

一种基于FCOWA-ER的SVM多分类方法

刘卫兵^a, 杨 艺^b, 韩德强^a

(西安交通大学 a. 智能网络与网络安全教育部重点实验室, b 机械振动与强度国家重点实验室, 西安 710049)

摘要: 支持向量机(SVM)在处理多分类问题时,需要综合利用多个二分类SVM,以获得多分类判决结果.传统多分类拓展方法使用的是SVM的硬输出,在一定程度上造成了信息的丢失.为了更加充分地利用信息,提出一种基于证据推理-多属性决策方法的SVM多分类算法,将多分类问题视为一个多属性决策问题,使用证据推理-模糊谨慎有序加权平均方法(FCOWA-ER)实现SVM的多分类判决.实验结果表明,所提出方法可以获得更高的分类精度.

关键词: 支持向量机; DS证据理论; 多属性决策

中图分类号: TP273

文献标志码: A

A multi-class SVM based on FCOWA-ER

LIU Wei-bing^a, YANG Yi^b, HAN De-qiang^a

(a. Ministry of Education Key Lab for Intelligent Networks and Network Security, b. State Key Laboratory for Strength and Vibration of Mechanical Structures, Xi'an Jiaotong University, Xi'an 710049, China. Correspondent: HAN De-qiang, E-mail: deqhan@mail.xjtu.edu.cn)

Abstract: Multiple bi-class SVMs are used together to obtain the final decision when the support vector machine(SVM) is applied to multi-class classification problems. The conventional methods of applying the SVM to multiple classification tasks are all based on the hard output of SVM, which can bring the loss of information to some extent. Therefore, a multi-class SVM based on an evidential reasoning based multiple attribute decision approach is proposed to use more information. The multi-class classification problem is modelled as a multi-criteria decision making problem. Then a fuzzy-cautious OWA(ordered weighted averaging) approach with evidential reasoning(FCOWA-ER) is used to implement multi-class classification and obtain the final decision. The simulation results show that the method proposed has better accuracy compared with conventional methods.

Keywords: support vector machine; DS evidence theory; multi-criteria decision making

0 引言

20世纪90年代, Vapnik等^[1-2]提出了支持向量机(SVM)这种机器学习方法.基础的SVM模型为软间隔支持向量机(C-SVM),随着对SVM的深入研究,提出了一些SVM的变形算法模型,包括 ν -SVM^[3]、最小二乘支持向量机(LS-SVM)^[4-5]和模糊支持向量机(FSVM)^[6-7]等.此类算法模型都是针对二分类问题而设立的,单个的SVM只能解决二分类问题,在解决多类分类问题时,需要构造多个子SVM并组合其输出,获得一个多分类支持向量机(MSVM).因此,如何组合多个子SVM的输出以获得最终的决策结果,是一

项必要且值得研究的工作.由二分类问题拓展到多类分类问题的传统方法包括投票法^[8]、决策树法^[9]、有向无环图法^[10]等,这些方法使用的都是SVM的硬输出,损失了一部分信息,且在使用决策树法时,其分层结构会使上层的误差累积到下层,影响总体的分类性能.

本文针对上述由二分类至多分类SVM拓展中的信息损失问题,将多分类问题视为一个多属性决策问题,提出一种利用SVM软输出信息、基于多属性决策的SVM多分类方法.该方法将多分类问题视为多属性决策问题,使用基于证据推理的模糊谨慎有序加权平均(FCOWA-ER)^[11]方法进行推理和决策.该推理方

收稿日期: 2014-08-05; 修回日期: 2014-11-19.

基金项目: 国家973计划项目(2013CB329405); 国家自然科学基金项目(61104214, 61203222); 陕西省科技计划项目(2013KJXX-46); 教育部博士点基金项目(20120201120036); 中央高校基本科研业务费专项资金项目(xjj2012104, xjj2014122).

作者简介: 刘卫兵(1989-),男,博士生,从事机器学习与信息融合的研究; 韩德强(1980-),男,副教授,博士生导师,从事信息融合与模式分类等研究.

法使用信息源的软输出,取悲观态度和乐观态度的折衷,且不存在分层结构,能有效利用信息,是一种行之有效的多属性决策方法.实验结果表明,相比传统方法,所提出的方法能够更有效地实现多分类SVM.

1 传统的多分类支持向量机

1.1 SVM简介

SVM是由线性可分情况下的最优超平面发展而来的.样本线性可分时,SVM按照结构风险最小化原则寻找一个最优分类超平面.该超平面能够尽可能正确地将样本划分,并使分类间隔最大.二分类线性可分SVM如图1所示.图1中:*点和+点分别代表两种类别的训练样本,被圈标记的点决定了两类样本的边界面 H_1 和 H_3 ,这些点即为支持向量. H_1 与 H_3 之间的距离为分类间隔,当此间隔最大时,它们的中间线 H_2 即为最优超平面,设其法向量为 w ,则其分类间隔可以表示为 $2/\|w\|$.

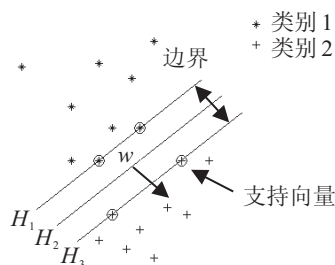


图1 支持向量机示意图

对于线性不可分的样本集,SVM通过添加惩罚因子使得训练时允许少量样本点被错分,或者使用非线性映射 $F: x \rightarrow \phi(x)$ 将样本映射到一个高维特征空间,以实现可分.以松弛变量软间隔SVM为例,其分类间隔最大化 $2/\|w\|$ (即最小化 $\|w\|$),加入惩罚因子 C 后,训练过程转化为求解如下二次优化问题:

$$\begin{aligned} \min & \frac{1}{2}w^T w + C \sum_{i=1}^m \xi_i; \\ \text{s.t.} & y_i(w^T \phi(x_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, i = 1, 2, \dots, m. \end{aligned} \quad (1)$$

其中: ξ_i 为松弛变量, $C > 0$ 为惩罚系数.优化问题的对偶问题为

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

其中 $k(x_i, x_j) = \phi^T(x_i)\phi(x_j)$ 为核函数.由此得到决策函数为

$$d(x) = \text{sgn}(w^T x + b) = \text{sgn}\left(\sum_{i=1}^m \alpha_i y_i x_i^T x + b\right). \quad (3)$$

根据决策函数的值是否大于零,可判定样本属于正类还是负类,进而解决二分类问题.若想将其应用于多分类问题,则需作进一步处理.

1.2 传统的多分类拓展方法

SVM最初是针对二分类问题提出的.目前,主要有两种思路构造多分类SVM:1)改变SVM的原始优化问题,在优化问题中考虑所有类别数据,使它能计算出多类分类决策函数^[12];2)通过构造多个两类SVM,组合其结果得到多分类SVM.前者的建模和计算非常复杂,制约了其实用性.后者是在实际应用中较常用的方法,包括“一对一”投票法^[8]、决策树法^[9]和有向无环图法^[10]等.

1.2.1 基于“一对一”投票法的多分类SVM

“一对一”投票法是一种基于二分类子SVM、使用“Max Wins”策略以获得决策输出的多分类方法.假设 k 分类问题,每次使用两类数据训练一个SVM分类器,遍历完所有组合,共构造 $k(k-1)/2$ 个子分类器.对测试样本 x 使用每个子分类器进行分类,当分类器判决 x 属于第 i 类时,第 i 类的得票数加1,遍历完所有的子分类器后,得票数最多的类别即为判定的 x 属于的类别.上述规则可以表述为

$$d(x) = \arg \max\{\text{vote}(i)\}, \quad i = 1, 2, \dots, k. \quad (4)$$

若由SVM $_l$ 可得 $x \in A_l, l = 1, 2, \dots, k(k-1)/2$,则令 $\text{vote}(i) = \text{vote}(i) + 1$.

1.2.2 基于决策树的多分类SVM

基于决策树的SVM首先将所有类别分为两个子类,再将两个子类进一步划分为两个次级子类.依此循环,直至获得一个单独的类别为止,如此便构建了一个倒立的二叉树.该方法将一个多分类问题转化为一系列的二分类问题,每个节点的分类函数采用SVM方法.该方法不存在拒绝识别的情况,但由于其分层结构,不可避免地存在误差累积的问题.因此,如何逐级划分子类至关重要.若根节点的分类性能不好,则会严重影响整体的分类性能.

1.2.3 基于有向无环图的多分类SVM

有向无环图法是由John提出的一种多分类SVM方法.该方法的训练过程类似于“一对一”投票法,对

于 k 分类问题, 构造 $k(k-1)/2$ 个子分类器. 在决策过程中, 使用一个包含 $k(k-1)/2$ 个节点的有向无环图, 每个节点代表判别第 i 类和第 j 类的 SVM. 给定一个测试样本, 从根节点开始计算相应 SVM 的决策函数值, 根据决策函数值判断转入下一层的方向, 直至得到判决结果. 实际上, 有向无环图法使用排除法对样本进行分类处理, 每经过一个节点的计算和判决, 即可排除剩余类别中最不可能的类别, 直到剩余一个类别为止. 与决策树法类似, 该方法也存在自上而下的误差累积. 因此, 如何逐层构造其分层结构将会对总体性能产生较大的影响. 与决策树法不同的是, 不是上层中每一个子 SVM 的误差都会累积到下一层, 只有当决策路线经过该节点时, 才会对总体的分类性能造成影响.

2 基于 FCOWA-ER 的 SVM 多分类方法

鉴于传统的 SVM 多分类拓展方法存在诸如信息损失和误差积累等问题, 本文提出一种基于 FCOWA-ER 的 SVM 多分类方法, 具体过程如下.

2.1 多分类 SVM 的多属性决策建模

假设有一个 k 类分类问题使用 SVM 来分类, 共需构建 $k(k-1)/2$ 个 OAO (one against one) SVM. 若将样本可能属于的类别视为多属性决策问题中的可选方案, 将每个二分类 SVM 视为一个专家 (该专家将对可选方案的某一属性进行评价), 则此 SVM 的输出可以视为该专家对可选方案的相应评价, 依此得到决策矩阵 C , 即

$$\begin{matrix} & S_1 & \cdots & S_j & \cdots & S_n \\ \begin{matrix} A_1 \\ \vdots \\ A_i \\ \vdots \\ A_k \end{matrix} & \begin{bmatrix} C_{11} & \cdots & C_{1j} & \cdots & C_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{i1} & \cdots & C_{ij} & \cdots & C_{in} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ C_{k1} & \cdots & C_{kj} & \cdots & C_{kn} \end{bmatrix} & = & C. \end{matrix} \quad (5)$$

其中: A_i 为样本可能属于的类别, S_i 为所有子 SVM, C_{ij} 为第 j 个子 SVM 对于样本是否属于第 i 类的评价值. 事实上, 决策矩阵的第 j 个子 SVM 只对与其相关的两个类别作评价, 对与其不相关的类别不予评价. 对于 SVM 对相关类别的评价, 希望得到一个 $0 \sim 1$ 之间的数值, 而非 0 或 1 的判决值. 因此, 使用 Sigmoid 函数模型^[13]将二分类 SVM 的输出转化为软输出, 其形式如下:

$$\begin{cases} p(y = 1|f) = \frac{1}{1 + \exp(Af + B)}, \\ p(y = -1|f) = 1 - p(y = 1|f). \end{cases} \quad (6)$$

该模型只需要 A 小于 0 即可满足 p 关于 f 的单调性, 并将 $f(x)$ 映射到 $[0, 1]$ 区间. 对参数 A 和 B 的优选可以通过训练数据的交叉熵最小化得到^[13].

得到决策矩阵 C , 即将多类分类问题建模为多属性决策问题, 据此可以使用多属性决策方法进行决策. 在所有多属性决策方法中, 有序加权平均方法 (OWA)^[14] 是非常重要的, 但 OWA 算子的结果与加权向量的选择关系非常大, 因此加权向量的选择是非常重要的, 在实际情况下也是困难的. FCOWA-ER 综合使用 OWA 算子、模糊隶属度函数和 DS 证据理论, 取悲观态度和乐观态度的折衷, 具有更好的鲁棒性, 因此, 本文选用 FCOWA-ER 来进行多属性决策.

2.2 基于 FCOWA-ER 的 SVM 多分类方法实现

基于证据推理的模糊谨慎有序加权平均 (FCOWA-ER)^[11] 是一种多属性决策框架下的不确定性推理方法. 该方法的基础是 DS 证据理论, 基本思想是在多属性决策框架下, 分别以最悲观和最乐观态度构造两个模糊隶属度函数, 通过 α -cut 获得对应两种态度的 mass 函数 $m_{\text{Pess}}()$ 和 $m_{\text{Opti}}()$, 最后基于 Dempster 组合规则和判决规则作出最后的决策.

据前文所述, 基于 FCOWA-ER 的 SVM 多分类方法的步骤如下.

Step 1: 对多分类 SVM 进行建模, 获得决策矩阵 C .

Step 2: 对决策矩阵 C 的每一行分别取悲观态度策略和乐观态度策略计算 OWA 算子, 即分别取该行评价值的最小值与最大值, 得到一个区间 (每一个专家对相应类别的评价都属于这一区间), 共得到 k 个区间, 表示如下:

$$E(C) = \begin{bmatrix} E(C_1) \\ E(C_2) \\ \vdots \\ E(C_k) \end{bmatrix} = \begin{bmatrix} [C_1^{\min}, C_1^{\max}] \\ [C_2^{\min}, C_2^{\max}] \\ \vdots \\ [C_k^{\min}, C_k^{\max}] \end{bmatrix}. \quad (7)$$

Step 3: 将区间 $E(C)$ 的左列 (下界) 视为一个信息源, 为悲观信息源, 其右列 (上界) 即为乐观信息源. 分别对其进行归一化, 可得到

$$E^{\text{Fuzzy}}(C) = \begin{bmatrix} [N_1^{\min}, N_1^{\max}] \\ [N_2^{\min}, N_2^{\max}] \\ \vdots \\ [N_k^{\min}, N_k^{\max}] \end{bmatrix}. \quad (8)$$

向量 $[N_1^{\min}, N_2^{\min}, \dots, N_k^{\min}]$ 、 $[N_1^{\max}, N_2^{\max}, \dots, N_k^{\max}]$ 分别代表悲观态度和乐观态度的模糊隶属度函数.

Step 4: 使用 α -cut 法^[15] 将模糊隶属度函数转化为 DS 证据理论中的基本信度分配. 假设 $\theta = \{A_1,$

$A_2, \dots, A_q\}$ 为辨识框架, $\mu(A_i)$ ($i = 1, 2, \dots, k$) 为模糊隶属度函数, 相应的基本概率分配可按如下方法获得. 根据文献 [15] 生成一个升序 α -cut ($0 < \alpha_1 < \alpha_2 < \dots < \alpha_M \leq 1$), 其中 $M \leq |\Theta| = n$. 则有

$$B_j = \{A_i \in \Theta | \mu(A_i) \geq \alpha_j\},$$

$$m(B_j) = \frac{\alpha_j - \alpha_{j-1}}{\alpha_M}. \quad (9)$$

其中: B_j ($j = 1, 2, \dots, M$) 为焦点, α_i ($i = 1, 2, \dots, M$) 为模糊隶属度函数的序列, $M = k$. 由此, 将两组证据转化为基本信度分配 $m_{\text{Pess}}()$ 和 $m_{\text{Opti}}()$.

Step 5: 对 $m_{\text{Pess}}()$ 和 $m_{\text{Opti}}()$ 使用 Dempster 组合规则进行组合, 并使用决策规则判定类别的最终归属.

基于上述步骤可实现对多个二分类 SVM 的综合利用, 并实现 SVM 的多分类输出.

由于 Step 5 中要进行 1 次证据组合, 当类别总数较多时, 会带来一定的运算复杂度, 这也是证据理论相关方法的缺陷所在. 其他一些基于证据理论的多属性决策方法(如 COWA-ER^[16]), 若类别个数为 n , 则最终组合次数为 $n-1$ 次, 相比之下, 选择 FCOWA-ER 要小很多. 此外, FCOWA-ER 中的证据焦点是嵌套结构, 证据组合运算复杂度会相对较小^[11]; COWA-ER 中的焦点并不存在这样的关系, 其单次证据组合时运算复杂度更高.

3 基于 FCOWA-ER 的 SVM 多分类算例

Soybean 的数据集^[17]是一个四分类问题, 设类别表示为 $\{A_1, A_2, A_3, A_4\}$, 以 A_1 类样本为例, 使用基于 FCOWA-ER 的 SVM 多分类方法对其进行计算. 在 SVM 多分类建模中, 构建二分类子 SVM, 将输出转化为类概率值, 数据如表 1 所示, 其中 N/A 表示该二分类 SVM 不作评价.

表 1 soybean 数据集的一次运算数据

	A_1	A_2	A_3	A_4
A_1 对 A_2	0.8501	0.1499	N/A	N/A
A_1 对 A_3	0.4693	N/A	0.5307	N/A
A_1 对 A_4	0.8885	N/A	N/A	0.1115
A_2 对 A_3	N/A	0.1624	0.8376	N/A
A_2 对 A_4	N/A	0.2994	N/A	0.7006
A_3 对 A_4	N/A	N/A	0.4133	0.5867
票数	2	0	2	2

分别取悲观态度和乐观态度对表 1 的每一列计算 OWA 算子, 可得

$$E(C) = \begin{bmatrix} [0.4693, 0.8885] \\ [0.1499, 0.2994] \\ [0.4133, 0.8376] \\ [0.1115, 0.7006] \end{bmatrix}. \quad (10)$$

对式 (10) 的每一列进行归一化, 可得

$$E^{\text{Fuzzy}}(C) = \begin{bmatrix} [1, 1] \\ [0.3194, 0.3373] \\ [0.8807, 0.9427] \\ [0.2377, 0.7885] \end{bmatrix}. \quad (11)$$

得到两个模糊隶属度函数为

$$\mu_1 = [1, 0.3194, 0.8807, 0.2377],$$

$$\mu_2 = [1, 0.3373, 0.9427, 0.7885].$$

使用 α -cut 将其转化为 bba, 得到 $m_{\text{Pess}}()$ 和 $m_{\text{Opti}}()$ 如表 2 所示.

表 2 转化后的两组 bba

焦点	$m_{\text{Pess}}()$	$m_{\text{Opti}}()$
A_1	0.1193	0.0572
$A_1 \cup A_3$	0.5614	0.1543
$A_1 \cup A_2 \cup A_3$	0.0817	—
$A_1 \cup A_3 \cup A_4$	—	0.4514
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.2377	0.3370

对 $m_{\text{Pess}}()$ 和 $m_{\text{Opti}}()$ 使用 DS 证据理论组合规则进行组合后得到 $m_{\text{Dempster}}()$, 如表 3 所示.

表 3 组合后的 mass 函数

焦点	$m_{\text{Dempster}}()$
A_1	0.4576
$A_1 \cup A_3$	0.3777
$A_1 \cup A_2 \cup A_3$	0.0180
$A_1 \cup A_3 \cup A_4$	0.0382
$A_1 \cup A_2 \cup A_3 \cup A_4$	0.0285

pignistic 概率如表 4 所示.

表 4 最终的单焦点 pignistic 概率

焦点	A_1	A_2	A_3	A_4
BetP()	0.4278	0.0854	0.2799	0.1473

可见, 样本被判定为 A_1 类, 该结果与直观感觉相符, 也符合样本的真实情况. 本例使用的是常规证据组合方法和 pignistic 概率转换^[18], 其他证据组合和推理判决方法见文献 [18].

使用“一对一”投票法对该样本进行计算, 得到投票情况和结果如表 1 最后一行所示, A_1 、 A_3 和 A_4 类获得的票数相同, 因此无法获得决策结果. 使用决策树法对 A_1 类样本进行运算时, 假设顶节点为 A_1 对 A_2 、 A_3 对 A_4 的 SVM, 其下层节点为 A_1 对 A_3 、 A_2 对 A_4 , 则分类正确要求每一个分类器的结果都是正确的, 且顶节点的结果更重要. 使用有向无环图法判决 A_1 类样本时, 需经过 3 次运算, 每次排除一个最不可能的类别, 与决策树法类似, 要求判决路径上的每一个分类结果都正确, 否则最终判决结果错误. 从本

例的计算结果看, 基于 FCOWA-ER 多属性决策的多分类 SVM 结果更合理.

4 仿真分析

4.1 基于 UCI 数据集的仿真实验

Iris、Wine、Soybean、Glass、Balancescale 和 Seeds 数据集^[17]均来源于加州大学欧文分校的机器学习数据库, 分别使用投票法^[8]、决策树法^[9]、有向无环图法^[10]和本文提出的 SVM 多分类方法对其进行分类.

使用 UCI 数据集时, 将数据集随机分为 5 份, 每次选择 2 份作为训练集, 其余 3 份作为测试集, 在一轮计算中遍历所有组合共运行 10 次, 作交叉验证, 循环运行 100 轮. SVM 选用高斯径向基核函数

$$k(x, x_i) = \exp\{-\|x - x_i\|^2 / 2\sigma^2\}, \sigma > 0,$$

其中参数使用网格法寻优获得. 平均分类精度如表 5 所示, 其中 balancescale 数据集简写为 Bs.

表 5 平均分类精度

	Iris	Wine	Soybean	Bs	Glass	Seeds
投票法	94.91	97.77	98.27	91.01	67.86	92.78
决策树	95.56	97.51	98.73	87.35	69.69	93.10
有向无环图	96.18	97.17	98.46	93.63	68.34	92.62
FCOWA-ER	96.37	97.88	99.13	95.55	70.02	93.14

4.2 基于人工三螺旋的仿真实验

利用人工数据集对本文算法作进一步验证, 由下式构造 3 条螺旋线:

$$\begin{cases} x = A \cos(A + \theta), \\ y = A \sin(A + \theta). \end{cases} \quad (12)$$

其中: θ 分别取 $0, 2\pi/3, 4\pi/3, A \in [1, 5\pi]$. 自变量 A 以 0.2 为间隔进行采样, 获得 3 组数据如图 2 所示. 加入标准差为 0.2 的高斯白噪声, 获得数据如图 3 所示.

训练集和测试集由间隔选取获得, 实验参数由网格法验证得到, 运行 100 次, 得到平均分类精度如表 6 所示.

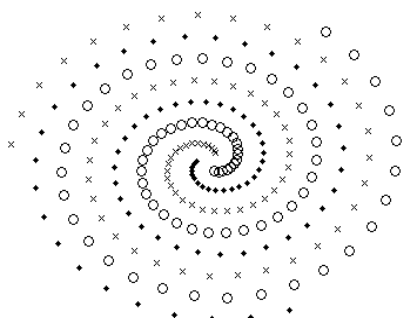


图 2 螺旋线数据

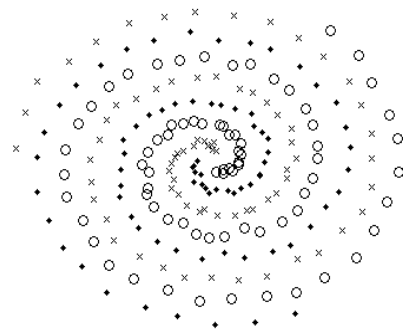


图 3 加入噪声的螺旋线数据

表 6 螺旋线数据的分类精度

	投票	决策树	有向无环图	FCOWA-ER
分类精度	93.12	91.67	92.97	93.78

4.3 仿真结果分析

由表 5 可以看出, 本文所提出的方法在各个数据集上的表现都优于其他 3 种方法或与其相当. 在使用 Balancescale、Iris 和 Glass 数据集时, 较为明显地提高了分类精度; 在使用 Wine、Seeds 和 Soybean 数据集时, 分类精度略有提升; 有向无环图法在分类精度上略微优于投票法; 决策树法的分类精度在各个数据集上波动较大. 使用螺旋线数据集作进一步试验, 验证其对一般规则数据的分类性能. 由表 6 可知, 决策树法具有最差的分类精度; 投票法与有向无环图法大致相当; 基于 FCOWA-ER 的分类方法具有最高的分类精度, 与表 5 中 UCI 数据集的实验结果相符.

由于前 3 种方法都使用的是子 SVM 的硬判决输出, 丢失了一部分信息, 当被测样本接近子 SVM 的分类超平面时, 表示该子 SVM 不能很好地划分该样本. 但其使用的是非此即彼的硬输出, 该子 SVM 对最终判决结果的影响与其他子 SVM 是相同的, 有可能对最终结果造成不利的影响. 对于决策树法和有向无环图法, 若该子 SVM 位于分层结构上层, 且分类错误, 则会导致最终分类结果的错误.

本文提出的方法使用的是子 SVM 的软输出, 且没有分层结构, 在多属性决策建模时, 每一个子 SVM 的地位相当. 在 FCOWA-ER 的推理过程中, 使用了悲观态度与乐观态度, 分类超平面离样本点较远的子 SVM 更容易对决策结果造成影响, 超平面离样本点较近的子 SVM 更容易在推理过程中被剔除. 因此, 该方法在处理被测样本位于训练集交界处的情况时, 具有更大的优势. 总体而言, 使用本文提出的方法在分类精度上优于其他 3 种传统方法.

5 结 论

本文提出了一种新的基于 FCOWA-ER 的 SVM 多分类方法. 该方法将一个多分类问题视为一个多

属性决策问题,将SVM的硬判决输出转化为软判决输出,再使用悲观和乐观策略计算OWA算子,获得两组证据,使用DS证据理论进行推理和决策.与传统的SVM多分类拓展方法相比,一方面,该方法无分层结构,不会出现误差累积的现象,且避免了拒绝识别的出现;另一方面,使用软输出进行推理和决策,更加充分地利用了信息.实验结果表明,本文所提出的方法总体上具有更高的分类精度.

DS证据理论的一个重要缺陷是高运算复杂度问题,特别是在辨识框架较大(类别数目较多)时尤为突出.研究DS证据理论的近似计算方法以降低运算复杂度,进而提升本文所提出方法的运算效率是未来研究的重要方向之一.

参考文献(References)

- [1] Vapnik V. The nature of statistical learning theory[M]. New York: Springer-Verlag, 1995: 138-154.
- [2] Cortes C, Vapnik V. Support-vector networks[J]. Machine Learning, 1995, 20(3): 273-297.
- [3] Schölkopf B, Smola A J, Williamson R C, et al. New support vector algorithms[J]. Neural Computation, 2000, 12(5): 1207-1245.
- [4] Suykens J A K, Vandewalle J. Least square support vector machine classifiers[J]. Neural Processing Letters, 1999, 9(3): 293-300.
- [5] 邢永忠, 吴晓蓓, 徐志良. 基于柯西加权的最小二乘支持向量机[J]. 控制与决策, 2009, 24(6): 937-940. (Xing Y Z, Wu X B, Xu Z L. Weighted least square support vector machine based on Cauchy distribution[J]. Control and Decision, 2009, 24(6): 937-940.)
- [6] Lin C F, Wang S D. Fuzzy support vector machine[J]. IEEE Trans on Neural Networks, 2002, 13(2): 464-471.
- [7] 温祥西, 孟相如, 马志强. 基于双重支持向量机的网络故障诊断[J]. 控制与决策, 2013, 28(4): 506-510. (Wen X X, Meng X R, Ma Z Q. Network fault diagnosis based on Dual-SVM[J]. Control and Decision, 2013, 28(4): 506-510.)
- [8] Hsu C W, Lin C J. A comparison of methods for multiclass support vector machines[J]. IEEE Trans on Neural Networks, 2002, 13(2): 415-425.
- [9] Takahashi F, Abe S. Decision-tree-based support vector machines[C]. The 10th Int Conf on Neural Information Processing. Singapore: IEEE Press, 2002: 1418-1422.
- [10] Platt J C, Cristianini N. Large margin DAGs for multiclass classification[C]. Advances in Neural Information Processing Systems. Denver: MIT Press, 2000, 12: 547-553.
- [11] Han D Q, Dezert J, Tacnet J M, et al. A fuzzy-cautious OWA approach with evidential reasoning[C]. The 15th Int Conf on Information Fusion. Singapore: IEEE Press, 2012: 278-285.
- [12] Weston J, Watkins C. Multi-class support vector machines[R]. Egham: Royal Holloway University of London, 1998.
- [13] Platt J C. Probabilistic outputs for support vector machines and comparison to regularized likelihood methods[M]. Cambridge: MIT Press, 2000: 61-74.
- [14] Yager R. On ordered weighted averaging operators in multi-criteria decision making[J]. IEEE Trans on Systems, Man, and Cybernetics, 1988, 18(1): 183-190.
- [15] Florea M C, Jusselme A L, Grenier D, et al. Approximation techniques for the transformation of fuzzy sets into random sets[J]. Fuzzy Sets and Systems, 2008, 159(3): 270-288.
- [16] Tacnet J M, Dezert J. Cautious OWA and evidential reasoning for decision making under uncertainty[C]. Proc of Fusion 2011 Conf. Chicago: IEEE Press, 2011: 1-8.
- [17] Fisher R A, Forina M, Asuncion A, et al. UCI machine learning repository[DB/OL]. (1987-01-01)[2014-08-02]. <http://archive.ics.uci.edu/ml/datasets.html>.
- [18] Dezert J, Wang P, Tchamova A. On the validity of Dempster-Shafer theory[C]. Proc of Fusion 2012 Conf. Singapore: IEEE Press, 2012: 655-660.

(责任编辑: 郑晓蕾)