

文章编号: 1000-8152(2002)01-0131-04

一种基于模糊感知器的模糊神经分类器*

许丽净 余英林

(华南理工大学电子与通信工程系·广州, 510640)

摘要: 隶属函数应能客观反映模式的不确定性分布情况, 模糊神经分类器中为每个输入特征划分相同数量的模糊变量显然是不合理的. 针对这一问题, 根据模糊积分的思想, 本文通过分析隶属函数的变化曲线来确定模糊密度, 以此为依据为输入特征划分合适的模糊变量, 并以 Iris 数据集为例验证了该方法的有效性.

关键词: 模糊感知器; 模糊积分; 模糊密度

文献标识码: A

A Neuro-fuzzy Classifier Based on Fuzzy Perceptron Model

XU Lijing and YU Yinglin

(Department of Electronic and Communication Engineering, South China University of Technology · Guangzhou, 510640, P. R. China)

Abstract: For a neuro-fuzzy classifier, it's unreasonable to partition the fuzzy sets of the same number for the different input features. In order to address this issue, fuzzy density is used here to partition the input feature reasonably. Iris data sets are used to illustrate the effectiveness of the proposed scheme.

Key words: fuzzy perceptron; fuzzy integral; fuzzy density

1 引言 (Introduction)

在模糊分类系统中, 模糊规则的形式如下:

if x_1 is μ_1 and x_2 is μ_2 and \dots and x_n is μ_n

then pattern (x_1, x_2, \dots, x_n) belongs to class i .

其中 μ_1, \dots, μ_n 是模糊集合. 专家的推理知识可以表达成用自然语言描述的规则, 而基于模糊集的可能性理论能良好地表征规则中的自然语言特征, 由此可以提高识别系统的智能度、识别过程的可理解性和可靠性. 但在模糊模式识别中, 寻找合适的隶属函数和模糊规则是一个很复杂的过程, 因此考虑将神经网络的学习能力引入到模糊系统的开发过程中构成模糊神经网络. 从模糊分类系统的角度来看, 模糊神经网络是一个利用神经网络学习算法来训练的模糊系统, 学习过程是基于数据, 不是基于知识的^[1].

2 模糊神经分类器学习算法 (Learning algorithm of neuro-fuzzy classifier)

本文讨论基于三层模糊感知器的模糊神经分类器 NEFLASS. 模糊感知器是多层模糊神经网络的一般模型, 结构类似一般的多层感知器, 权值被表示成模糊集合的形式, 并对激励函数、输出函数及传递

函数作相应改动. 三层模糊感知器定义如下^[2]:

1) $U = \cup U_i (i \in \{1, 2, 3\})$ 是非空集合. 对于 $i, j \in \{1, 2, 3\}, U_i \neq \phi, U_i \cap U_j = \phi (i \neq j)$. 其中, U_1 是输入层, U_2 是规则层, U_3 是输出层.

2) 设 $F(IR)$ 是论域 IR 的模糊子集, 权值被定义为: $W: U \times U \rightarrow F(IR)$. 当 $u \in U_i, v \in U_{i+1} (i \in \{1, 2\})$ 时, 存在权值 $W(u, v)$.

3) A_u 表示激活函数. 对于 $u \in U_1 \cup U_2, A_u: IR \rightarrow IR, a_u = A_u(\text{net}_u) = \text{net}_u$; 对于 $u \in U_3, A_u: F(IR) \rightarrow F(IR), a_u = A_u(\text{net}_u) = \text{net}_u$.

4) O_u 表示输出函数. 对于 $u \in U_1 \cup U_2, O_u: IR \rightarrow IR, O_u = O_u(a_u) = a_u$; 对于 $u \in U_3, O_u: F(IR) \rightarrow IR, O_u = O_u(\text{net}_u) = \text{DEFUZZ}_u(\text{net}_u)$, 其中 DEFUZZ_u 是去模糊函数.

5) NET_u 表示传递函数. 对于 $u \in U_1, \text{NET}_u: IR \rightarrow IR, \text{net}_u = ex_u$; 对于 $u \in U_2, (IR \times F(IR))^{t_1} \rightarrow [0, 1], \text{net}_u = \bigwedge_{u' \in U_1} (W(u', u)(O_{u'}))$, 其中 T 是 t 范数; 对于 $u \in U_3, \text{NET}_u: ([0, 1] \times F(IR))^{s_2} \rightarrow F(IR), \text{net}_u: IR \rightarrow [0, 1], \text{net}_u(x) = \bigwedge_{u' \in U_2} (T(O_{u'}, W(u', u)(x)))$, 其中 \perp 是 s 范数.

* 基金项目: 国家自然科学基金(69772026)及广东省自然科学基金(970484)资助项目.

收稿日期: 2000-04-18; 收修改稿日期: 2001-05-30.

6) 对于 $u \in U_1, ex: U_1 \rightarrow IR$, 定义外部输入 $ex(u) = ex_u$.

设计模糊系统的最简单的方法是调整规则权值. 在模糊系统中, 模糊规则是一个模糊样本, 是数据集中的多维模糊集合. 对模糊规则加权, 就改变了多维模糊集合的表示, 等价于改变规则命题的前提和结论, 会导致异常的模糊集合, 并使得相同的语言变量在不同的规则中表达形式不同, 会破坏模糊系统的可说明性. 因此, 如果模糊系统的可说明性是必要的, 应避免使用规则权值^[3]. 根据这一思路, 本文研究了 NEFCLASS. NEFCLASS 通过调整隶属函数的参数来设计模糊分类系统, 并在某些连接处采用了共享权值, 使得每个语言变量在所有的规则节点上的表达式相同^[1].

本文采用三角隶属函数, 表示为

$$\mu: IR \rightarrow [0,1], \mu(x) = \begin{cases} \frac{x-a}{b-a}, & x \in [a,b), \\ \frac{c-x}{c-b}, & x \in [b,c], \\ 0, & \text{其它.} \end{cases} \quad (1)$$

左右边界的隶属函数为半梯形分布. 学习算法由规则学习算法和模糊集合学习算法组成^[1]. 以 (p, t) 表示训练样本, 其中 $p \in IR^n$ 是输入模式, $t \in \{0,1\}^m$ 是类别的理想输出. 在规则学习算法中, 设 k 表示创建的规则数, k_{max} 表示最大规则数. 输入特征 x_i 由 q_i 个模糊集合表示, 算法如图 1 所示. 在此基础上, 在模糊集合学习算法中, 采用模糊误差反传算法调整三角隶属函数的参数 a, b, c , 算法如图 2 所示.

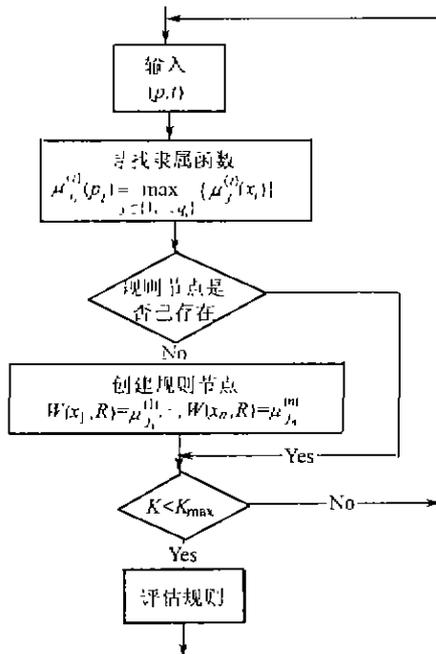


图 1 规则学习算法流程图

Fig 1 Flow chart of rule learning algorithm

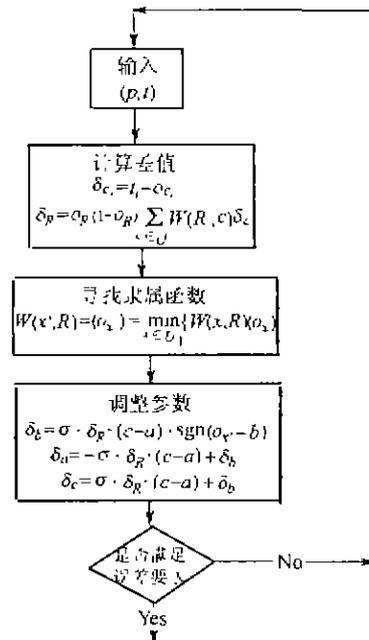


图 2 模糊集合学习算法流程图

Fig 2 Flow chart of fuzzy set learning algorithm

3 算法分析 (Algorithm analysis)

模糊模式分类中有基于模糊积分的直接匹配分类法和模糊神经分类器中采用的基于语言变量处理概念推理判别分类法. 尽管两者有很大区别, 我们仍可以利用直接匹配分类方法的思想, 利用模糊积分来分析上述学习算法. 下面首先简要介绍一下模糊积分的概念.

设 $(U, \bar{\lambda}, g)$ 是 F 测度空间, 其中 U 是全集, $\bar{\lambda}$ 是 U 上的 σ 域, $g(\cdot)$ 表示某集合的测试. $h: U \rightarrow [0,1]$ 是 U 上的可测函数, $A \in \bar{\lambda}$. h 在 A 上关于 g 的

模糊积分定义为^[4]:

$$\int_A h(u) \circ g(\cdot) = \sup_{\lambda \in [0,1]} (\lambda \wedge g(A \cap h_\lambda)). \quad (2)$$

其中 $h_\lambda = \{u | h(u) \geq \lambda\} (0 \leq \lambda \leq 1)$. 当 $A = U = \{u_1, u_2, \dots, u_n\}$, 且 $h(u_1) \geq h(u_2) \geq \dots \geq h(u_n)$ 时, 记 $A_i = \{u_1, u_2, \dots, u_i\}$, (2) 式变为:

$$\int h(u) \circ g(\cdot) = \bigvee_{i=1}^n (h(u_i) \wedge g(A_i)). \quad (3)$$

令 $g^i = g_\lambda(|x_i|)$, 称集合 $\{g^1, \dots, g^n\}$ 为 g_λ 的模糊密度函数, 其中第 i 个模糊密度 g^i 可以看做是 x_i

($i = 1, 2, \dots, n$) 的重要程度。

模糊积分可以用于输入模式的主观评估, 主观性包含在模糊测度中。正是这一非线性主观评价特点, 使得模糊积分在融合不同特征进行模式识别时具有潜在的应用价值。从模式分类过程来评价, 模糊积分可以看成是一种基于特征层次并结合主观经验的非线性分类器。令 X 代表输入模式, 由 n 个特征来刻画, $X = \{x_1, x_2, \dots, x_n\}$, 其中输入特征 x_i 用 q_i 个模糊集合表示。输出有 m 类, $\omega = [\omega_1, \dots, \omega_m]$ 。在直接匹配分类法中, 每一特征可以看作是识别某一类别的一个证据, 每一证据都具有鉴别某一类别的一种重要程度, 用主观测度表示。令 $\mu_j: X \rightarrow [0, 1]$, $\mu_j(x_i)$ 表示从单个特征值的角度来衡量 X 在 ω_j 类中的隶属程度, 这是客观的部分评估。通过模糊积分将客观的部分评估和主观测度综合在一起, 给出了模式 X 属于 ω_j 类程度的一个非线性估计结果。

在推理判别分类法中, 对于每个输入特征, 隶属函数不再是从单个特征值的角度直接来衡量模式隶属于某一类的隶属程度, 而是用来衡量输入特征隶属于某一语言变量的隶属程度, 再将其作为模糊规则的条件, 通过模糊推理得到分类结果。因此可以认为隶属函数是间接地从单个特征值的角度来衡量模式隶属于某一类的隶属程度。将推理系统视为一个函数, 对于输入特征 x_i , 将输出类别 $\bar{\omega} = [\bar{\omega}_1, \dots, \bar{\omega}_m]$ 近似映射成 $\bar{\omega}'_i = [\bar{\omega}'_{i1}, \dots, \bar{\omega}'_{iq}]$, 与输入特征的模糊集合相对应。经过这种变换后, 就可以采用直接匹配分类法思想, 利用模糊积分来研究在神经模糊分类器中客观的部分评估和主观测度是如何融合在一起的。

在模糊神经分类器中, 调整之前的隶属函数体现出客观的部分评估。无论调整规则权值或是调整隶属函数的参数都是对主观测度的评估。两种算法的不同之处在于如何将客观的部分评估和主观测度综合在一起。在调整规则权值的算法中, 利用模糊算子将二者结合起来, 数学概念清晰。而在调整隶属函数参数的算法中, 主观测度体现在对隶属函数的调整中, 在此基础上求输入特征的隶属程度、结合方式的数学概念不清晰, 但同样体现了模糊积分融合主客观特征进行模式识别的思想。

4 模糊划分的选择(Choice of fuzzy partition)

隶属函数应能客观地反映模式的不确定性分布情况, NEFCLASS 中为每个输入特征划分相同数量的模糊变量显然是不合理的, 本节利用模糊密度的概念来为每个输入特征量选择合适的模糊划分。以 Iris 数据集为例, 该数据集包括 150 个模式, 分别对应于 3 个类别, 每组模式包括 4 个输入特征。实验中利用前 75 组数据训练, 后 75 组数据检验分类效果。利用识别正确率和输出误差来说明分类效果。

当为每个输入特征划分相同数量和模糊变量时, 分类结果如表 1 所示, 其中 $[n_1, n_2, n_3, n_4]$ 中的 n_i ($i = 1, 2, 3, 4$) 表示为第 i 个输入特征划分的模糊变量的数目。由表 1 可见, 当划分过粗时, 模糊规则过少, 识别率降低; 而当模糊划分过细时, 输出误差较大。

表 1 分类结果

Table 1 Classification result

模糊划分	[2 2 2 2]	[3 3 3 3]	[4 4 4 4]	[5 5 5 5]	[6 6 6 6]
识别率	93.3%	97.3%	97.3%	97.3%	97.3%
输出误差	12.3253	9.2797	11.6863	12.1639	14.4937

本文通过判断输入特征模糊密度的大小程度来为每个特征量选择合适的模糊划分。一般而言, 特征域上不同类目标的模糊划分交叠越小, 特征就越典型, 所确定的模糊密度就越大, 反之越小^[5]。根据这一思想, 本文通过分析隶属函数的变化曲线来确定模糊密度的大小程度, 以此为根据来为每个输入特征量选择合适的模糊划分。

当模糊划分是 $[3 3 3 3]$ 时, 对应的隶属函数的变化曲线如图 3 所示, 图 (a), (b), (c), (d) 分别对应四个输入特征, 其中虚线表示调整之前的隶属函数曲线, 实线表示调整之后的隶属函数曲线。在图 3 (a), (b) 中, 位于中间的模糊变量的隶属函数完全重叠于前后两个模糊变量的隶属函数中, 说明模糊密度小, 特征不够典型, 因此为第一、第二输入特征划分三个模糊变量是不必要的。在图 3 (c) 中, 模糊划分的交叠小, 说明该特征需要较多的模糊划分来描述。通过调整, 当模糊划分为 $[2 2 4 3]$ 时, 识别率达 98.7%, 优于表 1 的结果。

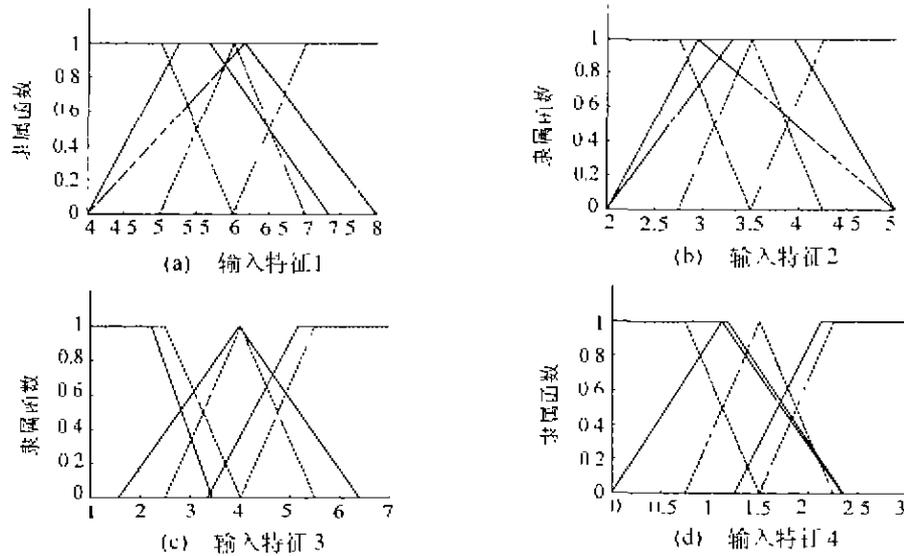


图3 模糊划分为[3 3 3 3]时的隶属函数变化曲线

Fig 3 The curve of membership function when the fuzzy partition is [3 3 3 3]

5 结论(Conclusion)

本文讨论了基于模糊感知器的模糊神经分类器 NEFCLASS. 实验表明利用模糊密度的概念来划分模糊变量可以提高识别率. 与调整规则权值相比, 调整隶属函数的参数避免了相同的语言变量在不同的规则中的表达形式不同这一语义问题, 但调整后隶属函数本身仍有语义问题. 模糊神经网络是基于数据而不是基于知识的, 是导致该问题的根本原因, 也是下一步工作中需要改进的地方.

参考文献(References)

- [1] Nauck D and Kruse R. What are neuro-fuzzy classifiers [A]. Proc. Seventh Int. Fuzzy System Association World Congress IFSA '97 [C]. Czech Republic: Academia Prague, 1997, 228 - 233
- [2] Nauck D. A fuzzy perceptron as a genetic model for neuro-fuzzy ap-

proaches [A]. Proc. Fuzzy-System '94, 2nd GI-Workshop [C]. Siemens Corporation, Munich, Germany, 1994

- [3] Nauck D and Kruse R. How the learning of rule weights affects the interpretability of fuzzy systems [A]. Proc. IEEE Int. Conference on Fuzzy Systems 1998 (FUZZ-IEEE '98) [C]. Anchorage, AK, USA, 1998, 1235 - 1240
- [4] Yang Lunbiao and Gao Yingyi. Theory and Application of Fuzzy Mathematics [M]. Guangzhou: South China University of Technology Publishing House, 1992 (in Chinese)
- [5] Guo Guorong. Fuzzy Pattern Recognition [M]. Changsha: National Defence University of Technology Publishing House, 1992 (in Chinese)

本文作者简介

许丽净 女, 1974年生, 华南理工大学电子与通信工程系博士生. 研究方向为模糊神经网络和模糊图象处理.

余英林 1932年生, 华南理工大学电子与通信工程系教授, 博士生导师. 研究方向为神经网络和图象处理.