

基于模糊神经网络的综合评判方法

李翔, 苏成, 王韶君

(中国矿业大学计算机学院, 徐州 221008)

摘要: 针对综合评判模型中主观因素影响评判精度问题, 引入模糊神经网络优化评判模型的方法, 采用模糊神经网络的训练过程逼近综合评判模型的权重集和隶属度函数。为了提高该方法的有效性, 对网络训练算法进行多方面的优化, 以便神经网络更快更稳定地收敛。结果证明该方法能较好地去除综合评判模型中的主观成分, 有效地提高评判的最终精度。

关键词: 综合评判; 神经网络; 权重; 评价矩阵

Comprehensive Evaluation Method Based on Fuzzy Neural Network

LI Xiang, SU Cheng, WANG Shao-jun

(College of Computer, China University of Mining and Technology, Xuzhou 221008)

【Abstract】 To solve the problem of getting a more precise result of comprehensive evaluations, fuzzy neural network method is introduced. This method uses a network training algorithm to cope with subjective factors in evaluations model. The algorithm can efficiently approximate the weight set and degree of membership function. To enhance its validity, many skills are adopted to optimize the network training process. The results show that the strategy can efficiently improve the final precisions. It is a better solution for synthesis evaluations problem.

【Key words】 comprehensive evaluations; neural network; weight; evaluation matrix

1 综合评判模型及其面临的问题

人们在社会实践中往往需要比较各种事物, 评价事物的优劣, 以便作出相应的处理。由于事物往往有多种属性或受多种因素影响, 因此在评价事物时, 必须对多个相关的因素进行综合考虑, 这便是综合评判。

取综合评判的模型为 $W \circ R = B$, 其中, W 代表权重集, 用来反映各因素的重要程度; R 代表单因素评判集; $R = r_{ij}$ 表示 i 种因素对 j 种评价的模糊隶属度; B 代表评价集; \circ 代表 W 与 R 的一种合成方式, 比如取最大最小合成算子。

以上为一般的模糊综合评价数学模型, 但权重集 W 的确定一般通过某领域的专家给出, 含有较多的人为因素。

为得到更准确的权重集, 可把神经网络引入计算^[1]。神经网络特有的逼近特性可通过对样本学习来自适应不确定的系统, 下面通过引入模糊神经网络得到较精确的评判权重集。

2 综合评判的模糊神经网络

2.1 模糊神经网络的结构

取神经元定义如下:

(1) 设 x_i 是模糊取大逻辑神经元的 n 个输入, $w_i \in [0,1]$ 是对应的权值, $i=1,2,\dots,n$, θ 为阈值, 作用函数 f 为

$$y = f(g(x)) = f(\bigvee_{i=1}^n (w_i \wedge x_i) - \theta) = \begin{cases} 1 & g(x) \geq 1 \\ g(x) & 0 < g(x) < 1 \\ 0 & g(x) < 0 \end{cases}$$

模糊取大逻辑神经元结构如图 1 所示。

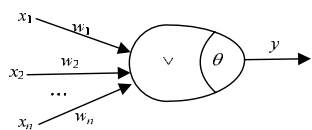


图 1 模糊取大逻辑神经元

(2) 设 x_i 是模糊取小逻辑神经元的 n 个输入, $w_i \in [0,1]$ 是对应的权值, $i=1,2,\dots,n$, θ 为阈值, 作用函数 f 为

$$y = f(g(x)) = f(\bigwedge_{i=1}^n (w_i \wedge x_i) - \theta) = \begin{cases} 1 & g(x) \geq 1 \\ g(x) & 0 < g(x) < 1 \\ 0 & g(x) < 0 \end{cases}$$

模糊取小逻辑神经元结构如图 2 所示。

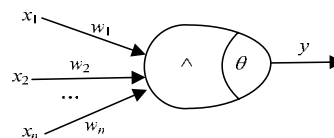


图 2 模糊取小逻辑神经元

对于一个综合评判模型, $W \circ R = B$, 其中, $W = (w_1, w_2, \dots, w_n)$ 为权向量, $R = (r_{ij})$ 为 n 行 m 列的评价矩阵, $0 \leq r_{ij} \leq 1$, $B = (b_1, b_2, \dots, b_m)$ 为评价结果向量^[2-3]。网络结构如图 3 所示。

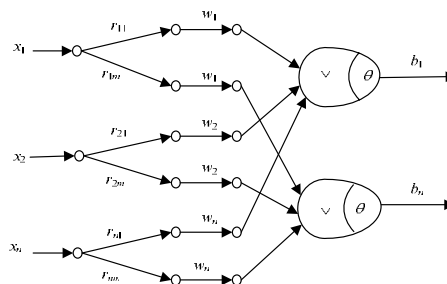


图 3 网络结构

作者简介: 李翔(1981-), 男, 硕士研究生, 主研方向: 神经网络, 数据挖掘; 苏成, 副教授; 王韶君, 硕士研究生

收稿日期: 2008-08-23 **E-mail:** lixiang1263@sina.com

在图 3 中, $X = \{x_1, x_2, \dots, x_n\}$ 是因素集, 被视为输入的第 1 层, 第 2 层输入 r_{ij} 为第 1 层输入的模糊隶属度, 用于表示第 i 个因素对第 j 个评价的隶属程度。第 2 层与第 3 层之间的连接 w_i 即为所要求的权重, 第 4 层是模糊逻辑运算, 最后是输出层 $b_j, j=1, 2, \dots, m$ 。

在进行样本训练时, 首先要对样本进行模糊化, 得到网络的输入集合 R 。关键是要选取合适的模糊隶属函数, 本文选用 $r_{ij} = u_{ij}(x_i)$ 。

样本模糊化后按如下网络训练方法得到权重集 $W^{[4]}$ 。

2.2 基本的网络学习算法

设 $b_j^*, j=1, 2, \dots, m$ 为期望输出向量。 $b_j, j=1, 2, \dots, m$ 为实际输出向量, $b_j = f(\bigwedge_{i=1}^n (w_{ij} \wedge x_i) - \theta)$ 。给定误差函数

$$E = \frac{1}{2} \sum_{j=1}^m (b_j^* - b_j)^2$$

(1) 给定初始值 $w_i(0)$, 评价矩阵的学习样本 R^1, R^2, \dots, R^p , 期望输出向量 $B_1^*, B_2^*, \dots, B_p^*$, 并给定误差 ε , 记 $w_i(t)$ 为 t 时刻第 i 个输入的权重向量, $i=1, 2, \dots, n$ 。

(2) 用学习样本对网络进行训练, 调整网络权值为

$$\Delta w_i(t) = w_i(t+1) - w_i(t) = \sum_{j=1}^m \frac{\partial E}{\partial b_j} \frac{\partial b_j}{\partial w_i} = \sum_{j=1}^m (b_j^* - b_j) \frac{\partial b_j}{\partial w_i}$$

其中, $\frac{\partial b_j}{\partial w_i}$ 为选取的模糊算子, 本文选用模糊取大算子。

(3) 按新的权值计算网络的输出和总误差, 若对每个样本都满足

$$\|B_p - B_p^*\| = \sqrt{\sum_{i=1}^m (b_i^{p_i} - b_i^{*p_i})^2} < \varepsilon$$

即对每个样本均稳定, 则停止迭代, 网络节点的值即为权重集 W 。否则, 跳转到(2)进行再训练。

2.3 学习过程的优化

为了更精确地得到需要的权重集, 需要对网络的学习算法在如下几个方面进行必要的优化:

(1) 确定学习率

在基本网络算法中, $\Delta w_i = w_i(t+1) - w_i(t)$ 引用了固定的学习率 $\eta=1$, 为了更好地控制网络的收敛性和学习速度, 可以根据需要选择自适应的变学习率, 一般取 $0 < \eta < 1$, 其权值调整公式如下:

$$\Delta w_i(t) = \eta(w_i(t+1) - w_i(t)) = \eta \left(\sum_{j=1}^m \frac{\partial E}{\partial b_j} \frac{\partial b_j}{\partial w_i} \right)$$

学习步长 η 是网络学习的一个重要参数, 在一定程度上决定了网络的收敛速度。学习步长过小会导致权重值更新量过小, 使收敛非常缓慢; 学习步长过大又会导致在极值点附近振荡的可能性加大, 乃至反复振荡而难以收敛。具体数值应根据对误差函数的影响决定。

(2) 加入动量项

这种措施又称为惯性校正法^[5], 其权值调整公式如下:

$$\Delta w_i(t) = \alpha \Delta w_i(t-1) + \eta \left(\sum_{j=1}^m \frac{\partial E}{\partial b_j} \frac{\partial b_j}{\partial w_i} \right)$$

其中, α 为动量系数。该方法在加速收敛的同时抑制了振荡。

(3) 改进误差函数

在基本网络算法中, 引用的函数是 $E = \frac{1}{2} \sum_{j=1}^m (b_j^* - b_j)^2$ 。

可以看到, 这是一个误差的绝对量, 不能有效地表征样本的相对误差程度。为了避免某个节点会出现误差的绝对量很大但比例却很小的情况, 引入相对误差函数:

$$\text{首先, 对 } E = \frac{1}{2} \sum_{j=1}^m (b_j^* - b_j)^2 \text{ 变形, } E = \frac{1}{2} \sum_{j=1}^m (b_j^* - b_j)^2 = \frac{1}{2} \sum_{j=1}^m (b_j^*)^2 (1 - b_j/b_j^*)^2 \text{, 把不含有实际输出的 } (b_j^*)^2 \text{ 提出, 得到}$$

$$\text{相对误差函数 } E_p = \frac{1}{2} \sum_{j=1}^m (1 - b_j/b_j^*)^2 = \frac{E}{\sum_{j=1}^m (b_j^*)^2} \text{。}$$

相对误差函数 E_p 使用误差量的相对比例来表征样本误差 ε , 具有更好的效果。

2.4 隶属函数的调整

当隶属函数不能精确给出而需要根据实际情况确定时, 需要对隶属函数进行调节。

同样, 对于上述 3 层前馈网络, 通常把输出层反传到中间层的误差作为中间层输出误差, 仍采用误差反传学习算法, 可以对隶属函数进行调整。

已知隶属度 $r_{ij} = u_{ij}(x_i)$, 其中, $u_{ij}(x_i)$ 为第 i 个因素对第 j 个指标的隶属函数。

$u_{ij}(x_i)$ 由函数本身的一些参数决定, 调整就归结为这些参数的调整。例如三角模糊数 $u_{ij}(x_i)$ 由中心 s_{ij} 、左移量 s_{ij}^L 和右移量 s_{ij}^R 决定。

$$u_{ij}(x_i) = \begin{cases} 1 - \frac{x_i - s_{ij}}{s_{ij}^R} & x_i \in [s_{ij}, s_{ij} + s_{ij}^R] \\ 1 + \frac{x_i - s_{ij}}{s_{ij}^L} & x_i \in [s_{ij} - s_{ij}^L, s_{ij}] \\ 0 & \text{其他} \end{cases}$$

由误差反传算法, 3 个参量的调整量为

$$\Delta s_{ij} = \frac{\partial E}{\partial s_{ij}}, \Delta s_{ij}^L = \frac{\partial E}{\partial s_{ij}^L}, \Delta s_{ij}^R = \frac{\partial E}{\partial s_{ij}^R}$$

当隶属函数取其他类型时, 调整的方式也会随之变化。

3 结束语

本文的模糊神经网络模型可以用于解决各种评判、预测问题及模糊系统中的主观因素问题, 较好地优化了权重向量和模糊隶属度函数, 并且网络的优化方法也可以保证网络更快更稳定地收敛。

参考文献

- [1] 仇春光, 刘玉树. 基于神经网络的多属性决策[J]. 北京理工大学学报, 2000, 20(1): 65-68.
- [2] 杜华英, 赵跃龙. 神经网络典型模型比较研究[J]. 计算机技术与发展, 2006, 12(5): 97-99.
- [3] 丛爽, 向微. BP 网络结构、参数及训练方法的设计与选择[J]. 计算机工程, 2001, 27(10): 36-38.
- [4] 侯祥林, 张春晖, 徐心和. 多层神经网络共轭梯度优化算法及其在模式识别中的应用[J]. 东北大学学报: 自然科学版, 2002, 23(1): 20-23.
- [5] 李晓峰, 刘光中. 神经网络 BP 算法的改进及其应用[J]. 四川大学学报: 工程科学版, 2000, 32(2): 105-109.

编辑 张正兴