

基于模糊区域分布的分类规则提取及推理算法

李 洁^{1),2)} 邓一鸣²⁾ 沈士团²⁾

¹⁾(信息产业部电信研究院通信标准研究所 北京 100045)

²⁾(北京航空航天大学电子信息工程学院 北京 100083)

摘 要 基于不同分类的样本在各规则对应模糊区域的隶属度分布,定义了一种规则相对匹配度,比分类匹配度更能体现样本在不同模糊区域的分布对比.设计了模糊区域分布矩阵,由该矩阵可以算得规则相对匹配度和分类匹配度,并提出了基于规则相对匹配度的分类规则提取算法,同分类匹配度算法相比,该算法充分考虑了每条规则之间的隶属度分布对比,同时以各分类样本的相对数量作为加权系数,从而兼顾了学习空间的全局密度优势和局部数量优势.通过解模糊器实现了基于规则的分类推理,其推理过程比以往算法具有更好的解释性和简洁性.最后,由 Iris 数据和 Wine 数据的分类实验证明:无论样本数量均衡与否,由规则相对匹配度提取规则都具有更好的分类效果.

关键词 分类规则;分类推理;解模糊器;模糊区域分布
中图法分类号 TP181

Classification Rule Extraction Based on Fuzzy Area Distribution and Classification Reasoning Algorithm

LI Jie^{1),2)} DENG Yi-Ming²⁾ SHEN Shi-Tuan²⁾

¹⁾(Institute of Communication Standards Research, China Academy of Telecommunication Research of MII, Beijing 100045)

²⁾(School of Electronics and Information Engineering, Beijing University of Beihang, Beijing 100083)

Abstract A definition of rule relative compatibility is proposed based on membership degree distribution on the fuzzy area corresponding to every rule. Compared with the class compatibility, rule relative compatibility is able to show more information about the samples distribution difference among every fuzzy area. The Fuzzy Area Distribution Matrix is designed, by which the rule relative compatibility and class compatibility are calculated respectively. Moreover, the algorithm of classification rule extraction is put forward by rule relative compatibility. Different from the class compatibility approach, the one based on rule relative compatibility contains the membership degree distribution comparison of every rule. Furthermore, this compatibility is weighted by the relative amount of every class samples so that it can take the consideration of the global density dominance as well as the local quantity dominance for study space. In addition, the classification reasoning based on fuzzy rule is implemented by the defuzzifier algorithm. The procedure is better than the previous algorithm due to its interpretability and simplicity. In the end, Iris data and Wine data are used to validate the proposed algorithm of fuzzy classification rule extraction. The testing results prove that whether the sample is distributed homogeneously or not, the rule extraction approach based on rule relative compatibility attains higher classification rate.

Keywords classification rule; classification reasoning; defuzzifier; fuzzy area distribution

收稿日期:2007-04-26;最终修改稿收到日期:2008-04-09. 本课题得到“十一五”国防预研基金项目资助. 李 洁,女,1979 年生,博士,工程师,主要研究方向为网络安全、人工智能、模式识别、故障诊断等. E-mail: lucy_lj@hotmail.com. 邓一鸣,男,1979 年生,硕士,主要研究方向为网络性能分析和人工智能等. 沈士团,男,1940 年生,教授,博士生导师,主要研究领域为武器装备自动测试与智能故障诊断、卫星导航与通信等.

1 引言

近年来,模糊理论在模式分类中的应用得到越来越多学者的关注.作为人工智能的一个分支,模式分类需要计算机具有类似人类的推理和思维模式,模糊集合和模糊命题都具有模糊性,基于它们的推理可以很大程度上反映人类对客观事物的认知和判断过程;除此之外,模糊规则具有较强的解释性,利用模糊规则进行模式分类便于人们对推理过程的深刻理解.

如何构建初始规则集是设计模糊分类系统的关键,国内外学者为此进行了深入而广泛的研究.文献[1]将神经网络应用于模糊分类规则集的产生,该网络是一个四层网络,利用裁剪算法完成对隶属度函数和属性的优选,由于神经网络本身具有知识存储和推理能力,增加提取规则环节降低了神经网络的计算效率;Chen 和 Wang^[2]分析了基于模糊规则的分类系统与核学习机(kernel machine)之间的关系,运用支持向量机(support vector machine)从训练样本中抽取规则,建立正定模糊分类器(positive definite fuzzy classifier),然而,如何根据不同的工程应用选择合适的核函数参数仍然存在困难;文献[3-5]通过聚类算法生成模糊规则,可以减少规则条件部分的属性数量及规则总数,但在精简模糊规则集的同时,也增加了计算复杂度;群体智能和遗传算法也应用于模糊分类规则集的优化;文献[6]将规则编码为粒子,应用粒子群优化算法进行分类规则的提取;文献[7]通过模糊聚类和模糊相似性分析得到约简的模糊分类系统,再由遗传算法对规则集进行优化,提高系统的分类精度;文献[8]结合 k -最近邻法(k -nearest neighbor),应用遗传算法优选分类对象的属性,降低高维分类问题的复杂性.虽然利用群体智能和遗传算法可以达到规则优化的效果,但是这种效果仍然要受到初始规则集优劣程度的制约.

为了在保证分类准确率和分类效率的同时,使规则抽取更易于理解和实现,Ishibuchi 等人设计了模糊分类规则的权值启发算法^[9-10],该算法不需要调整隶属度函数形状和特征参数,通过适当设置规则权值提高分类正确率.由于 Ishibuchi 等人在确定规则结论和规则权值时,仅强调各分类样本在同一规则指定模糊区域的分布对比,因此,无法全面衡量

分类与规则的匹配程度.本文定义了一种规则相对匹配度,在考虑单个模糊区域上各分类样本分布的同时,更强调了学习空间各模糊区域的全局分布.通过对 Iris 数据和 Wine 数据进行分类实验,证明基于该匹配度的初始规则集生成算法要优于 Ishibuchi 等人的算法.

针对模糊分类系统相对以往模糊系统的特殊性,Cordon 等人总结了之前各种基于规则的分类推理算法,并提出了一种通用的模糊分类推理模型^[12];基于这种模型,文献[13]提出了两种简单的推理方法:单一优胜法(single winner)和投票加权法(weighted vote),这些方法都属于 Cordon 通用推理模型的具体实现;van den Berg 等人应用概率论进行分类推理公式的推导^[14],得到了与文献[13]类似的结果;为改进最大匹配度原则在充分利用各规则信息上的不足,一致理论(consensus theory)也被用于模式分类的推理^[15-16].本文运用解模糊器分析模糊分类的推理过程,可以得到与文献[13-14]相似的推理公式,从经典模糊论角度给出了一种简单的分类推理算法.

2 基于规则的模糊分类系统

基于规则的模糊分类需要规则提取和模糊推理两个环节,对于有 K 个模式类的分类问题,规则提取要对训练样本空间 $S = \{s^1, s^2, \dots, s^M\}$ 进行学习,得到初始规则集 $R = \{R_1, R_2, \dots, R_N\}$,形式如下:

$$R_j: \text{IF } s_1 \text{ is } A_j^1 \text{ and } \dots \text{ and } s_n \text{ is } A_j^n \\ \text{THEN } y \text{ is } C_j \text{ with } r_j.$$

其中, $s_i (i=1, 2, \dots, n)$ 表示样本的第 i 个属性,样本 $s^t = \{s_1^t, s_2^t, \dots, s_n^t, y^t\}$ 是一个 $n+1$ 维向量,第 $n+1$ 个元素 y^t 为该样本所属的模式类别; $A_j^i (i=1, 2, \dots, n)$ 表示第 j 条规则中第 i 个属性上的模糊集; C_j 表示第 j 条规则的结论对应的模式类别; r_j 表示第 j 条规则的权值.

规则集构建完成之后,要根据模糊推理算法对测试样本空间 T 中的元素进行分类;在规则集产生和利用规则集进行推理的过程中,根据分类的效果对规则权值进行自学习操作^[17],使规则参数不断得到修正,从而达到优化规则集的目的.整个过程如图 1 所示.本文将在下面章节分别讨论初始分类规则集的产生(rules generation)以及基于规则集的模糊分类推理算法(fuzzy reasoning).

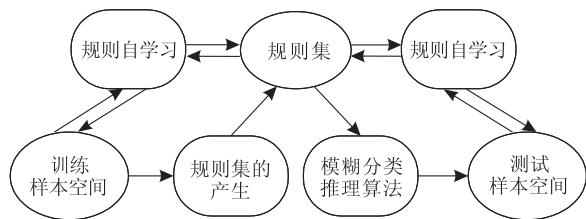


图 1 基于规则的模糊分类过程

3 模糊区域分布矩阵

3.1 规则相对匹配度

对于某一个训练样本 s^t , 它与第 j 条规则的匹配度可由规则条件部分各属性的隶属度函数进行 t -范数运算得到, 这里的 t -范数选用代数积算子, 如式(1):

$$\mu_j(s^t) = \prod_{i=1}^n \mu_j^i(s_i^t) \quad (1)$$

定义 1. 训练样本空间中分类 h 的所有样本在第 j 条规则上的匹配度之和称为分类 h 关于规则 j 的匹配度, 记作 $\gamma_{\text{Class } h}(R_j)$:

$$\gamma_{\text{Class } h}(R_j) = \sum_{\substack{t=1 \\ y^t = \text{Class } h}}^M \mu_j(s^t), \quad h=1, 2, \dots, K \quad (2)$$

定义 2. 分类 h 关于规则 j 的匹配度在各分类关于规则 j 的匹配度上进行归一化得分匹配度^[9-10], 记作 CD_j^h :

$$CD_j^h = \gamma_{\text{Class } h}(R_j) / \sum_{i=1}^K \gamma_{\text{Class } i}(R_j) \quad (3)$$

针对分类匹配度仅局限于一条规则的各类样本分布对比, 不足以体现学习空间的全局分布, 本文提出另一种匹配度——规则相对匹配度, 如下定义。

定义 3. 分类 h 关于规则 j 的匹配度在分类 h 关于各规则的匹配度上进行归一化得 RD_j^h (本文提出的 RD_j^h 概念更能体现学习空间的全局分布), 再由各分类样本的数量比例对 RD_j^h 进行赋权得规则相对匹配度, 记作 WD_j^h :

$$WD_j^h = \frac{\text{COUNT}_h}{M} RD_j^h \quad (4)$$

其中, COUNT_h ($h=1, 2, \dots, K$) 为各分类的样本数,

且 $M = \sum_{i=1}^K \text{COUNT}_i$, $RD_j^h = \gamma_{\text{Class } h}(R_j) / \sum_{i=1}^N \gamma_{\text{Class } h}(R_i)$.

WD_j^h 不仅通过权值体现了样本的区域数量分布, 而且由 RD_j^h 更加突出了样本的全局密度分布。

3.2 模糊区域分布矩阵的计算

通过分析模糊分类规则的形式, 本文将规则条

件部分视为对一个 n 维模糊区域的描述. 例如, 某分类问题的样本存在 2 个属性 (其中不包括样本所属类别). 对各属性均分成 4 个标准三角形模糊区间, 则 2 维学习空间被分成 $16 = 4 \times 4$ 个模糊区域, 如图 2 所示, 其中规则 R_8 条件部分的模糊区域对应于 “ s_1 is SM and s_2 is L”.

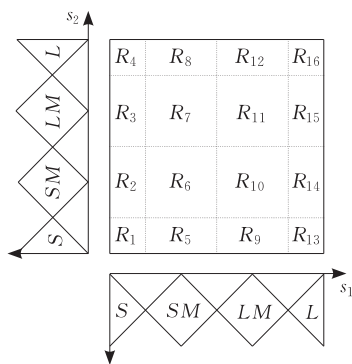


图 2 2 维学习空间的模糊区域划分

为表达样本在学习空间上的全局分布, 本文给出如下定义。

定义 4. 以各分类样本在每一个模糊区域上的分布强度 (即各分类关于每一条规则的匹配度) 为元素构成的一个 $N \times K$ 维矩阵称为模糊区域分布矩阵, 记作 $\text{DIS}_{N \times K}$, 如式(5):

$$\text{DIS}_{N \times K} = \begin{pmatrix} d_{11} & \cdots & d_{1K} \\ \vdots & \ddots & \vdots \\ d_{N1} & \cdots & d_{NK} \end{pmatrix} \quad (5)$$

$$d_{ij} = \gamma_{\text{Class } j}(R_i) = \sum_{\substack{t=1 \\ y^t = \text{Class } j}}^M \mu_i(s^t) \quad (6)$$

由定义 2 和定义 3 可知, 分类匹配度和规则相对匹配度可以分别对模糊区域分布矩阵的行和列进行归一化得到, 如式(7)和式(8)。

由式(7)可知, 分类匹配度表征了第 j 条规则所描述的模糊区域内各分类样本的相对分布, 仅针对某一条指定的规则, 而不考虑其它规则所描述的模糊区域内样本的分布, 因此, 某分类样本可以仅凭借其在局部模糊区域内的数量优势, 获得相对其它分类样本较高的匹配度, 从而使该分类成为规则结论, 因此, 分类匹配度强调学习空间的局部数量优势。

由式(4)和式(8)可知, 规则相对匹配度定义中的 RD_j^h , 表征了分类 h 的所有样本在各条规则所指模糊区域上的相对分布, 有别于分类匹配度仅局限于某一模糊区域. 因此, RD_j^h 体现了学习空间中样本的全局分布对比, 即 RD_j^h 强调样本的全局密度优势. 当某分类样本在学习空间中具有较大的相对数

量优势时,如果仿效分类匹配度,适当增大该分类成为相应模糊区域所对应规则的结论的可能性,可以进一步提高系统的分类正确率,因此,由各分类样本的数量比例对 RD_j^h 赋权可得规则相对匹配度,如式(8),从而兼顾学习空间的全局密度优势和局部数量优势.

$$CD_j^h = d_{jh} / \sum_{i=1}^K d_{ji} \quad (7)$$

$$WD_j^h = \frac{COUNT_h}{M} \left[d_{jh} / \sum_{i=1}^N d_{ih} \right] \quad (8)$$

4 分类规则提取算法

对于有 n 个属性 K 个模式类的分类问题,在提取规则之前,首先确定关于各属性的模糊集,设各属性划分成 L 个标准三角形模糊区间,为尽可能缩短规则长度(规则条件部分包含的有效属性数量),为各属性增加一个特殊模糊区间^[9]: $\mu(s_i^t) = 1$. 如果规则条件部分出现该模糊区间,表示相应的属性对规则条件不加限制.

结合上文定义的规则相对匹配度,本文提出如下分类规则提取算法:

1. 归一化各样本属性^[10],使模糊隶属度函数在数值上具有相同的论域 $[0,1]$,归一化公式如式(9):

$$\bar{s}_i^t = \frac{s_i^t - \min_{j=1}^M (s_j^t)}{\max_{j=1}^M (s_j^t) - \min_{j=1}^M (s_j^t)} \quad (9)$$

2. 初始化完备规则集,在 n 维学习空间中划分模糊区域,确定规则与模糊区域的映射关系. 对于规则条件部分所有属性都取特殊模糊区间: $\mu(s_i^t) = 1$ 的规则,没有意义,不选入完备规则集,因此,完备规则集的规则数量为 $N = (L+1)^n - 1$ 条. 设规则 R_j 条件部分(s_1 is A_1^1 and...and s_n is A_n^q)中的 A_i^j 对应第 F_j^i 个模糊区间,其中: $j = 1, 2, \dots, N$, $i = 1, 2, \dots, n$, $F_j^i = 0, 1, \dots, L$, 则规则索引 j 与条件各属性的模糊区间取值序列 $F_j^n F_j^{n-1} \dots F_j^1$ 之间是 10 进制与 $L+1$ 进制的转换关系^[18], 如式(10). 当 $F_j^i = 0$ 时,取特殊模糊区间.

$$F_j^n \cdot (L+1)^{n-1} + F_j^{n-1} \cdot (L+1)^{n-2} + \dots + F_j^2 \cdot (L+1) + F_j^1 = j \quad (10)$$

3. 由初始完备规则集和属性归一化的学习空间,算得模糊区域分布矩阵 $DIS_{N \times K}$.

4. 根据分类匹配度或者规则相对匹配度确定规则结论,每条规则的结论对应一个分类. 第 j 条规则的结论为 IF 部分所描述的模糊区域对应的分类 $C_j = \text{Class } h$, $h = 1, 2, \dots, K$. 由模糊区域分布矩阵算得分类匹配度和规则相对匹配度,基于这两种匹配度,分别对应两种结论确定方法,如式(11)和式(12),匹配度最大的分类成为规则结论.

$$CD_j^{C_j} = \max\{CD_j^1, CD_j^2, \dots, CD_j^K\} \quad (11)$$

$$WD_j^{C_j} = \max\{WD_j^1, WD_j^2, \dots, WD_j^K\} \quad (12)$$

5. 根据步 4 确定规则结论的不同方法,选用相应的计算公式获得规则权值. 文献[10]基于分类匹配度提出了 4 种权值算法,如式(13);本文基于规则相对匹配度提出 4 种新的权值算法,如式(14). 第 6 节将结合实验结果重点讨论应用本文提出的基于式(14)的规则权值算法在分类效果上比式(13)的优越性.

$$\begin{cases} r_j^I = CD_j^{C_j} \\ r_j^{II} = CD_j^{C_j} - \frac{1}{K-1} \sum_{i=1, i \neq C_j}^K CD_j^i \\ r_j^{III} = CD_j^{C_j} - \frac{1}{\text{MAX}_{i=1, i \neq C_j}^K (CD_j^i)} \sum_{i=1, i \neq C_j}^K CD_j^i \\ r_j^{IV} = CD_j^{C_j} - \sum_{i=1, i \neq C_j}^K CD_j^i \end{cases} \quad (13)$$

$$\begin{cases} r_j^I = WD_j^{C_j} / \sum_{i=1}^K WD_j^i \\ r_j^{II} = \left(WD_j^{C_j} - \frac{1}{K-1} \sum_{i=1, i \neq C_j}^K WD_j^i \right) / \sum_{i=1}^K WD_j^i \\ r_j^{III} = \left(WD_j^{C_j} - \frac{1}{\text{MAX}_{i=1, i \neq C_j}^K (WD_j^i)} \sum_{i=1}^K WD_j^i \right) / \sum_{i=1}^K WD_j^i \\ r_j^{IV} = \left(WD_j^{C_j} - \sum_{i=1, i \neq C_j}^K WD_j^i \right) / \sum_{i=1}^K WD_j^i \end{cases} \quad (14)$$

5 分类推理算法

5.1 通用分类推理模型

Cordon 等人在文献[12]中提出了一种用于模式分类的通用模糊推理模型,分为如下 5 步:

1. 计算模式 s^q 关于所有规则的匹配度,如式(15),其中, $t[\cdot]$ 表示算子 t -范数;

$$M^j(s^q) = t[\mu_1^j(s_1^q), \mu_2^j(s_2^q), \dots, \mu_n^j(s_n^q)], \quad j = 1, 2, \dots, N \quad (15)$$

2. 计算模式 s^q 与各分类在每一条规则上的关联度,如式(16);

$$b_j^h(s^q) = a[M^j(s^q), r_j^h], \quad h = 1, 2, \dots, K, \quad j = 1, 2, \dots, N \quad (16)$$

3. 计算加权函数,如式(17);

$$B_j^h(s^q) = g(b_j^h(s^q)), \quad h = 1, 2, \dots, K, \quad j = 1, 2, \dots, N \quad (17)$$

4. 计算各分类的模式分类健全度,如式(18);

$$Y^h = f(B_j^h(s^q), j = 1, 2, \dots, N \text{ 且 } B_j^h(s^q) > 0), \quad h = 1, 2, \dots, K \quad (18)$$

5. 给出模式 s^q 的分类结果,如式(19).

$$Y^{s^q} = \text{MAX}_{h=1}^K (Y^h) \quad (19)$$

当 $f[\cdot]$ 取最大值运算, $t[\cdot]$ 和 $a[\cdot]$ 都取代数积

运算,且 $g(x)=x$,可得 Y^h 如式(20),即文献[13]提出的单一优胜法(single winner).

$$Y^h = \text{MAX}_{j=1}^N (r_j^h \cdot \mu_j(s^q)), \quad h=1,2,\cdots,K \quad (20)$$

如果 $f[\cdot]$ 取代数和运算,可得 Y^h 如式(21),即文献[13]中的投票加权法(weighted vote).

$$Y^h = \sum_{j=1}^N r_j^h \cdot \mu_j(s^q), \quad h=1,2,\cdots,K \quad (21)$$

5.2 基于解模糊器的分类推理

在模糊控制系统中,输入变量模糊化后,进行模糊推理,得到输出变量的模糊值,经过解模糊器运算可输出一个与实际结果最为贴近的清晰值.

本文将样本各属性作为分类系统的输入值,不同规则结论对应的权值作为输出值,则上文所述分类系统可视为 n 输入 K 输出的模糊系统, K 个输出值中的最大者为分类结果.

首先,将 n 输入 K 输出系统分解成 K 个 n 输入单输出系统^[19],如图 3(a)所示.对于第 h 个模糊子系统,完备规则集中结论不为第 h 个分类的规则,其权值设为零,从而保证各模糊子系统规则集的完备性,各子系统的规则数量都为 N .应用规则集进行模糊推理后,经解模糊器可以得到模式 s^q 属于分类 h 的权值 \bar{r}^h ,如图 3(b)所示.对于第 h 个模糊子系统,其第 j 条规则的结论可视为中心等于 r_j^h 的模糊集,运用乘积推理机^[19]进行推理,由中心平均解模糊器^[20]可得 r^h ,如式(22). $r^h(h=1,2,\cdots,K)$ 中最大值对应的分类为模式 s^q 的分类结果.由于各模糊子系统规则集的完备性,式(22)中的分母对于各分类都相等,因此,式(22)分母可以忽略;结合式(1)可得简化后的式(23).

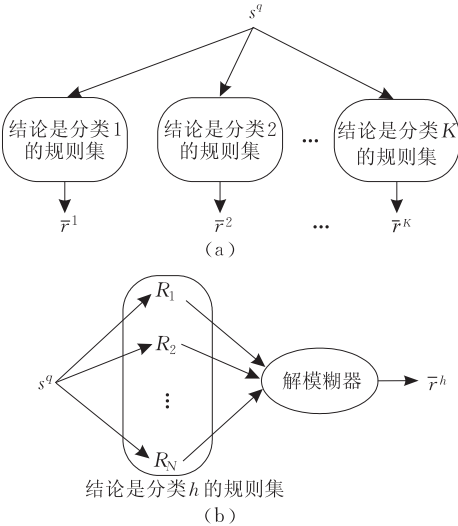


图 3 应用解模糊器进行分类推理示意图

$$r^h = \frac{\sum_{j=1}^N r_j^h \cdot \left(\prod_{i=1}^n \mu_j^i(s_i^q) \right)}{\sum_{j=1}^N \left(\prod_{i=1}^n \mu_j^i(s_i^q) \right)} \quad (22)$$

$$\bar{r}^h = \sum_{j=1}^N r_j^h \cdot \mu_j(s^q) \quad (23)$$

如果采用最大值解模糊器^[19],模糊子系统的输出为 r_l^h ,再由 $\mu_l(s^q)$ 对其加权,可得 \bar{r}^h :

$$\bar{r}^h = r_l^h \cdot \mu_l(s^q) = \text{MAX}_{j=1}^N (r_j^h \cdot \mu_j(s^q)) \quad (24)$$

可见,上述两种解模糊器的推理结果分别对应文献[13]中的投票加权法和单一优胜法.与通用推理模型、概率论^[14]和一致性理论^[15-16]相比,本文基于解模糊器的分类推理过程具有更好解释性和简洁性.

6 实验分析

应用 Iris 数据和 Wine 数据验证分析基于不同匹配度的分类规则提取算法,这两个数据集来源于应用广泛的标准模式识别 UC Irvine 数据库,是评价分类算法优劣的通用数据.

实验方法采用 LV1 法^[10](Leave-One-Out):选择样本空间中的一个样本作为测试数据,其它样本为训练数据,做一次分类实验;遍历样本空间中的每一个样本,使所有样本都能作为测试数据参加实验,共完成 M (样本总数)次实验,统计分类正确的实验次数在实验总数中的比例可得最终实验结果.

由于完备规则集的规则数量 $N=(L+1)^n-1$ 随着样本属性和模糊分区数量的增加变得异常庞大,降低了规则集的解释性和系统的分类效率.因此,在应用上文算法产生规则集之后,由式(25)计算各规则的 CS 值^[21],CS 值较大的规则被优先选用,抽取相同数量的以各分类为结论的规则,组成优选规则集,基于优选规则集进行中心平均解模糊分类推理.

$$CS(R_j) = \frac{[\gamma_{\text{Class } C_j}(R_j)]^2}{(M-1) \cdot \sum_{i=1}^K \gamma_{\text{Class } i}(R_j)} \quad (25)$$

6.1 Iris 数据实验

Iris 数据分为 3 类,共有 150 个样本,每类各包含 50 个样本,每个样本有 4 个属性.实验中,各属性划分 3 个标准三角形模糊区间,附加一个特殊模糊区间 $\mu(s_i^1)=1$,完备规则集的规则数为 $N=(3+1)^4-1=255$.计算每条规则的 CS 值,优选一定数量的规则,应用上文的规则结论和权值算法对

Iris 数据进行分类实验,两种匹配度算法的分类正确率见表 1.

对于各分类样本数量均衡的 Iris 数据,各分类样本在全局空间等概率发生,在确定规则结论时,无需根据样本数量适当提高相应分类成为规则结论的可能性. 分类样本在各模糊区域的全局密度优势成为判定规则结论的主要依据,由式(8)可知:规则相对匹配度中的因子 RD_j^b 强调了样本的全局密度优势,当各分类的样本数量相等时,规则相对匹配度的比例加权对各分类相等,由规则相对匹配度确定规

则结论等效于由强调全局密度优势的 RD_j^b 来确定规则结论.

文献[9-10]基于式(11)和式(13)的规则提取算法忽略了样本在各规则所描述模糊区域的隶属度分布对比;由式(4)可知:规则相对匹配度体现了同一分类样本在各模糊区域之间的隶属度分布对比,其全局密度优势有利于更加全面的分析样本数量均衡的分类规则提取问题.

因此,基于本文提出的规则相对匹配度算法比文献[9-10]具有更好的分类效果,如表 1 所示.

表 1 各规则提取算法对 Iris 数据的分类正确率/%

被选 规则数目	规则权值算法							
	I		II		III		IV	
	文献[9-10]算法	本文算法	文献[9-10]算法	本文算法	文献[9-10]算法	本文算法	文献[9-10]算法	本文算法
9	92.67	93.33	95.33	96.00	96.00	96.00	95.33	95.33
12	93.33	93.33	95.33	96.00	96.00	96.00	96.00	96.00
15	92.67	93.33	94.00	94.00	95.33	96.00	96.00	96.00
60	90.00	90.67	92.00	92.67	93.33	94.00	94.67	95.33
90	90.00	90.67	93.33	93.33	95.33	96.00	94.67	95.33

6.2 Wine 数据实验

Wine 数据分为 3 类,共有 178 个样本,各分类样本数量分别为 59,71 和 48,每个样本都有 13 个属性.与 Iris 数据实验相同,各属性划分 3 个标准三角形模糊区间,附加一个特殊模糊区间,因此,完备规则集的规则数为 $N=(3+1)^{13}-1=67108863$.为提高规则的生成效率,并缩短规则长度,实验中,规则的条件部分最多包含三个不取特殊模糊区间的属性^[10],则规则总数减少到 $N=C_{13}^3 \cdot 3^3+C_{13}^2 \cdot 3^2+C_{13}^1 \cdot 3=8463$.基于不同匹配度对 Wine 数据进行分类实验,结果见表 2.

对于样本数量失衡的 Wine 数据,各分类样本在全局空间不具有相同的发生概率,需要根据各分类样本的相对数量适当提高相应分类成为规则结论

的可能性.分类匹配度的局部数量优势得以体现,但从分类结果看:基于规则相对匹配度提取规则的分类正确率仍然要高于分类匹配度算法.这是由于后者仅仅考虑了分类样本的局部数量优势;而前者在全局学习空间中,除体现了某一分类在各模糊区域的隶属度对比(全局密度优势)外,还通过各分类样本数量比例的加权,使出现概率较大的分类拥有较高的优先级成为规则结论,这在一定程度上,等效于从分类匹配度局部数量优势的角度分析分类规则提取问题.从而达到了规则相对匹配度兼顾全局密度优势和局部数量优势的效果.

因此,由规则相对匹配度提取 Wine 数据的分类规则仍然要优于文献[9-10]中的算法,如表 2 所示.

表 2 各规则提取算法对 Wine 数据的分类正确率/%

被选 规则数目	规则权值算法							
	I		II		III		IV	
	文献[9-10]算法	本文算法	文献[9-10]算法	本文算法	文献[9-10]算法	本文算法	文献[9-10]算法	本文算法
6	93.82	94.38	93.82	93.82	91.57	91.57	91.57	93.26
9	92.13	92.13	93.26	93.82	91.01	92.70	93.26	94.38
12	92.70	93.26	93.26	93.26	92.70	92.70	93.26	93.82
60	94.94	94.94	94.38	94.38	94.94	95.51	95.51	95.51
120	96.07	96.07	94.94	94.94	94.94	95.51	95.51	95.51

7 结 论

鉴于模糊规则的可解释性和模糊推理的简洁性,本文讨论了基于不同匹配度的分类规则提取算法,并应用解模糊器实现了分类推理.由于分类匹配

度没有考虑全局模糊区域的样本分布,本文定义了一种规则相对匹配度,基于该匹配度确定规则结论和规则权值,不仅通过各分类样本的数量比例进行加权运算,使其具备了分类匹配度算法处理局部样本分布的优势,而且,还全面衡量了样本在各模糊区域上的隶属度分布情况.同时,本文以解模糊器算法

实现了基于模糊规则的分类推理,较以往算法易于实现,并具有更好的解释性.最后,由 Iris 数据和 Wine 数据的实验结果证明:本文基于规则相对匹配度的规则提取算法优于已有的分类匹配度算法.

参 考 文 献

- [1] Chakraborty D, Pal N R. A neuro-fuzzy scheme for simultaneous feature selection and fuzzy rule-based classification. *IEEE Transactions on Neural Networks*, 2004, 15(1): 110-123
- [2] Chen Y X, Wang J Z. Support vector learning for fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 2003, 11(6): 716-728
- [3] Castellano G, Fanelli A M. A staged approach for generation and compression of fuzzy classification rules//*Proceedings of the 9th IEEE International Conference on Fuzzy Systems*. 2000, 1(1): 42-47
- [4] Roubos Johannes A, Setnes Magne, Abonyi Janos. Learning fuzzy classification rules from label data. *Information Sciences*, 2003, 150(1-2): 77-93
- [5] Yang Ai-Min, Hu Yun-Fa. Method of dynamic cluster generating fuzzy classifier rules. *Mini-Micro Systems*, 2005, 26(9): 1540-1545(in Chinese)
(阳爱民, 胡运发. 一种基于动态聚类的模糊分类规则的生成方法. *小型微型计算机系统*, 2005, 26(9): 1540-1545)
- [6] Gao Liang, Gao Hai-Bing, Zhou Chi et al. Acquisition of pattern classification rule based on particle swarm optimization. *Journal of Huazhong University of Science and Technology(Nature Science)*, 2004, 32(11): 24-26(in Chinese)
(高亮, 高海兵, 周驰等. 基于粒子群优化算法的模式分类规则获取. *华中科技大学学报(自然科学版)*, 2004, 32(11): 24-26)
- [7] Xing Zong-Yi, Zhang Yong, Hou Yuan-Long et al. Design of interpretable and precise fuzzy classification system based on fuzzy clustering and genetic algorithm. *Acta Electronica Sinica*, 2006, 34(1): 83-88(in Chinese)
(邢宗义, 张永, 侯远龙等. 基于模糊聚类 and 遗传算法的具备解释性和精确性的模糊分类系统设计. *电子学报*, 2006, 34(1): 83-88)
- [8] Casillas J, Cordon O, del Jesus M J, Herrera F. Genetic feature selection in a fuzzy rule-based classification system learning process for high-dimensional problems. *Information Sciences*, 2001, 136(1-4): 135-157
- [9] Ishibuchi H, Nakashima T, Murata T. Performance evaluation of fuzzy classifier systems for multidimensional pattern

classification problems. *IEEE Transactions on Systems, Man, and Cybernetics—Part B: Cybernetics*, 1999, 29(5): 601-618

- [10] Ishibuchi H, Yamamoto T. Rule weight specification in fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 2005, 13(4): 428-435
- [11] Ray Kumar S, Ghoshal J. Approximate reasoning approach to pattern recognition. *Fuzzy Sets and Systems*, 1996, 77(2): 125-150
- [12] Cordon O, del Jesus M J, Herrera F. A proposal on reasoning methods in fuzzy rule-based classification systems. *International Journal of Approximate Reasoning*, 1999, 20(1): 21-45
- [13] Ishibuchi H, Nakashima T, Morisawa T. Voting in fuzzy rule-based systems for pattern classification problems. *Fuzzy Sets and Systems*, 1999, 103(2): 223-238
- [14] van den Berg J, Kaymak U, van den Bergh W M. Fuzzy classification using probability-based rule weighting//*Proceedings of the 2002 IEEE International Conference on Fuzzy Systems*. 2002, 2: 991-996
- [15] Benediktsson J A, Swain P H. Consensus theoretic classification methods. *IEEE Transactions on System, Man, and Cybernetics*, 1992, 22(4): 688-704
- [16] Kuncheva L I, Krishnapuram R. A fuzzy consensus aggregation operator. *Fuzzy Sets and Systems*, 1996, 79(3): 347-356
- [17] Nozaki K, Ishibuchi H, Tanaka H. Adaptive fuzzy rule-based classification systems. *IEEE Transactions on Fuzzy Systems*, 1996, 4(3): 238-250
- [18] Ishibuchi H, Nozaki K, Yamamoto N, Tanaka H. Selecting fuzzy if-then rules for classification problems using genetic algorithms. *IEEE Transactions on Fuzzy Systems*, 1995, 3(3): 260-270
- [19] Wang Li-Xin. *A Course In Fuzzy Systems & Control*. Beijing: Tsinghua University Press, 2003(in Chinese)
(王立新. *模糊系统与模糊控制教程*. 北京: 清华大学出版社, 2003)
- [20] Wang L X, Mendel J M. Generating fuzzy rules by learning from examples. *IEEE Transactions on System, Man, and Cybernetics*, 1992, 22(6): 1414-1427
- [21] Ishibuchi H, Yamamoto T. Heuristic extraction of fuzzy classification rules using data mining techniques: An empirical study on benchmark data sets//*Proceedings of the 2004 IEEE International Conference on Fuzzy Systems*. 2004, 1(1): 161-166



LI Jie, born in 1979, Ph. D., engineer. Her main research interests include network security, artificial intelligence, pattern recognition and fault diagnosis etc.

DENG Yi-Ming, born in 1979, M. S.. His main research interests include network performance analysis and artificial intelligence etc.

SHEN Shi-Tuan, born in 1940, professor, Ph. D. supervisor. His main research interests include auto-testing of weapons and equipment, satellite navigation and communication etc.

Background

This work is supported by the Defense Pre-Research Project of the 11th Five-Year-Plan of China.

As a branch of Artificial Intelligence, Pattern Classification is a sort of process of human thinking and reasoning by using computer. Based on fuzzy set and fuzzy proposition, Pattern Classification is just like the human reasoning due to their fuzzy characteristic. Furthermore, the reasoning process of Pattern Classification is comprehensive because of fuzzy rule's interpretability.

As one of the most promising research domains of fuzzy classification, initial fuzzy rule base generation has been widely studied recently. Neural Network has been applied for extraction of fuzzy rules, which consists of four layers and optimizes the membership function by employing pruning algorithm. Since the Neural Network has the capability of knowledge storage and reasoning, it is inefficient to append a step of rules generation for Neural Network. Kernel Machine has been used to fuzzy classification rule extraction, by which the rule base is generating from case space based on Support Vector Machine and then the Positive Definite Fuzzy Classifier is constructed. The Kernel Function Parameter, however, does depend on different applications, which is hard to be implemented by now. The Clustering algorithm is also applied for fuzzy rule generation to reduce the amount of rule base

and the amount of rule attribute. While the rule base is simplified, the complexity is increased. The Genetic Algorithm is also a method to generate fuzzy rule, of which performance is determined by initial rule-base though.

This research aims at the key techniques of fault classification for automatic fault diagnosis, which carries out the comprehensive and efficient classification rule base generation and classification reasoning by fuzzy theory. To guarantee both performance of classification and interpretability of reasoning, authors of this paper propose fuzzy rule extraction method using the rule relative compatibility defined in this paper. Rule Relative Compatibility considers both the relative amount of every class's case and the global distribution in the case space. The Fuzzy Area Distribution Matrix is designed to calculate all sorts of compatibility, and then the algorithm of the initial complete rule base generation is designed based on this matrix. Also, based on the defuzzifier method, a classification reasoning approach is designed to simplify the procedure of fuzzy classification. The experiment results form the benchmark data demonstrates that the algorithm proposed by this contribution is able to obtain more good classification performance for both homogeneous cases and heterogeneous cases.