

# 一种基于可信度的不确定性推理及其神经网络实现<sup>\*</sup>

施明辉<sup>1</sup>, 周昌乐<sup>1</sup>, 吴清锋<sup>1</sup>, 吴芸<sup>1</sup>, 张志枫<sup>2</sup>

(1. 厦门大学人工智能研究所, 福建厦门 361005; 2. 上海中医药大学研究生院, 上海 201203)

**摘要:** 提出了基于可信度因子和可信度区间的不确定性推理模型, 并用改进的 BP 神经网络实现其推理过程, 最后利用 MATLAB 神经网络工具箱给出仿真示例。改进的 BP 神经网络在实现不确定性推理方面有效避免了沿用传统方法所带来的规则数激增及推理缓慢等缺陷, 并提高了网络的泛化能力。仿真示例表明, 它不仅可以自动学习和模拟专家的典型经验, 而且还可以将专家的典型经验推广应用到一般情形。

**关键词:** 神经网络; 专家系统; 不确定性推理 机器学习

中图法分类号: TP311 文献标识码: A 文章编号: 1001-3695(2007)01-0241-03

## Uncertain Inference Based on Certainty Degree and Its Implementation by Artificial Neural Network

SHI Ming-hui<sup>1</sup>, ZHOU Chang-le<sup>1</sup>, WU Qing-feng<sup>1</sup>, WU Yun<sup>1</sup>, ZHANG Zhi-feng<sup>2</sup>

(1. Institute of Artificial Intelligence, Xiamen University, Xiamen Fujian 361005, China; 2. Graduate School, Shanghai University of T. C. M., Shanghai 201203, China)

**Abstract:** In this paper, we proposed an uncertain inference based on certainty degree (including certainty factor and certainty interval). A modified Back-Propagation (BP) Artificial Neural Network (ANN) is applied to realizing the uncertainty inference using two kinds of knowledge representation methods: certainty factor method and certainty interval method. First, the two methods of uncertainty knowledge representation are proposed. Next, BP network and its modified train function are presented. Then, how to apply the modified BP network to the uncertainty inference is illustrated in detail. Finally, an example for the certainty interval method using the MATLAB neural toolbox is given and analyzed. The simulation results show that ANN can play simple, but important and effective role in the uncertainty inference and that ANN can not only learn automatically experts' experiences, but also has the capability of generalizing the learned experience into more general situations according with experts' minds.

**Key words:** Neural Network; Expert System; Uncertainty Reasoning; Machine Learning

现实世界中客观事物或现象的不确定性导致了在各认识领域中的信息和知识大多是不确定的。自从 1967 年第一个专家系统 MYCIN 出现以来, 对于各种不确定性的研究就已引起了人们的重视。如何表示和处理不确定性知识成为人工智能研究的重要课题之一。相关的理论和应用层出不穷。二十世纪六七十年代出现了主观 Bayes 方法、确定性理论、可能性理论和证据理论等;八十年代以后, 又提出了灰色系统理论、粗糙集理论、可拓理论、集对分析等。这些理论从不同角度对不同类型的不确定性进行了研究, 并应用于各种场合。

### 1 基于可信度因子和可信度区间的知识表示

可信度是指人们对某一事物(或事情)的某方面性质在主观上确信的程度。传统的基于规则的知识表达方法如规则 1 所示。

规则 1: IF  $X_1$  AND  $X_2$  AND… AND  $X_n$  THEN  $Y_j$   
其中,  $X_i$  ( $i = 1, 2, \dots, n$ ) 表示前提;  $Y_j$  ( $j = 1, 2, \dots, m$ ) 表示结论。

这种表示方法虽然简单明了, 但在不确定性领域却显得捉

襟见肘。为了使基于规则表示的知识更为实用, 可以为规则中的带有不确定性的概念引入标志其可信程度的可信度因子或可信度区间。基于可信度因子的不确定性知识的表达方式可采用规则 2 所示。

规则 2: IF  $X_1(CF_{X_1})$  AND  $X_2(CF_{X_2})$  AND… AND  $X_n(CF_{X_n})$  THEN  $Y_j(CF_{Y_j})$

其中,  $X_i$  ( $i = 1, 2, \dots, n$ ) 表示前提;  $Y_j$  ( $j = 1, 2, \dots, m$ ) 表示结论;  $CF_{X_i}$  ( $i = 1, 2, \dots, n$ ) 是可信度因子, 表示相关前提的可信程度, 其取值区间为  $[0, 1]$ , 取值方法为

$$CF_{X_i} \begin{cases} 0 & \text{不可信} \\ 1 & \text{完全可信} \\ \text{其他} & \text{可信度的程度} \end{cases}$$

其中,  $CF_{Y_j}$  表示相关结论的可信程度, 其取值方法与  $CF_{X_i}$  类似。

规则的意义: 当前提  $X_i$  ( $i = 1, 2, \dots, n$ ) 的可信度因子为  $CF_{X_i}$  时, 则具有结论  $Y_j$  的可信度因子为  $CF_{Y_j}$ 。

更一般地, 采用可信度区间表示推理中“七八成”、“十有八九”等不确定性概念。可信度区间由两个可信度因子表示, 这两个可信度因子的取值分别是可信程度的下限和上限。基于可信度区间的不确定性知识的表达方式可采用如下规则:

规则 3: IF  $X_1(CF_{X_1}^L, CF_{X_1}^U)$  AND  $X_2(CF_{X_2}^L, CF_{X_2}^U)$

收稿日期: 2005-08-09; 修返日期: 2005-11-10

基金项目: 国家自然科学基金资助项目(60373000)

AND…AND  $X_n(CF_{X_n}^L, CF_{X_n}^U)$  THEN  $Y_j(CF_{Y_j}^L, CF_{Y_j}^U)$

其中,上标  $L, U$  分别表示下限和上限,如  $(CF_{X_1}^L, CF_{X_1}^U)$  表示的可信度区间是  $[CF_{X_1}^L, CF_{X_1}^U]$ 。

规则的意义:当前提  $X_i (i=1, 2, \dots, n)$  的可信度取值属于区间  $[CF_{X_i}^L, CF_{X_i}^U] (i=1, 2, \dots, n)$  时,则具有引发这些前提的结论  $Y_j$  的可信度取值属于区间  $[CF_{Y_j}^L, CF_{Y_j}^U]$ 。

由于对不确定性概念引入了可信度,使得原本模糊的概念定量化、清晰化,因而基于可信度的规则表示法更能反映中医思维方式,也更具有实用性。但是这种表示方法也有显著的缺陷:①对于可信度因子和可信度区间的取值不可能穷尽;②知识库中的规则数随着可信度因子和可信度区间的取值不同而成倍地增长,导致知识难以管理以及推理效率低下。然而,如果利用神经网络来实现基于可信度因子或可信度区间的不确定性推理,则可借助于神经网络良好的泛化能力和并行计算特性,有效地避免这些缺陷。

## 2 BP 神经网络及其改进

BP 神经网络是一种多层前馈神经网络,名字源于网络连接权值和阈值的调整规则采用的是误差反向传播(Back Propagation)学习算法,即 BP 学习算法。该算法是 Rumelhart 等人<sup>[1]</sup>在 1986 年提出的。由于 BP 神经网络结构简单、可塑性强,以及 BP 学习算法数学意义明确、步骤分明,BP 神经网络在函数逼近、模式识别、信息分类、数据压缩等领域得到了广泛的应用。

BP 算法是一种有监督学习算法,即对于一组确定的训练样本,给定了网络的期望输出值(目标输出值)。当样本中的输入作为网络的输入后,网络将产生实际输出。如果实际输出与样本的期望输出有误差,则根据 BP 算法,调整网络的连接权值和阈值,连接权值和阈值的不断调整过程就是使得网络的实际输出值逐步逼近期望输出值的过程。经过学习后的神经网络将知识分布于各连接权值和阈值上。神经网络将利用知识的过程转换为各神经元的并行计算过程。对于新的输入,神经网络将利用并行计算的特性迅速地在输出层得出响应。

在 BP 网络的训练过程中经常会出现网络泛化能力差的情况,即网络对于训练样本中的输入可以产生与期望输出误差很小的实际输出,但对于训练样本之外的新的输入却可能产生与相应的目标输出有较大误差的实际输出。

改进 BP 神经网络泛化能力的一种方法是规则化调整方法<sup>[2]</sup>。普通的 BP 神经网络都采用网络误差的均方根之和作为性能函数,如下所示:

$$mse = \frac{1}{N} \sum_{i=1}^N (e_i)^2 = \frac{1}{N} \sum_{i=1}^N (t_i - y_i)^2$$

其中,  $e_i, t_i, y_i$  分别表示第  $i$  个训练样本的训练误差、目标输出和网络实际输出。而规则化调整方法采用的网络性能函数如下式所示:

$$msereg = \gamma \times mse + (1 - \gamma) \times msw$$

其中,  $\gamma$  是性能参数,  $msw$  是网络中权值和阈值的均方和,其取值为  $msw = \frac{1}{n} \sum_{j=1}^n (w_j)^2$ , 其中  $n$  是网络中权值和阈值的个数。

采用规则化调整方法的性能函数训练神经网络可以减少

网络的有效权值和阈值,并且使网络的训练输出更加平滑,从而增强网络的泛化能力。但是性能参数  $\gamma$  的选择是一个不好解决的问题,如果太大,网络的泛化能力不好;如果太小,网络的训练精度又太低。在 MATLAB 神经网络工具箱中,提供了自动设置最优性能参数的函数 trainbr。该函数使用了 Bayesian 框架结构,假设网络的权值和阈值是特殊分布的随机变量,可用统计学的方法估计出  $\gamma$  的值。

## 3 用神经网络实现不确定性推理

用神经网络实现基于可信度因子的不确定性推理的具体算法步骤如下:

(1) 获取知识并形式化编码。从专家经验中总结出知识,并用规则表示。将规则中的前提和结论形式化编码,用变量表征每一个前提或结论。例如:用变量  $X_i (i=1, 2, \dots, n)$  表征各种前提,用变量  $Y_j (j=1, 2, \dots, m)$  表征各种结论,从而获得由形如规则 1 的规则组成的知识库。

(2) 获取典型的可信度组合。依据专家经验对变量取值。各变量的取值即专家认为的典型的前提与结论的可信度因子或可信度区间的组合,即

①对于可信度因子的组合,取  $X_i = CF_{X_i}, Y_j = CF_{Y_j}$

从而产生形如规则 2 的规则;

②对于可信度区间的组合,取

$$X_i = [CF_{X_i}^L, CF_{X_i}^U], Y_j = [CF_{Y_j}^L, CF_{Y_j}^U]$$

从而产生形如规则 3 的规则;

(3) 获取训练样本和测试样本。将  $X_i (i=1, 2, \dots, n)$  的值作为神经网络的输入,  $Y_j (j=1, 2, \dots, m)$  的值作为神经网络的期望输出,从而获得建立神经网络结构所需的训练样本和测试样本。

(4) 设计神经网络。

(5) 训练神经网络。以训练样本训练神经网络,训练过程中采用规则化调整方法。

(6) 测试神经网络。以测试样本测试神经网络。

在步骤(5)中,具体设计神经网络时,可采用基于可信度因子的方法,也可采用基于可信度区间的方法。根据结论间的关联程度,这两种方法又可采用以下两种方式来设计:①每一种结论的推断用一个神经网络来实现。②具有相同或相似前提的多个结论的推断用一个神经网络来实现。

对于方式①,神经网络与结论是一对一的关系,它有基于可信度因子与基于可信度区间两种推理模型,如图 1、图 2 所示。图 1 是基于可信度因子的推理模型,此时每一前提的可信度因子作为输入层的一个神经元的输入,输出层的神经元的输出作为结论的可信度因子;图 2 是基于可信度区间的推理模型,此时每一前提的可信度区间的下限和上限分别作为输入层的两个神经元的输入,输出层的两个神经元的输出分别作为结论的可信度区间的下限和上限。

对于方式②,神经网络与结论是一对多的关系,它也有基于可信度因子和基于可信度区间两种推理模型,如图 3、图 4 所示。图 3 是基于可信度因子的推理模型,此时每一前提的可信度因子作为输入层的一个神经元的输入,输出层的每一个神经元的输出作为一种结论的可信度因子;图 4 是基于

可信度区间的推理模型,此时每一前提的可信度区间的下限和上限分别作为输入层的两个神经元的输入,输出层中每两个神经元的输出分别作为一种结论的可信度区间的下限和上限。



#### 4 仿真示例

利用 MATLAB 7.0 提供的神经网络工具箱,可以方便地进行仿真实验。实际上,方式②是方式①的推广,基于可信区间的模型是基于可信度因子的模型的推广。限于篇幅,仅给出方式①的基于可信度区间的方法。具体步骤如下:

(1) 获取知识并形式化编码。以规则 1 作为示例规则。

(2) 获取典型的可信度组合。假设根据专家经验获得的前提与结论的可信度区间组合如表 1 所示。从表 1 中可以很容易人为地总结出前提与结论的可信度组合的规律。以下仿真实验将表明:利用表 1 中的数据训练神经网络,神经网络可以“学习”到前提与结论的可信度组合的规律,并能够将此规律推广应用。

表 1 前提与结论的可信度组合

编号	前提		结论			
	$X_1$		$Y_1$			
	$CF_{X_1}^L$	$CF_{X_1}^U$	$CF_{X_2}^L$	$CF_{X_2}^U$	$CF_{Y_1}^L$	$CF_{Y_1}^U$
1	0.2	0.3	0.2	0.3	0.2	0.3
2	0.3	0.4	0.3	0.4	0.3	0.4
3	0.5	0.6	0.5	0.6	0.5	0.6
4	0.6	0.7	0.6	0.7	0.6	0.7
5	0.7	0.8	0.7	0.8	0.7	0.8
6	0.8	0.9	0.8	0.9	0.8	0.9

(3) 获取训练样本和测试样本。将表 1 中编号为 1,3,5,6 的可信度组合作为神经网络的训练样本,编号为 2,4 的可信度组合作为测试样本。即

训练样本为

输入向量  $P = [0.2 \ 0.3 \ 0.2 \ 0.3; 0.5 \ 0.6 \ 0.5 \ 0.6; 0.7 \ 0.8 \ 0.7 \ 0.8; 0.8 \ 0.9 \ 0.8 \ 0.9]^T$

目标向量  $T = [0.2 \ 0.3; 0.5 \ 0.6; 0.7 \ 0.8; 0.8 \ 0.9]^T$

测试样本为

输入向量  $test\_input = [0.3 \ 0.4 \ 0.3 \ 0.4; 0.6 \ 0.7 \ 0.6 \ 0.7]^T$

目标向量  $test\_target = [0.3 \ 0.4; 0.6 \ 0.7]^T$

(4) 设计神经网络。BP 神经网络的设计要求确定网络的层数和各层中神经元的个数,以及各层神经元的传递函数等参数。Hornik 等人<sup>[3]</sup>已证明,三层 BP 网络模型结构可以很好地解决一般的函数逼近和模式识别问题。因此,这里采用三层 BP 网络结构。

根据图 2 所示的基于可信度区间的推理模型(一),可确定网络的输入层神经元个数为 4;网络的输出层神经元的个数为 2。隐含层神经元个数定为 20。隐含层神经元的传递函数

采用 S 型正切函数 tansig;输出层神经元的传递函数采用 S 型对数函数 logsig。最终确定的神经网络结构如图 5 所示。

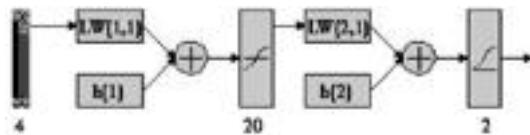


图 5 最终确定的神经网络结构

(5) 训练神经网络。神经网络的训练参数中除训练函数采用 trainbr 外,均采用 MATLAB 神经网络工具箱的默认参数。仿真程序中建立神经网络的语句为

```
net = newff([0 1;0 1;0 1;0 1],[20 2],{'tansig' 'purelin'},'trainbr')
训练曲线如图 6 所示。
```



图 6 神经网络训练过程曲线

(6) 测试神经网络。首先以训练样本的输入向量  $P$  作为网络输入进行测试。网络的实际输出向量为

$[0.1998 \ 0.2998; 0.5001 \ 0.6001; 0.7001 \ 0.8001; 0.7997 \ 0.8997]^T$  这与目标向量  $T$  非常接近。然后,以测试样本测试网络的泛化能力,即以测试样本的输入向量  $test\_input$  作为网络的输入进行测试。网络的实际输出向量为

$[0.2998 \ 0.3998; 0.6002 \ 0.7002]^T$

可以看出,网络的实际输出与目标输出  $test\_target$  非常接近,说明神经网络已具有一定的泛化能力,它所产生的输出符合它在训练阶段所获得的样本中输入与输出的规律。

#### 5 结束语

本文提出的基于可信度因子和可信度区间的不确定性推理反映了一种不确定性思维方式。由于继续沿用一般的规则表示法来实现这种推理具有规则数激增、推理缓慢等缺陷,以及传统的 BP 神经网络具有泛化能力不强的缺点,所以采用改进的 BP 神经网络来实现这种推理。MATLAB 仿真示例表明,采用改进的 BP 神经网络能够实现基于可信度因子和可信度区间的不确定性推理,并且可以有效避免上述缺陷;神经网络可以“学会”专家的典型经验,并将专家的典型经验进行推广应用,而这种推广应用是与专家思维规律相一致的;专家不再需要穷举各种前提与结论的可信度组合,只需给出若干典型组合样本即可,从而更具可操作性和实用性。此外,基于可信度因子与基于可信度区间的两种方法还可以相互结合,如前提的不确定性用可信度因子度量,结论的不确定性用可信度区间度量;或反之。用神经网络来实现不确定性推理的关键是需要能够正确反映前提与结论的可信度组合规律的典型样本。只要有了足够多的典型样本,神经网络就可以自动学习其中的规律,并利用其良好的泛化特性,将这些规律推广应用到更一般的推断情形中。

(下转第 312 页)