

由最大同类球提取模糊分类规则

徐明亮^{*①②} 王士同^①

^①(江南大学数字媒体学院 无锡 214122)

^②(无锡环境科学与工程研究中心 无锡 214063)

摘要: 为提高模糊分类规则的有效性和可解释性, 该文提出一种基于最大同类球的模糊规则提取方法。首先, 每个样本根据与最近异类之间的距离确定一个最大同类球。然后根据各个同类球中样本之间的包含关系和独有性对同类球进行约简。再根据约简后的同类球建立 MA 分类器的模糊规则前件。MA(Mamdani-Assilan)二分类器的模糊规则后件参数学习以加权分类错误平方最小化为目标函数, 采用共轭梯度法求解后件参数。KEEL 标准数据集中的 12 个 10 折交叉验证数据集的对比分类实验验证了该方法的有效性。

关键词: 模糊规则; 分类; 约简; Mamdani-Assilan (MA); 同类

中图分类号: TP18

文献标识码: A

文章编号: 1009-5896(2017)05-1130-06

DOI: 10.11999/JEIT160779

Extracting Fuzzy Rules from the Maximum Ball Containing the Homogeneous Data

XU Mingliang^{①②} WANG Shitong^①

^①(The School of Digital Media, Jiangnan University, Wuxi 214122, China)

^②(The Center of Wuxi Environmental Science & Engineering, Wuxi 214063, China)

Abstract: In order to improve the interpretability and effectiveness of the fuzzy classifier rules, this paper presents a new method to extract the fuzzy rules based on the maximum ball only containing the homogeneous data. At first, every sample constructs a maximum ball in the light of the shortest distance to heterogeneous samples. Then those balls are reduced according to the relation of inclusion and the unique among the samples that the ball encloses. Then the fuzzy rules are constructed with the reserved balls. The parameters learning of the antecedent part of the classifier are based on the minimization of the weight misclassification quadratic error and resolved with the conjugate gradient algorithm. The experiments on 12 benchmark datasets with 10 folds are performed to demonstrate the validity of the classifier.

Key words: Fuzzy rule; Classifier; Reduction; Mamdani-Assilan (MA); Homogeneous data

1 引言

分类一直是模式识别中的研究热点, 在这个领域中, 应用最普遍的方法有 SVM, Adaboost, KFD (Kernel Fisher Discriminant) 和模糊分类方法。与其他方法相比, 模糊分类方法具有可解释性的优点, 因此受到众多研究者的关注。在模糊分类系统的建立过程中, 模糊分类规则构建是其中的一个重要环节。模糊分类规则可以由领域专家直接给出。然而在多数情况下, 并没有直接经验可用, 只能从拥有

的样本集中获取模糊分类规则。为此, 研究者们提出许多行之有效的办法, 如文献[1]采用强化学习奖惩机制来对模糊规则进行调整和选择。该方法具有一定的自适应能力, 但学习性能对初始规则的选择较为敏感, 且必须经过多轮学习才有可能达到较好的学习效果, 因此学习周期较长。文献[2]先对输入空间进行简单划分, 然后通过竞争交互推理和强化学习对模糊规则进行调整。同样, 由于引入强化学习, 该方法的学习效率较低。近年来, 结合全局优化算法进行模糊规则学习的相关文献较为丰富, 如文献[3]先采用形式概念分析(FCA)从样本中提取模糊规则, 然后用遗传算法(GA)进行模糊规则的优化。由于形式概念分析离不开领域专家的经验, 因此该方法使用范围受到限制。文献[4]用模糊自适应共振网络的增量学习机制和 Q 学习的节点评价机制进行网络结构和参数学习, 然后利用遗传算法从中

收稿日期: 2016-07-22; 改回日期: 2017-01-09; 网络出版: 2017-02-28

*通信作者: 徐明亮 xml1973@126.com

基金项目: 国家自然科学基金(61170122, 61202311, 61272210), 江苏省自然科学基金(BK2012552), 江苏省青蓝工程资助项目(2014) Foundation Items: The National Natural Science Foundation of China (61170122, 61202311, 61272210), The Natural Science Foundation of Jiangsu Province (BK2012552), The Qing Lan Project of Jiangsu Province (2014)

抽取模糊规则。文献[5-7]采用遗传算法直接对模糊规则进行编码和优化。遗传算法具有全局搜索能力强的优点，但在用于模糊规则优化时，编码较为复杂，实现较为困难，参数选择范围大，求解结果与初始规则的选择存在一定的关联性。文献[8-10]则分别采用磷虾优化算法、粒子群优化算法、量子蚁群优化算法对随机产生的模糊规则进行优化和选择。与采用遗传算法一样，这些优化算法具有较好的全局搜索能力，但这些方法也存在求解效果与初始规则选择相关，参数选择范围宽，评价函数只考虑经验风险，没有考虑泛化风险的不足。另外应用较为广泛的方法则是根据学习样本的聚类结果来构建模糊分类规则。如文献[11]采用减法聚类(SCM)和差分进化算法(DE)确定模糊规则。文献[12]采用模糊聚类算法(FCM)建立初始模糊规则，然后利用遗传算法对模糊分类系统进行优化。文献[13]也是先用FCM对数据进行聚类，然后根据聚类结果构造出模糊分类规则。文献[14]提出一种能使异类聚类中心尽可能相互远离F(C+P)M聚类方法，并利用该方法获得模糊分类规则。不论是将各类数据分开聚类，还是将各类数据混合在一起进行聚类，所得到的模糊划分往往存在较多的异类样本，这样不仅会降低模糊规则的有效性，同时也使模糊规则的可解释性变差。此外聚类结果存在一定的随机性，因此分类器性能稳定性较差。这些不足在前述方法中也同样存在。为此，本文以MA(Mamdani-Assilan)模糊分类器为基础，尝试采用一种既具有一定稳定性，又能保证每一个划分中的数据都是同类的数据划分方法，即基于最大同类球的数据划分方法。并以此划分构建模糊分类规则前件，从而改善模糊规则的有效性和可解释性，以及模糊分类器的稳定性。为表述方便，将该模糊分类器记为MBESL-MA。

2 MA 模糊分类系统

MA 模型系统由若干条模糊规则构成，模糊推理以“IF ... THEN ...”的形式把输入空间的模糊集映射到输出空间的模糊集。对于第 m 条模糊规则 R^m ，其描述如式(1)：

$$\begin{aligned} R^m : & \text{IF } x_1 \text{ is } A_1^{(m)}, \dots, x_D \text{ is } A_D^{(m)} \\ & \text{THEN } y \text{ is } B^{(m)}(b_m, v_m) \end{aligned} \quad (1)$$

其中， x_i ($i = 1, 2, \dots, D$) 为输入向量 $\mathbf{x} = [x_1, \dots, x_i, \dots, x_D]^T$ 的第 i 个分量， $A_i^{(m)}$ 为第 m 条模糊规则中第 i 个分量上的输入空间的模糊子集。 $B^{(m)}(b_m, v_m)$ 表示第 m 条模糊规则中以 b_m 为质心和以 v_m 为模糊系数的输出空间的模糊集。当合取、蕴含和析取操作算子分别采用乘算子、乘算子和加算子时，且输出空间

模糊集为单点集(即输出为精确值)，则 MA 模糊系统最终输出为

$$y_o = \frac{\sum_{m=1}^M (b_m \mu_m(\mathbf{x}))}{\sum_{k=1}^M \mu_k(\mathbf{x})} \quad (2)$$

式中， m 为模糊规则总数， $\mu_m(\mathbf{x})$ 为第 m 条模糊规则的模糊隶属度函数，由 IF 部分中各维对应的隶属度值相乘获得，即

$$\mu_m(\mathbf{x}) = \prod_{i=1}^D \mu_{A_i^{(m)}}(x_i) \quad (3)$$

当模糊分类系统的模糊隶属度函数采用高斯函数时，第 k 条模糊规则中第 i 个分量上模糊集 A_i^k 高斯隶属度函数为

$$\mu_{A_i^k}(x_i) = \exp\left\{-\frac{(x_i - c_i^{(k)})^2}{2\lambda r^{(k)}}\right\}, \quad i \in \{1, 2, \dots, D\}, k \in \{1, 2, \dots, K\} \quad (4)$$

式中， λ 为重叠系数，该系数可以根据经验来确定，也可以用交叉验证的方式进行确定。 $c_i^{(k)}$ ， $r^{(k)}$ 分别为模糊集 A_i^k 高斯隶属度函数的中心和宽度。则第 k 条模糊规则的模糊隶属度函数为

$$\mu_k(\mathbf{x}) = \exp\left\{\sum_{i=1}^D \left[-\frac{(x_i - c_i^{(k)})^2}{2\lambda r^{(k)}}\right]\right\}, \quad k \in \{1, 2, \dots, K\} \quad (5)$$

令：

$$\mu_m(\mathbf{x}) = \mu_m(\mathbf{x}) / \sum_{k=1}^M \mu_k(\mathbf{x}) \quad (6)$$

由式(2)有

$$y_o = \tilde{\mu}^T(\mathbf{x})\mathbf{b} \quad (7)$$

其中， $\tilde{\mu}^T(\mathbf{x}) = [\mu_1(\mathbf{x}), \mu_2(\mathbf{x}), \dots, \mu_K(\mathbf{x})]$ ， $\mathbf{b} = [b_1, b_2, \dots, b_K]^T$ 。当将该 MA 模糊系统用于二分类时，通过在给定数据集上的学习，确定 MA 模糊系统后件参数 b_k ($k = 1, 2, \dots, K$)。对于数据 x ，其所属类别 l 由式(8)进行预测。

$$l = \begin{cases} +1, & y_o > 0 \\ -1, & y_o < 0 \end{cases} \quad (8)$$

3 MBESL-MA 分类系统模糊规则前件提取

为表述方便，记样本集为 $\{\mathbf{x}^{(n)}, y^{(n)} | n = 1, 2, \dots, N\}$ ，其中 $\mathbf{x}^{(n)} = [x_{n1}, \dots, x_{nd}, \dots, x_{nD}]$ ， $y^{(n)} \in \{1, -1\}$ 。对任意样本 $\mathbf{x}^{(n)}$ ，存在一个以其为球心的，半径最大，且其所包含的样本标签均为 $y^{(n)}$ 的超球体。根据该超球体的球心，半径，所包含的样本以及样本的标签，定义一个四元组 $\Phi^{(n)}$ ，即 $\Phi^{(n)} = (\mathbf{c}^{(n)}, r^{(n)}, \Theta^{(n)}, l^{(n)})$ ，其中 $\mathbf{c}^{(n)}$ 为超球体中心， $\mathbf{c}^{(n)} = \mathbf{x}^{(n)}$ ， $\Theta^{(n)}$ 为该超球体所包含的全部样本所构成的集合，即 $\Theta^{(n)} = \{\mathbf{x}^{(t)} | \mathbf{x}^{(t)} - \mathbf{c}^{(n)} \leq r^{(n)}, y^{(n)} = y^{(t)}\}$ 。 $r^{(n)}$ 为超球

体半径, 其值按式(9)确定, 即

$$r^{(n)} = \min \|\mathbf{x}^{(n)} - \mathbf{x}^{(j)}\|, y^{(j)} \neq y^{(n)} \quad (9)$$

式中 $y^{(n)}$, $y^{(j)}$ 分别为 $\mathbf{x}^{(j)}$, $\mathbf{x}^{(n)}$ 的类别标签。由式(1)可以看出, $r^{(n)}$ 为样本 $\mathbf{x}^{(n)}$ 到各个异类样本之间的最近距离, 也就是以样本 $\mathbf{x}^{(n)}$ 为球心的包含同类样本的最大超球体的半径。即对样本 $\mathbf{x}^{(k)}$, 如果 $\mathbf{x}^{(k)} - \mathbf{x}^{(n)} \leq r^{(n)}$, 则 $y^{(n)} = y^{(k)}$, $\mathbf{x}^{(k)} \in \Theta^{(n)}$ 。 $l^{(n)}$ 为超球体内样本的类标签。

记所有四元组 $\Phi^{(n)}$ 所构成的集合为 $\Omega = \{\Phi^{(n)} | n = 1, 2, \dots, N\}$ 。构造集合 Ω 的流程如表 1 的算法 1 所示。

表 1 构造集合 Ω 的流程

算法 1: 构造 Ω
步骤 1 初始化 $\Omega = \emptyset$;
步骤 2 令 $i = 1$;
步骤 3 $\Theta^{(i)} = \emptyset$, $\mathbf{c}^{(i)} = \mathbf{x}^{(i)}$;
步骤 4 按照公式(1)确定 $r^{(i)}$;
步骤 5 令 $k = 1$;
步骤 6 如果 $y^{(k)} = y^{(i)}$, 并且 $\mathbf{x}^{(k)} - \mathbf{x}^{(i)} \leq r^{(i)}$, 则 $\Theta^{(i)} = \Theta^{(i)} \cup \{\mathbf{x}^{(k)}\}$;
步骤 7 $k = k + 1$ 如果 $k < N$ 转到步骤 6, 否则 $\Omega = \Omega \cup \{\Theta^{(i)}\}$;
步骤 8 $n = n + 1$, 如果 $n < N$, 转到步骤 3;
步骤 9 结束。

由四元组的定义, 可以得出以下结论。

结论 1: 对于 Ω 中任意两个四元组 $\Phi^{(p)} = \{\mathbf{c}^{(p)}, r^{(p)}, \Theta^{(p)}, l^{(p)}\}$, $\Phi^{(q)} = \{\mathbf{c}^{(q)}, r^{(q)}, \Theta^{(q)}, l^{(q)}\}$, 如果样本 $\Theta^{(p)} \subseteq \Theta^{(q)}$, 则 $l^{(p)} = l^{(q)}$ 。

证明 因为 $\Theta^{(p)} \subseteq \Theta^{(q)}$, 所以 $\Theta^{(p)}$ 与 $\Theta^{(q)}$ 至少有一个共同的样本 $\mathbf{x}^{(k)}$, 根据四元组定义, 有 $l^{(p)} = y^{(k)}$, $l^{(q)} = y^{(k)}$, 则 $l^{(p)} = l^{(q)}$ 。 证毕

结论 2: 对于 Ω 中任意两个四元组 $\Phi^{(p)} = \{\mathbf{c}^{(p)}, r^{(p)}, \Theta^{(p)}, l^{(p)}\}$, $\Phi^{(q)} = \{\mathbf{c}^{(q)}, r^{(q)}, \Theta^{(q)}, l^{(q)}\}$, 如果 $l^{(p)} \neq l^{(q)}$, 则 $\Theta^{(p)} \cap \Theta^{(q)} = \emptyset$, 即 $\Theta^{(p)}$ 与 $\Theta^{(q)}$ 必定互不包含。

证明 如果 $\Theta^{(p)} \cap \Theta^{(q)} \neq \emptyset$, 则至少存在一个样本 $\mathbf{x}^{(i)}$, $\mathbf{x}^{(i)} \in \Theta^{(p)}$ 且 $\mathbf{x}^{(i)} \in \Theta^{(q)}$, 则 $y^{(i)} = y^{(q)}$ 且 $y^{(i)} = y^{(p)}$, 则 $y^{(p)} = y^{(q)}$, 与 $y^{(p)} \neq y^{(q)}$ 矛盾, 则假设不成立, 命题成立。

四元组的各个元素给出了样本空间的一个划分。如果采用高斯函数为模糊规则的隶属度函数, 那么根据四元组 $\Phi^{(n)}$ 中的 $\mathbf{x}^{(n)}$ 和 $r^{(n)}$ 就可以确定高斯函数的中心和宽度。

在模糊系统中, 模糊规则的选择一般要遵循两

个原则, 一个是覆盖原则, 即所有模糊规则对样本空间的划分要尽可能覆盖大部分样本, 另外一个就是最小冗余原则, 即模糊规则数尽可能少, 以降低系统复杂性, 提高可解释性。由算法 1 可知 $\bigcup_{q=1}^N \Theta^{(q)}$ 包含所有样本, 因此满足覆盖性要求。同时, 多数样本会被多个四元组中的样本集包含, 因此直接根据 Ω 建立模糊规则, 那么模糊规则数量会有大量冗余, 增加学习空间复杂度和时间复杂度, 也容易造成过拟合, 因此有必要对集合 Ω 中的四元组进行删减, 只需保留 1 部分四元组即可。

集合 Ω 中的四元组删减分为两步, 第 1 步是根据各四元组中的样本集合 $\Theta^{(i)}$ 之间的包含关系进行约简。具体约简流程如表 2 的算法 2 所示。

表 2 约简 Ω 流程

算法 2: 根据样本集合包含关系来约简 Ω
步骤 1 初始化 $\text{res} = \text{ones}(N, 1)$, $i = 1$, $j = i + 1$;
步骤 2 如果 $\Theta^{(i)} \subseteq \Theta^{(j)}$, 则 $\text{res}(i, 1) = 0$;
步骤 3 如果 $\Theta^{(i)} \supset \Theta^{(j)}$, 则 $\text{res}(j, 1) = 0$;
步骤 4 $j = j + 1$, 如果 $j > N$ 转到步骤 6, 否则转到步骤 2;
步骤 5 $i = i + 1$, 如果 $i > N$ 转到步骤 6, 否则 $j = i + 1$, 转到步骤 2;
步骤 6 对于任意 $k \in \{1, 2, \dots, N\}$, 如果 $\text{res}(k, 1) = 0$, 则从 Ω 中删除 $\Phi^{(k)}$ 。

经过算法 2 处理之后, 虽然 Ω 中任意两个元素所含的样本集合都不存在包含关系, 但还可能存在部分四元组, 其样本集中的所有样本同时也包含在其他四元组所对应的样本集中, 这样也存在重复包含, 因此可以将其删除。即对某个四元组 $\Phi^{(i)}$ ($i \in \{1, 2, \dots, N\}$), 如果 $\Theta^{(i)} - \bigcup_{j \neq i} \Theta^{(j)} = \emptyset$, 则将该四元组删除。而当 $\Theta^{(i)} - \bigcup_{j \neq i} \Theta^{(j)} \neq \emptyset$, 即该四元组所对应的样本集中含有独有样本时, 则保留 $\Phi^{(i)}$ 。为保证模糊规则的覆盖性, 每次只能删除 1 个不含独有样本的四元组, 然后继续检查各个四元组的样本集是否都含有独有样本, 直至所有四元组的样本集均含有独有样本。每次删除时, 依据一定的法则从不含独有样本的四元组中选择 1 个进行, 则继续依据法则从这些非独有样本的四元组中挑选出 1 个予以删除, 直至集合 Ω 中的四元组均含有独有样本。挑选的法则可以是半径最大的四元组, 也可以是半径最小的四元组, 分别记为 max_r 和 min_r 。为进一步减少模糊规则冗余, 将独有样本数量少于预定的阈值 omit_num 的四元组也进行删除。此外, 独

有样本数量较少的四元组往往位于两类边界区域，将其删除有利于降低过拟合的可能性，提高模糊分类系统的泛化性能。

由结论 1 和结论 2，在约简时无需对每两个四元组中的同类集合 $\Theta^{(i)}$ 进行子集合并操作，只需对具有相同类标签的四元组进行子集合并操作即可。据此可以得到如表 3 的算法 3 所示的四元组约简算法。

表 3 四元组约简算法

算法 3	无独有元素 Ω 约简
步骤 1	初始化 $\text{monopoly}=\text{zeros}(1,N)$, $\text{reduction_regulation}$, omit_num ;
步骤 2	令 $j=1$;
步骤 3	如果 $\text{monopoly}(j)=0$, 转到步骤 4, 否则转到步骤 5;
步骤 4	如果 $ \Theta^{(j)} - \bigcup_{j \neq i, i \neq j} \Theta^{(i)} > \text{omit_num}$, 则令 $\text{monopoly}(j)=1$;
步骤 5	$j=j+1$, 如果 $j < N$, 则转到步骤 3;
步骤 6	如果 $\text{reduction_regulation}=1$, 则 $\text{index} = \text{argmax}_{\text{monopoly}(k)=0} r^{(k)}$;
步骤 7	如果 $\text{reduction_regulation}=2$, 则 $\text{index} = \text{argmin}_{\text{monopoly}(k)=0} r^{(k)}$;
步骤 8	令 $\text{monopoly}(\text{index})=1$;
步骤 9	如果 $\text{any}(\text{monopoly})=0$, 转到步骤 2;
步骤 10	对于任意 $t \in \{1, 2, \dots, N\}$, 如果 $\text{monopoly}(t)=1$, 则保留 $\Phi^{(t)}$.

记约简后的四元组集 $\Omega_R = \{\Phi^{(k)} \mid k=1, 2, \dots, K\}$ 。根据 Ω_R 内四元组中的 $\mathbf{c}^{(k)}$, $r^{(k)}$ 就可以确定模糊规则前件参数。

4 MA 模糊二分类系统规则后件参数学习

文献[14]在模糊后件参数求解中，为提高分类系统泛化性能，在学习目标函数中引入了 SVM 最大分类间隔的思想，并采用迭代最小二乘算法进行后件参数求解^[15]。即后件参数 b 应满足式(10)条件：

$$y_o = \begin{cases} \mu(\mathbf{x}^{(i)})^T b > 1, & y^{(i)} = 1 \\ \mu(\mathbf{x}^{(i)})^T b < -1, & y^{(i)} = -1 \end{cases} \quad (10)$$

即

$$y_o = y^{(i)} \mu(\mathbf{x}^{(i)})^T b > 1 \quad (11)$$

模糊系统后件参数 b 的加权平方分类损失目标函数为

$$\min J(y) \triangleq \sum_{i=1}^N \frac{h_i}{2} \left(y^{(i)} \mu(\mathbf{x}^{(i)})^T b - 1 \right)^2 \quad (12)$$

其中，参数 h_i 为第 i 个训练样本所对应的加权因子，

该因子是由学习样本所对应的分类决策函数的误差来决定。即当样本 $\mathbf{x}^{(i)}$ 被正确分类时，加权因子 h_i 为 0，分类误差为 0，否则 h_i 为 1，对应的分类误差不为 0。

目标函数的向量表达式为

$$\min J(b) \triangleq \frac{1}{2} (\tilde{\mu}(\mathbf{x})b - 1)^T \mathbf{H} (\tilde{\mu}(\mathbf{x})b - 1) \quad (13)$$

其中， $\mathbf{H} = \text{diag}(h_1, h_2, \dots, h_N)$ 。由于参数 h_i 是与模糊系统后件参数 b 相关，为此引入迭代加权方法进行求解。则目标函数和对角矩阵 \mathbf{H} 中 h_i 的迭代公式分别为

$$\min J^{(k)}(b^{(k)}) \triangleq \frac{1}{2} (\tilde{\mu}(\mathbf{x})b^{(k)} - 1)^T \mathbf{H}^{(k)} (\tilde{\mu}(\mathbf{x})b^{(k)} - 1) \quad (14)$$

$$h_i^{(k+1)} = \begin{cases} 0, & y^{(i)} \mu(\mathbf{x}^{(i)})^T b^{(k)} - 1 \geq 0 \\ 1, & y^{(i)} \mu(\mathbf{x}^{(i)})^T b^{(k)} - 1 < 0 \end{cases} \quad (15)$$

模糊规则后件参数 b 加权迭代最小平方误差求解流程如表 4 的算法 4 所示。

表 4 b 加权迭代最小平方误差求解流程

算法 4	模糊规则后件参数 b 加权迭代最小平方误差求解
步骤 1	初始化 $\mathbf{H}^{(0)} = \mathbf{I}$, $\text{loop}=0$;
步骤 2	使用共轭梯度法求解方程 $\tilde{\mu}(\mathbf{x})^T \mathbf{H}^{(k)} \tilde{\mu}(\mathbf{x}) b = \tilde{\mu}(\mathbf{x})^T \mathbf{H}^{(k)} \mathbf{1}_{N \times 1}$;
步骤 3	按照式(15)更新矩阵 $\mathbf{H}^{(k)}$ 元素 $h_i^{(k+1)}$, 其中 $i=1, 2, \dots, N$;
步骤 4	如果 $\text{loop} > 1$, 并且 $b^{(k+1)} - b^{(k)} < 10^{-3}$, 则迭代结束, 否则 $\text{loop} = \text{loop} + 1$, 转到步骤 2。

5 分类实验结果及讨论

为验证本方法的有效性，从 <http://www.keel.es/> 中选取 `australian`, `Bands`, `breast_cancer`, `bupa`, `ionosphere`, `haberman`, `mammographic`, `monk-2`, `pima`, `spectfheart`, `wdbc`, `wisconsin` 这 12 个 10 折验证数据集进行分类实验，同时与文献[8]中的 FCM-IRLS, F(C+P)-IRLS 的分类系统，以及 SVM 方法进行对比分类实验。所选数据集的特征如表 5 所示。实验中，FCM-IRLS 方法中的聚类数的选择集为 $\{5, 10, 15, 20, 25, 30, 40, 50\}$, F(C+P)-IRLS 中的各类的聚类数的选择集为 $\{5, 7, 9, 11, 13, 15, 17, 19\}$ 。与文献[8]相同，聚类操作不区分样本类别，即是在所有学习样本上进行聚类，而不是按照样本类别分别进行聚类。

MBESL-MA 方法中的约简原则 `method` 和约简数量 `omit_num` 设置如表 6 所示。FCM-IRLS, F(C+P)-IRLS, MBESL-MA 中的宽度因子选择集

初始值为 0.1, 结束值为 5, 步长为 0.1。实验中 SVM 算法采用学者林智仁所提供的代码(网址: <https://www.csie.ntu.edu.tw/~cjlin/libsvm/>), 该算法所用核函数为高斯核函数, 参数 C 和 g 的选择集初始值均为 -10, 结束值为 10, 步长为 1。各个算法中的参数选择是在训练集上以 20 折交叉验证的方式进行。相关参数获取后即在测试集上进行分类试验, 除 SVM 是以 MATLAB 和 C 混编外, 其余代码均以 MATLAB 语言进行编写。表 6 给出了相关算法在数据集上的 10 折分类错误平均值和方差。

从表 6 可以看出, 在这 12 个数据集中, MBESL-IRLS 模糊系统虽然在 bands 和 bupa 数据这两个数据集上效果最差, 但在 breast_cancer, ionosphere, haberman, mammographic, pima, spectfheart,

wdbc, wisconsin 这 8 个数据集上获得最好的分类效果。在 monk-2 数据集上, SVM 方法分类错误为零, MBESL-IRLS 模糊系统虽然不能取得相同的分类效果, 但要优于 FCM-IRLS 模糊系统和 F(C+P)-IRLS 模糊系统。三者模糊分类系统的差别仅在于模糊规则初始化方法不同, 也说明 MBESL-IRLS 模糊系统所构建模糊规则更为有效。在 australian 数据集上虽然优于 SVM 方法, 但比 FCM-IRLS 模糊系统和 F(C+P)-IRLS 模糊系统要差, 原因可能在于 FCM-IRLS 模糊系统和 F(C+P)-IRLS 模糊系统的规则前件划分是椭圆形, 而 MBESL-IRLS 模糊系统是球形, 而这两个数据集分布更适合用椭球划分, 也可能模糊规则约简机制不适用于该数据集。

为考察约简方法 method 和忽略数量 omit_num 对系统分类性能的影响, 选取 ionosphere 数据集采用不同的约简方法和忽略数量进行分类实验。10 折分类实验平均模糊规则数和分类错误分别如图 1, 图 2 所示。

从图 1 可以看出, 不论是采用何种约简方法, 随着忽略数量的增加, 模糊规则数随之减少, 并且 max_r 约简方法所对应的模糊规则数要略少于 min_r 约简方法。从图 2 可知, 相同的忽略数量下, min_r 约简方法分类错误要好于 max_r 约简方法。此外, 不论采用何种约简方法, 随忽略数量的增加, 分类误差先下降再增加。特别当忽略数量为 0 时, 就会有只含单个样本的同类球, 这些样本往往位于两类边界中, 保留这些同类球会造成过拟合, 造成模糊分类系统泛化性能下降, 同时还使模糊规则过多, 影响学习效率。当忽略数量过大时会造成欠拟合, 导致分类性能下降。

表 5 实验数据集列表

数据集名称	维数	样本数
pima	9	768
ionosphere	33	351
haberman	3	306
wisconsin	9	683
mammographic	5	830
bupa	6	345
bands	19	365
australian	14	690
breast_cancer	9	277
wdbc	30	569
spectfheart	44	267
monk-2	6	432

表 6 MBESL-MA 参数设置与对比算法分类性能列表

数据集名称	MBESL-IRLS 算法参数		Ce(分类错误)			
	method	omit_num	MBESL-MA	FCM-IRLS	F(C+P)M-IRLS	SVM
australian	3	2	31.45±3.2	16.96±4.4	15.47±4.5	46.38±3.6
bands	2	3	36.69±3.7	36.13±2.9	36.68±3.8	32.26±0.9
breast_cancer	3	3	26.91±2.4	34.94±3.0	31.22±2.7	33.33±1.4
bupa	3	4	27.30±3.5	21.43±4.1	22.62±4.2	22.86±2.5
ionosphere	3	2	6.68±2.5	14.86±1.0	15.23±1.8	8.57±1.3
haberman	3	3	24.87±2.9	30.00±2.9	25.47±2.3	27.00±4.7
mammographic	3	3	16.71±3.1	22.09±3.4	23.46±1.6	22.09±2.2
monk-2	3	3	3.21±1.8	18.18±5.6	15.18±3.2	0±0
pima	3	2	21.35±3.6	27.40±1.5	26.79±1.7	28.57±0.5
spectfheart	3	4	22.07±2.3	22.69±2.2	23.74±2.6	26.92±1.7
wdbc	3	4	6.68±1.4	21.03±0.9	19.24±3.9	10.71±2.1
wisconsin	3	3	2.47±1.1	4.48±1.3	3.91±1.1	5.97±0.4

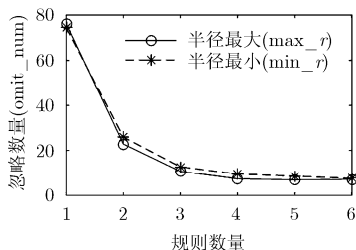


图1 不同约简方法和忽略数量设置时

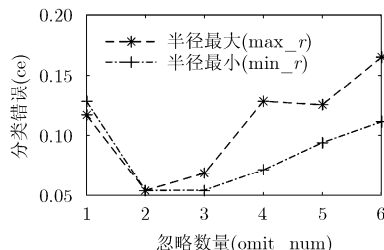


图2 不同约简方法和忽略数量设置时的平均分类错误

6 结束语

从分类实验总体结果来看, MBESL-MA 分类器在大多数数据集上能取得更好的分类效果。该分类器的特点一是模糊规则所对应的模糊划分中只包含同类样本, 因此解释性更好; 二是通过忽略数量参数设置可以在一定程度上避免过拟合的出现; 三是当约简方法和忽略数量一定时, 模糊划分是确定不变的。虽然 MBESL-MA 分类器有两个参数, 但是这两个参数都是有限可列的。MBESL-MA 分类器中的模糊划分是球形划分, 如果能将其推广到椭圆划分, 也许能取得更好的分类效果。此外在规则约简中, 是根据集合包含的特有元素来进行约简, 也许还存在更好的约简方法, 这些将是今后工作中的需要进一步探索的问题。

参考文献

- [1] HARANDI F A and DERHAMI V. A reinforcement learning algorithm for adjusting antecedent parameters and weights of fuzzy rules in a fuzzy classifier[J]. *Journal of Intelligent & Fuzzy Systems*, 2016, 30(4): 2339-2347. doi: 10.3233/IFS-152004.
- [2] JAMALABADI H, NASROLLAHI H, ALIZADEH S, et al. Competitive interaction reasoning: A bio-inspired reasoning method for fuzzy rule based classification systems[J]. *Information Sciences*, 2016, 352: 35-47. doi: 10.1016/j.ins.2016.02.052.
- [3] CINTRA M E, CAMARGO H A, and MONARD M C. Genetic generation of fuzzy systems with rule extraction using formal concept analysis[J]. *Information Sciences*, 2016, 349: 199-215. doi: 10.1016/j.ins.2016.02.026
- [4] POURPANAHA F, LIM C P, and MOHAMAD SALEHA J. A hybrid model of fuzzy ARTMAP and genetic algorithm for data classification and rule extraction[J]. *Expert Systems With Applications*, 2016, 49: 74-85. doi: 10.1016/j.eswa.2015.11.009.
- [5] 李继东, 张学杰. 基于遗传算法的多维模糊分类器构造的研究[J]. *软件学报*, 2005, 16(5): 779-785.
- [6] LI J D and ZHANG X J. Research on the construction of fuzzy classifier system for multidimensional pattern classification using genetic algorithms[J]. *Journal of Software*, 2005, 16(5): 779-785.
- [7] RUDZINSKI F. A multi-objective genetic optimization of interpretability-oriented fuzzy rule-based classifiers[J]. *Applied Soft Computing*, 2016, 38: 118-133. doi: 10.1016/j.asoc.2015.09.038.
- [8] MARIAN B G and RUDZINSKI F. A multi-objective genetic

- optimization for fast, fuzzy rule-based credit classification with balanced accuracy and interpretability[J]. *Applied Soft Computing*, 2016, 40: 206-220. doi: 10.1016/j.asoc.2015.11.037.
- [9] SHANGHOOSHABAD A M and ABADEH M S. Robust, interpretable and high quality fuzzy rule discovery using krill herd algorithm[J]. *Journal of Intelligent and Fuzzy Systems*, 2016, 30(3): 1601-1612. doi: 10.3233/IFS-151867
- [10] GARCÍA-GALÁN S, PARDO P R, and MUNOZ EXPÓSITO J E. Rules discovery in fuzzy classifier systems with PSO for scheduling in grid computational infrastructures[J]. *Applied Soft Computing*, 2015, 29: 424-435. doi: 10.1016/j.asoc.2014.11.064.
- [11] WU Jue, YANG Lei, LI Tianrui, et al. Rule-based fuzzy classifier based on quantum ant optimization algorithm[J]. *Journal of Intelligent & Fuzzy Systems*, 2015, 29 (6): 2365-2371. doi: 10.3233/IFS-151935.
- [12] MAHDIZADEH M and EFTEKHARI M. Generating fuzzy rule base classifier for highly imbalanced datasets using a hybrid of evolutionary algorithms and subtractive clustering[J]. *Journal of Intelligent and Fuzzy Systems*, 2014, 27(6): 3033-3046. doi: 10.3233/IFS-141261.
- [13] 邢宗义, 张永, 侯远龙, 等. 基于模糊聚类和遗传方法的具备解释性和精确性的模糊分类系统设计[J]. *电子学报*, 2006, 34(1): 83-88.
- [14] XING Zongyi, ZHANG Yong, HOU Yuanlong, et al. Design of interpretable and precise fuzzy classification system based on fuzzy clustering and genetic algorithm[J]. *Acta Electronica Sinica*, 2006, 34(1): 83-88.
- [15] 王莉, 周献中, 李华雄. 基于决策粗糙集的模糊分类模型[J]. *信息与控制*, 2014, 43(1): 24-29. doi: 10.3724/SP.J.1219.2014.00024.
- [16] WANG Li, ZHOU Xianzhong, and LI Huaxiong. Fuzzy classification model based on decision-theoretic rough set[J]. *Information and Control*, 2014, 43(1): 24-29. doi: 10.3724/SP.J.1219.2014.00024.
- [17] JACK M L. Fuzzy ($c + p$)-means clustering and its application to a fuzzy rule-based classifier: Towards good generalization and good interpretability[J]. *IEEE Transactions on Fuzzy Systems*, 2015, 23(4): 802-812. doi: 10.1109/TFUZZ.2014.2327995
- [18] LESKI J M. Iteratively reweighted least squares classifier and its ℓ_2 - and ℓ_1 -regularized kernel versions[J]. *Bulletin of the Polish Academy of Sciences: Technical Sciences*, 2010, 58(1): 171-182. doi: 10.2478/v10175-010-0018-2.

徐明亮: 男, 1973 年生, 博士, 副教授, 研究方向为人工智能与模式识别。

王士同: 男, 1964 年生, 教授, 博士生导师, 研究方向为人工智能与模式识别、数据挖掘、模糊系统等。