

Research and Application of SVM Model Based on Privileged Information

Ying Cao, Fenyan Wang, Changjing Lu

School of Mathematics and Physics, China University of Geosciences, Wuhan Hubei
Email: CY520526@cug.edu.cn, wangfenyan@cug.edu.cn, luchangjing@cug.edu.cn

Received: Dec. 5th, 2017; accepted: Dec. 22nd, 2017; published: Dec. 29th, 2017

Abstract

When using Support Vector Machine (SVM) to training classification model, we may encounter additional information in training samples. Because it will seriously affect classification accuracy of test samples, it really can't be ignored. In this paper, a SVM model based on privilege information is proposed, which contains many cases of privilege information and can solve many kinds of distribution problems effectively. This paper first described the basic principles of SVM classification, and then it introduced the concept of privilege information. Next this paper proposed a SVM model based on privilege information and gave its special cases. Then, the application of SVM model based on privilege information was introduced. Finally, the existing problems and the development direction of this research were summarized.

Keywords

Privilege Information, Support Vector Machine (SVM), Application

基于特权信息的SVM模型研究及应用

曹颖, 王芬艳, 卢常景

中国地质大学数学与物理学院, 湖北 武汉
Email: CY520526@cug.edu.cn, wangfenyan@cug.edu.cn, luchangjing@cug.edu.cn

收稿日期: 2017年12月5日; 录用日期: 2017年12月22日; 发布日期: 2017年12月29日

摘要

当使用支持向量机(Support Vector Machines, SVM)进行分类时, 可能会遇到训练样本中存在额外信息的情况。它是不可忽视的, 因为它会严重影响测试样本的分类准确率。本文提出了一种基于特权信息的

SVM模型, 其包含了特权信息分布的多种情况, 能有效地解决多种特权信息分布问题。本文首先叙述了**SVM**分类的基本原理, 在简单介绍特权信息概念后, 提出了基于特权信息的**SVM**模型及其特例。接着介绍了基于特权信息的**SVM**模型应用, 最后对该研究领域存在的问题及发展方向进行了总结。

关键词

特权信息, 支持向量机, 应用

Copyright © 2017 by authors and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

支持向量机(SVM)是一种基于统计学习理论的机器学习方法[1]。SVM在一定程度上能够解决“维数灾难”与“过学习”等传统学习器难以解决的问题, 同时它有直观的几何形式与完美的数学解释。由于其良好的性质, SVM被广泛应用于文本分类、目标检测、字符识别、回归预测等领域[2] [3] [4]。SVM也因此成为模式识别、机器学习中的热门课题。

在数据挖掘过程中, 基于特权信息的学习方法(Learning Using Privileged Information, 简称LUPI)是Vladimir Vapnik在近几年正在研究的一种新的学习方法。在这种学习方法中, 除了标准的训练数据集以外, 还具有有一些额外信息, 即特权信息。特权信息指的是在学习过程中对一个样本给出的解释、评价和比较等有价值性的信息, 其作用不可忽视[5]。此外, 基于特权信息的学习方法, 在一定程度上能够对SVM模型起到补充作用, 使得对待分类的测试样本点有更小的误差率。本文的主要研究是针对特权信息的分布情况提出了一种基于特权信息的SVM综合模型, 该模型能够很好的处理多种特权信息分布模型。

2. 支持向量机(SVM)分类原理

SVM对数据的学习是把分类问题转化为一个有约束的二次规划问题进行求解, 得到最优解以构造最优决策分类超平面来实现分类[6]。以二分类样本集的学习问题为例, 已知样本集为 $\{(x_i, y_i), i=1, 2, \dots, l\}$, 其中 $x_i \in R^l$ 表示输入变量, $y_i \in \{-1, 1\}$ 表示输出变量, l 为样本个数。则得到SVM模型为

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

其中, $\phi(\cdot)$ 为非线性映射, ξ_i 为非负的松弛变量, C 为惩罚参数。 ξ_i 用于软化约束条件, C 表示对错误的惩罚代价。在实际情况中, 往往都是通过求解其对偶问题来获得最优解。对于该二次规划问题, 采用拉格朗日乘法将其转换为对偶形式:

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

其中, $K(\cdot, \cdot)$ 为核函数, 且 $K(x_i, x_j) = \phi(x_i) \cdot \phi(x_j)$ 。

可以证明, 上述优化问题的最优解 $\alpha^* = (\alpha_1^*, \alpha_2^*, \dots, \alpha_l^*)^T$ 中只有一部分(通常是少部分) α_i^* 的分量不为零。这时, 我们就把不为零的 α_i^* 对应的样本点定义为支持向量, 这就是 SVM 的基本思想。选择向量 α^* 中一个小于 C 的正分量 α_j^* , 并据此计算

$$b^* = y_j - \sum_{i=1}^l y_i \alpha_i^* K(x_i, x_j), \forall j \in \{j | \alpha_j^* > 0\}$$

从而得到最优决策函数:

$$f(x) = \text{sgn} \left\{ \sum_{i=1}^l \alpha_i^* y_i K(x_i, x) + b^* \right\}$$

3. 具有特权信息的 SVM 模型

3.1. 基于特权信息的监督学习

通常来说, 可以将监督学习描述为: 给定一个概率分布 $P(x, y)$ 未知的训练集 $(x_1, y_1), \dots, (x_l, y_l)$ 。对这个训练集进行学习, 在给定的一系列函数 $f(x, a), a \in \Lambda$ 中寻找一个函数 $y = f(x, \alpha^*)$ 使得待分类的测试样本点被错误分类的概率达到最小。

在数据挖掘过程中, 基于特权信息的学习方法除了需要处理标准的训练数据集 $(x, y) \in X \times \{\pm 1\}$ 以外, 还需要处理一些特权信息 $x^* \in X^*$ 。这些特权信息能应用到训练样本中, 却不能应用于测试样本。所以该方法要求: 给定一个训练集 $\{(x_i, x_i^*, y_i)\}_{i=1}^l$, 寻找一个函数 $f: X \rightarrow \{-1, 1\}$, 使得对待分类的测试样本点有较小的误差率。此外, 基于特权信息的学习方法, 在一定程度上能够对 SVM 模型起到补充作用。基于特权信息的模型和一般的 SVM 模型的不同之处就在于, 在训练阶段, 给出的是带有特权信息的样本集 (x, x^*, y) 而不是 (x, y) 。这些特权信息 $x^* \in X^*$ 是属于一个不同于 X 空间的 X^* 空间。

在现实生活中, 特权信息是普遍存在的。而对于几乎所有的机器学习问题, 都会存在着一些可应用的不同种类的特权信息。基于特权信息的学习方法的主要目标是通过使用特权信息来增加算法的收敛速度[7][8]。同时, 这些特权信息都或多或少的与预测的结果相关联, 所以在机器学习中加入这样的特权信息, 有望使得数据的类型得到丰富, 预测的结果得到进一步的精准。但是, 值得注意的是, 这些特权信息在训练集上是可提供的, 但是对于测试集来说就难以提供了。

3.2. 具有特权信息的 SVM 模型及其特例

3.2.1. 具有特权信息的 SVM 模型

给定 $l(l \in N^+)$ 个样本, 假设:

- (1) 其中 $n(n \in N^+)$ 个样本无特权信息, 其余 $(l-n)$ 个样本具有特权信息;
- (2) 具有特权信息的 $(l-n)$ 个样本来自于 $t(t \in N^+)$ 个不同的特权空间。

即样本集合可表示为: $(x_1, y_1), \dots, (x_n, y_n), (x_{n+1}, x_{n+1}^1, y_{n+1}), \dots, (x_m, x_m^1, y_m), \dots, (x_A, x_A^j, y_A), \dots, (x_B, x_B^j, y_B), \dots, (x_d, x_d^j, y_d), \dots, (x_l, x_l^j, y_l), 0 \leq n \leq m \leq A \leq B \leq d \leq l, n \in N^+, m \in N^+, A \in N^+, B \in N^+, d \in N^+$ 。

对训练集 (x, y) 与 (x, x^j, y) , 我们将向量 $x \in X$ 映射到空间 Z , 向量 $x^j \in X^j$ 映射到空间 Z^j 。第 $j(j=1, 2, \dots, t)$ 个特权空间中的松弛变量 ξ_i^j 可用校正函数 $\varphi_i^j = (w \cdot z_i^j) + b^j$ 与变量 ξ_i^j 混合模拟, 即

$\xi_i^j = (w^j \cdot z_i^j) + b^j + \xi_i^j$, $(w^j \cdot z_i^j) + b^j \geq 0$, $\xi_i^j \geq 0$, $1 \leq i \leq l$ 。求解如下线性规划问题:

$$\begin{aligned} \min_{w, w^1, \dots, w^l, b, b^1, \dots, b^l, \xi, \xi^1, \dots, \xi^l} & \frac{1}{2} \left[(w \cdot w) + \gamma \left((w^1 \cdot w^1) + \dots + (w^j \cdot w^j) + \dots + (w^l \cdot w^l) \right) \right] + C \sum_{i=1}^n \xi_i \\ & + C_1 \sum_{n+1}^m \left[(w^1 \cdot z_i^1) + b^1 \right] + \theta_1 C_1 \sum_{i=n+1}^m \xi_i^1 + \dots + C_j \sum_A^B \left[(w^j \cdot z_i^j) + b^j \right] \\ & + \theta_j C_j \sum_{i=A}^B \xi_i^j + \dots + C_k \sum_d^l \left[(w^l \cdot z_i^l) + b^l \right] + \theta_k C_k \sum_{i=d}^l \xi_i^l \\ \text{s.t.} & y_i \left[(w \cdot z_i) + b \right] \geq 1 - \xi_i, i = 1, \dots, n \\ & y_j \left[(w \cdot z_i) + b \right] \geq 1 - \left[(w^j \cdot z_i^j) + b^j + \xi_i^j \right], j = 1, 2, \dots, t \\ & \xi_i \geq 0, i = 1, \dots, n \\ & \xi_i^1, \xi_i^2, \dots, \xi_i^l \geq 0 \\ & (w^j \cdot z_i^j) + b^j \geq 0, j = 1, 2, \dots, t \end{aligned}$$

其中 i 为对应特权空间的样本序号, γ, C, C_j 为惩罚参数, 参数 θ_j 的作用是调节线性函数 $\varphi_i^j = (w^j \cdot z_i^j) + b^j$ 与变量 ξ_i^j 两者之间的惩罚程度。

求解此二次优化问题的标准方法是构造拉格朗日函数, 引入拉格朗日乘子 α_i 与 β_i , $\alpha_i \geq 0, \beta_i \geq 0$, $0 \leq i \leq l$ 。首先极小化 $w, w^1, \dots, w^l, b, b^1, \dots, b^l, \xi, \xi^1, \dots, \xi^l$, 然后极大化拉格朗日乘子 α 与 β , 可得到原始模型的对偶问题为:

$$\begin{aligned} \max_{\alpha, \beta} & \sum_{i=1}^l \alpha_i - \frac{1}{2} \sum_{i,k=1}^l \alpha_i \alpha_k y_i y_k K(x_i, x_k) \\ & - \frac{1}{2\gamma} \sum_{i,k=n+1}^m (\alpha_i + \beta_i - C_1) (\alpha_k + \beta_k - C_1) K^1(x_i^1, x_k^1) \\ & \dots - \frac{1}{2\gamma} \sum_{i,k=A}^B (\alpha_i + \beta_i - C_j) (\alpha_k + \beta_k - C_j) K^j(x_i^j, x_k^j) \\ & \dots - \frac{1}{2\gamma} \sum_{i,k=d}^l (\alpha_i + \beta_i - C_l) (\alpha_k + \beta_k - C_l) K^l(x_i^l, x_k^l) \\ \text{s.t.} & \sum_{i=1}^l \alpha_i y_i = 0 \\ & \sum_{i=1}^n (\alpha_i + \beta_i - C) = 0 \\ & \sum_{i=n+1}^l (\alpha_i + \beta_i - C_j) = 0, j = 1, \dots, t; \\ & 0 \leq \alpha_i \leq \theta_j C_j, i = 1, \dots, l; j = 1, \dots, t \\ & \alpha_i \geq 0, \beta_i \geq 0 \end{aligned}$$

其中 A, B 分别为相应特权空间的起始样本与终止样本序号。

在这种情况下, 对偶空间解决方案定义了决策函数 $y = \text{sgn}[f(x)]$, 其中

$$f(x) = (w \cdot z) + b = \sum_{i=1}^l \alpha_i y_i K(x_i, x)。$$

3.2.2. 具有特权信息的 SVM 模型特例

1) 全部样本存在特权信息的 SVM 模型

给定的训练集 (x, x^*, y) 包括了来自空间 X 的向量 x 和来自空间 X^* 的向量 x^* ，我们做映射 $x \rightarrow z \in Z, x^* \rightarrow z^* \in Z^*$ 。由此，我们分别定义线性的决策函数和校正函数为 $(w \cdot z) + b$ 和 $(w^* \cdot z^*) + b^*$ 。求解如下的最优化问题：

$$\begin{aligned} \min_{\omega, \omega^*, b, b^*} & \frac{1}{2} [(w \cdot w) + \gamma (w^* \cdot w^*)] + C \sum_{i=1}^l [(w^* \cdot z_i^*) + b^*] \\ \text{s.t.} & y_i [(w \cdot z_i) + b] \geq 1 - [(w^* \cdot z_i^*) + b^*], i=1, \dots, l \\ & (w^* \cdot z_i^*) + b^* \geq 0, i=1, \dots, l \end{aligned}$$

其中， γ 与 C 为惩罚系数。

2) 全部样本存在特权信息且松弛变量改动的 SVM 模型

我们通过线性函数 $\phi(x_i) = (w^* \cdot z_i^*) + b^*$ 和变量 ξ_i^* 混合的函数来定义松弛变量 ξ_i ，由此得出：

$$\begin{aligned} \xi_i &= [(w^* \cdot z_i^*) + b^*] + \xi_i^*, i=1, \dots, l \\ (w^* \cdot z_i^*) + b^* &\geq 0, \xi_i^* \geq 0, i=1, \dots, l \end{aligned}$$

求解如下的最优化问题：

$$\begin{aligned} \min_{\omega, \omega^*, b, b^*, \xi^*} & \frac{1}{2} [(w \cdot w) + \gamma (w^* \cdot w^*)] + C \sum_{i=1}^l [(w^* \cdot z_i^*) + b^*] + \theta C \sum_{i=1}^l \xi_i^* \\ \text{s.t.} & y_i [(w \cdot z_i) + b] \geq 1 - [(w^* \cdot z_i^*) + b^*] - \xi_i^* \\ & (w^* \cdot z_i^*) + b^* \geq 0 \\ & \xi_i^* \geq 0 \end{aligned}$$

其中， γ 与 C 为惩罚系数，参数 θ 的作用是用来调节线性函数和变量 $\phi(x_i) = (w^* \cdot z_i^*) + b^*$ 和变量 ξ_i^* 两者之间的惩罚程度。

3) 只有部分训练样本存在特权信息的 SVM 模型

当样本中只有一部分存在可提供的特权信息，而有 n 个样本是没有特权信息时，样本集合表示为： $(x_1, y_1), \dots, (x_n, y_n), (x_{n+1}, x_{n+1}^*, y_{n+1}), \dots, (x_l, x_l^*, y_l)$ 。在这种情况下，只需要对有特权信息的样本做些处理，这种情况的最优化模型为：

$$\begin{aligned} \min_{\omega, \omega^*, b, b^*} & \frac{1}{2} [(w \cdot w) + \gamma (w^* \cdot w^*)] + C \sum_{i=1}^n \xi_i + C^* \sum_{i=n+1}^l [(w^* \cdot z_i^*) + b^*] \\ \text{s.t.} & y_i [(w \cdot z_i) + b] \geq 1 - \xi_i, i=1, \dots, n \\ & y_i [(w \cdot z_i) + b] \geq 1 - [(w^* \cdot z_i^*) + b^*], i=n+1, \dots, l \\ & \xi_i \geq 0, i=1, \dots, n \\ & (w^* \cdot z_i^*) + b^* \geq 0, i=n+1, \dots, l \end{aligned}$$

其中， γ ， C 与 C^* 为惩罚系数。

4) 来自多空间特权信息的 SVM 模型

假设所给的特权信息来自于不同的空间，不失一般性，我们可以考虑两个校正空间： X^* 和 X^{**} 空间。令所给的样本集中部分样本点的形式为： $(x_i, x_i^*, y_i), i=1, \dots, n$ ，剩下部分样本点的形式为： $(x_i, x_i^{**}, y_i), i=n+1, \dots, l$ 。

为了让样本集线性可分,我们把向量 $x \in X$ 映射到空间 Z , 向量 $x^* \in X^*$ 映射到空间 Z^* , 向量 $x^{**} \in X^{**}$ 映射到空间 Z^{**} , 由此组成三个线性函数 $(w \cdot z) + b$, $(w^* \cdot z^*) + b^*$, $(w^{**} \cdot z^{**}) + b^{**}$ 。由此我们可以得到来自多空间的特权信息学习模型的最优化问题为:

$$\begin{aligned} \min_{\omega, \omega^*, \omega^{**}, b, b^*, b^{**}} \quad & \frac{1}{2} \left[(w \cdot w) + \gamma \left((w^* \cdot w^*) + (w^{**} \cdot w^{**}) \right) \right] \\ & + C \sum_{i=1}^n \left[(w^* \cdot z_i^*) + b^* \right] + C^* \sum_{i=n+1}^l \left[(w^{**} \cdot z_i^{**}) + b^{**} \right] \\ \text{s.t.} \quad & y_i \left[(w \cdot z_i) + b \right] \geq 1 - \left[(w^* \cdot z_i^*) + b^* \right], i = 1, \dots, n \\ & y_i \left[(w \cdot z_i) + b \right] \geq 1 - \left[(w^{**} \cdot z_i^{**}) + b^{**} \right], i = n+1, \dots, l \\ & (w^* \cdot z_i^*) + b^* \geq 0, i = 1, \dots, n \\ & (w^{**} \cdot z_i^{**}) + b^{**} \geq 0, i = n+1, \dots, l \end{aligned}$$

其中, γ , C 与 C^* 为惩罚系数。

4. 具有特权信息的 SVM 模型应用

1) 具有特权信息的 SVM 模型在恶意软件检测中的应用

工程、金融、医药等方面的一些重要应用问题可以作为基于分类的异常检测问题[9]。解决这个问题一个经典办法是使用 SVM 来描述正常状态。然后为了检测异常, Burnaev 提出了一种新的分类方法。同时制定了一个新的问题陈述及相应的算法, 可以在训练阶段考虑到特权信息。使用合成数据集以及恶意软件分类数据集来评估新方法的性能。

Burnaev 提供了改进分类问题的方法即允许合并特权信息。从实验结果可以看出, 在某些情况下, 特权信息可以显著提高异常检测的准确性。在特权信息对于相关问题的结构没有用的情况下, 特权信息不会对分类功能产生重大影响[10]。

2) 基于特权信息的 SVM 模型在高级学习范式中的应用

在 Vapnik 的“基于经验数据的依赖估计”一书后记中, 引入了一种名为学习使用隐藏信息的高级学习范式(LUHI) [11]。这个后记还提出了 SVM 的扩展算法来解决 LUHI 示例。与现有的机器学习模式中教师不起作用相比, 在新的范式中, 教师可以为学生提供解释, 评论, 比较等方面存在的隐藏信息。Vapnik 讨论了新范式的细节和相应算法, 同时考虑了特定信息的几种具体形式, 在解决实际问题时展示了新范式优于经典范式。

这种新的学习范式使得科学与人文和情感等元素进行整合。在数字识别问题中这样整合的元素描述为特权信息, 这种特权信息有助于新范式的学习[12]。新范式是广泛适用的, 它不仅成为机器学习的一个重要分析方向, 也可以成为统计学, 认知科学和哲学领域的重要分析方向。

3) 基于特权信息的 SVM 模型与经验风险最小化算法的综合应用

在使用特权信息的学习范式中, 除了决策空间中的标准训练数据之外, 教师还向学习者提供特权信息[13]。学习者的目标是在决策空间中找到一个泛化误差较小的分类器。Pechyony 制定了具有特权信息的经验风险最小化算法, 称为特权 ERM, 并给出了风险约束且叙述了修正空间的条件。即使决策空间中的原始学习问题非常困难, 特权 ERM 也允许进行快速学习。结果显示在决策空间中特权 ERM 比常规 ERM 的学习速率快得多[14] [15]。

5. 总结与讨论

本文的主要研究内容是: 针对特权信息的分布情况, 提出了一种基于特权信息的 SVM 综合模型, 其

中包含了多种特权信息分布的模型。此外,本文总结了国内外研究者对基于特权信息的 SVM 在几个领域的应用,并加以讨论。总结发现,基于特权信息的 SVM 模型以统计学习理论为基础,存在全局泛化性能好、收敛速度较快等优点,但同时也存在诸多缺陷,有很多问题需深入研究。

(1) SVM 算法的核心是核函数及其参数,它们的正确选取对 SVM 的预测及泛化性能影响很大[16]。对于具体问题,基于特权信息的 SVM 模型究竟选择哪种核函数并找到最优的参数对求解问题至关重要。因此,如何快速准确地选择核函数及对应的参数是亟待解决的问题。

(2) 在大规模及实时性要求较高的系统中,基于特权信息的 SVM 算法受制于求解问题的收敛速度和系统规模的复杂程度。尤其要处理大规模数据时,基于特权信息的 SVM 算法需要解决样本规模和速度间的矛盾,提高训练的效率和精度。

(3) 如何有效地将二分类有效地扩展到多分类问题上,基于特权信息的多分类 SVM 模型的优化设计也是今后研究的内容。

(4) 针对特定问题如何实现基于特权信息的 SVM 模型与其他算法的融合,从而顺利地解决问题也是今后需要研究的方向。

基金项目

国家自然科学基金项目(11301492); 中国地质大学(武汉)基础研究基金项目(CUGL140420)。

参考文献 (References)

- [1] Vapnik, V.N. (1998) *Statistical Learning Theory*. John Wiley & Sons, Hoboken.
- [2] 熊平. 数据挖掘算法与 Clementine 实践[M]. 北京: 清华大学出版社, 2011.
- [3] Bishop, C.M. (2006) *Pattern Recognition and Machine Learning*. Springer, Berlin.
- [4] 朱明. 数据挖掘[M]. 合肥: 中国科学技术大学出版社, 2002.
- [5] Burnaev, E. and Smolyakov, D. (2016) One-Class SVM with Privileged Information and Its Application to Malware Detection. ArXiv, Los Alamos.
- [6] Gine, E. and Koltchinskii, V. (2006) Concentration Inequalities and Asymptotic Results for Ratio Type Empirical Processes. *Annals of Probability*, **34**, 1143-1216. <https://doi.org/10.1214/009117906000000070>
- [7] 邓乃扬, 田英杰. 数据挖掘的新方法—支持向量机[M]. 北京: 科学出版社, 2004.
- [8] Manevitz, L.M. and Yousef, M. (2001) One-Class SVMs for Document Classification. *Journal of Machine Learning Research*, **2**, 139-154.
- [9] Vapnik, V. and Vashist, A. (2009) A New Learning Paradigm: Learning Using Privileged Information. *Neural Networks*, **22**, 544-557. <https://doi.org/10.1016/j.neunet.2009.06.042>
- [10] Boser, G. (1992) A Training Algorithm for Optimal Margin Classifiers. *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, Pittsburgh, 27-29 July 1992, 144-152. <https://doi.org/10.1145/130385.130401>
- [11] Cristianini, N., Shawe-Taylor, J. 支持向量机导论[M]. 李国正, 王猛, 曾华军, 译. 北京: 电子工业出版社, 2004.
- [12] Pechyony, D. and Vapnik, V. On the Theory of Learning with Privileged Information. John Wiley & Sons, Hoboken.
- [13] Boucheron, S., Bousquet, O. and Lugosi, G. (2005) Theory of Classification: A Survey of Some Recent Advances. *ESAIM: Probability and Statistics*, **9**, 329-375. <https://doi.org/10.1051/ps:2005018>
- [14] Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. *Machine Learning*, **20**, 273-297. <https://doi.org/10.1007/BF00994018>
- [15] Schölkopf, B., Smola, A.J., Williamson, R.C. and Bartlett, P.L. (2000) New Support Vector Algorithms. *Neural Computation*, **12**, 1207-1245. <https://doi.org/10.1162/089976600300015565>
- [16] Micchelli, C. and Pontil, M. (2005) Learning the Kernel Function via Regularization. *Journal of Machine Learning Research*, **6**, 1099-1125.

知网检索的两种方式：

1. 打开知网页面 <http://kns.cnki.net/kns/brief/result.aspx?dbPrefix=WWJD>
下拉列表框选择：[ISSN]，输入期刊 ISSN：2324-7991，即可查询
2. 打开知网首页 <http://cnki.net/>
左侧“国际文献总库”进入，输入文章标题，即可查询

投稿请点击：<http://www.hanspub.org/Submission.aspx>

期刊邮箱：aam@hanspub.org