

鲁棒的光滑支持向量机

胡金扣 邢红杰

(河北大学数学与计算机学院河北省机器学习与计算智能重点实验室 保定 071002)

摘要 光滑支持向量机(Smooth Support Vector Machine, SSVM)是传统支持向量机的一种改进模型,它利用光滑方法将传统支持向量机的二次规划问题转化成无约束优化问题,并使用 Newton-Armijo 算法求解该无约束优化问题。在光滑支持向量机的基础上提出了鲁棒的光滑支持向量机(Robust Smooth Support Vector Machine, RSSVM),其利用 M-estimator 代替 SSVM 中基于 L2 范数的正则化项,并利用半二次最小化优化方法求解相应的最优化问题。实验结果表明所提方法可以有效地提高 SSVM 的抗噪声能力。

关键词 光滑支持向量机,半二次最小化,核函数

中图分类号 TP391.4 **文献标识码** A **DOI** 10.11896/j.issn.1002-137X.2015.10.048

Robust Smooth Support Vector Machine

HU Jin-kou XING Hong-jie

(Hebei Province Key Laboratory of Machine Learning and Computational Intelligence, College of Mathematics and Computer Science, Hebei University, Baoding 071002, China)

Abstract Smooth support vector machine (SSVM) is regarded as an improved model of the traditional support vector machine. SSVM utilizes the smooth technique to reformulate the quadratic programming problem of the traditional support vector machine as an unstrained optimization one. Moreover, the Newton-Armijo algorithm is used to solve the unstrained optimization problem. In the paper, on the basis of SSVM, robust smooth support vector machine (RSSVM) was proposed by utilizing M-estimator to substitute the L2-norm based regularization term of SSVM. Furthermore, the half-quadratic minimization method is used to solve the corresponding optimization problem of RSSVM. Experimental results demonstrate that the proposed method can efficiently enhance the anti-noise capability of SSVM.

Keywords Smooth support vector machine, Half-quadratic minimization, Kernel function

1 引言

支持向量机(SVM)是基于统计学习理论的一种机器学习方法^[1-3]。由于具有较强的泛化性能,近年来它得到了广泛的关注。在实际应用中,SVM表现出优于其他传统学习方法的性能^[1]。此外,SVM在模式识别^[4,5]和函数回归中取得了较好的效果^[6,7],主要是因为它能够同时最小化训练误差和模型的复杂度^[8]。

SVM需要对带有线性约束的凸二次规划问题进行求解,在求解过程中,利用标准二次型优化技术求解二次规划的对偶问题^[8]。光滑支持向量机(SSVM)^[9]是标准支持向量机的一种改进方法,它由 Lee 和 Mangasarian 提出。SSVM 利用光滑技术,将 SVM 的约束最优化问题转化为光滑无约束优化问题,并利用 Newton-Armijo 算法求解^[9]。SSVM 具有强凸性、可微性和更优的性能^[10]。可以证明,当 SSVM 中的参数 λ 趋于无穷时,SSVM 的最优解收敛于原优化问题的最优解。Lee 等将 SSVM 推广到 ϵ 非灵敏回归中并取得了较好的

效果^[11]。Chang^[12]等将 SSVM 应用于处理多类分类问题。Ji 等^[13]在 ϵ 平滑支持向量回归的基础上进一步提出了基于增益光滑支持向量回归机的 Takagi-Sugeno 模糊模型。

虽然 SSVM 取得了优于 SVM 的性能,但是它对抗噪声较为敏感。为了提高 SSVM 的抗噪声能力,本文利用 M-estimator 代替 SSVM 中基于 L2 范数的正则化项,并利用半二次最小化求解相应的优化问题。半二次最小化的乘法和加法形式由 Geman 等提出^[14,15]。Nikolova 和 Ng 从理论和实验两方面验证了半二次最小化方法在运行时间和迭代次数上均优于其他相关的优化方法,并证明了半二次最小化方法具有较好的收敛性^[16]。He 等^[17]提出了基于半二次最小化的鲁棒稀疏表示方法,并利用半二次框架解决了鲁棒稀疏表示问题。

与利用 Newton-Armijo 算法求解的 SSVM 相比,所提的鲁棒平滑支持向量机(RSSVM)具有以下改进之处:

(1) RSSVM 利用 M-estimator 代替 SSVM 中基于 L2 范数的正则化项,避免了 L2 范数对噪声的敏感性,提高了抗噪声能力;

收稿日期:2014-09-10 返修日期:2014-12-15 本文受国家自然科学基金项目(60903089, 61473111),河北省自然科学基金项目(F2013201060),河北大学基金项目(3504020)资助。

胡金扣(1988—),女,硕士,主要研究方向为模式识别、机器学习,E-mail:hukouzilove@163.com;邢红杰(1976—),男,博士,教授,主要研究方向为模式识别、机器学习,E-mail:hjxing@hbu.edu.cn(通信作者)。

(2) RSSVM 利用半二次最小化方法对所提模型的优化问题进行求解,当数据集比较大时,半二次最小化方法的运行效率优于 Newton-Armijo 算法。

本文第 2 节简要地回顾了传统的 SSVM 方法;第 3 节详细阐述了所提出的 RSSVM 方法;第 4 节通过实验验证了所提方法在人工数据集和 UCI 标准数据集上能够取得优于 SSVM 的性能;最后给出了概括性总结。

2 SSVM

在 n 维欧氏空间中, A 表示 $m \times n$ 阶矩阵, 矩阵 A 的第 i 行元素表示第 i 个样本的特征, D 为 m 阶对角矩阵, 主对角上的元素为 1 或 -1。SVM 的优化问题如下:

$$\begin{aligned} \min_{\xi, \alpha, b} \quad & \frac{c}{2} \xi' \xi + \frac{1}{2} (\alpha' \alpha + b^2) \\ \text{s. t.} \quad & D(K(A, A') D \alpha - e b) + \xi \geq e \\ & \xi \geq 0 \end{aligned} \quad (1)$$

$C > 0$ 为折衷参数, 它控制对错分样本的惩罚程度, ξ 为松弛变量。 α 是分类超平面的法向量。 $K(A, A')$ 是 $R^{m \times n} \times R^{n \times m}$ 到 $R^{m \times m}$ 的核函数。根据优化问题(1), 可以得出非线性分类超平面的表示如下:

$$K(x', A') D \alpha = b \quad (2)$$

根据式(1), ξ 可以改写为

$$\xi = (e - D(K(A, A') D \alpha - e b))_+ \quad (3)$$

将式(3)代入式(1), 即可得到 SVM 的无约束最优化问题:

$$\min_{\alpha, b} \quad \frac{c}{2} \| (e - D(K(A, A') D \alpha - e b))_+ \|_2^2 + \frac{1}{2} (\alpha' \alpha + b^2) \quad (4)$$

上述问题是严格的无约束最优化问题, 具有唯一解。然而, 该目标函数不光滑, 文献[10]应用光滑策略将非光滑的无约束问题式(4)转化为光滑无约束问题。引入函数 $p(\Delta, \lambda)$ 近似替代 Δ_+ , 当 $\lambda \rightarrow \infty$ 时, $p(\Delta, \lambda)$ 收敛于 Δ_+ 。 $p(\Delta, \lambda)$ 的表达式如下:

$$p(\Delta, \lambda) = \Delta + \frac{1}{\lambda} \ln(1 + e^{-\Delta \lambda}), \lambda > 0 \quad (5)$$

因此, 无约束优化问题(4)可以近似于光滑无约束的最优化问题, 从而得到 RSSVM 的优化问题为:

$$\min_{\alpha, b} \quad \frac{c}{2} \| p(e - D(K(A, A') D \alpha - e b), \lambda) \|_2^2 + \frac{1}{2} (\alpha' \alpha + b^2) \quad (6)$$

式(6)可进一步表示为:

$$\min_u \quad \frac{c}{2} \| p(e - DKD_1 u, \lambda) \|_2^2 + \frac{1}{2} \| u \|_2^2 \quad (7)$$

其中, $K = (K(A, A'), -1)$, $D_1 = \begin{pmatrix} 1 & 0 \\ 0 & D \end{pmatrix}$, $u = (b, \alpha)$ 。

3 鲁棒的光滑支持向量机(RSSVM)

为了提高 SSVM 的抗噪声能力, 提出了鲁棒的 SSVM (RSSVM), 即利用 M-estimator^[18] 代替 SSVM 中基于 L2 范数的正则化项, 并利用半二次最小化方法^[19] 求解相应的优化问题。

M-estimators 可用势函数 Φ 表示, 并代替式(7)中基于 L2 范数的正则化项, 则有

$$L(u) = \min_u \quad \frac{c}{2} \| p(e - DKD_1 u, \lambda) \|_2^2 + \frac{1}{2} \Phi(u) \quad (8)$$

本文利用半二次最小化方法求解优化问题(8)。

命题 1 根据共轭凸函数理论^[20, 21], 对于势函数 Φ , 存在一个共轭凸函数 φ , 使得:

$$\Phi(u) = \min_v \{ (u - v)^2 - \varphi(v) \} \quad (9)$$

其中, $v = (v_0, v_1, \dots, v_m)^T$ 为辅助变量。

当选取不同的 M-estimators 即 Φ 函数时, 式(9)取得最小值的 v 也不同, 如表 1 所列。

表 1 不同 M-estimators(势函数 Φ) 对应的最小值函数 v

M-estimators	势函数 Φ	$v(u)$
L1-L2	$\Phi(u) = \sqrt{\theta + u^2}$	$v(u) = u - u / \sqrt{\theta + u^2}$
Fair	$\Phi(u) = u / \theta - \log(1 + u / \theta)$	$v(u) = u - u / \theta(1 + u)$
Log-cosh	$\Phi(u) = \log(\cosh(\theta u))$	$v(u) = u - \theta \tanh(\theta u)$
Welsch($\sigma^2 = 0.5$)	$\Phi(u) = 1 - \exp(-u^2 / \sigma^2)$	$v(u) = u - u \exp(-u^2 / \sigma^2)$

注: θ 是一个常量。

根据命题 1, 最优化问题(8)可改写为:

$$L(u, v) = \min_{u, v} \quad \frac{c}{2} \| p(e - DKD_1 u, \lambda) \|_2^2 + \frac{1}{2} [(u - v)^2 - \varphi(v)] \quad (10)$$

求解优化问题(10), 需要对目标函数 L 关于 u_i 求导, 令导数等于 0, 结合式(5), 可得

$$\frac{\partial L}{\partial u_j} = c \sum_{i=0}^l (\Delta_i)_+ \times \frac{\partial \Delta_i}{\partial u_j} \times \frac{e^{\Delta_i}}{1 + e^{\Delta_i}} + (u_j - v_j) \quad (11)$$

其中, $\Delta_i = 1 - y_i (\sum_{j=0}^l y_j u_j k(x_j, x_i))$ 。

当 $\lambda \rightarrow \infty$ 时, $p(\Delta, \lambda)$ 收敛于 Δ_+ , 所以当 $\lambda \rightarrow \infty$ 时,

$$\frac{\partial L}{\partial u_j} = \begin{cases} \sum_{i=0}^l \Delta_i \frac{\partial \Delta_i}{\partial u_j} + (u_j - v_j), & \Delta_i \geq 0 \\ u_j - v_j, & \Delta_i < 0 \end{cases} \quad (12)$$

$$\text{令 } h_j = \begin{cases} \sum_{i=0}^l \Delta_i \frac{\partial \Delta_i}{\partial u_j}, & \Delta_i \geq 0 \\ 0, & \Delta_i < 0 \end{cases}, \text{ 可得 } \frac{\partial L}{\partial u_j} = h_j + (u_j - v_j), \text{ 令}$$

导数等于 0, 得到 u 的表达式为: $u = v - h$ 。

因此, 可通过下面的迭代过程^[22] 求解最优化问题(10):

$$v^r = u^{r-1} - \frac{u^{r-1}}{\sqrt{\theta + (u^{r-1})^2}} \quad (13)$$

$$u^r = v^r - h \quad (14)$$

其中, τ 表示第 τ 次迭代, $h = (h_0, h_1, \dots, h_m)^T$, v 由势函数 Φ 决定, 本文选取 $\Phi(u) = \sqrt{\theta + u^2}$, 则使势函数 Φ 取得最小值的 $v(u) = u - u / \sqrt{\theta + u^2}$, u 通过对式(10)求导得到。

所提 RSSVM 的构造过程如算法 1 所示。

算法 1 RSSVM 算法

输入: 输入矩阵 $A(m \times n)$, 目标矩阵 $D(m \times m)$

输出: 最优的权重向量 u

初始化: 惩罚参数 c , 最大迭代次数 N , 终止值 ϵ

步骤 1: 随机初始化权重向量 u

步骤 2: 计算核矩阵 K

步骤 3: 更新权重向量 u

For $= 1, \dots, N$

根据式(13), 更新辅助变量 v

根据式(14), 更新权重向量 u

判断终止条件: 当 $|L^r - L^{r-1}| \leq \epsilon$ 时终止, 否则继续更新。

在 RSSVM 算法中, 第 2 步的计算复杂度为 $m^2 + 2m$, m

为样本个数。对于第 3 步,根据式(14)可以得到,在每次迭代过程中,计算权重向量的复杂度为 $m(m+1)+2m(m+1)^2+m^3+m^2+2m+1$ 。因此,在步骤 3 中总的计算复杂度为 $O(N[3m^3+6m^2+5m+1]+2m^3+3m^2+m)$, N 为迭代次数。则算法 1 的计算复杂度是 $O((3N+2)m^3)$ 。SSVM 的计算复杂度是 $O(4Nm^3)$,而 N 的取值范围一般是 $5 \leq N \leq 8$,因此,与 SSVM 相比,RSSVM 的计算复杂度并未增加。

4 实验结果

为了验证所提 RSSVM 的抗噪声能力,本节将它与 SSVM 进行了比较。在实验比较中,使用了 2 个人工数据集和 10 个标准数据集。

4.1 人工数据集

本节在 2 个人工数据集上验证了所提方法的有效性。2 个人工数据集的描述如下。

Two-Moon:该数据集含有 10 个二维样本。令 $z \sim U(0, \pi)$,则 100 个正类样本由函数 $\begin{cases} x_1 = \cos z \\ x_2 = \sin z \end{cases}$ 生成,剩余的 100 个负类样本由函数 $\begin{cases} x_1 = 1 + \cos z \\ x_2 = \frac{1}{2} - \sin z \end{cases}$ 产生。为了验证 RSSVM 的抗噪声能力,从每类样本中选取 20% 的样本,改变它们的类别标号,即由 -1 改为 +1,由 +1 改为 -1。

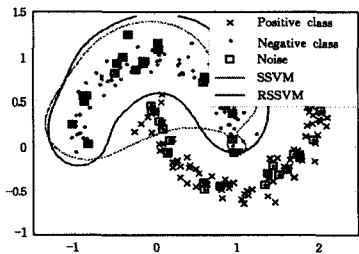
Ripley: Ripley 人工数据集^[23]由两个高斯分布混合生成。其训练集含有 250 个二维样本,测试集含有 1000 个样本,噪声的生成方法与 Two-Moon 相同。

SSVM 和 RSSVM 的核函数均采用高斯核函数,两种方法的参数均在训练集上采用五折交叉验证方法进行选取。两种方法在 2 个人工数据集上的参数设置如表 2 所列。

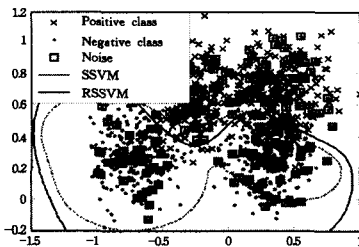
表 2 SSVM 和 RSSVM 在 2 个人工数据集上的参数设置

数据集	SSVM		RSSVM		
	c	σ	c	θ	σ
Two-Moon	0.001	0.005	0.001	10^{-3}	0.25
Ripley	0.02	0.01	0.006	10^{-2}	0.25

注:c 为惩罚参数, θ 为大于 0 的参数, σ 为宽度参数。



(a) Two-Moon



(b) Ripley

图 1 SSVM 和 RSSVM 在 2 个人工数据集上的分类效果

SSVM 和 RSSVM 在 Two-Moon 数据集上的分类效果如图 1(a) 所示,它们在该数据集上相应的分类准确率分别为 86% 和 99%。在 Ripley 测试集上的分类效果如图 1(b) 所示,它们在测试集上的分类准确率分别为 80.4% 和 85.1%。从图 1 即可看出,在两个人工数据集上,RSSVM 比 SSVM 具有更优的泛化性能。

4.2 标准数据集

为了进一步验证 RSSVM 的有效性,下面在标准数据集上对 SSVM 和 RSSVM 进行了测试。所使用的 10 个标准数据集均取自 UCI 机器学习数据库^[24],对数据的描述如表 3 所列。对于每个数据集,从所有样本中随机选取 50% 作为训练样本,剩余的 50% 作为测试样本,两种方法在每个数据集上均重复 100 次实验,分别使用 100 次实验所得平均训练准确率和平均测试准确率作为最终的训练准确率和测试准确率。

表 3 标准数据集的相关信息

数据集	样本数	特征数
Banana	5300	2
Banknote authentication	1372	4
Breast Cancer	263	9
Checkdata	345	7
Curldata	35	25
Fertility	100	9
Titanic	24	3
Transfusion	748	4
Planning relax	182	12
Wpbc	194	33

在以下的实验中,SSVM 和 RSSVM 的核函数亦采用高斯核函数,两种方法的最优参数亦使用五折交叉验证方法选取,所得参数值概括在表 4 中。两种方法的平均训练准确率及相应的标准差如表 5 所列,平均测试准确率及相应的标准差则概括在表 6 中。为了验证两种方法具有显著性差异,利用 SSVM 和 RSSVM 在每个数据集的 100 个训练准确率和 100 个测试准确率,使用配对 t 检验进行显著性检验,检验结果分别概括在表 5 和表 6 中。

表 4 SSVM 和 RSSVM 在标准数据集上的参数设置

数据集	SSVM		RSSVM		
	c	σ	c	θ	σ
Banana	0.05	0.001	0.001	0.01	0.01
Banknote Authentication	0.02	0.01	0.005	0.001	0.25
Breast Cancer	0.001	0.001	0.01	0.005	0.01
Checkdata	0.25	0.05	0.025	0.01	0.025
Curldata	0.1	0.25	0.005	0.001	0.75
Fertility	0.005	0.1	0.05	0.0001	0.05
Titanic	0.025	0.01	0.075	0.5	0.0025
Transfusion	0.0001	0.01	0.025	0.05	0.1
Planning relax	0.001	0.05	0.05	0.1	0.05
Wpbc	0.001	0.2	0.0025	0.001	0.005

从表 5、表 6 的结果可以发现,本文所提 RSSVM 在所有标准数据集上的分类性能均优于 SSVM。具体而言:

(1) 在所有标准数据集上,RSSVM 的平均测试准确率均高于 SSVM,取得了更优的泛化性能。

(2) 从实验结果中的标准差数据可以看到,所提出的 RSSVM 方法具有更优的稳定性。

(3) 在所有标准数据集上,通过配对 t 检验的结果可以发现,RSSVM 与 SSVM 具有显著性差异。

表5 SSVM 和 RSSVM 在标准数据集上的训练准确率及标准差

数据集	SSVM	RSSVM
	训练准确率	训练准确率和 p-value
Banana	0.8856±0.0504	0.9052±0.0448 2.9133e-5
Banknote Authentication	0.9914±0.0032	0.9955±0.0024 6.5471e-9
Breast Cancer	0.7734±0.0247	0.7846±0.0139 0.1954e-10
Checkdata	0.9450±0.1668	0.9540±0.0136 0.0469
Curieldata	0.9021±0.0452	0.9376±0.0154 7.4382e-5
Fertility	0.8875±0.0559	0.9400±0.0158 4.9516e-8
Titanic	0.7806±0.0512	0.8413±0.0615 7.5364
Transfusion	0.7737±0.0072	0.8315±0.0209 0.6153e-2
Planning relax	0.7180±0.0142	0.7247±0.0035 3.7142e-13
Wpbc	0.8298±0.0064	0.8412±0.0129 2.7710e-4

表6 SSVM 和 RSSVM 在标准数据集上的测试准确率及标准差

数据集	SSVM	RSSVM
	测试准确率	测试准确率和 p-value
Banana	0.8702±0.0533	0.8898±0.0053 9.5582e-5
Banknote Authentication	0.9859±0.0053	0.9888±0.0030 0.4981
Breast Cancer	0.7151±0.0469	0.7299±0.0456 4.6227e-6
Checkdata	0.9130±0.0404	0.9216±0.0115 -2.0720e-5
Curieldata	0.7553±0.1076	0.7647±0.0015 5.0479e-15
Fertility	0.8676±0.0273	0.8758±0.0082 8.3782e-4
Titanic	0.7668±0.0366	0.7737±0.0119 0.0747
Transfusion	0.7720±0.0092	0.7731±0.0083 0.2217
Planning relax	0.6958±0.0308	0.7143±0.0008 3.2233e-8
Wpbc	0.7683±0.0481	0.7797±0.0130 0.0139

结束语 本文在 SSVM 基础上提出了鲁棒的 SSVM,即 RSSVM。所提方法结合了 M-estimator 和半二次最小化优化方法的优点,主要体现在两方面:(1)M-estimator 的使用减弱了 L2 范数正则化项对噪声的敏感性,提高了抗噪声能力。(2)在保持或优于 SSVM 分类性能的前提下,利用半二次最小化方法代替 Newton-Armijo,提高了优化问题的求解效率。实验结果表明,与 SSVM 相比,RSSVM 在对含有噪声的样本进行分类时,具有更优的泛化性能。

参考文献

- [1] Burges C. A tutorial on support vector machines for pattern recognition[J]. Data Mining and Knowledge Discovery, 1998, 2(2):121-167
- [2] Vapnik V N. The Nature of Statistical Learning Theory [M]. New York: Springer-Verlag, 1995
- [3] Vapnik V N. Statistical Learning Theory [M]. New York: Wiley, 1998
- [4] Tsai C F. Training support vector machines based on stacked generalization for image classification [J]. Neurocomputing, 2005, 64(2):497-503
- [5] Zhang Y Q, Shen D G. Design efficient support vector machine for fast classification[J]. Pattern Recognition, 2005, 38(1):157-161
- [6] Vapnik V N, Golowich S, Smola A. Support vector method for function approximation, regression estimation and signal processing[M]//Advances in Neural Information Processing Systems, 1997:281-287
- [7] Tsoing J J. Hybrid approach of selecting hyper-parameters of support vector machine for regression[J]. IEEE Transactions on Systems, Man, and Cybernetics, Part B: Cybernetics, 2006, 36(3):699-709
- [8] Celikyilmaz A, Türksen I B. Fuzzy functions with support vector machines[J]. Information Sciences, 2007, 177(23):5163-5177
- [9] Lee Y J, Mangasarian O L. SSVM: A smooth support vector machine for classification[J]. Computational Optimization and Applications, 2001, 20(1):5-22
- [10] Lin Y. Support vector machine and the Bayes rule in classification[J]. Data Mining and Knowledge Discovery, 2002, 6(3):259-275
- [11] Lee Y J, Hsieh W F, Huang C M. ϵ -SSVR: A Smooth Support Vector Machine for ϵ -Insensitive Regression[J]. IEEE Transactions on Knowledge and Data Engineering, 2005, 17(5):678-685
- [12] Chang C C, Chien L J, Lee Y J. A novel framework for multi-class classification via ternary smooth support vector machines [J]. Pattern Recognition, 2011, 44(6):1235-1244
- [13] Ji R, Yang Y, Zhang W. Incremental smooth support vector regression for Takagi-Sugeno fuzzy modeling [J]. Neurocomputing, 2014, 123:281-291
- [14] Geman D, Reynolds G. Constrained restoration and the recovery of discontinuities [J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 1992, 14(3):367-383
- [15] Geman D, Yang C. Nonlinear imagerecovery with half-quadratic regularization [J]. IEEE Transactions on Image Processing, 1995, 5(7):932-946
- [16] Nikolova M, Ng M. Analysis of half-quadratic minimization methods for signal and image recovery[J]. SIAM Journal on Scientific Computing, 2005, 27(3):937-966
- [17] He R, Zheng W, Hu B. Maximum correntropy criterion for robust face recognition[J]. IEEE Transactions on Pattern Analysis and Machine Intelligence, 2011, 33(8):1561-1576
- [18] He R, Zheng W. A regularized correntropy framework for robust pattern recognition[J]. Neural Computation, 2011, 23(8):2074-2100
- [19] He R, Hu B. Robust Principal Component Analysis Based on Maximum Correntropy Criterion[J]. IEEE Transactions on Image Processing, 2011, 20(6):1485-1494
- [20] Mackay D J C. A practical Bayesian framework for backpropagation networks[J]. Neural Computation, 1992, 4(3):448-472
- [21] Boyd S, Vandenberghe L. Convex Optimization [M]. Cambridge: Cambridge University Press, 2004
- [22] Yuan X, Hu B. Robust Feature Extraction via Information Theoretic Learning[C]//Proceedings of the 26th Annual International Conference on Machine Learning, 2009:1193-1200
- [23] Ripley B D. Pattern Recognition and Neural Networks [M]. Cambridge: Cambridge University Press, 1996
- [24] Frank A, Asuncion A. UCI machine learning repository [EB/OL]. <http://archive.ics.uci.edu/ml>