

文章编号: 1001-0920(2008)09-1015-06

一种新的快速模糊规则提取方法

陈铁明^{1,2}, 龚荣盛³, SAMUEL H Huang³

(1. 浙江工业大学 软件学院, 杭州 310014; 2. 北京航空航天大学 软件开发环境国家重点实验室, 北京 100086; 3. 辛辛那提大学 工程学院, 美国 辛辛那提 45221)

摘 要: 提出一种高效的规则提取算法, 采用熵测量改进 Chi-merge 特征区间离散化方法, 模糊划分输入空间. 先为每个数据生成单条规则, 再聚集相同前项的单条规则产生带概率属性的分类规则. 提取的规则无需任何调整, 应用模糊推理便可获得较理想的分类效果, 同时支持增量式规则更新. 最后给出了新方法的性能测试结果.

关键词: 模糊规则; 规则提取; Chi-merge 方法; 离散化; 模糊推理

中图分类号: TP182

文献标识码: A

A novel and quick fuzzy rule extraction method

CHEN Tie-ming^{1,2}, GONG Rong-sheng³, SAMUEL H Huang³

(1. College of Software Engineering, Zhejiang University of Technology, Hangzhou 310014, China; 2. State Key Laboratory of Software Development Environment, Beihang University, Beijing 100086, China; 3. School of Engineering, University of Cincinnati, Cincinnati 45221, USA. Correspondent: CHEN Tie-ming, E-mail: tmchen@zjut.edu.cn)

Abstract: This paper presents a novel method for classification problems. Firstly, entropy is introduced to select the best discretization solution from Chi-merge for each feature to guide grid partition. Afterwards one simple rule is generated from each data point, and the combination of simple rules with the same antecedent yields the consequent of an aggregated rule for each grid. Because grid partition is based on rational discretization, the extracted rules are not only easy to understand but also produce accurate results even without tuning. In addition to quick learning, the method accommodates incremental learning, namely, rules can be updated quickly once new data are obtained. A benchmark case is studied to demonstrate the effectiveness of the proposed method.

Key words: Fuzzy rules; Rule extraction; Chi-merge method; Discretization; Fuzzy inference

1 引 言

模糊规则系统最初由 Zadeh^[1] 提出, 它是基于模糊集合和模糊逻辑的一种系统. 目前主要有两大类模糊规则, 即 Mamdani^[2] 和 Takagi-Sugeno^[3]. 二者的最大区别在于规则的后项 (IF-THEN 规则的结论部分) 不同, Takagi-Sugeno 规则后项为实值函数, Mamdani 后项则为一个模糊集合.

业界最早提出从数据中直接提取模糊规则而与先验知识无关的是 Wang-Mendel (WM) 方法^[4]. WM 方法已得到广泛应用, 成为领域内的基准方法^[5]. 它主要通过输入空间的均匀划分, 采用查表策略为每个空间划分提取规则. 其最大问题是提取的

规则数量随特征向量维数的增大而快速增加, 从而使效率下降. 聚类方法产生的模糊规则^[6] 可在一定程度上降低规则提取的复杂度, 但提取的规则不易理解.

当前, 几乎所有的模糊规则提取方法都是先粗略划分输入空间, 生成模糊规则, 再采用神经网络等方法不断调整规则^[7], 以保证分类精度. 因此, 在处理数据规模较大且动态增长的问题时, 模糊规则提取的效率低下.

本文针对分类问题, 提出一种新的模糊规则提取方法. 该方法可从大规模数据中一次性快速挖掘分类规则, 不作任何修正即可获得准确率较高、效率

收稿日期: 2007-07-03; 修回日期: 2007-09-21.

基金项目: 国家自然科学基金项目 (60673080, 60773115); 国家 863 计划项目 (2006AA01Z235); 浙江省自然科学基金项目 (Y106290).

作者简介: 陈铁明 (1978 →), 男, 浙江诸暨人, 讲师, 博士生, 从事数据挖掘、信息安全等研究; SAMUEL H Huang (1969 →), 男, 美籍教授, 从事数据挖掘、智能建模等研究.

性能兼顾的规则,并支持规则对增量式数据的快速更新.

2 模糊规则提取

本文算法分别处理规则前项和规则后项.在规则前项处理阶段,对每个数据特征采用 Chi-merge 方法,将特征值的连续域离散化为多个区间,并基于最大熵方法确定最优的区间数.每个区间对应一个语言表达式,如小、中、大,其隶属函数由区间内数据的统计特性决定.所有特征的前项确定后,开始执行后项的计算.对于输入空间的每个划分区间,依据数据产生规则后项,并附带所属类别的概率值.

2.1 规则前项

2.1.1 离散化

无监督的离散化方法仅仅依赖于输入的数值信息,而带监督的方法还应考虑输出类别信息,往往比无监督方法更精确. Chi-merge^[8]是最常用的带监督离散化方法,它采用 χ^2 方法. χ^2 值计算如下:

$$\chi^2 = \sum_{i=1}^m \sum_{j=1}^k \frac{(A_{ij} - E_{ij})^2}{E_{ij}} \quad (1)$$

其中: A_{ij} 表示第 i 区间第 j 类的数据样本数, $i = 1, \dots, m, j = 1, \dots, k, k$ 表示分类数量; E_{ij} 表示对 $A_{ij} = R_i C_j / N$ 的期望频率; $R_i = \sum_{j=1}^k A_{ij}$ 表示第 i 区间的总数据样本; $C_j = \sum_{i=1}^m A_{ij}$ 表示第 j 类的总数据样本; N 表示总的的数据样本数.

Chi-merge 方法执行过程如下:视初始不同的输入值为区间界,合并 χ^2 值最小的相邻区间,直到 χ^2 值超出某个显著性水平下的阈值.显然,Chi-merge 方法通过显著性水平作为迭代合并的终止判定条件,不同的显著性水平将影响离散化效果.为了解决这一问题,本文引入熵方法^[9]设计区间数的判定准则.

熵是信息论中用于测量不确定性的最常用方法,一个随机变量 R 的熵计算如下:

$$H(R) = - \sum_r p_r \log p_r \quad (2)$$

其中: r 表示随机变量 R 的某个取值,相应的概率用 p_r 表示.显然, R 的概率分布越均匀, R 的熵 $H(R)$ 越大.熵越大表示越大的不确定性或不可预测性.

本文将分类问题的输出视为一个随机变量,对于每个区间每类输出的概率及熵都可估计.假设一个连续的特征量 X 被离散成 m 个区间 I_1, I_2, \dots, I_m , 输出 Y 拥有 n 个类 y_1, y_2, \dots, y_n , 则输出 Y 的熵为

$$H(Y) = \sum_{i=1}^m H(Y_{I_i}) = \sum_{i=1}^m \sum_{j=1}^n - p_{y_j}^{I_i} \log p_{y_j}^{I_i} \quad (3)$$

其中: $H(Y_{I_i})$ 是输出 Y 在区间 I_i 中的熵, $p_{y_j}^{I_i}$ 是区间

I_i 输出类 y_j 的估计概率.显然,输出熵最小值的离散化方案可认为具有最小的不确定性和最高的正确率.

将熵方法应用于 Chi-merge 方法,确定最优的区间数.具体作法如下:忽略 Chi-merge 方法的合并终止条件,合并每对具有 χ^2 最小值的相邻区间,直到输入域只包含两个区间时结束.该方法的最后 5 步迭代将 6 个区间合并成最后的 2 个,最后 5 步所确定的区间方案即为最优区间划分的备选.因此任何特征至少需要 2 个隶属函数区分不同的输出,而至多只需 6 个隶属函数遍历判别.针对每个备选方案,计算输出的熵,选择最小熵对应的方案作为 Chi-merge 离散化区间划分.

现以分类基准问题 IRIS^[10] 的 Pedal width 特征为例,说明改进的 Chi-merge 方法.根据输出的不同,一组 Pedal width 数据的初始划分为

$$(0.1, 1.0, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.5),$$

代表的区间为

$$[0.1, 1.0], [1.0, 1.4], \dots, [1.8, 1.9], [1.9, 2.5].$$

从左至右计算相邻区间的 χ^2 值分别为

$$78.0, 0.065, 0.137, 0.375, 2.431, 2.896.$$

将最小值 0.065 对应的区间对 $[1.4, 1.5], [1.5, 1.6]$ 合并,得到新的区间

$$(0.1, 1.0, 1.4, 1.6, 1.7, 1.8, 1.9, 2.5).$$

重复合并,直到最后剩下两个区间.表 1 显示了区间划分 $(0.1, 1.0, 1.8, 2.5)$ 中,每个区间内输出 3 个类别各自的数据样本数,其相应的熵计算如下:

$$H(Y_{I_1}) = 0, H(Y_{I_2}) = 0.308, H(Y_{I_3}) = 0.105, H(Y) = H(Y_{I_1}) + H(Y_{I_2}) + H(Y_{I_3}) = 0.413.$$

最后 5 步迭代所得的区间划分计算输出熵如表 1 区间 $(0.1, 1.0, 1.8, 2.5)$ 的数据样本数分布

区 间	分类样本数量			Total
	Setona	Versicolor	Virginica	
$I_1[0.1, 1.0]$	50	0	0	50
$I_2[1.0, 1.8]$	0	49	5	54
$I_3[1.8, 2.5]$	0	1	45	46

表 2 最后 5 步迭代所得区间的输出熵

区间数	区 间 分 界	熵
2	$(0.1, 1.0, 2.5)$	0.693
3	$(0.1, 1.0, 1.8, 2.5)$	0.413
4	$(0.1, 1.0, 1.4, 1.8, 2.5)$	0.594
5	$(0.1, 1.0, 1.4, 1.8, 1.9, 2.5)$	0.776
6	$(0.1, 1.0, 1.4, 1.7, 1.8, 1.9, 2.5)$	1.431

2 所示. 显然, 区间(0.1, 1.0, 1.8, 2.5) 的熵最小. 因此 Pedal width 特征离散成 3 个区间 [0.1, 1.0], [1.0, 1.8] 和 [1.8, 2.5].

2.1.2 模糊空间划分

连续特征量被离散成区间后, 即为每个区间对应的语言表达式. 例如上述对 Pedal width 的 3 个区间 I_1, I_2 和 I_3 , 可分别描述为: Small, Medium 和 Large. 采用不同的高斯隶属函数分别描述, 有

$$\mu_{Term}(x) = \begin{cases} e^{-(x-b)^2/2a^2}, & x > b; \\ 1, & \text{其他}, \end{cases} \quad (4)$$

$$\mu_{Term}(x) = e^{-(x-b)^2/2a^2}; \quad (5)$$

$$\mu_{Term}(x) = \begin{cases} e^{-(x-b)^2/2a^2}, & x < b. \\ 1, & \text{其他}, \end{cases} \quad (6)$$

根据统计学原理, 区间内数据样本的均值和方差分别是参数 b 和 a 的无偏估计. 所有特征量的隶属函数确定后, 便可产生特征空间的网格划分. 针对上例的 IRIS 问题, 考虑两个特征量 Petal width 和 Petal length, 且每个特征划分成 3 个区间, 即 Small, Medium 和 Large. IRIS 的空间划分如图 1 所示, 其中灰色区域表示隶属函数的模糊性产生的交叉区域. 空间划分的每个区域代表一个独立的规则前项, 例如右上角区域即覆盖前项为“if Petal width is Large and Petal length is Large”的所有规则.

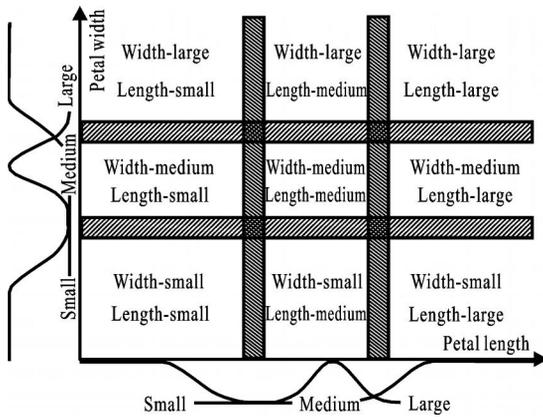


图 1 模糊空间划分

2.2 规则后项

首先给出用于描述规则后项的符号说明:

- I : 输入特征总数;
- X_i : 第 i 个输入特征, $i = 1, 2, \dots, I$;
- I_i : X_i 的语言表达术语数量, $i = 1, 2, \dots, I$;
- A_{ik} : X_i 的第 k 条术语, $k = 1, 2, \dots, I_i$;
- $\mu_{A_{ik}}$: A_{ik} 的隶属函数;
- L : 输出类别总数;
- Y : 输出变量;
- Y_l : 第 l 个输出类别, $l = 1, 2, \dots, L$;
- J : 数据样本数;

x_{ji} : 第 j 个数据样本点的第 i 个输入, $i = 1, 2, \dots, I, j = 1, 2, \dots, J$;

y_j : 第 j 个数据样本点的输出, $y_j \in \{Y_1, Y_2, \dots, Y_L\}, j = 1, 2, \dots, J$;

N : 最终产生的规则总数;

R_n : 第 n 条规则, $n = 1, 2, \dots, N$.

设 j_i 满足

$$\mu_{A_{i j_i}}(x_{ji}) = \max\{\mu_{A_{ik}}(x_{ji})\}, \quad k = 1, 2, \dots, I_i, \quad d_j = \prod_{i=1}^I \mu_{A_{i j_i}}(x_{ji}). \quad (7)$$

对每个数据样本点 $(x_{j1}, x_{j2}, \dots, x_{jI}, y_j)$, 产生如下规则:

IF X_1 is $A_{1 j_1}$ and X_2 is $A_{2 j_2}$ and ... and X_I is $A_{I j_I}$, THEN Y is y_j with Degree d_j .

不失一般性, 规则前项表示为

IF X_1 is $A_{1 n_1}$ and X_2 is $A_{2 n_2}$ and ... and X_I is $A_{I n_I}$, $n_i \in \{1, 2, \dots, I_i\}$.

记 T_l 表示后项为“ Y is Y_l ”($l = 1, 2, \dots, L$) 的规则数, d_{lr} 表示后项为“ Y is Y_l ”($r = 1, 2, \dots, T_l$) 第 r 条规则的度数, $S_l = \sum_{r=1}^{T_l} d_{lr}$. 聚集所有相同前项的单条规则, 产生单条规则 R_n 如下:

IF X_1 is $A_{1 n_1}$ and X_2 is $A_{2 n_2}$ and ... and X_I is $A_{I n_I}$, THEN Y is Y_l with Probability D_{n1} , Y_2 with Probability D_{n2} , ..., Y_L with Probability D_{nL} .

其中

$$D_{nl} = S_l / \sum_{l=1}^L S_l, \quad l = 1, 2, \dots, L. \quad (8)$$

当新的数据样本点加入时, 聚集的规则便可快速更新. 更新过程如下: 首先为新加入的数据点产生单条规则, 假设

IF X_1 is $A_{1 n_1}$ and X_2 is $A_{2 n_2}$ and ... and X_I is $A_{I n_I}$, THEN Y is Y_o with Degree d .

在聚集的规则中, 选取与新数据点具有相同前项的规则, 执行如下更新:

IF X_1 is $A_{1 n_1}$ and X_2 is $A_{2 n_2}$ and ... and X_I is $A_{I n_I}$, THEN Y is Y_l with Probability D_{n1} , Y_2 with Probability D_{n2} , ..., Y_L with Probability D_{nL} .

其中

$$D_{no} = \frac{d + S_o}{d + \sum_{l=1}^L S_l}, \quad D_{nl} = \frac{S_l}{d + \sum_{l=1}^L S_l}, \quad l = o.$$

回到 IRIS 实例,两个特征量 Pedal length 和 Pedal width 用 X_1 和 X_2 表示,分别用 3 个表达式描述特征 Small(S),Medium(M) 和 Large(L),对应的隶属函数表示为: $\mu_{A_{1S}}, \mu_{A_{1M}}, \mu_{A_{1L}}$ 以及 $\mu_{A_{2S}}, \mu_{A_{2M}}, \mu_{A_{2L}}$.

对于样本数据(4.8, 1.8, Vesicolor), 计算得

$$\mu_{A_{1S}}(4.8) = 0, \mu_{A_{1M}}(4.8) = 0.658,$$

$$\mu_{A_{1L}}(4.8) = 0.183, \mu_{A_{2S}}(1.8) = 0,$$

$$\mu_{A_{2M}}(1.8) = 0.077, \mu_{A_{2L}}(1.8) = 0.477.$$

因此产生的单条规则如下:

IF Petal length is M and Petal width is L,
THEN Vesicolor with Degree of
 $0.658 \times 0.477 = 0.3139$.

针对 IRIS 的 150 个数据样本,可以获得 150 条规则.其中:前项为“IF Petal length is M and Petal width is L”的有 10 条,后项及度的计算结果如表 3 所示.由表 3 得

$$S_1 = 0, S_2 = \prod_{i=1}^2 d_i = 0.4203,$$

$$S_3 = \prod_{i=3}^{10} d_i = 0.2853.$$

计算得

$$D_1 = S_1 \prod_{j=1}^3 S_j = 0\%,$$

$$D_2 = S_2 \prod_{j=1}^3 S_j = 14.0\%,$$

$$D_3 = S_3 \prod_{j=1}^3 S_j = 86.0\%.$$

上述前项的聚集规则为(分别用 Se, Ve, Vi 代表 Setosa, Vesicolor, Virginica)

IF Petal length is M and Petal width is L,
THEN Se with Probability 0%, Ve with
Probability 14.0%, Vi with Probability 86.0%.

聚集所有的规则,最终提取的 5 条模糊规则如下:

表 3 前项为“IF Length is M and Width is L”的模糊规则

规则	后项	度数	规则	后项	度数
1	Versicolor	0.3139	6	Virginica	0.5027
2	Versicolor	0.1064	7	Virginica	0.2447
3	Virginica	0.2548	8	Virginica	0.3913
4	Virginica	0.3062	9	Virginica	0.2548
5	Virginica	0.3139	10	Virginica	0.3139

Rule 1: IF PL is S and PW is S, THEN Se (100%), Ve (0%), Vi (0%);

Rule 2: IF PL is M and PW is M, THEN Se (0%), Ve (98.9%), Vi (1.1%);

Rule 3: IF PL is M and PW is L, THEN Se (0%), Ve (14.0%), Vi (86.0%);

Rule 4: IF PL is L and PW is M, THEN Se (0%), Ve (10.8%), Vi (89.2%);

Rule 5: IF PL is L and PW is L, THEN Se (0%), Ve (0%), Vi (100%).

2.3 模糊推理

本文采用如下 3 个步骤完成模糊推理:

步骤 1: 模糊化输入. 通过隶属函数确定输入属于每条语言表达式的度,如

$$x_i \text{ Belongs to } A_i \text{ to the Degree of } \mu_{A_i}(x_i).$$

步骤 2: 应用模糊算子乘积确定规则的权重,如规则 R_n 为

IF X_1 is A_{1n_1} and X_2 is A_{2n_2} and ...
and X_l is A_{ln_l} ,

THEN Y is Y_l with Probability D_{n1} , Y_2 with
Probability D_{n2} , ..., Y_L with Probability D_{nl} .

对于输入 (x_1, x_2, \dots, x_l) , 规则 R_n 的权重 $w_n = \prod_{i=1}^l \mu_{A_{in_i}}(x_i)$.

步骤 3: 聚集输出. 通过取权重均值的方法聚集所有规则的输出,即

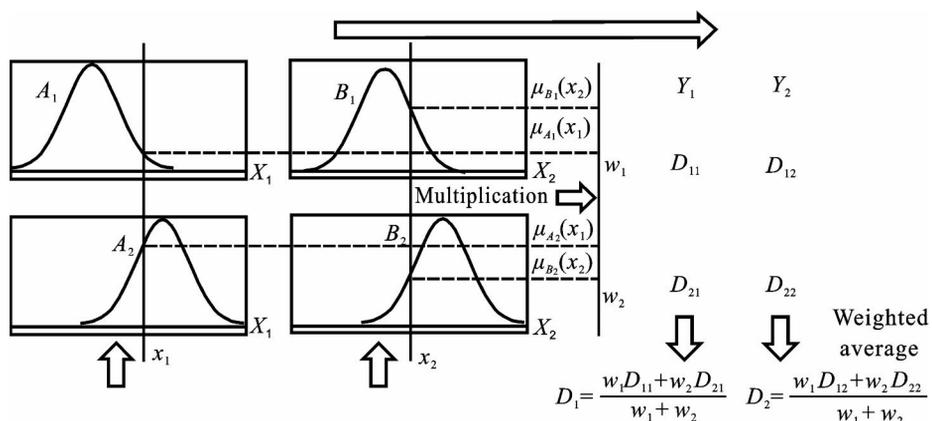


图 2 模糊推理过程示例

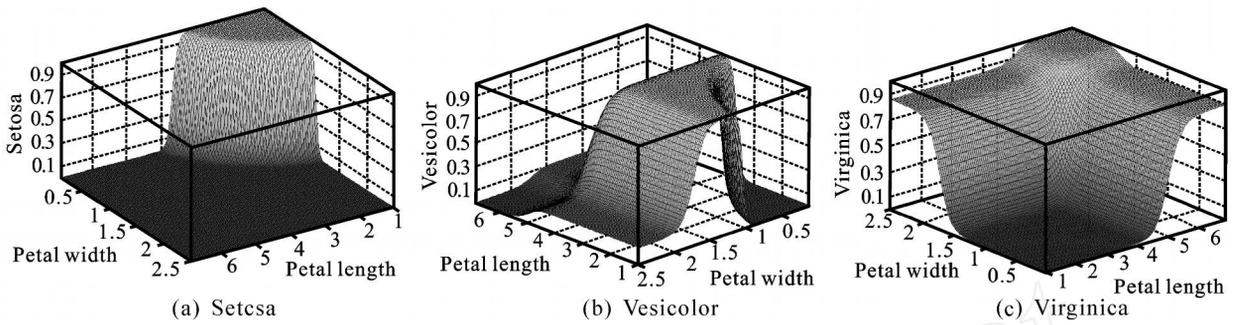


图 3 IRIS 问题分类曲面

$$D_l = \frac{\sum_{n=1}^N w_n D_{nl}}{\sum_{n=1}^N w_n}, l = 1, 2, \dots, L. \quad (9)$$

它表示给定数据属于输出类 Y_l 的概率.

下面以包含 2 个输入、2 个输出类以及 2 条规则的例子说明上述推理机制的工作过程. 假设两条规则如下:

Rule1: IF X_1 is A_1 and X_2 is B_1 , THEN Y is $Y_1(D_{11})$ and $Y_2(D_{12})$;

Rule2: IF X_1 is A_2 and X_2 is B_2 , THEN Y is $Y_1(D_{21})$ and $Y_2(D_{22})$.

则模糊推理的过程如图 2 所示.

目前,多数分类方法对给定输入一般仅提供确定性结论,如:“IF Input is (x_1, x_2, \dots, x_l) , THEN Y is Class Y_l ”.由于规则后项带概率,本文提出的算法能为每个输出类提供光滑的划分曲面.根据上述 IRIS 问题提取的 5 条模糊规则,图 3 给出了 3 个输出类的划分曲面.本文提出的模糊规则提取算法可直接产生分类划分曲面,效果直观,便于理解.例如,任何 Pedal length 和 Pedal width 都是 Large 时,花的归类趋势为 Virginica.

3 性能测试

本文选择 Bupa^[10] 数据分类问题,用于验证算法的分类效果.数据集包含 6 个连续特征量,其中前 5 个都是血液测试值,且其指标随着体内酒精含量的增高而上升,最后一个表示每天从酒精饮料中摄取的等效酒精含量.分析的数据集包含两组男性个体(肝病患者或健康者)345 个记录值.有别于 IRIS 问题,对 Bupa 数据的分类存在较大困难,目前大多数学习算法的错误率一般都在 30% ~ 40% 之间.

将数据集随机分成 10 组,轮流保留 10 组中的任何一组,用于测试从剩余 9 组数据集训练提取的模糊规则.表 4 给出了计算 10 次分类准确率的均值和标准方差,并与其他学习算法^[11] 的性能对比.由表 4 知,本文方法具有较高的准确率均值和较小的标准差,提取效果优于一般的确定性方法.

表 4 Bupa 数据处理的性能对比 %

指标	本文方法	NN	C5.0	CART	Logistic	CBR
均值	68.64	55.59	66.47	67.94	67.35	60.88
方差	5.52	6.86	6.23	4.89	5.27	9.20

为了说明算法的高效性,以 Pima indian diabetes 分类问题为例进行分析.根据文献[12]提供的方法,本文只用 F_2 (Plasma glucose concentration) 和 F_6 (Body mass index) 两个输入特征,最终提取 4 条模糊规则如下:

Rule1: IF F_2 is S and F_6 is S, THEN Healthy (86.38%) Diabetic (13.62%);

Rule2: IF F_2 is S and F_6 is L, THEN Healthy (67.33%) Diabetic (32.67%);

Rule3: IF F_2 is L and F_6 is S, THEN Healthy (33.13%) Diabetic (66.87%);

Rule4: IF F_2 is L and F_6 is L, THEN Healthy (15.65%) Diabetic (84.35%).

分类曲面如图 4 所示.经验证,上述分类规则可达到 75% 的正确率.尽管比文献[13]提出的

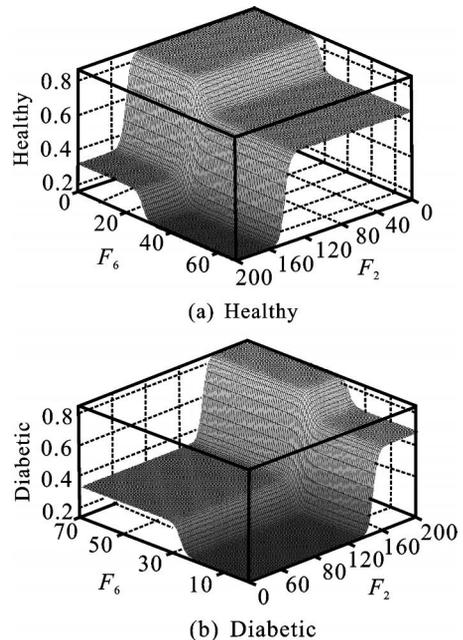


图 4 Pima 问题的分类曲面

Logistic 判别式方法的正确率(77.7%)略低,但本文提取的规则简单易懂,执行效率高。

4 结 论

本文提出一种新的面向分类问题的模糊规则提取方法,可快速提取易于理解的简单规则,并得到分类趋势曲面。不同于当前的模糊规则提取算法,新方法采用基于最小熵的 Chi-merge 方法划分输入空间,基于空间划分产生相应的分类规则,每条规则带有概率属性,以提高分类精度。提取的规则无需任何调整即可获得较高的准确性,从而大大提高了分类效率。

本文的模糊规则提取算法支持对新增数据的简单更新。下一步工作将结合前项和后项研究规则更新算法,以提高模糊规则的增量式学习更新能力。

致 谢 特别感谢浙江工业大学海外人才引进计划提供的合作研究机会;对于美国辛辛那提大学智能系统实验室的研究支持和帮助深表谢意!

参考文献(References)

- [1] Zadeh L A. Fuzzy sets[J]. Information and Control, 1965, 8(3): 338-353.
- [2] Mamdani E H, Assilian S. An experiment in linguistic synthesis with a fuzzy logic controller[J]. Int J of Man-machine Studies, 1975, 7(1): 1-13.
- [3] Takagi T, Sugeno M. Fuzzy identification of systems and its applications to modeling and control[J]. IEEE Trans on Systems, Man and Cybernetics, 1985, 15(1): 116-132.
- [4] Wang L X, Mendel J. Generating fuzzy rules by learning from examples[J]. IEEE Trans on Systems, Man and Cybernetics, 1992, 22(6): 1414-1427.
- [5] Wang L X. WM method completed: A flexible fuzzy system approach to data mining[J]. IEEE Trans on Fuzzy Systems, 2003, 11(6): 880-904.
- [6] Yage R R, Filev D P. Generation of fuzzy rules by mountain clustering[J]. J of Intelligent and Fuzzy System, 1994, 2(3): 209-219.
- [7] Kim J, Kasabov N K. HyFIS: Adaptive neuro-fuzzy inference systems and their application to nonlinear dynamical systems[J]. Neural Networks, 1999, 12(9): 1301-1319.
- [8] Kerber R. Chi-merge: Discretization of numeric attributes[C]. Proc of AAAI-92. San Jose: MIT Press, 1992: 129-134.
- [9] Shannon C, Weaver W. The mathematical theory of information[M]. Urban: University of Illinois Press, 1948.
- [10] Merz C J, Murphy P M. UCI repository of machine learning database[R]. Irvine: University of California, 1996.
- [11] Park Y J, Kim B C, Chun S H. New knowledge extraction technique using probability for case-based reasoning: Application to medical diagnosis[J]. Expert Systems, 2006, 22(1): 2-20.
- [12] Duch W, Adamczak R, Grabczewski K. New methodology of extraction, optimization and application of crisp and fuzzy logic rules[J]. IEEE Trans on Neural Networks, 2000, 11(2): 1-31.
- [13] Michie D, Spiegelhalter D J, Taylor C C. Machine learning, neural and statistical classification [M]. London: Elis Horwood, 1994.

(上接第 1014 页)

- [8] 于金盈, 许宏光, 赵克定, 等. 全阶滑模变结构解耦控制器在液压转台设计中的应用[J]. 航空学报, 2005, 26(3): 376-379.
(Yu J Y, Xu H G, Zhao K D, et al. Application of full-order sliding mode variable structure decoupling controller to hydraulic simulator design [J]. Acta Aeronautica et Astronautica Sinica, 2005, 26(3): 376-379.)
- [9] Chern T L, Wu Y C. An optimal variable structure control with integral compensation for electro hydraulic position servo control systems [J]. IEEE Trans on Industrial Electronics, 1992, 39(5): 460-463.
- [10] 刘金琨. 滑模变结构控制 Matlab 仿真[M]. 北京: 清

华大学出版社, 2005.

- (Liu J K. Matlab simulation for sliding mode control [M]. Beijing: Tsinghua University Press, 2005.)
- [11] Han J, Micheline K. Data mining: Concepts and techniques[M]. Beijing: China Machine Press, 2001: 225-278.
- [12] 李博群, 张克君, 傅剑, 等. 活套高度和张力系统的神经网络自适应解耦控制[J]. 控制与决策, 2006, 21(1): 46-50.
(Li B Q, Zhang K J, Fu J, et al. Adaptive neural network decoupling control for the loopers' height and tension system [J]. Control and Decision, 2006, 21(1): 46-50.)