

基于核相似性的模糊多核最小二乘支持向量机

宋菲菲^{1,2}, 何强^{1,2*}, 王恒友^{1,2}, 张长伦^{1,2}, 陈琳琳^{1,2}

¹北京建筑大学理学院, 北京

²北京建筑大学大数据建模理论与技术研究所, 北京

收稿日期: 2022年3月2日; 录用日期: 2022年3月31日; 发布日期: 2022年4月8日

摘要

最小二乘支持向量机(Least Squares Support Vector Machine, LSSVM)由于同等对待所有样例, 从而易受噪声干扰, 影响分类性能。模糊LSSVM的提出一定程度上克服了以上问题。本文给出了一种新的样例隶属度计算方法, 其在特征空间中, 利用每一样例与其他样例核相似性获得隶属度, 并将其应用于模糊多核LSSVM (Fuzzy Multi-Kernel LSSVM, FMK-LSSVM), 得到具有强鲁棒性的基于核相似性的模糊多核LSSVM。实验结果验证该方法的可行性与有效性。

关键词

最小二乘支持向量机, 隶属度, 多核学习, 核相似性

Kernel Similarity-Based Fuzzy Multi-Kernel Least Squares Support Vector Machine

Feifei Song^{1,2}, Qiang He^{1,2*}, Hengyou Wang^{1,2}, Changlun Zhang^{1,2}, Linlin Chen^{1,2}

¹School of Science, Beijing University of Civil Engineering and Architecture, Beijing

²Institute of Big Data Modelling and Technology, Beijing University of Civil Engineering and Architecture, Beijing

Received: Mar. 2nd, 2022; accepted: Mar. 31st, 2022; published: Apr. 8th, 2022

Abstract

Least squares support vector machine (LSSVM) is vulnerable to noise and affects the classification performance because it treats all samples equally. Fuzzy LSSVM overcomes the above problems by introducing membership. In this paper, we develop a new method to compute membership. In the feature space, the membership degree is obtained by using the kernel similarity be-

*通讯作者。

tween each sample and other samples, and applied to fuzzy multi-kernel LSSVM (FMK-LSSVM) to obtain a strong robust FMK-LSSVM. Experimental results verify the feasibility and effectiveness of this method.

Keywords

LSSVM, Membership, Multi-Kernel Learning, Kernel Similarity

Copyright © 2022 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

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



Open Access

1. 引言

支持向量机(Support Vector Machine, SVM)以统计学理论为基础[1] [2], 是一种性能良好的机器学习方法, 有效地解决了局部极小值点、小样本、非线性等实际问题, 目前已成功应用于模式识别、回归预测、聚类问题等诸多领域, 有力地推动了机器学习理论与应用的发展。但是 SVM 由于需要求解二次优化问题, 使其具有高时间复杂度。1999 年 Suykens 等提出最小二乘支持向量机(LSSVM) [3], 通过求解线性方程组来得到决策模型, 大大降低了算法计算量。LSSVM 利用核技巧解决非线性可分性问题, 但同时也带来了核选择难题, 为解决这一问题, 各类多核学习方法被引入 LSSVM, 即多核组合代替单核[4] [5] [6] [7], 即多核 LSSVM (Multi-Kernel LSSVM, MK-LSSVM) [8]。

由于同等对待每一样例, 包括 LSSVM 在内的各类 SVM 算法都易受噪音数据影响。模糊隶属度的引入, 得到模糊 SVM (Fuzzy SVM, FSVM) [9], 一定程度上缓解了这一问题。当然这一思想同样适用于 LSSVM, 从而有了模糊 LSSVM (Fuzzy LSSVM, F-LSSVM) [10]。对于 FSVM 与 F-LSSVM 来说, 求得样例隶属度都是至关重要的。现有各类隶属度计算方法, 都是基于对噪音的检测。一般而言, 噪音应该被赋予较小的隶属度, 从而降低其对最终决策函数的影响。有基于 Kohonen 自组织映射、模糊性指数和模糊 c-maan 算法等技术的隶属度计算方法[9]。也有学者将隶属度设置为点与其类中心之间距离的函数[11] [12] [13]。此外, 通过考虑样例的属性和决策之间的不一致性, He 等提出了一类基于模糊粗糙集的隶属度计算方法[14]。上述隶属度的引入, 一定程度上提高了 SVM 的鲁棒性。但所有隶属度计算方法都未结合 SVM 算法本身特点, 只是在样例所在原始空间中度量其相似性程度, 从而计算隶属度。然而 SVM 是在特征空间中实现对样例的分割。原始空间与特征空间的不一致可能造成传统隶属度并不能很好反映样例在特征空间中的重要性程度。本文中, 针对 MK-LSSVM, 通过其组合核获得特征空间中的样例相似性, 给出一种基于样本核相似性的隶属度计算方法。从而得到了基于核相似性的模糊 MK-LSSVM (Kernel Similarity-Based FMK-LSSVM, KS-FMKLSSVM)。

本文后续内容安排如下, 第 2 节介绍 LSSVM、MK-LSSVM 和 FMK-LSSVM 的相关知识; 基于核相似性的隶属度计算方法和 KS-FMKLSSVM 算法在第 3 节中详细说明; 第 4 节给出实验结果, 第 5 节为结论。

2. 相关知识

2.1. 最小二乘支持向量机(LSSVM)

与经典 SVM 相同, LSSVM 模型通过最优超平面解决分类问题, 其主要思想是通过核函数实现在高

维特征空间中样例的线性分割。特征空间中的最优分割超平面对应于原始空间中的非线性分割函数。

设有训练集 $\{(x_1, y_1), (x_2, y_2), \dots, (x_N, y_N)\}$ ，其中 $x_i \in R^n, y_i \in \{1, -1\}$ 。 $\phi(\cdot)$ 为某非线性映射，将样本由原空间映到高维特征空间，SVM 在特征空间中构造最优分割超平面[1]。设 ω 与 b 为特征空间中分割超平面的法向量与偏置。超平面为 $\omega^T \phi(x) + b = 0$ 。LSSVM 需要求解以下二次规划问题：

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

其中 C 为惩罚系数、 ξ_i 为松弛变量。引入拉格朗日乘子，可得对偶问题[3]：

$$\begin{pmatrix} 0 & \tilde{1}^T \\ \tilde{1} & \Omega + \frac{I}{C} \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ Y \end{pmatrix} \quad (2)$$

式中： $Y = [y_1, y_2, \dots, y_N]$ ， $\tilde{1}$ 是所有元素为 1 的列向量， I 是 N 阶单位矩阵，其中

$\Omega_{i,j} = y_i y_j \phi(x_i)^T \phi(x_j) = y_i y_j k(x_i, x_j)$ ， $k(x_i, x_j)$ 为核函数[2]。通过求解上述方程组得到 LSSVM 决策函数：

$$f(x) = \text{sign} \left[\sum_{i=1}^N \alpha_i y_i k(x_i, x) + b \right] \quad (3)$$

2.2. 多核 LSSVM

核函数度量了样本在特征空间中的相似性，但是单个核函数对样本相似性的刻画能力往往有限，核函数与核参数的选择对 LSSVM 性能有关键影响。然而，当前并不存核函数选择的通用方法，大多靠经验或试算，效果往往不理想且效率低下。作为核方法的重要成果，多核学习近年来已成为机器学习领域广大学者的研究热点[4] [5] [6] [7]。多核的意义主要体现在通过多个核函数组合可以更准确地表达数据间的相似性，同时也克服了核函数选择这一难题。多核学习成果的引入将使我们可以通过使用多个核函数，而不是一个核函数来表达样本间的相似性，即通过多个核函数的融合来得到样本间最终的相似性[15]。将多核学习引入到 LSSVM，即得到多核 LSSVM (MK-LSSVM)。有效避免了核函数及其参数选择难题。线性加权组合是最常用的多核组合方式：

$$k_\eta(x_i, x_j) = \sum_{l=1}^M \eta_l k_l(x_i, x_j), \eta_l \geq 0 \quad (4)$$

其中， $k_l, \eta_l, l = 1, 2, \dots, M$ 分别是基础核与核权重，本文采用该种加权求和的组合方式。在 LSSVM 模型中，用多核替代单核，得到 MK-LSSVM 决策函数为：

$$f(x) = \text{sign} \left[\sum_{i=1}^N \alpha_i y_i \sum_{l=1}^M \eta_l k_l(x_i, x) + b \right] \quad (5)$$

MK-LSSVM 性能很大程度上依赖于组合核，而其关键就是核权重。本文中，利用核对齐方法[16] [17] [18]，通过求解以下优化问题来得到核权重 η_l ：

$$\begin{aligned} \max_{\eta} \quad & A(K, yy^T) = \frac{\left\langle \sum_{l=1}^M \eta_l k_l, yy^T \right\rangle_F}{\sqrt{\left\langle \sum_{l=1}^M \eta_l k_l, \sum_{k=1}^M \eta_k k_k \right\rangle_F} \left\langle yy^T, yy^T \right\rangle_F} = \frac{\sum_{l=1}^M \sum_{i,j=1}^N \eta_l K_{l,i,j}^l}{N \sqrt{\sum_{l,s=1}^M \eta_l \eta_s \langle K_l, K_s \rangle_F}} \\ \text{s.t.} \quad & \eta_l \geq 0, l = 1, 2, \dots, M \end{aligned} \quad (6)$$

这里 K 是组合核矩阵, $\langle \cdot, \cdot \rangle_F$ 是两个矩阵间的 Frobenius 内积, 其中

$$\langle K_1, K_2 \rangle_F = \sum_{i=1}^N \sum_{j=1}^N k_1(x_i, x_j) k_2(x_i, x_j) \quad (7)$$

显然, A 越大, 则组合核矩阵 K 与理想核矩阵 yy^T 间的距离越小, 即组合核对数据相似性刻画能力越强[16]。为方便求解, (6)式可变形为下述优化问题[17]:

$$\begin{aligned} \max_{\eta} \quad & \sum_{m=1}^P \sum_{i,j=1}^l \eta_m k_{ij}^m - \sum_{m,s=1}^P \eta_m \eta_s \langle K_m, K_s \rangle_F - \delta \sum_{m=1}^P \eta_m^2 \\ \text{s.t.} \quad & \eta_m \geq 0, m = 1, 2, \dots, P \end{aligned} \quad (8)$$

其中, $\delta > 0$ 是正则项系数。

2.3. 模糊多核 LSSVM

LSSVM 继承了 SVM 的优点, 实现了结构风险最小化, 但也继承了对噪声敏感的缺点。为此, 引入隶属度得到模糊 LSSVM [10], 提升了 LSSVM 鲁棒性。设 s_i 为样本 x_i 的隶属度, 则训练集可表示为 $\{(x_1, y_1, s_1), (x_2, y_2, s_2), \dots, (x_N, y_N, s_N)\}$, 其中 $x_i \in R^n, y_i \in \{1, -1\}, s_i \in [0, 1]$, 模糊多核 LSSVM 的模型为:

$$\begin{aligned} \min_{\omega, b, \xi} \quad & \frac{1}{2} \sum_{l=1}^M \eta_l \omega_l^T \omega_l + \frac{1}{2} C \sum_{i=1}^N s_i \xi_i^2 \\ \text{s.t.} \quad & y_i \left[\sum_{l=1}^M \eta_l \omega_l^T \phi(x_i) + b \right] = 1 - \xi_i \\ & i = 1, 2, \dots, N; \eta_l \geq 0, l = 1, 2, \dots, M \end{aligned} \quad (9)$$

相应对偶问题为:

$$\begin{pmatrix} 0 & \tilde{\Gamma}^T \\ \tilde{\Gamma} & \Omega + \frac{I}{Cs} \end{pmatrix} \begin{pmatrix} b \\ \alpha \end{pmatrix} = \begin{pmatrix} 0 \\ Y \end{pmatrix} \quad (10)$$

(10)式中, $\Omega_{i,j} = y_i y_j \sum_{l=1}^M \eta_l k_l(x_i, x_j)$, $s = \{s_1, s_2, \dots, s_N\}$, 其他符号含义与(2)式相同。通过求解上述方程组得到 LSSVM 决策函数:

$$f(x) = \text{sign} \left[\sum_{i=1}^N \alpha_i y_i \sum_{l=1}^M \eta_l k_l(x, x_i) + b \right] \quad (11)$$

显然, 当 $s_i = 1, i = 1, 2, \dots, N$ 时, FMK-LSSVM 模型退化为 MK-LSSVM。

3. 基于核相似性的 FMK-LSSVM

标准的 LSSVM 因为平等的对待每个样例, 所以易受噪音影响而性能变差。为了缓解这种现象, 模糊隶属度被引入, 得到模糊 LSSVM, 使每一训练样例因为重要性不同而对分割超平面产生不同的影响, 从而提升算法鲁棒性。模糊隶属度的计算是模糊 LSSVM 的关键。本节通过数据间的核相似性衡量每个样本的隶属度, 为与其核相似性高的样例赋予较高的隶属度, 为噪音赋予较低的隶属度, 来降低噪音的影响。

SVM 中的核函数不仅将原空间中的问题转换为特征空间的问题, 还有效衡量了数据间的相似性。此处, 使用 FMK-LSSVM 的组合核来度量样本间相似性, 两个样例间的核函数值越大, 则表明这两个样例间的相似性越高。对于某一样例, 与其相似性大且标签一致的样例越多, 则该样例越能代表此类数据的特征, 应赋予较大的隶属度, 反之亦然。基于上述思想, 本节提出了一种基于核相似性的隶属度计算方法。主要过程如下。

首先利用组合核函数计算训练集俩俩样本之间的核函数值, 并从大到小排序。通过参数 $\alpha \in [0, 1]$ 确

定衡量相似性大小的阈值 θ ，若两样例间核函数值大于 θ ，则认为其相似性大；从而可以得到每个样例与其相似度大的同类与异类样例数，最终通过 S 型函数计算得到每个样例的隶属度值[19]。S 型函数表达式为 $f(x) = 1/(1 + e^{-\tau x})$ ，其可使正常样例与噪音的隶属度差异更大，最大程度降低噪音的干扰，其中 S 型函数中的参数 τ 用来控制这种差异性的程度。本文中，我们取 $\tau = 10$ 。隶属度函数可被写为：

$$s(x_i) = \frac{1}{1 + \exp\left(-\tau y_i \left(\frac{p_i u_i}{PU} - \frac{q_i v_i}{QV}\right)\right)} \quad (12)$$

其中 u_i, v_i 分别是与样例 x_i 类别相同且相似性达到阈值 θ 的样例个数， p_i, q_i 分别为与样例 x_i 类别相同且相似性达到阈值 θ 的核函数值的和，计算公式如下：

$$p_i = \sum_{y_j=y_i=+1, k_{ij} \geq \theta} k_{ij} \quad (13)$$

$$q_i = \sum_{y_j=y_i=-1, k_{ij} \geq \theta} k_{ij} \quad (14)$$

此外， $U = \max_i \{u_i\}$, $V = \max_i \{v_i\}$, $P = \max_i \{p_i\}$, $Q = \max_i \{q_i\}$ 。

主要计算过程如算法 1 所示：

算法 1: 基于核相似性的隶属度

输入: 训练样例 $(x_i, y_i), i = 1, 2, \dots, N$

组合核 $k_\eta(x, z) = \sum_{j=1}^M \eta_j k_j(x, z), \eta_l \geq 0$

参数 $\alpha \in [0, 1], \tau > 0$

输出: 训练样例隶属度 $s_i, i = 1, 2, \dots, N$

- 1: 训练数据归一化;
 - 2: 利用(4)计算得到合成核矩阵 K_η ;
 - 3: 将 K_η 带入(8)式, 求得核权重 $\eta_1, \eta_2, \dots, \eta_M$;
 - 4: 将核权重 $\eta_1, \eta_2, \dots, \eta_M$ 带入, 求得 K_η ;
 - 5: 取 K_η 上三角部分(不含对角线)所有元素, 按从大到小排序, 得到向量 $G = \{g_1, g_2, \dots, g_t\}, t = N(N-1)/2$;
 - 6: 令 $h = [\alpha t]$ (向下取整), 得阈值 $\theta = g_h$;
 - 7: for $i=1:N$ do
 - 8: $u_i = 0, v_i = 0, p_i = 0, q_i = 0$;
 - 9: for $i=1:N$ do
 - 10: if $k_{ij} \geq \theta \ \&\& \ y_j = y_i$
 - 11: if $y_i = 1$
 - 12: $u_i = u_i + 1, p_i = p_i + k_{ij}$;
 - 13: end if
 - 14: if $y_i = -1$
 - 15: $v_i = v_i + 1, q_i = q_i + k_{ij}$;
 - 16: end if
 - 17: end if
 - 18: end for
 - 19: 令 $U = \max_i \{u_i\}, V = \max_i \{v_i\}, P = \max_i \{p_i\}, Q = \max_i \{q_i\}$;
 - 20: for $i=1:N$ do
 - 21: 由(12)式得到所有样例隶属度 s_i ;
 - 22: end for
 - 23: end
-

将基于核相似性引入 FMK-LSSVM，即得到 KS-FMKLSSVM。下节中，将通过实验验证 KS-FMKLSSVM 的可行性与有效性。

4. 实验

论文中的数据选自 UCI 中的二分类数据集。本节中将 KS-FMKLSSVM 与其他 3 类 LSSVM 方法进行对比。分别是普通 MK-LSSVM、基于模糊粗糙集隶属度的模糊多核 LSSVM (Fuzzy Rough Set Based FMK-LSSVM, FR-FMKLSSVM)和基于欧氏距离隶属度[13]的模糊多核 LSSVM (Euclidean Distance Based FMK-LSSVM, ED-FMKLSSVM)。所选数据集的信息如表 1 所示。

Table 1. The information of UCI dataset in experiment

表 1. 实验中的 UCI 数据集信息

序号	数据集	正类样本	负类样本	属性个数
1	Abalonel	1341	2835	8
2	bupa	145	200	6
3	hepatitis	85	70	19
4	lymphography	73	75	18
5	parkinsons	147	48	22
6	sonar	111	97	60
7	Wilt	313	187	5
8	Winequality	918	681	11

实验中均是选择 70%作为训练集，30%作为测试集，重复 10 次试验求平均得到最终结果。为检验算法鲁棒性，每次训练集中随机选取 20%样例改变其标签，作为噪音放回训练集中。此外，本节中多核算法中组合核都由 7 个高斯核作为基本核组合而成，核参数分别为 $[2^{-3}, 2^{-2}, 2^{-1}, 1, 2, 4, 8]$, $\alpha = 0.5, \tau = 10$ 。各个算法测试精度与训练时间结果表 2 和表 3 所示。

Table 2. Testing accuracy (%) values of 4 algorithms

表 2. 4 种方法的测试精度(%)

数据集	MK-LSSVM	FR-FMKLSSVM	ED-FMKLSSVM	KS-FMKLSSVM
Abalonel	63.14	78.81	81.17	82.26
bupa	56.79	56.55	60.36	62.74
hepatitis	63.42	65.00	60.79	67.89
lymphography	55.00	62.22	55.83	64.17
Parkinsons	67.50	75.63	83.75	85.63
sonar	57.25	66.86	67.84	71.37
Wilt	66.75	62.17	74.28	75.83
Winequality	64.27	67.97	76.33	77.42
平均	61.77	66.90	70.04	73.41

Table 3. Training time (s) of 4 algorithms
表 3. 4 种方法的训练时间(s)

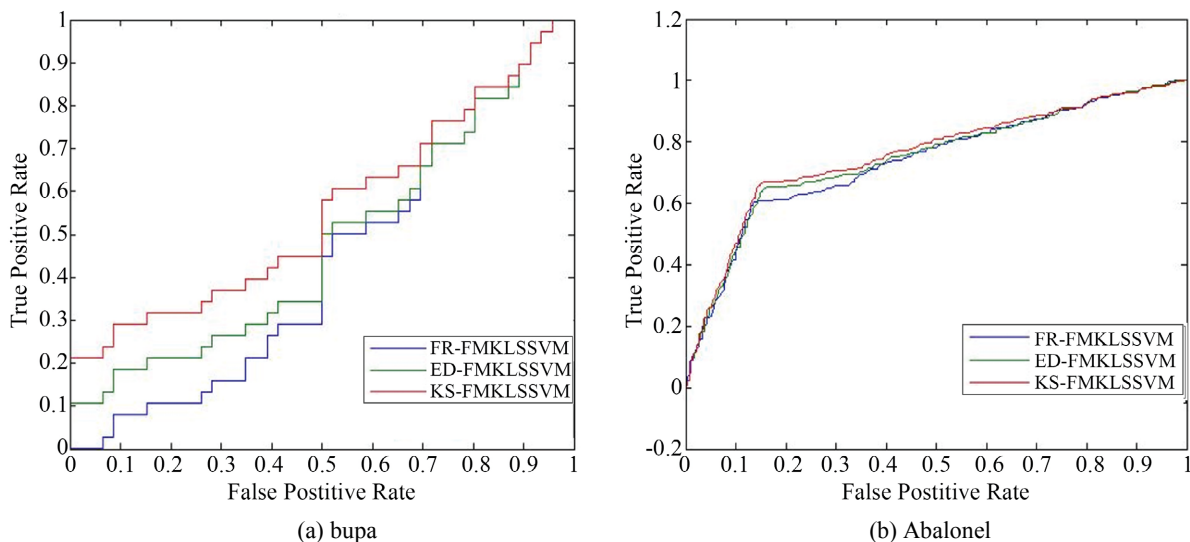
数据集	MK-LSSVM	FR-FMKLSSVM	ED-FMKLSSVM	KS-FMKLSSVM
Abalanel	66.72	97.11	94.79	119.86
bupa	0.44	0.63	0.70	0.78
hepatitis	0.34	0.52	0.53	0.57
lymphography	0.34	0.54	0.49	0.49
Parkinsons	0.47	0.67	0.63	0.69
sonar	0.95	1.01	1.32	1.09
Wilt	1.00	1.49	1.58	2.07
Winequality	9.69	14.5	14.86	18.96
平均	9.99	14.56	14.36	18.06

实验结果显示, 本文提出的 KS-FMKLSSVM 在 UCI 的 8 个数据集上, 分类精确都较其他算法有较明显提高, 而时间消耗与其他方法相比并没有显著增加, 尤其是与其他 FMK-LSSVM 算法相比, 时间支出基本相当。这也验证了本文所提隶属度计算方法的可行性与有效性。

ROC 曲线是机器学习常用的模型性能比较指标, 曲线纵横坐标分别是假正例率和真正利率。A 算法的 ROC 曲线被 B 算法的曲线“包住”, 则 B 算法性能优于前者; 如果曲线交叉, 可以根据 ROC 曲线下面积大小进行比较, 也即 AUC 值[20]。图 1 是部分数据集上 4 中算法的 ROC 曲线。从图中我们可以看出, 在各个数据集上, KS-FMKLSSVM 算法的 ROC 曲线都“包住”了其他三种算法的曲线, 即 KS-FMKLSSVM 的综合性能是最佳的, 进一步验证了本文给出的隶属度计算方法的有效性。

为进一步检验所提 KS-FMKLSSVM 的鲁棒性, 我们以 sonar 数据为例, 展示不同噪声水平下, 本文所提方法与其他方法的分类效果的变化情况, 结果如图 2 所示。

从图 2 可以看出, 随着噪音水平的升高, KS-FMKLSSVM 的性能在所有四种算法中表现最佳, 分类精度最高, 且受噪音影响相对较小, 这表明了本文提出的隶属度方法可以最大程度提高 FMK-LSSVM 的鲁棒性。



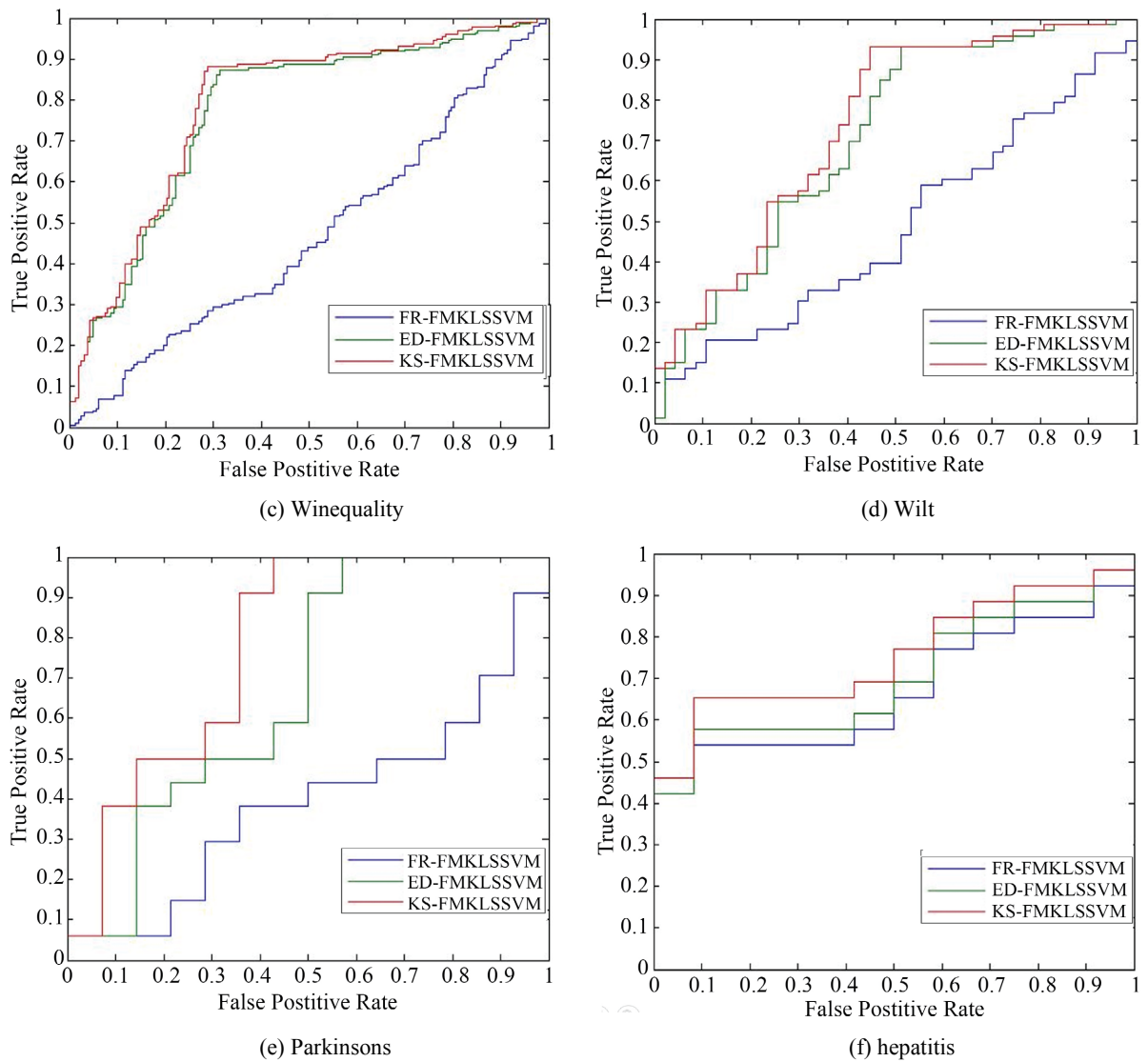


Figure 1. ROC graphs of some data sets

图 1. 部分数据集的 ROC 曲线

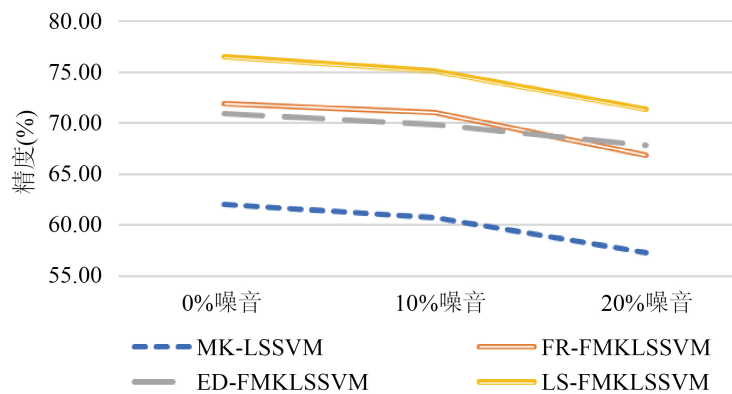


Figure 2. Comparison of robustness of 4 algorithms on sonar

图 2. Sonar 数据集上 4 种算法鲁棒性比较

5. 结论

本文针对 FMK-LSSVM, 提出一种新的隶属度计算方法。该方法以 MK-LSSVM 中的组合核函数来度量样例间的相似性程度, 通过每一样例附近与其相似的同类与异类样例比例的差值来获得该样例的隶属度, 并引入到 FMK-LSSVM。该隶属度区别于已有的各类隶属度计算方法, 其直接以各样例在 MK-LSSVM 特征空间中的相似性来获得隶属度。LSSVM 决策函数同样是在特征空间中构造产生, 从而使得该隶属度与 LSSVM 保持一致性, 仿真实验结果也验证了所提方法的有效性。

上述工作虽然取得了一定成果, 但仍存在不完善之处, 后续相关研究, 拟从以下几个方面展开:

1) 模糊隶属度算法中的阈值仍是经验值, 后续将进一步研究, 设计启发式阈值选取方法, 以进一步提高 FMK-LSSVM 的鲁棒性;

2) 可将线性加权求和的多核模型扩展到局部多核学习模型, 使最终的分类器模型进一步利用到样本的信息, 提升算法性能。

基金项目

国家自然科学基金(62072024, 61473111); 北京市教委一般项目(KM202110016001); 北京建筑大学科学研究基金(KYJJ2017017, Y19-19, Y18-11, PG2021094); 住房和城乡建设部科学技术计划北京建筑大学北京未来城市设计高精尖创新中心开放课题(NO. UDC2019033324, UDC201703332); 北京建筑大学课程建设重点培育项目“高等数学 ZDXX202008”。

参考文献

- [1] Cortes, C. and Vapnik, V. (1995) Support-Vector Networks. *Machine Learning*, **20**, 273-297. <https://doi.org/10.1007/BF00994018>
- [2] Vapnik, V. (1998) Statistical Learning Theory. Wiley, New York.
- [3] Suykens, J.A.K. and Vandewalle, J. (1999) Least Squares Support Vector Machine Classifiers. *Neural Processing Letters*, **9**, 293-300. <https://doi.org/10.1023/A:1018628609742>
- [4] Zhao, Y.G., Song, Z., Zheng, F. and Shao, L. (2018) Learning a Multiple Kernel Similarity Metric for Kinship Verification. *Information Sciences*, **430-431**, 247-260. <https://doi.org/10.1016/j.ins.2017.11.048>
- [5] Mehmet, G. and Ethem, A. (2011) Multiple Kernel Learning Algorithms. *Journal of Machine Learning Research*, **12**, 2211-2268.
- [6] Wang, T., Lu, J. and Zhang, G. (2018) Two-Stage Fuzzy Multiple Kernel Learning Based on Hilbert-Schmidt Independence Criterion. *IEEE Transactions on Fuzzy Systems*, **26**, 3703-3714. <https://doi.org/10.1109/TFUZZ.2018.2848224>
- [7] Wang, Z., Zhu, Z. and Li, D. (2020) Collaborative and Geometric Multi-Kernel Learning for Multi-Class Classification. *Pattern Recognition*, **99**, 107050. <https://doi.org/10.1016/j.patcog.2019.107050>
- [8] Gautam, C. and Tiwari, A. (2018) Localized Multiple Kernel Support Vector Data Description. *2018 IEEE International Conference on Data Mining Workshops (ICDMW)*, Singapore, 17-20 November 2018, 1514-1521. <https://doi.org/10.1109/ICDMW.2018.00224>
- [9] Lin, C. and Wang, S. (2002) Fuzzy Support Vector Machines. *IEEE Transactions on Neural Networks*, **13**, 464-471. <https://doi.org/10.1109/72.991432>
- [10] Song, F.F. and He, Q. (2021) Kernel-Target Alignment based Fuzzy Multi-Kernel Least Squares Support Vector Machine. *International Journal of Intelligent Information and Management Science*, **10**, 113-119.
- [11] Huang, H.P. and Liu, Y.H. (2001) Fuzzy Support Vector Machines for Pattern Recognition and Data Mining. *International Journal of Fuzzy Systems*, **4**, 826-835.
- [12] Liu, S.Y. and Du, Z.U. (2007) An Improved Fuzzy Support Vector Machine Method. *CAAI Transactions on Intelligent Systems*, **2**, 30-33.
- [13] 曾山, 同小军, 桑农, 等. 基于对应分析的冗余模糊 C 均值聚类算法研究. *华中科技大学学报(自然科学版)*, 2012(40), 107-111.

- [14] He, Q. and Wu, C.X. (2011) Membership Evaluation and Feature Selection for Fuzzy Support Vector Machine Based on Fuzzy Rough Sets. *Soft Computing*, **15**, 1105-1114. <https://doi.org/10.1007/s00500-010-0577-z>
- [15] De Diego, I.M., Muñoz, A. and Moguerza, J.M. (2010) Methods for the Combination of Kernel Matrices within a Support Vector Framework. *Machine Learning*, **78**, Article No. 137. <https://doi.org/10.1007/s10994-009-5135-5>
- [16] Wang, T.H., Zhao, D.Y. and Tian, S.F. (2015) An Overview of Kernel Alignment and Its Applications. *Artificial Intelligence Review*, **43**, 179-192. <https://doi.org/10.1007/s10462-012-9369-4>
- [17] Cortes, C., Mohri, M. and Rostamizadeh, A. (2012) Algorithms for Learning Kernels Based on Centered Alignment. *Journal of Machine Learning Research*, **13**, 795-828.
- [18] Wang, T., Qiu, Y. and Hua, J. (2020) Centered Kernel Alignment Inspired Fuzzy Support Vector Machine. *Fuzzy Sets and Systems*, **394**, 110-123. <https://doi.org/10.1016/j.fss.2019.09.017>
- [19] He, Q., Zhang, Q.S., Wang, H.Y. and Zhang, C.L. (2020) Local Similarity Based Fuzzy Multiple Kernel One-Class Support Vector Machine. *Complexity*, **2020**, Article ID: 8853277. <https://doi.org/10.1155/2020/8853277>
- [20] 周志华. 机器学习[M]. 北京: 清华大学出版社, 2016.