

适合大规模数据集且基于LLM的0阶TSK模糊分类器

李滔[†], 王士同

(江南大学 数字媒体学院, 江苏 无锡 214122)

摘要: 针对传统分类器的泛化性能差、可解释性及学习效率低等问题, 提出0阶TSK-FC模糊分类器. 为了将该分类器应用到大规模数据的分类中, 提出增量式0阶TSK-IFC模糊分类器, 采用增量式模糊聚类算法(IFCM($c+p$))训练模糊规则参数并通过适当的矩阵变换提升参数学习效率. 仿真实验表明, 与FCPM-IRLS模糊分类器、径向基函数神经网络相比, 所提出的模糊分类器在不同规模数据集中均能保持很好的性能, 且TSK-IFC模糊分类器在大规模数据分类中尤为突出.

关键词: TSK-FC; TSK-IFC; 最小学习机; TSK型模糊分类器; 大规模数据集

中图分类号: TP181

文献标志码: A

Zero-order TSK fuzzy classifier based on LLM for large-scale data sets

LI Tao[†], WANG Shi-tong

(School of Digital Media, Jiangnan University, Wuxi 214122, China)

Abstract: In order to overcome the shortcoming that traditional classifiers cannot achieve satisfactory generalization performance, good interpretability and fast learning efficiency for datasets, the zero-order TSK fuzzy classifier called TSK-FC is proposed to solve the classification problem of middle-scale datasets. In order to make the TSK-FC suitable for large-scale data sets, its incremental version called TSK-IFC is developed, in which the incremental fuzzy clustering algorithm called incremental fuzzy ($c+p$)-means clustering(IFCM($c+p$)) is used to train antecedent parameters of fuzzy rules while fast consequent parameter learning is achieved through an appropriate matrix computation trick for the least learning machine. The proposed fuzzy classifiers, the TSK-FC and the TSK-IFC are experimentally compared with the conventional fuzzy classifier called FCPM-IRLS and the RBF neural network, and the results show the power of the proposed fuzzy classifiers, especially the great applicability of the TSK-IFC for large-scale data sets.

Keywords: TSK-FC; TSK-IFC; least learning machine; TSK fuzzy classifier; large-scale data sets

0 引言

模糊分类作为机器学习的一个重要领域, 使用语言规则和模糊集对知识进行表达, 因此模糊系统往往具有很好的可解释性, 并已在图像处理、文字识别、语音识别等领域得到了广泛应用. 模糊分类一般包含以下过程: 1) 将输入样本映射到模糊子空间, 即模糊划分过程; 2) 建立与模糊子空间相适应的模糊规则; 3) 在模糊规则之上对每个输入样本进行分类判别. 可见, 模糊规则是模糊分类系统中相当重要的一部分, 广大学者在设计模糊分类器时围绕模糊 if-then 规则的学习方式进行了广泛而深入的研究^[1-11]. 考虑不同的知识表达形式, 常用的有 Mamdani 型和 TSK

型模糊系统, 本文研究工作是在 TSK 型模糊系统的基础之上展开的.

在训练模糊 if-then 规则时, 常常使用模糊聚类方法学习模糊 if-then 规则的前件^[7-8]. 文献[7]中采用了 FCPM 模糊聚类算法学习模糊规则的前件, 研究发现, 该算法在大规模数据集中性能较差; 为了解决该问题, 文献[12]提出了增量式模糊聚类算法(IFCM($c+p$)), 该算法能够实现快速聚类且有较好的聚类性能, 可以很好地解决大规模数据集的聚类问题. 本文将分别采用 FCPM 和 IFCM($c+p$) 模糊聚类算法训练模糊 if-then 规则的前件.

人工神经网络的众多优点使其能够在众多领

收稿日期: 2015-12-22; 修回日期: 2016-03-10.

基金项目: 国家自然科学基金项目(61170122, 61272210); 江苏省自然科学基金项目(BK20130155).

作者简介: 李滔(1990—), 男, 硕士, 从事人工智能与模式识别、机器学习的研究; 王士同(1964—), 男, 教授, 博士生导师, 从事人工智能与模式识别、机器学习、深度学习等研究.

[†]通信作者. E-mail: chasingdream119@163.com

域得以应用,其中BP神经网络最具代表性,然而该神经网络的学习速度太慢且易陷入局部最优^[13-14].为了解决上述问题,文献[15-16]通过揭示单隐层前馈神经网络(SLFN)的极限学习机(ELM)与岭回归以及中心化的岭回归之间的关系,提出了SLFN的最小学习机(LLM)理论,通过证明核化的中心岭回归与KPCA之间的关系,提出了以无限可微的核函数为激励函数的多层前馈神经网络(MLFN)的LLM^[17].LLM具有训练速度快,需要人工干预少,泛化能力强等优点,可以很好地解决传统神经网络难以适用于大规模样本的问题.同时,LLM理论的提出为MLFN的深入研究提供了新思路,在回归学习、深度学习上也都得到了很好的应用^[17-18].因此,本文将LLM理论引入到模糊分类器的设计中,用于训练模糊if-then规则的后件.

为了得到令人满意的泛化性能、很好的可解释性及快速的学习效率,本文通过前面提到的模糊规则的训练方法,提出0阶TSK-FC模糊分类器,从而快速解决较大规模数据集的分类问题.为了有效提高该模糊分类器处理大规模数据集的效率,本文又在前者的基础上提出增量式0阶TSK-IFC模糊分类器.该分类器在处理大规模数据集的分类问题时能够更快建立模糊if-then规则,同时保证设计出的模糊分类器具有更好的性能.

1 TSK型模糊分类器简介

TSK型模糊系统模型因其输出简洁且具有不错的逼近性能而被广泛应用^[19],本文选择0阶TSK型模糊系统进行研究,其第*i*条模糊if-then规则表示为

$$R_i: \text{if } \mathbf{x}_1 \text{ is } A_{i,1}^l \ \& \ \mathbf{x}_2 \text{ is } A_{i,2}^l \ \& \ \cdots \ \& \ \mathbf{x}_n \text{ is } A_{i,n}^l, \\ \text{then } f_i^l(\mathbf{x}) = y_i^l, \quad i \in [1, \gamma]. \quad (1)$$

其中: $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_n$ 表示模糊分类器的输入, n 表示输入样本的特征数; l 表示训练集的实际类别数; γ 表示模糊规则数; $A_{i,n}^{(l)}$ 表示在第 l 类中第 i 条输入规则下的第 n 个特征所对应的模糊子集,其对应的隶属函数为 $\mu_{A_{i,n}^{(l)}}(\mathbf{x}_n)$; $f_i^l(\mathbf{x})$ 表示在第 l 类中第 i 条输入规则下的输出变量, y_i^l 表示其相应的输出值. 式(1)中对应的隶属函数表示为

$$\mu_{A_i^l}(\mathbf{x}) = \exp\left(-\frac{1}{2} \sum_{d=1}^n \frac{(x_d - v_{i,d}^l)^2}{\delta^2 s_{i,d}^l}\right). \quad (2)$$

其中: $v_{i,d}^l$ 表示在第 $l \in \{1, 2\}$ 类中第 $i \in [1, \gamma]$ 簇的第 d 个特征的聚类中心, $s_{i,d}^l$ 表示其离散度, δ 表示离散度参数.

$$s_i^l(\mathbf{x}) = \frac{\sum_{k \in \varpi_i} u_{ik}^l ((\mathbf{x}_k - \mathbf{v}_i^l)^T (\mathbf{x}_k - \mathbf{v}_i^l))}{\sum_{k \in \varpi_i} u_{ik}^l}, \quad i \in [1, \gamma]. \quad (3)$$

其中: $\mu_{i_k}^l$ 表示在第 l 类中第 k 个样本隶属于第 i 簇的程度; ϖ_1, ϖ_2 分别表示在二分类训练集 (\mathbf{x}_j, θ_j) , $\theta_j = \begin{cases} +1, & \mathbf{x}_j \in \varpi_1, \\ -1, & \mathbf{x}_j \in \varpi_2, \end{cases} j \in [1, N]$ 中属于两个类的样本集合.

与传统的高斯型隶属函数不同,式(2)所示的隶属函数可以忽略样本的具体分布.在实际情况下,样本的分布往往是未知的,即使该数据集不服从正态分布,只要通过式(3)就可以计算样本的离散情况,同时可以在训练模糊规则的过程中通过调节离散度参数 δ 校正模糊系统的学习误差.

式(1)的输出值及其决策函数表示为

$$y_{od} = f(\mathbf{x}_d) = \frac{\sum_{l=1}^2 \sum_{i=1}^{\gamma} \mu_{A_i^l}(\mathbf{x}_d) y_i^l}{\sum_{l=1}^2 \sum_{i=1}^{\gamma} \mu_{A_i^l}(\mathbf{x}_d)}, \quad d \in [1, n]; \quad (4)$$

$$\hat{f}(\mathbf{x}) = \arg \max_s \{f_s(\mathbf{x})\}. \quad (5)$$

模糊if-then规则是模糊分类器的重要组成部分,通过对其前件和后件进行学习才能建立相应的模糊分类器模型.

2 0阶TSK模糊分类器设计

2.1 前件的学习方法

前件的学习实际上是对输入样本进行模糊划分的过程.文献[7]中提出的FCPM-IRLS模糊分类器采用FCPM模糊聚类算法训练模糊if-then规则的前件,该算法对某一类已知的小规模数据集有不错的聚类性能,然而当样本规模很大时,其聚类效果明显下降,时间花费大幅增加.为了能够解决大规模数据集的聚类问题,即要在大幅缩短聚类时间的情况下尽量提高聚类性能,文献[12]提出了增量式模糊聚类算法(IFCM($c+p$)),该算法对数据集进行分块聚类,将最后一个数据块的聚类结果作为整个数据集的聚类结果.在聚类的过程中,为了增大数据块间的相互影响程度,该算法在FCPM算法的基础上添加了平衡项 $\alpha \sum_{i=1}^c \|\mathbf{v}_i - \mathbf{v}_i^o\|^2$.其中: α 被称为平衡因子,通过调节该参数即可调整算法的聚类性能; \mathbf{v}_i^o 表示经过IFCM($c+p$)算法得到的上一个数据块的聚类中心.除此之外,该算法还将每一个数据块的聚类中心及其周围的一些样本加入到下一个数据块中进行聚类,通过这两种方式加强对每次新增数据的约束.由此可得如下IFCM($c+p$)算法的目标函数:

$$J(\mathbf{U}, \mathbf{T}, \mathbf{V}, \alpha) = J(\mathbf{U}, \mathbf{T}, \mathbf{V}) + \alpha \sum_{i=1}^c \|\mathbf{v}_i - \mathbf{v}_i^o\|^2 =$$

$$\sum_{i=1}^c \sum_{k=1}^{N_s} \mu_{ik}^m \|\mathbf{x}_k - \mathbf{v}_i\|^2 + \sum_{j=1}^p \sum_{k=1}^{N_s} \zeta_{jk}^m \|\mathbf{x}_k - \mathbf{z}_j\|^2 + \alpha \sum_{i=1}^c \|\mathbf{v}_i - \mathbf{v}_i^o\|^2. \quad (6)$$

其中: $J(\mathbf{U}, \mathbf{T}, \mathbf{V})$ 表示FCPM算法的目标函数; c 表示未知类中的聚类个数; p 表示已知类中的聚类个数; μ 和 ζ 分别表示其对应的隶属度; m 表示模糊指数; N_s 表示第 s 个数据块的样本数; \mathbf{x}_k 表示在每个数据块中的第 k 个输入样本, $k \in [1, N_s]$. 该算法的详细求解过程如文献 [12] 所示.

2.2 后件的学习方法

后件的学习实际上是根据模糊 if-then 规则的前件进行模糊推理得到模糊输出的过程. FCPM-IRLS 模糊分类器采用基于梯度下降的迭代重加权最小二乘 (IRLS) 优化算法训练模糊规则的后件, 具有较高的时空花费, 显然该算法难以适用于大规模数据集的应用场景. 而 Wang 等^[17-18]提出的 LLM 理论具有学习速度快、泛化性能好等优势, 还能够解决 MLFN 问题, 显然可以很好地解决大规模数据带来的高时空花费问题, 同时还具有很好的泛化性能.

在多输入单输出 SLFN 的 LLM 中, 假设训练集 $D = \{(\mathbf{x}_j, t_j) | \mathbf{x}_j \in R^n, t_j \in R, j \in [1, N]\}$, 且 $\mathbf{X} = [\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N]^T$, 则 n 维的训练样本在无限可微的核函数 $G(\mathbf{x}_i, \boldsymbol{\Theta}_i)$ (在 SLFN 的 ELM 中也称为激励函数) 作用下经过 LLM 隐空间最终映射到 L 维的隐特征空间中, 在最后一个隐含层中求逆或求伪逆即可得到输出权重 y_i ($i \in [1, L]$). 因此, 在无限可微的核函数作用下, 输入样本与期望输出之间的逼近问题可以转化为隐特征空间中的逼近问题. 为了防止输入样本在尽可能逼近期望输出时出现过拟合的现象, 可以在最小化输出权重的范数时添加正则项, 于是样本逼近问题即可转换为岭回归问题, 也被称为 HFSR^[17], 即

$$\min \left(\frac{1}{2} \|\mathbf{y}\|^2 + \rho \frac{1}{2} \sum_{j=1}^N \xi_j^2 \right); \quad (7)$$

s.t. $\mathbf{h}(\mathbf{x}_j)\mathbf{y} = t_j - \xi_j, j \in [1, N].$

其中: ρ 表示正则参数, 由用户指定; ξ_j 表示训练误差; t_j 表示期望输出.

为了快速确定模糊规则的后件, 根据岭回归问题的求解方法^[16, 20]可得式 (7) 所示的 HFSR 问题的解

$$\mathbf{y} = \left(\frac{\mathbf{I}}{\rho} + \mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{T}. \quad (8)$$

其中: $\mathbf{I}_{L \times L}$ 表示在实数集上的单位矩阵; $\mathbf{H}_{N \times L} = [\mathbf{h}(\mathbf{x}_1), \mathbf{h}(\mathbf{x}_2), \dots, \mathbf{h}(\mathbf{x}_N)]^T$ 表示特征映射矩阵, $\mathbf{h}(\mathbf{x}_j) = [G(\mathbf{x}_j, \boldsymbol{\Theta}_1), \dots, G(\mathbf{x}_j, \boldsymbol{\Theta}_i), \dots, G(\mathbf{x}_j, \boldsymbol{\Theta}_L)]$ 表示特征映射函数, $\boldsymbol{\Theta}_i$ 表示核参数, 可以随机指定; $\mathbf{T}_{N \times 1} = [t_1, t_2, \dots, t_N]^T$ 表示期望输出矩阵; $\mathbf{y}_{L \times 1} = [y_1, y_2,$

$\dots, y_L]^T$ 表示输出权重矩阵. 本文选用 Gaussian 型的核函数, 如式 (2) 所示.

2.3 0阶TSK-FC模糊分类器设计

为了能够快速解决较大规模数据集的分类问题, 同时保证不错的泛化性能, 需要快速建立模糊规则, 即快速完成从输入样本到模糊子空间的模糊划分过程和快速进行模糊推理完成模糊输出的过程. 文献 [7] 提出的 FCPM-IRLS 模糊分类器采用基于梯度的 IRLS 训练后件, 这无疑将花费较多的时间, 因此本文首先考虑改进模糊分类器后件的学习方法. 而 LLM 理论具有极高的学习效率, 可以很好地解决不同规模样本的分类问题. 因此, 采用 LLM 理论训练模糊 if-then 规则的后件, FCPM 模糊聚类算法训练前件, 即可得 0 阶 TSK-FC 模糊分类器, 该分类器能够大幅提高学习效率, 同时保证不错的泛化性能.

定义归一化激活公式对式 (2) 所示的前件的隶属度进行归一化处理, 即

$$\overline{\mu_{A_i^d}}(\mathbf{x}_d) = \frac{\mu_{A_i^d}(\mathbf{x}_d)}{\sum_{l=1}^2 \sum_{i=1}^{\gamma} \mu_{A_i^d}(\mathbf{x}_d)}, d \in [1, n]. \quad (9)$$

由式 (9) 可知, 式 (4) 所示的模糊分类器的输出可以等价地表示为

$$f(\mathbf{x}) = \mathbf{h}(\mathbf{x})\mathbf{y}. \quad (10)$$

其中: $\mathbf{h}(\mathbf{x}_d) = [\overline{\mu_{A_1^d}}(\mathbf{x}_d), \overline{\mu_{A_2^d}}(\mathbf{x}_d), \dots, \overline{\mu_{A_n^d}}(\mathbf{x}_d), \overline{\mu_{A_1^d}}(\mathbf{x}_d), \overline{\mu_{A_2^d}}(\mathbf{x}_d), \dots, \overline{\mu_{A_n^d}}(\mathbf{x}_d)]$ 表示前件的隶属度矩阵, \mathbf{y} 表示模糊分类器的输出权重矩阵. 由此可以得到该模糊分类器的输出函数和决策函数

$$f(\mathbf{x}) = \mathbf{h}(\mathbf{x}) \left(\frac{\mathbf{I}}{\rho} + \mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{T}, \quad (11)$$

$$\hat{f}(\mathbf{x}) = \text{sign} \left(\mathbf{h}(\mathbf{x}) \left(\frac{\mathbf{I}}{\rho} + \mathbf{H}^T \mathbf{H} \right)^{-1} \mathbf{H}^T \mathbf{T} \right). \quad (12)$$

从式 (11) 和 (12) 可以看出, 分类器只需要在前件中快速确定特征映射矩阵即可快速确定输出, 通过 LLM 的矩阵变换避免了大量的迭代更新操作, 可见 TSK-FC 模糊分类器具有很高的学习效率, 在较大规模的数据集中也能得以应用.

通过上面的分析可以得到 0 阶 TSK-FC 模糊分类器的基本设计思路, 其具体训练过程如下.

输入: 训练集 \mathbf{X} , 迭代终止条件 ξ , 模糊规则数 c , 离散度参数 δ , 模糊指数 m 以及正则参数 ρ ;

输出: 输出权重矩阵 \mathbf{y} .

Step 1: 在训练集 \mathbf{X} 上确定 ϖ_1 和 ϖ_2 类的样本;

Step 2: 利用文献 [7] 中的算法 3 获取 ϖ_1 和 ϖ_2 类的聚类中心 $\mathbf{V}_1, \mathbf{V}_2$ 和离散度 $\mathbf{S}_1, \mathbf{S}_2$;

Step 3: 利用式 (2) 计算在 ϖ_1 和 ϖ_2 类中的前件

的隶属度矩阵 μ_1, μ_2 ;

Step 4: 利用式 (9) 对 Step 3 中得到的矩阵进行归一化处理;

Step 5: 将 Step 4 的结果合并, 得到在训练集 X 上的前件的隶属度矩阵;

Step 6: 根据式 (8) 计算输出权重 y .

通过使用上述算法对模糊 if-then 规则参数进行训练, 即可快速建立 0 阶 TSK-FC 模糊分类器模型. 其中: Step 1 ~ Step 5 属于前件的训练过程, Step 2 可以对输入的训练样本进行划分, Step 3 是对每一个划分结果进行模糊化处理的过程, Step 6 是后件的训练过程. 可见其并不需要任何迭代, 只需确定前件的隶属度矩阵, 再将其代入式 (8) 即可完成模糊分类器的训练. 最后将上述训练结果代入式 (12) 即可确定 TSK-FC 模糊分类器的实际输出, 再将其与期望输出进行比较即可得到测试精度, 该值越接近 1 表明系统泛化性能越好.

此外, 可以看出, TSK-FC 模糊分类器易受前件 FCPM 模糊聚类算法的约束, 其随着数据集规模的增加时空花费快速增加, 通过使用 LLM 快速训练后件, 能够在一定程度上降低前件带来的高时空花费, 因此 TSK-FC 模糊分类器可以较好地解决较大规模数据集的分类问题. 为了将其应用到大规模数据集的分类问题中, 本文对其进行了改进, 提出了增量式 0 阶 TSK-IFC 模糊分类器.

2.4 增量式 0 阶 TSK-IFC 模糊分类器设计

TSK-IFC 模糊分类器采用增量式模糊聚类算法 (IFCM ($c+p$)) 训练模糊 if-then 规则的前件. 该算法通过对输入样本进行分块聚类的思想, 不但将每一个数据块的聚类中心及其临近点加入下一个数据块进行聚类, 还加入了平衡项以加强对新增数据的约束力, 通过这种方法对每个数据块进行聚类, 直至获得最后一个数据块的聚类中心, 这就是所求的整个数据集的聚类中心. 研究表明, 这种方法可以很好地提高聚类性能和聚类效率^[12], 显然使用该算法可以快速获得模糊规则的前件. 为了在每个类中确定更加合理、更具解释性的规则, 在前件学习时, 需要在相应的类中得到更加稳定的聚类中心, 即更加准确地对输入样本进行划分. 本文通过循环迭代的思想获取两个类中的聚类中心, 将其带入前件的隶属度函数即可完成模糊化过程, 即完成前件的学习过程. 使用 LLM 理论的快速矩阵变换训练后件, 从而加快输出权重的求取. 由此可得另一种效率更高的增量式 0 阶 TSK-IFC 模糊分类器, 显然该分类器能够最大程度地降低模糊规则的训练时间, 同时还可以保证良好的泛化性能.

通过上面的分析可以得到增量式 0 阶 TSK-IFC 模糊分类器的基本设计思路, 其具体训练过程如下.

输入: 训练集 X , 迭代终止条件 ξ , 模糊规则数 c , 离散度参数 δ , 模糊指数 m , 平衡因子 α , 每个数据块中聚类中心的临近样本数 n_0 以及正则参数 ρ ;

输出: 权重矩阵 y .

Step 1: 在训练集 X 上确定 ϖ_1 和 ϖ_2 类的样本.

Step 2: 利用文献 [7] 中的算法 1 获取 ϖ_1 和 ϖ_2 类的初始聚类中心 V_{uk1}, V_{uk2} .

Step 3: 定义一个空集 V_{null} 作为已知类的聚类中心.

Step 4: 对 ϖ_1 和 ϖ_2 类中的训练样本进行随机均等分块.

Step 5: 使用 IFCM ($c+p$) 模糊聚类算法确定 ϖ_1 类的聚类中心 V_1, V_{uk1} 作为其第 1 个数据块中未知类的初始聚类中心, V_{null} 作为其已知类的聚类中心.

1) 使用 IFCM ($c+p$) 模糊聚类算法的聚类中心和样本隶属度更新公式计算当前数据块的聚类中心;

2) 找出距离当前数据块的聚类中心最近的 n_0 个样本;

3) 将 2) 中找到的最近邻样本和聚类中心加入到下一个数据块中;

4) 使用文献 [7] 中的算法 1 计算下一个数据块的初始聚类中心;

5) 使用 V_{null} 作为下一个数据块的已知类的聚类中心;

6) 重复 1) ~ 5) 的过程直至获得最后一个数据块的聚类中心, 将其作为 ϖ_1 类的聚类中心.

Step 6: 使用 Step 5 中的算法计算属于 ϖ_2 类的聚类中心 V_2, V_{uk2} 作为其第 1 个数据块的未知类的初始聚类中心, V_{null} 作为其已知类的聚类中心.

Step 7: 使用 Step 5 中的算法更新 ϖ_1 类的聚类中心 V_1 , 将 V_1 和 V_2 分别作为未知类和已知类的初始聚类中心.

Step 8: 使用 Step 5 中的算法更新 ϖ_2 类的聚类中心 V_2 , 将 V_2 和 V_1 分别作为未知类和已知类的初始聚类中心.

Step 9: 计算 ϖ_2 类的聚类中心 V_2 相对于上一次聚类中心的变化量的 Frobenius 范数, 若该值小于 ξ , 则停止迭代; 否则, 回到 Step 7 继续更新两个类的聚类中心.

Step 10: 使用样本隶属度矩阵公式确定 ϖ_1, ϖ_2 类的样本隶属度矩阵 U_1, U_2 .

Step 11: 使用式 (3) 计算 ϖ_1, ϖ_2 类的离散度 S_1, S_2 .

Step 12: 使用式(2)计算 ω_1, ω_2 类中前件所对应的隶属度矩阵 μ_1, μ_2 .

Step 13: 使用式(9)将Step 11中得到的隶属度矩阵 μ_1, μ_2 归一化,再合并得到在训练集 X 上的前件的隶属度矩阵 $H_{N \times 2c}$.

Step 14: 根据式(8)计算输出权重 y .

上面所示的算法流程即是TSK-IFC模糊分类器的完整训练过程。Step 1~Step 12都属于前件的训练过程,其中Step 5是本文提到的增量式模糊聚类算法(IFCM($c+p$))的算法步骤。为了得到更加合理的规则,需要准确地对输入样本进行划分,即需要在训练前件时获取更加稳定的聚类中心,如上述算法的Step 7~Step 9所示。由于在每次聚类时都进行分块聚类,显然在整个训练集上的聚类效率很高。为了提高聚类性能,在每个数据块聚类时(从第2个数据块开始),都加入来自上一个数据块的聚类中心及其临近点,利用平衡因子加强对每次新增数据的约束,并把最后一个数据块的聚类中心作为整个数据集的聚类中心。通过以上步骤即可快速获得前件的隶属度矩阵。Step 14为后件的训练过程,其简洁性显而易见,整个求解过程并不需要任何迭代更新,将 H 代入式(8)即可建立模糊分类器模型。通过式(12)所示的决策函数即可得到TSK-IFC模糊分类器的实际输出,再将其与期望输出进行比较即可得到其测试精度,该值越接近1表明分类器的泛化性能越好。

由于在训练前件时采用了增量式的模糊聚类算法(IFCM($c+p$)),使其能够快速准确地对输入样本进行划分,后件采用LLM理论能够高效地确定模糊输出。可见它们都能够很好地处理大规模的样本,因此所提到的增量式0阶TSK-IFC模糊分类器能够高效地解决大规模数据集的分类问题,同时保持不错的性能。为了得到分类器的最佳性能,本文采取交叉验证的策略对分类器进行训练以获得最优模型。

2.5 0阶TSK模糊分类器时空复杂度分析

本文提出的两种模糊分类器的时空花费由两部分构成,即训练前件和后件时分别花费的时间和占用的空间。假设 t 表示FCPM聚类算法的迭代次数, t' 表示IFCM($c+p$)聚类算法中每个数据块的平均迭代次数, n_0 表示在IFCM($c+p$)算法中距每个数据块的聚类中心最近的样本点个数,则TSK-FC模糊分类器在训练前件时的时空花费分别可以表示为 $O(tnd(c+p)+tc)$, $O(n(d+c+p))$,TSK-IFC模糊分类器在训练前件时的时空花费分别可以表示为 $O(t'nd(c+p)+t'c)$, $O((d+c+p+n_0)n/s)$ 。而两种模糊分类器均采用LLM的快速矩阵变换对后件进行训练,唯一的时空花费主要是在计算输出权重时的求逆运算,但也

只会占用 $O((2c)^2)$ 的存储空间,花费 $O((2c)^3)$ 的时间。这样,TSK-FC模糊分类器的时空花费分别可以表示为 $O(tnd(c+p)+tc+(2c)^3)$, $O(n(d+c+p)+(2c)^2)$,而TSK-IFC模糊分类器的时空花费分别可以表示为 $O(t'nd(c+p)+t'c+(2c)^3)$, $O((d+c+p+n_0)n/s+(2c)^2)$ 。可见,模糊规则数(即隐节点数)对模糊分类器的时空花费有很大影响。如果模糊规则数设置得过大,则会增加模型的复杂度,不利于实际应用;若设置得过少,则会降低系统的泛化性能。因此,本文的模糊规则数将进行合理的设置,从而保证所提出的模糊分类器具有很少的时空花费。比较发现,TSK-IFC模糊分类器具有更少的时空花费。

当训练集规模很大时,有 $c \ll N$,此时TSK-IFC模糊分类器将比TSK-FC模糊分类器更高效,在大规模数据集的应用场景中将有更好的性能。与以往的模糊分类器和传统的神经网络相比,本文提出的模糊分类器具有更少的时空花费。

3 实验研究与分析

本节将在7个不同规模的二分类数据集上对本文提到的模糊分类器TSK-FC、TSK-IFC与文献[7]中的模糊分类器FCPM-IRLS以及目前广泛应用的径向基神经网络RBF分类器在泛化性能和学习效率上进行比较。

3.1 实验设置

3.1.1 数据集描述及实验环境设置

为了验证本文提出的两种模糊分类器的泛化性能及学习效率,将采用7个不同规模的二分类数据集对其进行验证。其中:Svmguide1、Ijcnn1和Cod-rna来自LIBSVM;Ringnorm和Twonorm来自MLDATA;Magic和Skin-nonskin来自UCI。Skin-nonskin一共有245057个样本,随机抽取其中100000个样本用于本文的验证实验。为了保证仿真实验结果的合理性和可靠性,本文所有的模糊分类器都进行5重交叉验证实验,并对所有的数据集按特征归一化到[0,1],各数据集具体信息如表1所示。

表1 实验数据集

| 数据集 | 训练集 | 测试集 | 特征数 |
|--------------|--------|-------|-----|
| Svmguide1 | 5672 | 1417 | 4 |
| Ringnorm | 5920 | 1480 | 20 |
| Twonorm | 5920 | 1480 | 20 |
| Magic | 15216 | 3804 | 10 |
| Ijcnn1 | 39992 | 9998 | 22 |
| Skin-nonskin | 80000 | 20000 | 3 |
| Cod-rna | 125931 | 31482 | 8 |

Svmguide1、Ringnorm、Twonorm、Magic的数据集在硬件平台为Intel Core i3-3240×2 CPU,主频为3.4 GHz,内存为4 GB,且编程环境为Matlab 2015a的系统上进行仿真实验;Ijcnn1、Skin-nonskin、Cod-rna

数据集在硬件平台为Intel E5-2407×2 CPU, 主频为2.4 GHz, 内存为64 GB, 编程环境为Matlab 2013b的系统上进行仿真实验。

3.1.2 相关参数设置

在本文所有实验中, 最大迭代次数均为100, 在Cod-rna数据集中迭代终止条件为 $\xi = 1e-2$, 其他数据集中的迭代终止条件为 $\xi = 1e-3$. 在TSK-IFC模糊分类器前件的训练中: 每个数据块的聚类中心附近的样本点个数 $n_0 = 5$; Skin-nonskin和Cod-rna数据集的数据块大小分别取1%、2.5%、5%、10%, 其余数据集数据块大小分别取10%、25%、35%、50%; 平衡因子 α 在取值时, 保证平衡项 $\alpha \sum_{i=1}^c \|v_i - v_i^o\|^2$ 与 $J(U,$

$T, V)$ 具有相同量纲时的 α 即可^[12]; 实际情况下模糊规则数 c 的取值可参照文献[21]中的方法, 为了获得系统最佳性能, 本文中的模糊规则数 c 和隶属度参数 δ 的取值参照文献[7], 分别从 $\{2, 3, \dots, 20\}$ 和 $\{0.5, 0.6, \dots, 2\}$ 中取值; 实际情况下模糊指数 m 可参照文献[22]进行取值, 本文从 $\{1.1, 1.2, \dots, 3\}$ 中取值; 正则参数 ρ 从 $\{2^{-12}, 2^{-11}, \dots, 2^{11}, 2^{12}\}$ 中取值^[20]. 通过对以上参数进行训练获得最佳分类器模型, TSK-IFC和TSK-FC模糊分类器在各数据集中的具体参数取值情况如表2所示。

FCPM-IRLS模糊分类器在各数据集中的取值如文献[7]所示. RBF神经网络分类器采用Matlab自带的函数进行仿真实验。

表2 TSK-FC和TSK-IFC分类器在各数据集中的参数设置

| 数据集 | TSK-IFC | | | | | | | | | | | | TSK-FC | | | | | |
|-----------|---------|----------|----------|-----|----------|----------|-----|----------|----------|-----|----------|----------|--------|----------|-----|----------|-----|----------|
| | m | α | 10% | | | 25% | | | 35% | | | 50% | | | m | ρ | c | δ |
| | | | ρ | c | δ | ρ | c | δ | ρ | c | δ | ρ | c | δ | | | | |
| Twonorm | 2 | 7.5 | 2^{12} | 16 | 1.9 | 2^{12} | 3 | 1.9 | 2^{10} | 3 | 2 | 2^{12} | 10 | 1.9 | 1.1 | 2^{-6} | 14 | 1.8 |
| Ringnorm | 2.5 | 10 | 2^{12} | 5 | 1.3 | 2^{12} | 4 | 1.8 | 2^{12} | 5 | 1.8 | 2^{12} | 6 | 1.7 | 2 | 2^{12} | 2 | 2 |
| Svmguide1 | 3 | 5 | 2 | 4 | 1.6 | 2 | 4 | 1.7 | 2^{-1} | 16 | 1.2 | 2^{-4} | 5 | 1.3 | 1.4 | 2^{-8} | 20 | 1.1 |
| Magic | 2 | 8 | 2^{12} | 5 | 1.9 | 2^4 | 5 | 1.9 | 2^4 | 7 | 2 | 2^{12} | 6 | 2 | 2 | 2^{-3} | 13 | 2 |
| Ijcn1 | 2.5 | 3 | 2^{12} | 3 | 1.6 | 2^{12} | 3 | 1.7 | 2^{12} | 3 | 1.6 | 2^{12} | 3 | 1.7 | 1.9 | 2^7 | 5 | 2 |

| 数据集 | TSK-IFC | | | | | | | | | | | | TSK-FC | | | | | |
|--------------|---------|----------|----------|-----|----------|----------|-----|----------|----------|-----|----------|----------|--------|----------|-----|----------|-----|----------|
| | m | α | 1% | | | 2.5% | | | 5% | | | 10% | | | m | ρ | c | δ |
| | | | ρ | c | δ | ρ | c | δ | ρ | c | δ | ρ | c | δ | | | | |
| Skin-nonskin | 2 | 8 | 2^{12} | 5 | 1.1 | 2^{12} | 4 | 1.7 | 2^{12} | 6 | 1.8 | 2^{12} | 5 | 1 | 1.1 | 2^{12} | 12 | 1.8 |
| Cod-rna | 2 | 90.5 | 2^5 | 4 | 2 | 2^{12} | 4 | 2 | 2^{12} | 4 | 2 | 2^{12} | 4 | 2 | 1.1 | 2^{-1} | 12 | 2 |

3.2 0阶TSK模糊分类器性能比较

3.2.1 模糊规则数和正则参数对0阶TSK模糊分类器的影响

对于一个模糊分类器而言, 为了寻求更好的性能, 需要选择合理的模糊规则数, 模糊规则数太少时系统可解释性降低、泛化性能不好, 而模糊规则数太多则会增加模型的训练难度, 不利于实际应用, 因此选择合理的模糊规则数对模糊分类器的性能有着至关重要的作用. 而正则参数可以对模糊分类器的泛化误差进行修正, 以避免发生过拟合的现象, 可见, 合理地选择正则参数对模糊分类器而言也是很重要的. TSK-FC模糊分类器随着模糊规则数和正则参数的变化, 分类器在Twonorm数据集上的测试精度和训练时间的变化情况如图1和图2所示。

采用增量式模糊聚类算法训练模糊规则前件的TSK-IFC模糊分类器, 随着模糊规则数和正则参数的变化, 分类器在Twonorm数据集上的测试精度和训练时间的变化情况如图3和图4所示. 在不同大小的数据块下, 模糊规则数和正则参数的变化对模糊分类器的性能影响也可以从图3和图4中直观地看出. 此时模糊规则数从 $\{2, 3, \dots, 30\}$ 中取值, 其他取值情况

见3.1.2节中的相关参数设置。

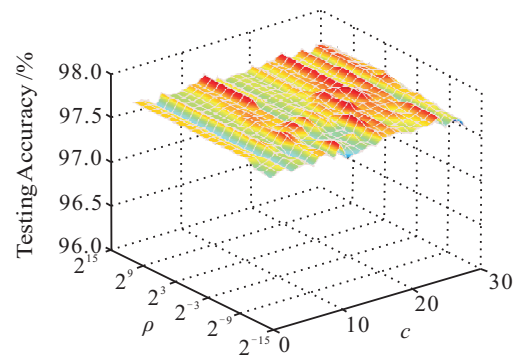


图1 TSK-FC模糊分类器在Twonorm数据集上的测试精度

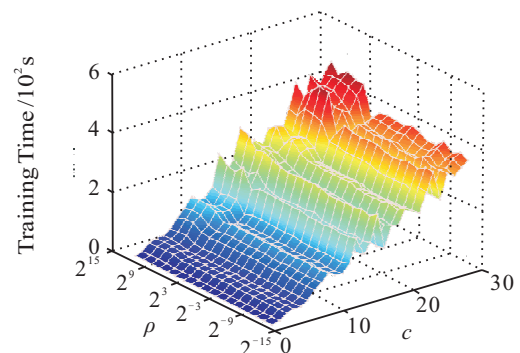
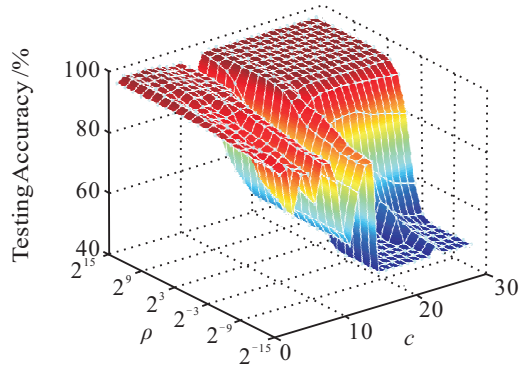
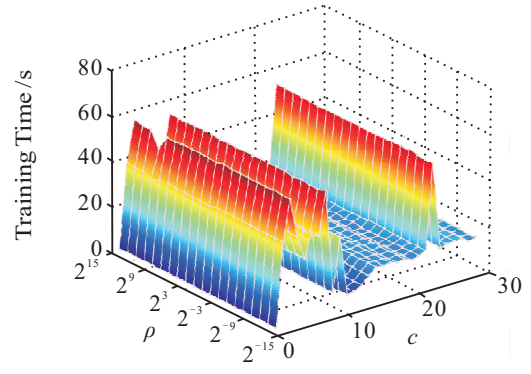


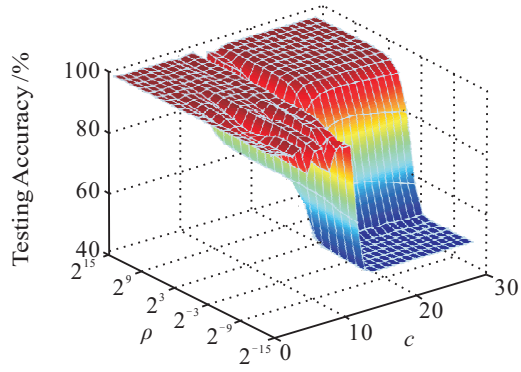
图2 TSK-FC模糊分类器在Twonorm数据集上的训练时间



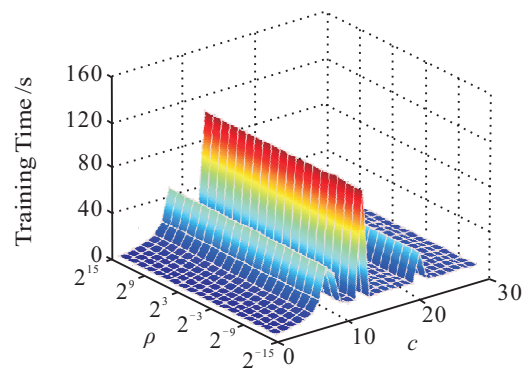
(a) chunk size = 10%



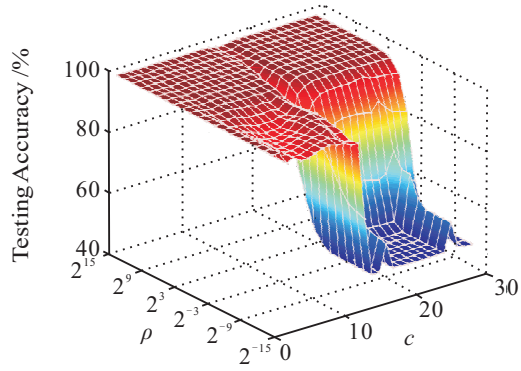
(a) chunk size = 10%



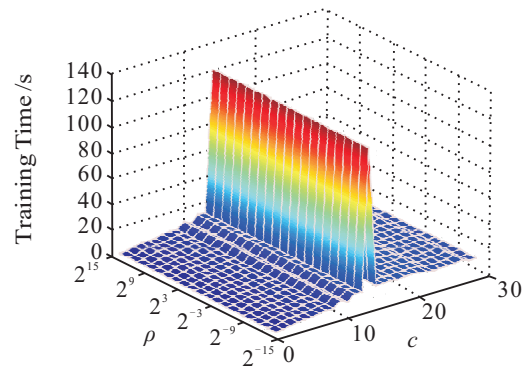
(b) chunk size = 25%



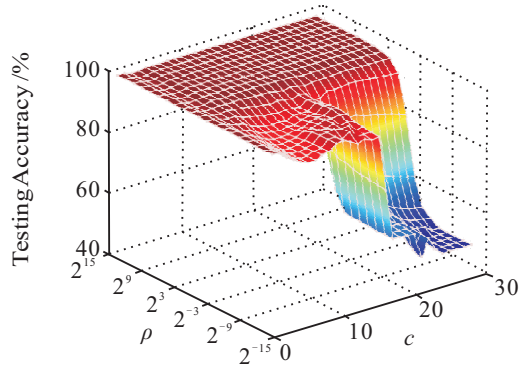
(b) chunk size = 25%



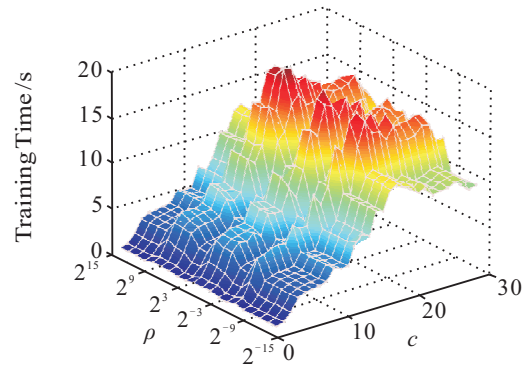
(c) chunk size = 35%



(c) chunk size = 35%



(d) chunk size = 50%



(d) chunk size = 50%

图 3 TSK-IFC 模糊分类器在 Twonorm 数据集上的不同数据块大小下的测试精度

图 4 TSK-IFC 模糊分类器在 Twonorm 数据集上的不同数据块大小下的训练时间

在 TSK-FC 模糊分类器中, 从图 1 可以看出: 随着正则参数和模糊规则数的变化, 分类器在 Twonorm 数据集上的测试精度小幅波动, 可见其泛化性能对正则参数和模糊规则数不敏感; 从图 2 可以看出: 随着模

糊规则数的增加, 其在 Twonorm 数据集上的训练时间迅速增加, 可见其对模糊规则数很敏感, 而对正则参数不是很敏感; 在 TSK-IFC 模糊分类器中, 从图 3 可以看出: 当正则参数取值较小而模糊规则数取值较

大时, 对其在 Twonorm 数据集上的测试精度有较大影响, 说明此时分类器的泛化性能受两个参数的影响较大, 而在其他情况下, 该分类器对两个参数均不敏感, 可见只要正则参数足够大就可以尽量保证分类器拥有很好的泛化性能; 从图 4 可以看出: 随着模糊规则数的增加, 其在 Twonorm 数据集上的训练时间缓慢增加, 而当正则参数增加时, 其训练时间基本不变, 可见其对两个参数都不是很敏感。

表 3 各分类器在不同规模数据集上的泛化性能及训练时间比较

| 分类器 | 性能 | Svmguide1 | Twonorm | Ringnorm | Magic | Ijcn1 | Skin-nonskin | Cod-rna | |
|-----------|---------------------|---------------------|--------------------|--------------------|--------------------|--------------------|-----------------|--------------------|-------------|
| RBF | Testing Accuracy /% | 96.75 ±0.65 | 90.7 ±0.45 | 86±0 | 82.75 ±0.63 | 90.64±0.17 | — | — | |
| | Training Time /s | 150.4 | 155.4 | 174.57 | 1526.3 | 25215 | — | — | |
| FCPM-IRLS | Testing Accuracy /% | 96.3 ±0.86 | 97.85 ±0.37 | 98.55±0.41 | 83.06 ±0.54 | 90.39±0.28 | 99.85 ±0 | 93.92 ±0.32 | |
| | Training Time /s | 51.35 | 46.07 | 10.18 | 330.70 | 249.81 | 192.17 | 1479.6 | |
| TSK-FC | Testing Accuracy /% | 96.29 ±0.53 | 97.89 ±0.33 | 92.32±0.92 | 82.2 ±0.65 | 90.81 ±0.24 | 99.68 ±0 | 92.01 ±0.42 | |
| | Training Time /s | 62.26 | 5.53 | 3 | 279.06 | 15.13 | 124.49 | 1178.9 | |
| TSK-IFC | 10% | Testing Accuracy /% | 95.96 ±0.66 | 97.82 ±0.43 | 97.8±0.44 | 80.06 ±0.3 | 90.76±0.25 | 1% 97.6 ±0.43 | 91.07 ±0.43 |
| | | Training Time /s | 0.81 | 10.84 | 2.29 | 9.39 | 6.54 | 25.76 | 26.9 |
| | 25% | Testing Accuracy /% | 95.96 ±0.66 | 97.82 ±0.43 | 97.8±0.44 | 80.62±0.52 | 90.73±0.13 | 2.5% 98.51 ±0.03 | 91.22 ±0.41 |
| | | Training Time /s | 1.65 | 0.36 | 18.18 | 34.07 | 7.41 | 19.90 | 30.7 |
| | 35% | Testing Accuracy /% | 95.60 ±0.64 | 97.81 ±0.35 | 98.38±0.48 | 80.42 ±0.36 | 90.77±0.16 | 5% 98.63 ±0 | 91.11 ±0.32 |
| | | Training Time /s | 0.97 | 0.78 | 3.36 | 19.60 | 7.34 | 15.19 | 32.4 |
| 50% | Testing Accuracy /% | 95.85 ±0.6 | 97.84 ±0.14 | 98.57 ±0.39 | 80.61 ±0.32 | 90.76±0.18 | 10% 98.93 ±0.15 | 91 ±0.32 | |
| | Training Time /s | 1.77 | 1.78 | 4.40 | 28.39 | 7.31 | 25.93 | 49.52 | |

从表 3 中可以看出: 在测试精度方面, 虽然 RBF、FCPM-IRLS 在某些数据集中具有较好的测试精度, 但是本文提出的 TSK-FC 和 TSK-IFC 模糊分类器在各个数据集中的测试精度总能最大程度地接近甚至超过 FCPM-IRLS。同时, TSK-IFC 模糊分类器的测试精度随着数据块大小的增加而缓慢增加; 在训练时间方面, 本文提出的两种模糊分类器拥有最短的训练时间, 其中 TSK-IFC 模糊分类器的训练时间更短, 且随着数据块大小的增加而缓慢增加。另外, 当数据集规模逐渐增大时, 前者的训练时间也逐渐增加, 其学习效率明显低于后者, 可见增量式 TSK-IFC 模糊分类器可以很好地处理大规模数据集的分类问题。

从以上实验结果中还可以看出: RBF 随着数据集规模的增大, 其训练时间急剧增加。在进行大规模数据集实验时总是出现内存溢出的现象, 如表 3 中的“—”所示, 尽管已经采用了高性能的服务器, 但是仍然不能避免这种现象, 因此可以发现 RBF 难以适用于大规模数据集。

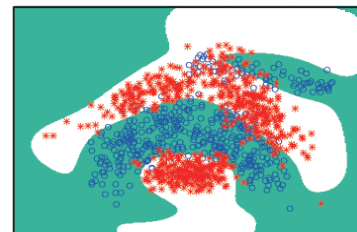
各分类器产生的分界线对 Banana (2 维 5 300 个样本两类) 数据集进行分类的结果如图 5 和图 6 所示。从图中可以很直观地看出: 基于 LLM 的分类器产生的分界线能够最大程度地将两类样本分开, RBF 则导致过多样本错分, FCPM-IRLS 模糊分类器的分类效果欠佳, 可见基于 LLM 的分类器具有最好的泛

化性能。对于本文提出的两种模糊分类器相比而言, 显然 TSK-IFC 模糊分类器在 Twonorm 数据集上具有更好的泛化性能以及更短的训练时间。

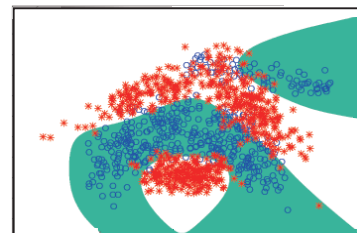
3.2.2 各分类器泛化性能及学习效率比较

本文提到的模糊分类器 TSK-FC 和 TSK-IFC 以及用于比较的分类器 FCPM-IRLS、RBF 在不同规模数据集上的泛化性能及训练时间如表 3 所示, 最优结果已用粗体标出。

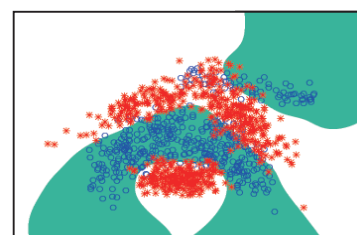
化性能。



(a) RBF



(b) FCPM-IRLS



(c) TSK-FC

图 5 RBF、FCPM-IRLS、TSK-FC 分类器在 Banana 数据集上的分类结果

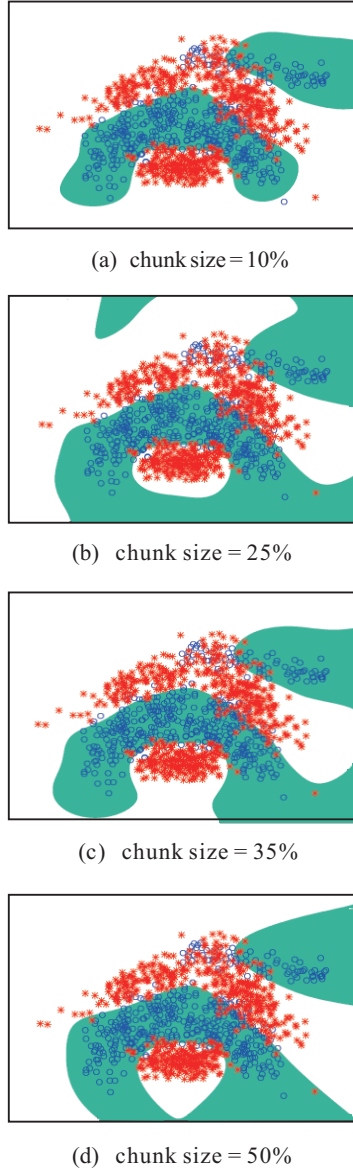


图 6 TSK-IFC 模糊分类器在 Banana 数据集上的不同数据块大小下的分类结果

进一步地, 为了通过仿真实验说明本文提出的两种 0 阶 TSK 模糊分类器的高可解释性, 以数据集 Svmguide1 为例, 这里列出两个模糊分类器得到的 4 条模糊规则:

1) TSK-FC 模糊分类器得到的部分模糊规则

$$R_1: \text{if } \mathbf{x}_1 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_1 - 0.7616}{0.1645}\right)^2\right) \&$$

$$\mathbf{x}_2 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_2 - 0.5669}{0.1273}\right)^2\right) \&$$

$$\mathbf{x}_3 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_3 - 0.3987}{0.0526}\right)^2\right) \&$$

$$\mathbf{x}_4 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_4 - 0.9420}{0.0654}\right)^2\right),$$

then y is 0.1123;

$$R_2: \text{if } \mathbf{x}_1 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_1 - 0.0847}{0.0414}\right)^2\right) \&$$

$$\mathbf{x}_2 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_2 - 0.2490}{0.0733}\right)^2\right) \&$$

$$\mathbf{x}_3 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_3 - 0.6576}{0.0629}\right)^2\right) \&$$

$$\mathbf{x}_4 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_4 - 0.5072}{0.0740}\right)^2\right),$$

then y is 0.4565.

2) TSK-IFC 模糊分类器得到的部分模糊规则 (以数据块取 10% 为例)

$$R_1: \text{if } \mathbf{x}_1 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_1 - 0.1181}{0.1808}\right)^2\right) \&$$

$$\mathbf{x}_2 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_2 - 0.2285}{0.2481}\right)^2\right) \&$$

$$\mathbf{x}_3 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_3 - 0.4447}{0.2617}\right)^2\right) \&$$

$$\mathbf{x}_4 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_4 - 0.5962}{0.2970}\right)^2\right),$$

then y is 0.5335;

$$R_2: \text{if } \mathbf{x}_1 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_1 - 0.1190}{0.1804}\right)^2\right) \&$$

$$\mathbf{x}_2 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_2 - 0.2300}{0.2473}\right)^2\right) \&$$

$$\mathbf{x}_3 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_3 - 0.4479}{0.2591}\right)^2\right) \&$$

$$\mathbf{x}_4 \text{ is } \exp\left(-\frac{1}{2}\left(\frac{\mathbf{x}_4 - 0.6004}{0.2932}\right)^2\right),$$

then y is 0.6899.

明显地, 这 4 条规则是简洁且容易解释的.

综上可知, RBF 的性能会随着数据集规模的增大而急剧下降, FCPM-IRLS 模糊分类器通过牺牲大量的训练时间来获取不错的泛化性能, 可见这两种分类器都难以保证快速的学习效率; 而本文提出的两种模糊分类器在具有高可解释性和快速的学习效率的同时还能保持不错的泛化性能. TSK-FC 模糊分类器可以快速解决较大规模数据集的分类问题, 而采用增量式模糊聚类算法训练前件的 TSK-IFC 模糊分类器可以更高效地解决大规模数据集的分类问题.

4 结 论

本文首先提出了 0 阶 TSK-FC 模糊分类器用于快速解决一般规模及较大规模数据集的分类问题, 但由于其采用 FCPM 算法训练模糊规则的前件致使其难以在大规模数据集中应用. 为此, 采用增量式模糊聚类算法 (IFCM($c+p$)) 训练前件, 采用 LLM 的快速矩阵变换训练后件, 在前者的基础上提出了增量式 0 阶 TSK-IFC 模糊分类器用于高效处理大规模数据集的分类问题. 因为本文提出的两种 0 阶 TSK 模糊分类器的每一条输入规则都是 TSK 形式的, 故都具有很好的可解释性; 然而, 本文提出的模糊分类器对参数较为敏感, 其模糊规则数以及正则参数与系统泛化性能之间存在一般规律, 这些问题的解决都将是今后研究的重点.

参考文献(References)

- [1] Junco L, Sánchez L. Using the Adaboost algorithm to induce fuzzy rules in classification problems[J]. Proc ESTYLF, 2000: 297-301.
- [2] 邢珍. 基于遗传算法的模糊分类器的研究及其在汉语元音识别中的应用[D]. 广州: 华南理工大学理学院, 2014. (Xing Z. The study of fuzzy classifier based on genetic algorithm and its application on chinese vowels recognition[D]. Guangzhou: School of Science, South China University of Technology, 2014.)
- [3] Ling W X, Xing Z, Chen Z M, et al. Designing a fuzzy classifier using genetic algorithm and application in chinese vowel recognition[C]. The 9th Int Conf on Computational Intelligence and Security. Leshan: IEEE, 2013: 344-348.
- [4] Nauck D, Kruse R. A neuro-fuzzy method to learn fuzzy classification rules from data[J]. Fuzzy Sets and Systems, 1997, 89(3): 277-288.
- [5] 刘淑英. 混合神经模糊分类器的实现[J]. 计算机技术与发展, 2013, 23(12): 113-115. (Liu S Y. Implementation of hybrid neuro-fuzzy classifier[J]. Computer Technology and Development, 2013, 23(12): 113-115.)
- [6] 梅蓉蓉, 吴小俊, 冯振华. 改进的动态模糊神经网络及其在人脸识别中的应用[J]. 计算机应用与软件, 2012, 29(1): 56-59. (Mei R R, Wu X J, Feng Z H. Improved dynamic fuzzy neural networks with its application to face recognition[J]. Computer Applications and Software, 2012, 29(1): 56-59.)
- [7] Leski J M. Fuzzy $(c+p)$ -means clustering and its application to a fuzzy rule-based classifier: Towards good generalization and good interpretability[J]. IEEE Trans on Fuzzy Systems, 2014, 23(4): 802-812.
- [8] Yang X, Zhang G, Lu J, et al. A kernel fuzzy c -means clustering-based fuzzy support vector machine algorithm for classification problems with outliers or noises[J]. IEEE Trans on Fuzzy Systems, 2014, 19(1): 105-115.
- [9] 阳爱民, 李心广, 周咏梅, 等. 一种基于支持向量机的模糊分类器[J]. 系统仿真学报, 2008, 20(13): 3414-3419. (Yang A M, Li X G, Zhou Y M, et al. Fuzzy classifier based on support vector machine[J]. J of System Simulation, 2008, 20(13): 3414-3419.)
- [10] 刘忠宝, 王士同. 基于边界的最大间隔模糊分类器[J]. 光学精密工程, 2012, 20(1): 140-147. (Liu Z B, Wang S T. Maximum-margin fuzzy classifier based on boundary[J]. Optics and Precision Engineering, 2012, 20(1): 140-147.)
- [11] Deng Z, Cao L, Jiang Y, et al. Minimax probability TSK fuzzy system classifier: A more transparent and highly interpretable classification model[J]. IEEE Trans on Fuzzy Systems, 2014, 23(4): 813-826.
- [12] 李滔, 王士同. 适合大规模数据集的增量式模糊聚类算法 IFCM $(c+p)$ [J]. 智能系统学报, 2016, 11(2): 188-199. (Li T, Wang S T. Incremental fuzzy $(c+p)$ -means clustering for large data[J]. CAAI Trans on Intelligent Systems, 2016, 11(2): 188-199.)
- [13] Cai H, Liu L, Cao Q. A novel BP algorithm based on self-adaptive parameters and performance analysis[C]. The 1st Int Conf on Innovative Computing, Information and Control. Beijing: IEEE, 2006: 383-387.
- [14] Tang C Z, Kwan H K. Parameter effects on convergence speed and generalization capability of back propagation algorithm[J]. Int J of Electronics, 1993, 74(1): 35-46.
- [15] Zhai C M, Du J X. Applying extreme learning machine to plant species identification[C]. Int Conf on Information and Automation. Changsha: IEEE, 2008: 879-884.
- [16] Suresh S, Babu V, Sundararajan N. Image quality measurement using sparse extreme learning machine classifier[C]. The 9th Int Conf on Control, Automation, Robotics and Vision. Singapore: IEEE, 2006: 1-6.
- [17] Wang S, Chung F L, Wu J, et al. Least learning machine and its experimental studies on regression capability[J]. Applied Soft Computing, 2014, 21(8): 677-684.
- [18] Wang S, Jiang Y, Chung F L, et al. Feedforward kernel neural networks, generalized least learning machine, and its deep learning with application to image classification[J]. Applied Soft Computing, 2015, 37(c): 125-141.
- [19] Jiang Y, Chung F L, Ishibuchi H, et al. Multitask TSK fuzzy system modeling by mining intertask common hidden structure[J]. IEEE Trans on Cybernetics, 2015, 45(3): 548-561.
- [20] Huang G B, Zhou H, Ding X, et al. Extreme learning machine for regression and multiclass classification[J]. IEEE Trans on Systems, Man, and Cybernetics — Part B: Cybernetics, 2012, 42(2): 513-529.
- [21] Sun H, Wang S, Jiang Q. FCM-based model selection algorithms for determining the number of clusters[J]. Pattern Recognition, 2004, 37(10): 2027-2037.
- [22] Yu J, Cheng Q, Huang H. Analysis of the weighting exponent in the FCM[J]. IEEE Trans on Systems, Man, and Cybernetics — Part B: Cybernetics, 2004, 34(1): 634-639.