_i^* x_i$ ， $b^* = y_j - \sum y_i \alpha_i^* x_i^T x_j$ ，其中， $\alpha_i^* > 0$ ，进而可以推导出最优决策超平面为

$$f(x) = \text{sgn}(w^{*T} x + b^*) = \text{sgn}(\sum y_i \alpha_i^* x_i^T x + b^*) \quad (3)$$

然而，在现实生活中，样本容易受多种因素的干扰，导致获取到的训练集都是有噪声的。为了缓解这个问题，Comes和Vapnik引入松弛因子 $\xi_i \geq 0$ ，进而式(2)中的对偶问题可以转化为

$$\begin{aligned} \min_{w, b, \xi} & \frac{1}{2} \|w\|^2 + C \sum_{i=1}^l \xi_i \\ \text{s.t.} & y_i (w^T x_i + b) \geq 1 - \xi_i, i = 1, \dots, l \\ & \xi_i \geq 0, i = 1, \dots, l \end{aligned} \quad (4)$$

其中, C 表示惩罚参数, 满足 $C > 0$ 。式(4)的对偶问题的具体形式为

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j \mathbf{x}_i^T \mathbf{x}_j - \sum_{i=1}^l \alpha_i \\ \text{s.t.} & \sum_{i=1}^l y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C, i=1, 2, \dots, l \end{aligned} \quad (5)$$

求解式(5)中的最优解 α^* 。因此, 最开始问题中的最优解为 $\mathbf{w}^* = \sum y_i \alpha_i^* \mathbf{x}_i$, $b^* = y_j - \sum y_i \alpha_i^* \mathbf{x}_i^T \mathbf{x}_j$, $\alpha_i^* \in (0, C)$, 进而推导出最优分化超平面为

$$f(x) = \text{sgn}(\mathbf{w}^{*T} \mathbf{x} + b^*) = \text{sgn}(\sum_{i=1}^l y_i \alpha_i^* \mathbf{x}_i^T \mathbf{x} + b^*) \quad (6)$$

3.1.2 非线性支持向量机

以上是在线性可分情况下进行的分析, 下面考虑非线性分类问题。非线性支持向量机的基本思想是将在低维空间线性不可分的训练样本点, 通过某种非线性映射 $\Phi(\cdot)$, 将其转化为高维特征空间中的线性可分问题。由于在高维的特征空间只需要考虑内积运算, 因此, 为了减少高维计算开销, 可以用原空间中的函数来代替这种内积运算^[39]。泛函的相关理论表明, 在不知道 $\Phi(\cdot)$ 具体表达形式的情况下, 可以通过使用满足 Mercer 条件^[51]的核函数 $K(\mathbf{x}_i, \mathbf{x}_j) = \langle \Phi(\mathbf{x}_i), \Phi(\mathbf{x}_j) \rangle$ 来代替高维空间的内积运算。因此, 对偶问题可以表示为

$$\begin{aligned} \min_{\alpha} & \frac{1}{2} \sum_{i=1}^l \sum_{j=1}^l y_i y_j \alpha_i \alpha_j K(\mathbf{x}_i, \mathbf{x}_j) - \sum_{i=1}^l \alpha_i \\ \text{s.t.} & \sum_{i=1}^l y_i \alpha_i = 0 \\ & 0 \leq \alpha_i \leq C, i=1, 2, \dots, l \end{aligned} \quad (7)$$

进而可以得到最优分化超平面为

$$f(x) = \text{sgn}(\mathbf{w}^{*T} \mathbf{x} + b^*) = \text{sgn}(\sum_{i=1}^l y_i \alpha_i^* K(\mathbf{x}_i, \mathbf{x}) + b^*) \quad (8)$$

从以上分析可以看出, 核函数在非线性支持向量机模型中有举足轻重的作用, 支持向量机的实际应用性能与不同定义的核函数密切相关。当前, 在非线性支持向量机研究中比较常用的核函数如下。

1) 线性核函数 $K(\mathbf{x}, \mathbf{x}') = \mathbf{x}^T \mathbf{x}'$

2) 多项式核函数

$$K(\mathbf{x}, \mathbf{x}') = (\mathbf{x}^T \mathbf{x}' + c)^d, c \geq 0$$

3) 高斯径向基 (RBF) 核函数

$$K(\mathbf{x}, \mathbf{x}') = \exp\left(-\frac{\|\mathbf{x} - \mathbf{x}'\|^2}{2\sigma^2}\right)$$

4) Sigmoid 核函数 (多层感知器)

$$K(\mathbf{x}, \mathbf{x}') = \tanh(k\mathbf{x}^T \mathbf{x}' + v), k > 0, v < 0$$

3.2 垂直分布数据的隐私保护支持向量机

针对垂直分布数据构造具有隐私信息保护能力的支持向量机这一课题, Yu 等展开了相关工作^[37,52~55]。2006 年, 他们在文献[52]中指出全局核函数的求解是垂直分布数据的隐私保护支持向量机的关键技术。基于此, 提出可以将全局核矩阵的求解分解成求局部核矩阵和的问题, 并通过在求解过程中运用安全多方计算中的安全求和, 这样, 在求得核函数后, 可在不泄露原始信息的情况下建立支持向量机模型并进行分类预测, 既确保了敏感数据的安全性, 也提高了模型分类的准确性。但不足之处在于该方法中涉及了多次闭环的串行交互, 导致在实际应用中实施起来比较困难, 效率不高。

针对垂直分布数据的隐私保护问题, 以 Mangasarian 为代表的研究小组也做了大量的研究工作, 其中, 比较有代表性的思想是将完全随机核引入到 1-范数隐私保护支持向量机。该方法提出将约简支持向量机 (RSVM)^[53,54]运用到具有隐私保护能力的支持向量机研究中。在基于完全随机核的 1-范数隐私保护支持向量机中, 所有参与者通过计算自己的随机矩阵来构造各自的局部核矩阵, 然后将所有这些参与方的局部核矩阵求和, 即可构建全局约简核矩阵, 进而得到隐私保护支持向量机模型。

近年来, Sun^[37]在垂直分布数据的隐私保护支持向量机已取得的研究成果基础上, 进一步将研究方向扩展到有监督分类和半监督分类问题的隐私保护中心支持向量机上 (P³SVM), 分别提出了带有扰动的 P³SVM、基于 JL 变换的 P³SVM (P³SVM-JLT)、保持垂直分布的 P³SVM-JLT (VP³SVM-JLT) 以及半监督隐私保护中心支持向量机 (P³S³VM) 模型。

一方面, 1-范数隐私保护支持向量机中采用的是完全随机核的思想, 导致其稳定性较低; 另一方面, 1-范数 PPSVM 的训练速度会随着被处理数据集规模的逐渐增大变得越来越慢。为了弥补以上这些不足, Sun 在 1-范数 PPSVM 的基础

上进行了改进和推广,提出了带有扰动的 P^3SVM 。该方法在 $RSVM$ 的基础上,引入带有扰动的全局约简核,采用具有速度优势的中心支持向量机模型代替原有的1-范数隐私保护支持向量机模型,进而构建全局分类器。因此,该方法不仅充分利用了中心支持向量机训练速度快的优点,而且在充分提高支持向量机分类准确度的同时保证了训练和预测的速度,发挥了约简核的优势。

为了解决以往研究中只有实验验证,缺少相关理论支持的不足, Sun 在 P^3SVM 的研究基础上进一步引入 JL 变换理论,提出了 $P^3SVM-JLT$ 算法。在该算法中, Sun 继续采用 $PSVM$ 为原型,每一个参与者分别采用各自满足 JL 性质的随机矩阵构造自己的局部安全核,进而求得全局安全核,最后构建具有隐私保护能力的中心支持向量机,从而达到了隐私保护的目的。因此, $P^3SVM-JLT$ 不仅保护了原始数据,提高了隐私保护能力和分类的准确度,且在理论支撑上也更有说服力。

此后, Sun 又针对 $P^3SVM-JLT$ 中存在的相同维数限制问题,仍从 JL 变换理论出发,提出了 $VP^3SVM-JLT$ 算法。与 $P^3SVM-JLT$ 方法不同, $VP^3SVM-JLT$ 从保持数据垂直分布形式的角度出发,重构了一种新的全局安全核,而且还提供了相关理论支持。 $VP^3SVM-JLT$ 解决了 $P^3SVM-JLT$ 中受相同维数限制的问题,具有更高的灵活性。

以往的研究方法大都是基于监督的隐私保护分类,但是针对现实生活中遇到的标签不一致或没有标签的情况,通过以上方法则无法解决分类问题。因此 Sun 针对半监督的隐私保护支持向量机分类问题又进行了研究,提出一种半监督隐私保护中心支持向量机 (P^3S^3VM)。该方法在构建 P^3S^3VM 模型的过程中引入协同训练的 Tri-training 思想,在训练阶段,同时采用有标签和没有标签的样本一起训练。 P^3S^3VM 可以在半监督学习过程中,将无标签样本潜在的有效数据传递到最终的分类器中,具有良好的分类效果。

3.3 水平分布数据的隐私保护支持向量机

在针对水平分布数据进行研究具有隐私信息保护能力的 SVM 课题上 Yu 等也做了很多的工作,并取得了一定的进展,其中,比较典型的是针对布尔数据的隐私保护支持向量机。与以往方

法不同, Yu 等利用散列函数的单向计算性质和加密技术的可交换性质,提出用计算集合交集代替以往支持向量机中的求布尔向量的内积问题,最后构建全局的具有隐私保护能力的支持向量机。由于 Yu 研究的算法主要是针对水平分布的布尔型数据,因此其实际应用范围较少。

2007 年, Mangasarian 等^[56]针对以上课题也做了一些研究。同样,他们将 $RSVM$ 引入具有隐私保护能力的 SVM,与基于完全随机核的1-范数隐私保护支持向量机不同,在水平分布数据中,每个参与者都产生相同的随机矩阵并分别求得各自的局部核矩阵,然后,将所有这些参与方的局部核矩阵求和就构成所有数据的简约核矩阵,进而得到所要的隐私保护支持向量机模型。虽然该方法可以在保护原始信息的前提下,取得比较准确的分类结果,但其在训练过程中采用的是完全随机矩阵,无法保证算法的稳定性,并缺少相关理论支撑。

一般情况下,局部信息中容易隐藏个体的敏感信息,而整体信息中则没有这些敏感数据。而且,从样本的局部信息容易重构整体信息,但相反过程则无法实现。因此,整体信息的泄露不会影响样本的敏感信息。基于此,一些学者们提出将样本的整体信息和局部信息同时引入到支持向量机模型中,通过两者之间的相互协作以提高分类的准确性。

受以上思想启发, Zhang^[57]提出一种新的按标签划分的协作式隐私保护分类机 LP^2M 。 LP^2M 中参与分类的2类样本分别计算各自的均值和协方差,并将这些数据作为样本的整体信息,同时参与的双方可以相互共享各自的整体信息,这样每个个体都可以使用自己的隐私数据和对方的整体信息分别训练获得一个具有隐私保护能力的分类器,最后由参与双方训练所得的2个分类器相互协作重构最终的分类器。由于该方法在训练过程中没有采用任何的机密技术,因此通信开销较小。同时, Zhang 针对测试过程的隐私保护也做了相关研究,在文献[57]中,利用 Paillier 同态加密和 Goethals 安全内积计算协议技术,设计了一种安全测试算法,该技术可以同时保护待测样本和分类规则的安全性和隐私性。另一方面,可以通过借助核技巧和 Vaidya 等提出的算法计算内积矩阵来实现 LP^2M 非线性识别的扩展,这样,在

训练的过程中, LP^2M 非核化的线性模型就可以不利用任何第三方和加密技术来实现保护隐私的功能。通过这种技巧, 既能够保证参与双方数据元的隐私, 又可以确保不泄露数据元的数量信息。同样, 将此技巧拓展到测试过程也能够保护待测试样本的隐私, 同时能够保护分类规则不泄露。

LP^2M 主要是针对在样本按照标签划分的情况下提出的, 仅适用于一些比较特殊的场合。为了拓展应用领域, Zhang 在 LP^2M 的研究基础上, 进一步提出一种针对水平划分数据的协作式隐私保护分类机制 HP^2M 。与 LP^2M 一样, 为了保护待测数据的隐私和分类规则的不泄露, HP^2M 中引入安全内积协议和同态加密算法, 可以比较准确地估算出隐藏个体真实信息的整体信息。且 HP^2M 在训练阶段没有引入任何的加密技术, 具有更好的数据适应性。

通常情况下, 支持向量机的分类效率在一定程度上容易受支持向量个数的制约, 向量个数越多, 分类速度越慢。因此, 为了同时解决支持向量机中的隐私保护问题和分类速度问题, Hu 基于最小包含球球心在原始空间中的代理原像, 提出一种隐私保护的快速 SVM 分类方法 ($FCA-SVM_{WPP}$), 而且, Hu 等在文献[58,59]中分别设计了 QP 解法和直接解法的新方法, 旨在通过这 2 种解法实现代理球心原像的求解。在实际应用中, $FCA-SVM_{WPP}$ 不仅具有良好的隐私保护能力, 且在保证较高分类准确率的条件下, 实现了具有隐私保护能力 SVM 的快速分类。

3.4 任意分布的隐私保护支持向量机

针对任意分布的数据, Vaidya 等^[60]于 2008 年提出一种具有隐私保护能力的支持向量机方法。与以往方法不同, 该方法在采用同态加密的安全多方计算协议前提下, 通过半可信的第三方完成保护敏感信息的内积矩阵计算, 并由这个第三方采用 SVM 算法进行分类的训练, 从而实现对任意分布数据的隐私保护和分类。由于这种方法可以在不泄露原始数据的情况下得到全局核矩阵, 因此, 具有较高的隐私保护能力, 但由于该方法中采用的是安全多方计算技术, 当参与方增多时, 容易导致较大的计算开销和通信开销。

4 结束语

研究具备隐私信息保护能力的学习器是数据

挖掘领域面临的挑战之一。同时, 它也是支持向量机在隐私保护方面遇到的新问题, 旨在保护用户数据中敏感信息的同时提高分类算法的准确度。随着信息时代大数据的飞速发展, 无论是理论上还是实际应用中, 进一步深入探讨具备隐私信息保护技术的学习器有非常大的研究价值与研究空间。目前, 虽然在学术界关于具备隐私信息保护能力的支持向量机研究已取得了一定的进展, 但总体上针对该课题的研究还处于发展阶段, 值得进一步深入的研究。

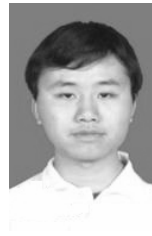
参考文献:

- [1] 刘忠宝, 王士同. 面向大规模数据的隐私保护学习机[J]. 电子科技大学学报, 2013, 42(2): 272-276.
- [2] LIU Z B, WANG S T. Privacy preserving learning machine for large scale datasets[J]. Journal of University of Electronic Science and Technology of China, 2013, 42(2): 272-276.
- [3] AGRAWAL R, SRIKANT R. Privacy-preserving data mining[J]. ACM Sigmod Record, 2000, 29(2): 439-450.
- [4] AGGARWAL C C, PHILIP S Y. A condensation approach to privacy preserving data mining[M]//Advances in Database Technology. Berlin Heidelberg: Springer, 2004: 183-199.
- [5] MOSKOWITZ L, CHANG I S. A decision theoretical based system for information downgrading[C]//The 5th Joint Conference on Information Sciences. c2000.
- [6] CHANG L W, MOSKOWITZ I S. An integrated framework for database privacy protection[C]//The 14th Working Conference on Database Security: Data and Application Security, Development and Directions. c2000: 161-172.
- [7] GREENBERG B G, KUEBLER R R J, ABERNATHY J R, et al. Application of the randomized response technique in obtaining quantitative data[J]. Journal of the American Statistical Association, 1971, 66(334): 243-250.
- [8] GOMATAM S, KARR A F, SANIL A P. Data swapping as a decision problem[J]. Journal of Official Statistics, 2005, 21(4): 635-655.
- [9] RIZVI S J, HARITSA J R. Maintaining data privacy in association rule mining[C]//The 28th VLDB Conference. c2002: 682-693.
- [10] KARGUPTA H, DATTA S, WANG Q, et al. On the privacy preserving properties of random data perturbation techniques[C]//IEEE International Conference on Data Mining, IEEE. c2003: 99-106.
- [11] SAYGIN Y, VERYKIOS V S, ELMAGARMID A K. Privacy preserving association rule mining[C]//ACM Siggraph Symposium on Geometry Processing. c2002: 97-115.
- [12] CHEN K, LIU L. Privacy preserving data classification with rotation perturbation[C]// The 5th IEEE International Conference on Data Mining (ICDM'05). c2005: 589-592.
- [13] EVFIMIEVSKI A, SRIKANT R, AGRAWAL R, et al. Privacy preserving mining of association rules[J]. Information Systems, 2004, 29(4): 343-364.
- [14] WARNER S L. Randomized response: a survey technique for eliminating evasive answer bias[J]. Journal of the American Statis-

- tical Association, 1965, 60(309): 63-69.
- [14] DU W, ZHAN Z. Using randomized response techniques for privacy-preserving data mining[C]//The 9th ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. c2003: 505-510.
- [15] DU W, ZHAN Z. Building decision tree classifier on private data[C]//IEEE International Conference on Privacy, Security and Data Mining. c2002:1-8.
- [16] PINKAS B. Cryptographic techniques for privacy-preserving data mining[J]. ACM Sigkdd Explorations Newsletter, 2002, 4(2): 12-19.
- [17] YAO C C. How to generate and exchange secrets[C]//Annual Symposium on Foundations of Computer Science. c1986:162-167.
- [18] CHAUM D, CREPEAU C, DAMGARD I. Multiparty unconditionally secure protocols[M]. Berlin Heidelberg: Springer, 1988: 462-462.
- [19] GOETHALS B, LAUR S, LIPMAA H, et al. On private scalar product computation for privacy-preserving data mining[C]// International Conference in Information Security & Cryptology. c2004: 104-120.
- [20] VAIDYA J, CLIFTON C. Privacy-preserving k -means clustering over vertically partitioned data[C]// The 9th ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. c2003:206-215.
- [21] YAO A C. Protocols for secure computations[C]//Annual Symposium on Foundations of Computer Science. c1982:160-164.
- [22] LINDELL Y, PINKAS B. Secure multiparty computation for privacy-preserving data mining[J]. Journal of Privacy and Confidentiality, 2009, 1(1):59-98.
- [23] FEIGENBAUM J, ISHAI Y, MALKIN T, et al. Secure multiparty computation of approximations[J]. Lecture Notes in Computer Science, 2002, 2(3):927-938.
- [24] CLIFTON C, KANTARCIOGLU M, VAIDYA J, et al. Tools for privacy preserving distributed data mining[J]. ACM SIGKDD Explorations Newsletter, 2002, 4(2):28-34.
- [25] LIU Y H, WEI Z J. Private-preserving naive bayesian classification[J]. Journal of Information Engineering University, 2003, 4(1):86-89.
- [26] SWEENEY L. K -anonymity: a model for protecting privacy[J]. International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, 2012, 10(5): 557-570.
- [27] SWEENEY L. Achieving k -anonymity privacy protection using generalization and suppression[J]. International Journal of Uncertainty, Fuzziness and Knowledge-based Systems, 2012, 10(5): 571-588.
- [28] MIKLAU G, SUCIU D. A formal analysis of information disclosure in data exchange[C]// ACM Sigmod International Conference on Management of Data. c2004:575-586.
- [29] YAO C, WANG X S, JAJODIA S. Checking for k -anonymity violation by views[C]//The 31th International Conference on Very Large Data Bases. c2005:910-921.
- [30] XIAO X, TAO Y. Dynamic anonymization: accurate statistical analysis with privacy preservation[C]//The ACM Sigmod International Conference on Management of Data. c2008:107-120.
- [31] MACHANAVAJJHALA A, KIFER D, ABOWD J, et al. Privacy: theory meets practice on the map[C]//IEEE 29th International Conference on Data Engineering (ICDE). c2008: 277-286.
- [32] DALVI N, SUCIU D. Answering queries from statistics and probabilistic views[C]// The 31th International Conference on Very Large Data Bases. c2005:805-816.
- [33] WONG R C, LI J, FU A W, et al. (α , k)-anonymity: an enhanced k -anonymity model for privacy preserving data publishing[C]//The 12th ACM Sigkdd International Conference on Knowledge Discovery and Data Mining. c2006:754-759.
- [34] LI J, WONG R, FU A, et al. Achieving k -anonymity by clustering in attribute hierarchical structures[M]// Data Warehousing and Knowledge Discovery. Berlin Heidelberg: Springer, 2006:405-416.
- [35] AGGARWAL G, FEDER T, KENTHAPADI K, et al. Approximation algorithms for k -anonymity[C]//The 8th Latin American Conference on Theoretical Informatics. c2010:81-93.
- [36] LEFEVRE K, DEWITT D J, RAMAKRISHNAN R. Incognito: efficient full-domain k -anonymity[C]// ACM Sigmod International Conference on Management of Data. c2005:49-60.
- [37] 孙立. 基于隐私保护技术的支持向量机研究[D]. 北京: 中国农业大学, 2014.
- SUN L. Research on privacy-preserving support vector machine[D]. Beijing: China Agricultural University, 2014.
- [38] 张战成. 基于统计学习的协作分类与隐私保护方法及应用研究[D]. 无锡: 江南大学, 2011.
- ZHANG Z C. Collaborative classification based on statistical learning and its application to privacy-preserving[D]. Wuxi: Jiangnan University, 2011.
- [39] 周水庚, 李电, 陶宇飞, 等. 面向数据库应用的隐私保护研究综述[J]. 计算机学报, 2009, 32(5): 847-861.
- ZHOU S G, LI F, TAO Y F, et al. Privacy preservation in database applications: a survey[J]. Chinese Journal of Computers, 2009, 32(5): 847-861.
- [40] PEI J, XU J, WANG Z, et al. Maintaining k -anonymity against incremental updates [C]//International Conference on Scientific and Statistical Database Management. IEEE Computer Society. c2007:5.
- [41] AGGARWAL C C. On k -anonymity and the curse of dimensionality[C]//International Conference on Very Large Data Bases. c2005:901-909.
- [42] DU Y, XIA T, TAO Y, et al. On multidimensional k -anonymity with local recoding generalization[C]//IEEE International Conference on Data Engineering. c2007:1422-1424.
- [43] MEYERSON A, WILLIAMS R. On the complexity of optimal k -anonymity[C]//The 23th ACM Sigmod-sigact-sigart Symposium on Principles of Database Systems. c2004:223-228.
- [44] EMAM K E, DANKAR F K. Protecting privacy using k -anonymity[J]. Journal of the American Medical Informatics Association, 2008, 15(5): 627-637.
- [45] MACHANAVAJJHALA A, KIFER D, GEHRKE J, et al. L -diversity: privacy beyond k -anonymity[J]. ACM Transactions on Knowledge Discovery from Data (TKDD), 2007, 1(1): 3.
- [46] ZHOU B, PEI J. The k -anonymity and L -diversity approaches for privacy preservation in social networks against neighborhood attacks[J]. Knowledge and Information Systems, 2011, 28(1): 47-77.
- [47] LI N, LI T, VENKATASUBRAMANIAN S. T -closeness: privacy beyond k -anonymity and L -diversity[C]//The 23th International Conference on Data Engineering (ICDE). c2007:106-115.

- [48] 刘晓红. 隐私保护支持向量机算法研究[D]. 山东: 山东科技大学, 2011.
LIU X H. Study on the algorithms of privacy preserving support vector machine[D]. Shandong: Shandong University of Science and Technology, 2011.
- [49] 李光. 分类挖掘中的隐私保护问题研究[D]. 哈尔滨: 哈尔滨工业大学, 2011.
LI G. Research on the privacy protection in classification mining[D]. Harbin: Harbin Institute of Technology, 2011.
- [50] 王健. 基于隐私保护的数据挖掘若干关键技术研究[D]. 上海: 东华大学, 2011.
WANG J. Study of several key issues on data mining based on privacy-preserving technology [D]. Shanghai: Donghua University, 2011.
- [51] 邓乃扬, 田英杰. 支持向量机: 理论、算法与拓展[M]. 北京: 科学出版社, 2009.
DENG N Y, TIAN Y J. New method in data mining: support vector machine [M]. Beijing: China Science Press, 2004: 5-16.
- [52] YU H, VAIDYA J, JIANG X Q. Privacy-preserving SVM classification on vertically partitioned data[J]. Journal of Jiamusi University (Natural Science Edition), 2011(3): 647-656.
- [53] LEE Y J, MANGASARIAN O L. RSVM: reduced support vector machines[C]//The 1st SIAM International Conference on Data Mining. c2001: 57-64.
- [54] LIN K M, LIN C J. A study on reduced support vector machines[J]. IEEE Transactions on Neural Network, 2003, 45(2): 199-204.
- [55] YU H, JIANG X Q, VAIDYA J. Privacy-preserving SVM using nonlinear kernels horizontally partitioned data[C]//The ACM Symposium on Applied Computing. c2006: 603-610.
- [56] MANGASARIAN O L, WILD E W. Privacy-preserving classification of horizontally partitioned data via random kernels[C]//The International Conference on Data Mining. c2008: 473-479.
- [57] 张战成, 王士同, 钟富礼. 具有隐私保护功能的协作式分类机制[J]. 计算机研究与发展, 2011, 48(6): 1018-1028.
ZHANG Z C, WANG S T, CHUNG F L. Collaborative classification mechanism for privacy-preserving[J]. Journal of Computer Research and Development, 2011, 48(6): 1018-1028.
- [58] 胡文军. 关于模式识别中大样本分类技术的几个关键问题研究[D]. 无锡: 江南大学, 2012.
HU W J. The study of several key issues on large data sets classification techniques in pattern recognition[D]. Wuxi: Jiangnan University, 2012.
- [59] 胡文军, 王士同. 隐私保护的 SVM 快速分类方法[J]. 电子学报, 2012, 40(2): 280-286.
HU W J, WANG S T. Fast classification approach of support vector machine with privacy preservation [J]. Acta Electronica Sinica, 2012, 40(2): 280-286.
- [60] VAIDYA J, YU H, JIANG X. Privacy-preserving SVM classification[J]. Knowledge and Information Systems, 2008, 14(2): 161-178.
- [61] 张学工. 关于统计学习理论与支持向量机[J]. 自动化学报, 2000, 26(1): 32-42.
ZHANG X G. Introduction to statistical learning theory and support vector machines [J]. Acta Automatica Sinica, 2000, 26(1): 32-42.
- [62] 刘向东, 陈兆乾. 一种快速支持向量机分类算法的研究[J]. 计算机研究与发展, 2004, 41(8): 1327-1332.
LIU X D, CHEN Z Q. A fast classification algorithm of support vector machines[J]. Journal of Computer Research and Development, 2004, 41(8): 1327-1332.
- [63] 汤琳, 何丰. 隐私保护的数据挖掘方法的研究[J]. 计算机技术与发展, 2011, 21(4): 156-159.
TANG L, HE F. Research on privacy-preserving data mining method[J]. Computer Technology and Development, 2011, 21(4): 156-159.

作者简介:



蒋亦樟 (1988-), 男, 江苏无锡人, 博士, 江南大学讲师, 主要研究方向为软计算、计算智能与信息安全。



王士同 (1964-), 男, 江苏扬州人, 江南大学教授、博士生导师, 主要研究方向为人工智能和模式识别。