

结构化加权最小二乘支持向量机

鲁淑霞 田如娜

(河北大学数学与计算机学院 保定 071002)

摘要 针对最小二乘支持向量机(LSSVM)没有考虑样例本身的结构信息和对异常点敏感,提出了一种新的分类器——结构化加权最小二乘支持向量机(SWLSSVM),SWLSSVM通过在目标函数中引入协方差矩阵考虑了样例的结构信息;为了减少异常点的影响,其根据本类样本点到该类中心的距离对误差项进行加权。实验表明,SWLSSVM与LSSVM和SVM相比具有更好的分类和泛化性能。

关键词 最小二乘支持向量机,结构化,权,协方差矩阵

中图法分类号 TP181 文献标识码 A

Structural Weighted Least Squares Support Vector Machine Classifier

LU Shu-xia TIAN Ru-na

(College of Mathematics and Computer Science, Hebei University, Baoding 071002, China)

Abstract The structure information in data has not been exploited in the Least Squares Support Vector Machine Classifier (LSSVM) and the LSSVM is sensitive to the outliers. Focused on the above issues of the LSSVM, this paper proposed a new classifier—structural weighted least squares support vector machine (SWLSSVM). The structure information is considered by incorporating the covariance matrix into the objective function, and in order to reduce sensitive to the outliers, according to difference of the distances from different types of samples to the center of the sample, the different weights are assigned to the different training samples in the error term of objective function. The experimental results show that the SWLSSVM is more superior to the LSSVM and the SVM in classification and generalization performances.

Keywords Least squares support vector machine, Structure, Weight, Covariance matrix

1 引言

Suykens^[1]于1999年提出了最小二乘支持向量机(LSSVM),它将支持向量机(SVM)中不等式约束问题转化成等式约束问题,通过求解方程组来获得最优解,这样简化了计算,加快了计算速度;LSSVM被广泛地应用到很多领域,如基于在线LSSVM实时预测船舶摇荡^[2],基于LSSVM方法解决核电厂环境辐射监测传感器网络数据的丢失^[3]等。但是LSSVM对异常点敏感,为了减轻LSSVM对异常点的敏感,对样本点进行加权,如基于密度加权^[4],本文采用距离加权,根据样本点到该样本中心的距离不同得到不同的权值,这种加权使得LSSVM在一定程度上解决了这个问题。

SVM更多考虑的是如何使得两类数据间隔最大,而没有考虑到类内数据结构信息,然而现实问题中,每个数据都隐含着一些信息;最近提出了一些把结构信息考虑进去的算法,这类算法主要分为两大类,一类是基于流行的分类器,如LapSVM^[5];另一类是基于簇假设的分类器,如最大间隔分类器(SLMM)^[6]、结构化支持向量机(SSVM)^[7]、新的半监督分类方法基于适当的聚簇假设^[8]、最大最小间隔分类器(M⁴)^[9]、结构正则化支持向量机-结构最大间隔分类器的框架^[10]等

等;这些算法的实验表明,把样本数据本身内部隐含的结构信息考虑进去对我们设计一个好的分类器起着至关重要的作用,使用更多的数据信息可以提高分类器的泛化能力。

为了提高泛化能力和减少异常点的影响,本文提出了一种基于距离加权的结构化最小二乘支持向量机,这种加权就是根据本类样本点到该类中心距离,对不同的距离给予不同的权值,该算法通过在目标函数中引入协方差矩阵来考虑数据的分布情况;实验结果表明该算法提高了分类器的泛化能力。

本文第1节是引言;第2节介绍最小二乘支持向量机;第3节提出结构化加权最小二乘支持向量机;第4节通过实验来说明该方法的优越性;最后是结论。

2 最小二乘支持向量机

设训练样例为 $\{x_i, y_i\}_{i=1}^N \in R^n \times \{\pm 1\}$; $\phi(\cdot)$ 是从 R^n 空间映射到高维空间的特征映射。

优化问题:

$$\begin{aligned} \min_{w, \xi} J(w, \xi) &= \frac{1}{2} w \cdot w + \gamma \sum_{i=1}^N \xi_i \\ \text{s. t. } y_i [\phi(x_i) \cdot w + b] &= 1 - \xi_i \end{aligned} \quad (1)$$

到稿日期:2013-05-21 返修日期:2013-07-95 本文受国家自然科学基金(61170040),河北省自然科学基金(F2011201063)资助。

鲁淑霞(1966—),女,博士,教授,CCF会员,主要研究领域为机器学习、支持向量机,E-mail:cmclusx@126.com(通信作者);田如娜(1988—),女,硕士生,主要研究领域为机器学习、支持向量机。

通过拉格朗日函数,利用 KKT 条件求解式(1)得优化问题:

$$L(w, b, \xi; \alpha) = \frac{1}{2} w \cdot w + \gamma \sum_{i=1}^N \xi_i - \sum_{i=1}^N \alpha_i \{ y_i [\phi(x_i) \cdot w + b] - 1 + \xi_i \}$$

$$\frac{\partial L}{\partial w} = 0 \rightarrow w = \sum_{i=1}^N \alpha_i y_i \phi(x_i)$$

$$\frac{\partial L}{\partial b} = 0 \rightarrow \sum_{i=1}^N \alpha_i y_i = 0$$

$$\frac{\partial L}{\partial \xi_i} = 0 \rightarrow \alpha_i = \gamma \xi_i$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \rightarrow y_i [\phi(x_i) \cdot w + b] - 1 + \xi_i = 0, i=1, \dots, N$$

上述问题可转化为求解如下方程组:

$$\begin{pmatrix} 0 & -Y^T \\ Y & \Omega + \gamma^{-1} I \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ \bar{1} \end{pmatrix} \quad (2)$$

其中

$$Y = [y_1, \dots, y_N], \bar{1} = [1, \dots, 1], \alpha = [\alpha_1, \dots, \alpha_N],$$

$$\Omega = ZZ^T, Z = [\phi(x_1) y_1, \dots, \phi(x_N) y_N]$$

最后求解出 α 和 b , 得到判别函数:

$$f(x) = \text{sgn} \left(\sum_{i=1}^N \alpha_i K(x_i, x) + b \right) \quad (4)$$

3 结构化加权最小二乘支持向量机

3.1 距离加权定义

SWLSSVM 主要是为了获得过类中心的超平面,离中心比较近的点对该超平面影响比较大,那些离类中心比较远的点和异常点对超平面的影响比较小;显然离类中心距离的远近对超平面的影响是不同的,对离类中心越近的样例给予越大的权值,反之,则给越小的权值;这样可以突出样例的重要性和减少异常点的影响;因此权值定义如下:

$$s_j = 1 - |m_i - x_j| / (r_i + \delta), \text{ if } y_j = i \quad (5)$$

$$\begin{bmatrix} 0 & & & & 1 \\ 1 & \varphi(x_1)(I + \lambda \Sigma)^{-1} \varphi(x_1) + \frac{1}{s_1 \gamma} & \dots & & \varphi(x_1)(I + \lambda \Sigma)^{-1} \varphi(x_N) \\ \vdots & \vdots & \ddots & & \vdots \\ 1 & \varphi(x_N)(I + \lambda \Sigma)^{-1} \varphi(x_1) & \dots & & \varphi(x_N)(I + \lambda \Sigma)^{-1} \varphi(x_N) + \frac{1}{s_N \gamma} \end{bmatrix} \times \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_N \end{bmatrix} \quad (9)$$

$$\begin{bmatrix} 0 & & & & 1 \\ 1 & K_{11} - \lambda K_1 P(P + \lambda PKP)^{-1} K_1 + \frac{1}{s_1 \gamma} & \dots & & K_{1N} - \lambda K_1 P(P + \lambda PKP)^{-1} K_N \\ \vdots & \vdots & \ddots & & \vdots \\ 1 & K_{M1} - \lambda K_N P(P + \lambda PKP)^{-1} K_1 & \dots & & K_{NN} - \lambda K_N P(P + \lambda PKP)^{-1} K_N + \frac{1}{s_N \gamma} \end{bmatrix} \times \begin{bmatrix} b \\ \alpha_1 \\ \vdots \\ \alpha_N \end{bmatrix} = \begin{bmatrix} 0 \\ y_1 \\ \vdots \\ y_N \end{bmatrix} \quad (10)$$

其中

$$P = \begin{pmatrix} M_1 - Q_1^T Q_1 & 0 \\ 0^T & M_2 - Q_2^T Q_2 \end{pmatrix},$$

$$M_1 = \begin{pmatrix} \frac{1}{n_1} & 0 & \dots & 0 \\ 0 & \frac{1}{n_1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{n_1} \end{pmatrix}_{(n_1 * n_1)},$$

其中

$$r_i = \max_{(x_j, y_j = i)} |m_i - x_j| \quad (6)$$

式中, m_i 为第 i 类的均值, $i=1, -1$; r_i 为第 i 类的半径, δ 是一个很小的正数,目的是为了避开 $s_i=0$, 离本类样本中心的距离比较近的样本就给它一个相对比较大的权重,反之就给它一个较小的权值。

3.2 结构化加权的最小二乘支持向量机(SWLSSVM)

SWLSSVM 通过引入协方差矩阵来考虑样例的结构信息,并且通过对误差项进行距离加权来减少异常点的影响,从而提高泛化能力。

SWLSSVM 的目标函数为:

$$\min_{w, \xi} \frac{\|w\|^2}{2} + \frac{\lambda}{2} w^T \Sigma w + \frac{\gamma}{2} \sum_{i=1}^N s_i \xi_i^2 \quad (7)$$

$$\text{s. t. } y_i [w^T \phi(x_i) + b] = 1 - \xi_i, i=1, \dots, N$$

其中, s_i 是 3.1 节中式(5)表示的权值。

通过引入拉格朗日函数,根据 KKT 条件,求解过程如下:

$$L(w, b, \alpha, \xi) = \frac{\|w\|^2}{2} + \frac{\lambda}{2} w^T \Sigma w + \frac{\gamma}{2} \sum_{i=1}^N s_i \xi_i^2 - \sum_{i=1}^N \alpha_i [y_i (w^T x_i + b) - 1 + \xi_i]$$

$$\frac{\partial L}{\partial w} = 0 \Rightarrow w + \lambda \Sigma w - \sum_{i=1}^N \alpha_i y_i \varphi(x_i) = 0 \Rightarrow w = (1 + \lambda \Sigma)^{-1} \sum_{i=1}^N \alpha_i y_i \varphi(x_i) \quad (8)$$

$$\frac{\partial L}{\partial b} = 0 \Rightarrow \sum_{i=1}^N \alpha_i y_i = 0$$

$$\frac{\partial L}{\partial \alpha_i} = 0 \Rightarrow y_i (w^T \varphi(x_i) + b) = 1 - \xi_i$$

$$\frac{\partial L}{\partial \xi_i} = 0 \Rightarrow \alpha_i = \gamma s_i \xi_i$$

问题转化为如下方程组的形式:

$$Q_1 = \begin{pmatrix} \frac{1}{n_1} & \dots & \dots & \frac{1}{n_1} \end{pmatrix}_{(1 * n_1)},$$

$$M_2 = \begin{pmatrix} \frac{1}{n_2} & 0 & \dots & 0 \\ 0 & \frac{1}{n_2} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \frac{1}{n_2} \end{pmatrix}_{(n_2 * n_2)},$$

$$Q_2 = \begin{pmatrix} \frac{1}{n_2} & \dots & \dots & \frac{1}{n_2} \end{pmatrix}_{(1 * n_2)},$$

$$0 = \begin{pmatrix} 0 & 0 & \dots & 0 \\ 0 & \vdots & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & 0 \end{pmatrix}_{(n_1 * n_2)}$$

n_1 是正类样例的个数, n_2 是负类样例的个数, K_i 指的是核矩阵 K 的第 i 列。

式(9)到式(10)的过程如下:

对于每一个协方差, 我们有下式

$$\begin{aligned} \Sigma_i^p &= \frac{1}{|C_i^p|} \sum_{\varphi(x_j) \in C_i^p} [\varphi(x_j) - \mu_i^p][\varphi(x_j) - \mu_i^p]^T \\ &= \frac{1}{|C_i^p|} T_i^p T_i^{pT} - T_i^p \bar{1}_{|C_i^p|} \bar{1}_{|C_i^p|}^T T_i^{pT} \end{aligned} \quad (11)$$

$$\begin{aligned} \Sigma^p &= \sum_{i=1}^2 \Sigma_i^p = \sum_{i=1}^2 T_i^p T_i^{pT} / |C_i^p| - T_i^p \bar{1}_{|C_i^p|} \bar{1}_{|C_i^p|}^T T_i^{pT} \\ &= [T_1^p T_2^p]^T \times \\ &\quad \begin{bmatrix} I_{|C_1^p|} / |C_1^p| - \bar{1}_{|C_1^p|} \bar{1}_{|C_1^p|}^T & \\ & I_{|C_2^p|} / |C_2^p| - \bar{1}_{|C_2^p|} \bar{1}_{|C_2^p|}^T \end{bmatrix} \\ &\quad \times \begin{bmatrix} T_1^p \\ T_2^p \end{bmatrix} \end{aligned} \quad (12)$$

$$N^p = [T_1^p \quad T_2^p], \Sigma^p = N^p \phi N^{pT},$$

$$\phi = \begin{bmatrix} I_{|C_1^p|} / |C_1^p| - \bar{1}_{|C_1^p|} \bar{1}_{|C_1^p|}^T & \\ & I_{|C_2^p|} / |C_2^p| - \bar{1}_{|C_2^p|} \bar{1}_{|C_2^p|}^T \end{bmatrix}$$

C_1^p 和 C_2^p 表示的是特征空间中的正类样例和负类样例, $|C_1^p|$ 和 $|C_2^p|$ 是正类样例和负类样例的个数; μ_i^p 和 μ_j^p 是正类样例和负类样例的均值, $\bar{1}_{|C_i^p|}$ 指的是 $|C_i^p|$ 个元素为 $\frac{1}{|C_i^p|}$ 的向量, $I_{|C_i^p|}$ 指的是 $|C_i^p| \times |C_i^p|$ 个元素的单位矩阵。 T_i^p 指的是正类样例和负类样例的子集, $i=1, 2$ 。

由 Woodbury 公式^[11] 知

$$(A + UB V)^{-1} = A^{-1} - A^{-1} U B (B + B V A^{-1} U B)^{-1} B V A^{-1} \quad (13)$$

根据式(13)可得

$$\begin{aligned} (I + \lambda \Sigma^p)^{-1} &= (I + \lambda N^p \phi N^{pT})^{-1} \\ &= I - \lambda N^p \phi (\phi + \lambda \phi N^{pT} N^p \phi)^{-1} \phi N^{pT} \end{aligned} \quad (14)$$

把式(14)代入式(9)可得到式(10)。

最后求解出 α 和 b , 得到判别函数为

$$f(x) = \text{sgn} \left[\sum_{i,j=1}^N \alpha_i y_i (K_{ij} - \lambda K_i \phi (\phi + \lambda \phi K \phi)^{-1} K_j) + b \right] \quad (15)$$

4 实验结果

本实验均在 Inter(R) Pentium(R) Dual T3400 2G 内存上进行; MATLAB 采用的是 MATLAB7.1 版本; 实验中 LSSVM 使用的是 LS-SVMlab1.8 版本, SWLSSVM 是在 LS-SVMlab1.8 版本上进行的改进; SVM 使用的是 LIBSVM 3.1^[12] 版本。

实验中, 我们均采用高斯核函数 $K(X_i, X_j) = \exp(-\frac{\|X_i - X_j\|_2^2}{\sigma^2})$, 其中高斯核参数为 σ^2 , SWLSSVM 优化问题式(7)的正则化参数 λ 是从 $\{2^{-10}, 2^{-9}, \dots, 2^9, 2^{10}\}$ 通过 10 次交叉验证的方法得到的。

为了验证本算法的优越性, 此部分将 SWLSSVM 与 LSSVM 和 SVM 进行比较, 首先在两个小的人工数据集上进

行比较, 表 1 给出了人工数据的均值与方差以及训练集和测试集的个数与属性, 表 2 给出了实验中人工数据所得到的参数, 表 3 给出了人工数据集上 3 种不同方法下所对应的测试精度。

表 1 人工数据

训练 * 测试 * 属性	类别 个数	均值	方差
人工数据 1 (400 * 200 * 2)	正类 300 个 负类 300 个	[5 5]	[6, 0; 0, 0.5]
人工数据 2 (120 * 80 * 2)	正类 100 个 负类 100 个	[5 0]	[6, 0; 0, 0.5]

表 2 人工数据参数

数据集	λ	SWLSSVM ($\gamma \sigma^2$)	LSSVM ($\gamma \sigma^2$)	SVM (σ^2 C)
人工数据 1	2^2	(1910, 12 0, 1325)	(663, 269 0, 1316)	(0.2 inf)
人工数据 2	2^{-2}	(2.596 0, 29714)	(11, 044 1023, 19)	(0.5 inf)

表 3 人工数据集的测试精度

数据集	SWLSSVM 测试精度	LSSVM 测试精度	SVM 测试精度
人工数据 1	94.5	93.5	86.5
人工数据 2	67.5	61.15	63.75

表 1 中生成的数据是互相垂直的两类数据; 从表 3 中可以看出人工数据 1 中 SWLSSVM 的测试精度均高于 LSSVM 和 SVM; 人工数据 2 中 SWLSSVM 的测试精度依旧高于 LSSVM 和 SVM, 而 SVM 低于 LSSVM。

因此可以得出 SWLSSVM 这种算法的测试精度较高, 而 LSSVM 和 SVM 在不同的数据集上的测试精度有大有小。为了进一步验证该算法的性能, 在 UCI 数据集上进行了实验, 实验中对于多类数据我们采用的是一对余的方法。表 4 给出了该数据集的描述。

表 4 UCI 数据集

数据集	训练样例个数	测试样例个数	属性个数	类别数
Credit	326	327	18	2
Cancer	341	342	13	2
Ionosphere	176	175	34	2
Iris	75	75	5	3
Liver	172	173	7	2
Sonar	150	58	61	2
Balance	312	313	5	3
Indata	250	50	11	2
Wis	284	285	32	2

表 5 给出的是 UCI 数据集实验中所用到的参数值。表 6 给出的是 UCI 数据集上的测试误差。实验结果显示对于绝大多数样例, SWLSSVM 比 LSSVM 和 SVM 的精度要高; 而对于 Statolag 这一样例来说, SWLSSVM 的测试精度低于 SVM。图 1 给出了 UCI 数据集中 Credit, Cancer, Iris, Wis, Ionosphere, Balance, Indata 这 7 个数据集的测试精度的直观图。从图中可以明显看出 SWLSSVM 的测试精度高于其他两种算法。

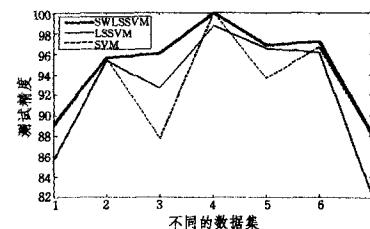


图 1 不同数据集上的测试精度

[7] 吕桃霞,刘培玉.一种基于矩阵的强关联规则生成算法[J].计算机应用研究,2011,28(4):1301-1303

[8] 张笑达,徐立臻.一种改进的基于矩阵的频繁项集挖掘算法[J].计算机技术与发展,2010,20(4):93-96

[9] 付沙,廖明华,宋丹.基于压缩矩阵方式的 Apriori 改进算法[J].微电子学与计算机,2012,29(6):28-32

[10] Khare N, Adlakha N, Pardasani K R. An Algorithm for Mining Multidimensional Association Rules using Boolean Matrix[C]//2010 International Conference in Recent Trends in Information, Telecommunication and Computing. Kochi, Kerala, 2010: 95-99

[11] Liu Hui-zhen, Dai Shang-ping, Jiang Hong. Quantitative association rules mining algorithm based on matrix[C]//2009 International Conference on Computational Intelligence and Software Engineering. Wuhan, 2009: 1-4

[12] 何丽.基于规模压缩的关联规则数据挖掘算法研究[J].计算机科学,2007,34(9):148-150

[13] 闫珍,皮德常,吴文昊.高维稀疏数据频繁项集挖掘算法的研究[J].计算机科学,2011,38(6):183-186

[14] 李瑞,康良玉,耿浩.基于数组的关联规则算法的改进[J].科学技术与工程,2008,8(21):5846-5849

(上接第 54 页)

表 5 UCI 数据集上实验参数

数据集	λ	SWLSSM ($\gamma \sigma^2$)	LSSVM($\gamma \sigma^2$)	SVM($\sigma^2 C$)
Credit	2^{-2}	(747.229 99.1677)	(8.6702 15.90)	(0.5 1000)
Cancer	2^{-5}	(747.229 99.1677)	(81.975 4.751)	(0.2 3000)
Ionosphere	2^{-5}	(14.876 7.4717)	(1.7253 21.893)	(0.5 3000)
Iris	2^0	(1.6396 1.5397)	(1.8295 1.4354)	(0.5 1000)
Liver	2^{-2}	(4627.87 339.238)	(278.8 242.25)	(5 3000)
Statolog	2^{10}	(0.801 12.4462)	(5.056 343.32)	(0.2 3000)
Balance	2^2	(3362.64 62.4015)	(331857 102)	(5 3000)
Wis	2^{-2}	(10.0714 104.853)	(180.0 412.07)	(0.2 3000)
Indata	2^{-2}	(0.4879 6.4259)	(0.32 13.835)	(0.3 5000)

表 6 UCI 数据集上 3 种方法的测试精度

数据集	SWLSSVM	LSSVM	SVM
Credit	89.156	85.627	88.888
Cancer	95.614	95.3216	95.609
Ionosphere	96.022	92.6136	87.73
Iris	100	98.7	100
Liver	69.364	66.438	68.076
Statolog	80.741	77.778	92.682
Balance	96.805	96.4856	93.617
Wis	97.193	96.1404	96.674
Indata	88	82	87.888

另外,本文从 ROC 曲线角度进一步分析,ROC 曲线是根据一系列不同的二分类方式,以真阳性率(灵敏度)为纵坐标、假阳性率(1-特异度)为横坐标绘制的曲线,ROC 曲线将灵敏度与特异性以图示方法结合在一起,可准确反映该方法特异性和敏感性的关系,曲线下的面积越大,判断价值越高;灵敏度:就是把实际为真值的判断为真值的概率。特异度:就是把实际为假值的判断为假值的概率。图 2 给出了 UCI 数据集 Ionosphere 数据的 ROC 曲线,从这个曲线中能看出 SWLSSVM 具有较高的精度。

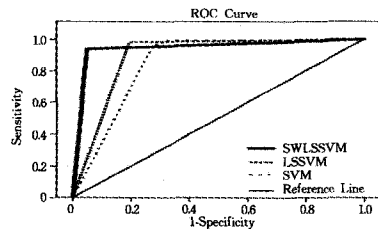


图 2 Ionosphere 数据集上的 ROC 曲线

SWLSSVM 的测试精度优于其他两种算法的原因如下:

(1) SVM 和 LSSVM 仅考虑了如何最大间隔地分离两类数据,而忽略了数据本身的结构信息。

(2) SWLSSVM 通过引入协方差矩阵考虑了样例的内部结构信息,在最大间隔的同时最小化类内紧密度;另外,又通

过引入距离加权,降低了对异常点的敏感度。

结束语 为了提高分类器的泛化能力和减少异常点的影响,提出一种新的分类器——结构化加权最小二乘支持向量机(SWLSSVM),其通过引入协方差矩阵将样例本身隐含的内部结构信息考虑到优化问题中,同时满足了最大化类间间隔和最小化类内紧密度,另外,通过加权减少了噪音点对分类器的影响,提高了分类器的泛化能力。在 UCI 数据库上的实验表明,SWLSSVM 能够更好地处理分类问题,具有较好的泛化能力。

参 考 文 献

[1] Suykens JAK, Vandewalle J. Least squares support vector machines classifiers [J]. Neural Network Letters, 1999, 19(3): 293-300

[2] Liu Sheng, Yang Zhen. Real-time online forecasting model of ship rolling motion based on chaotic online LSSVM; Industrial Engineering and Engineering Management[C]//IEEE 18th Conference. Chicago, Illinois, 2011: 1732-1736

[3] Gao Song, Tang Yao-geng. LSSVM based missing data imputation in nuclear power plant's environmental radiation monitor sensor network; Advanced Computational Intelligence [C] // IEEE fifth Conference. USA; Alaska, 2012: 479-484

[4] 王熙照,崔芳芳,鲁淑霞.密度加权近似支持向量机[J].计算机科学,2012,39(1):182-184

[5] Belkin M, Niyogi P, Sindhvani V. Manifold regularization: A geometric framework for learning from examples [J]. Journal of Machine Learning Research, 2006, 7: 2399-2434

[6] Yeung D, Wang D, Ng W, et al. Structured large margin machines; Sensitive to data distribution [J]. Mach. Learn., 2007, 68: 171-200

[7] Xue H, Chen S, Yang Q. Structural support vector machine [C]// Proc. 15th Int. Symp. Neural Netw. Beijing, 2008, 5263: 501-511

[8] Wang Y, Chen S, Zhou Z. New semi-supervised classification method based on modified cluster assumption [J]. IEEE Trans. Neural Netw. Learn. Syst., 2012, 23: 689-702

[9] Huang K, Yang H, King I, et al. Learning large margin classifiers locally and globally [C]// Proc. 21st Int. Conf. Mach. Learn. Canada; Banff, 2004: 1-8

[10] Xue H, Chen S, Yang Q. Structural Regularized Support Vector Machine: A Framework for Structural Large Margin Classifier [J]. Neural Networks, 2011, 22(4): 573-587

[11] Woodbury M A. Inverting modified matrices [M]. Princeton, NJ, Memo. Stat. Res. Group, Princeton Univ., 1950: 42

[12] Chang C C, Lin C J. LIBSVM; a library for support vector machines [CP]. <http://www.csie.ntu.edu.tw/~cjlin/libsvm>, 2013