

一种基于 ANN 的中医辨证不确定性推理模型研究

施明辉,周昌乐

SHI Ming-hui,ZHOU Chang-le

1.厦门大学 人工智能研究所,福建 厦门 361005

2.厦门大学 计算机科学系,福建 厦门 361005

1.Institute of Artificial Intelligence,Xiamen University,Xiamen,Fujian 361005,China

2.Department of Computer Science,Xiamen University,Xiamen,Fujian 361005,China

E-mail:smh@xmu.edu.cn

SHI Ming-hui,ZHOU Chang-le.Approach to uncertain inference for syndrome differentiation in TCM based on artificial neural network.Computer Engineering and Applications,2007,43(7):10-13.

Abstract: In this paper,a modified Back-Propagation(BP) Artificial Neural Network(ANN) is applied to realize the uncertainty inference of syndrome differentiation in Traditional Chinese Medicine (TCM),which uses two kinds of knowledge representation methods:certainly factor method and certainty interval method.First,the two methods of uncertainty knowledge representation of syndrome differentiation in TCM are proposed.Second,BP network and its modified train function are presented.Third,how to apply the modified BP network to the uncertainty inference of syndrome differentiation in TCM is illustrated in detail.Finally,two simulative examples using the MATLAB neural toolbox,which include both certainty factor method and certainty interval method for each example,are given and analyzed.The simulation results show that ANN can play an simple,but important and effective role in the uncertainty inference of syndrome differentiation in TCM,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;Traditional Chinese Medicine(TCM);syndrome differentiation

摘要:提出中医辨证中不确定性推理的基于可信度因子和可信度区间的模型,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 神经网络在实现中医辨证不确定性推理方面有效地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,它不仅自动学习和模拟专家的典型经验,而且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;中医;中医辨证

文章编号:1002-8331(2007)07-0010-04 **文献标识码:**A **中图分类号:**TP183

1 引言

中医药现代化已列入国家中长期科技发展战略规划。如何充分利用现代科技手段改造中医的运用方式,完善中医的内部整合,是提高中医疗效,扩大其影响力,增强其生命力,实现中医现代化的重要措施^[1]。随着中医现代化研究的进展,计算机和人工智能技术被应用于中医现代化研究。近来,已有一些关于人工神经网络(简称神经网络)在中医证候诊断方面的初步研究。例如:文献[2]中,医学诊断的过程被看作一个映象问题,设想医学诊断是通过症状(原象)找出对应的诊断(象);文献[3]将神经网络应用于类风湿性关节炎病情的分级研究;文献[4]利用神经网络对一批乙肝临床病例数据作了分析;文献[5,6]则将神经网络应用于脉象的研究;文献[7]初步分析了神经网络方法用

于证候研究的可行性;文献[8]则将神经网络应用于中医八纲辨证诊断。可见神经网络正被越来越广泛地应用于中医诊断的研究中。

中医医生临床辨证的基本过程是:通过四诊方法获得患者的症状和体征等指标(以下统称为症状),然后运用掌握的知识或经验判断出病因、病性、病位和病势,并归纳为某种证型(主证)或若干次要证型(兼证),最后根据证型确定治法、给出处方和医嘱。辨证论治是中医诊断的一大特色。

从信息处理的角度看,中医辨证的诊断过程可看作是一个映射过程,即通过症状(原象)找出对应的证型集(象)的过程。在目前的大多数有关神经网络在中医辨证的应用研究中,考虑的症状空间^[9]是二值空间,即用变量表征各症状,若症状群表征

基金项目:国家自然科学基金(the National Natural Science Foundation of China under Grant No.60275023, No.60672018);厦门大学科技创新项目(No.XDKJCX20063011)。

作者简介:施明辉(1973-),男,讲师,在职博士研究生,研究方向:神经网络、模糊理论、粗集理论、中医智能信息处理;周昌乐(1959-),男,教授,博士生导师,研究方向:计算语言学、理论脑科学和认知逻辑学。

为

$$X=\{x_i|i=1,2,\dots,n\}$$

其中

$$x_i=\begin{cases} 0 & \text{当第 } i \text{ 个症状出现时} \\ 1 & \text{当第 } i \text{ 个症状不出现时} \end{cases}$$

则 $X=(x_1, x_2, \dots, x_n)$ 为 n 维空间中的一个点,显然 $U=\{(x_1, x_2, \dots, x_n)\}$ 就是所考虑的论域,每一个患者的临床症状集与 U 中的一个元素相对应。

然而,中医辨证的过程中,病人出现的症状,以及医生所作的诊断具有很强的模糊性和不确定性,某一症状的出现具有程度上的差异,而并非仅有出现与不出现两种情况,二值法不能反映这种程度上的差异。因此,有些学者很自然地想到用模糊数学的方法来处理中医诊断中的模糊性问题^[9-12]。应用模糊数学方法需要确定模糊子集及模糊矩阵,实际推理时还要计算贴近度,而这些有时并不容易做到。

本文首先提出采用基于可信度因子和可信度区间的方法表示中医辨证过程中的不确定性知识,并指出继续沿用传统的产生规则方法将带来规则数激增及推理缓慢等缺陷,而利用神经网络,则可借助于其良好的泛化能力和并行计算特性,有效地避免这些缺陷;然后介绍 BP 神经网络及其改进方法;接着阐述用改进的 BP 神经网络实现中医辨证不确定性推理的具体算法;最后利用 MATLAB 神经网络工具箱给出仿真示例。

2 基于可信度因子和可信度区间的中医辨证知识表示

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。中医辨证中许多概念都不是精确的,而是一种大致的判断,例如,中医理论中的八纲:阴、阳、表、里、寒、热、虚、实都不是精确的概念,在某种程度上都带有模糊性和不确定性。因此,可信度的概念非常适合于中医知识的表达。

传统的基于规则的知识表达方法虽然简单明了,但在中医辨证的不确定性领域却显得捉襟见肘。例如,中医舌诊理论中有一条经验“舌黑而干燥者肾阴虚”,这条经验用规则可表示为:

Rule_1:IF 舌黑 AND 舌干燥 THEN 肾阴虚

但是实际应用中,“舌黑”、“舌干燥”、“肾阴虚”都是不确定的概念,因此这条规则在计算机模拟中医辨证时并不实用。

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识其可信程度的可信度因子或可信度区间。

一般地,基于可信度因子的中医辨证不确定性知识的表达方式可采用如下规则所示:

Rule_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, $X_i(i=1,2,\dots,n)$ 表示症状; $Y_j(j=1,2,\dots,m)$ 表示证候; $CF_{X_i}(i=1,2,\dots,n)$ 是可信度因子,表示相关症状的可信程度,其取值区间为 $[0,1]$,取值方法为

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

CF_{Y_j} 表示相关证候的可信程度,其取值方法与 CF_{X_i} 类似。

规则的意义:当症状 $X_i(i=1,2,\dots,n)$ 的可信度因子为 CF_{X_i}

时,则具有引发这些症状的证候 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中医辨证推理中“七八成”、“十有八九”等不确定性概念。某一不确定性概念的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。基于可信度区间的中医辨证不确定性知识的表达方式可采用如下规则所示:

Rule_3:IF $X_1(CF_{X_1}^L, CF_{X_1}^U)$ AND $X_2(CF_{X_2}^L, CF_{X_2}^U)$ 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”,例如: $(CF_{X_i}^L, CF_{X_i}^U)$ 表示的可信度区间是 $[CF_{X_i}^L, CF_{X_i}^U]$ 。

规则的意义:当症状 $X_i(i=1,2,\dots,n)$ 的可信度区间为 $[CF_{X_i}^L, CF_{X_i}^U](i=1,2,\dots,n)$ 时,则具有引发这些症状的证候 Y_j 的可信度区间为 $[CF_{Y_j}^L, CF_{Y_j}^U]$ 。

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:一方面对于可信度因子和可信度区间的取值不可能穷尽;另一方面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中医辨证不确定性推理,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。

3 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。BP 学习算法是 Rumelhart 等在 1986 年提出的^[13]。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 神经网络结构是一种单向传播的多层前馈神经网络结构。它是一种具有三层或三层以上的神经网络,包括输入层、中间层(隐含层)和输出层。前后层之间实现全连接,而每层神经元之间无连接。输入层和输出层神经元的个数可以由具体问题决定,而中间层的神经元个数目前尚无理论上的规定,一般根据经验或者反复实验确定。

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。神经网络把利用知识的过程转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^[4]。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,如下式表示:

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

其中, e_i 、 t_i 、 y_i 、 N 分别表示第 i 个训练样本的训练误差、目标输出、网络实际输出及输入层神经元个数。而规则化调整方法采用的网络性能函数如下式所示:

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

其中 γ 是性能参数, msw 是网络中权值和阈值的均方和, 其取

$$msw = \frac{1}{n} \sum_{j=1}^n (w_j)^2$$

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 γ 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, 然后用统计学的方法估计出 γ 的值。

4 用神经网络实现中医辨证不确定性推理

用神经网络实现基于可信度因子的中医辨证不确定性推理的具体算法如下:

(1) 获取知识

从中医理论或专家经验中总结出中医辨证知识, 并用不带可信度的规则表示 (如 Rule_1 所示), 从而形成一般的规则知识库。

(2) 形式化编码

将形如 Rule_1 的规则中的症状和证候形式化编码, 用变量表征每一个症状和证候。例如: 用变量 X_i ($i=1, 2, \dots, n$) 表征各种症状, 用变量 Y_j ($j=1, 2, \dots, m$) 表征各种证候。

(3) 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是典型的症状与证候的可信度因子或可信度区间的组合, 即

$$X_i = CF_{X_i}, Y_j = CF_{Y_j}$$

从而产生形如 Rule_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 Rule_3 的规则。

(4) 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。

(5) 设计神经网络

具体设计神经网络时, 可以采用基于可信度因子的方法, 也可以采用基于可信度区间的方法。根据证候间的关联程度, 这两种方法又都可以采用以下两种方式来设计。

方式 1 每一种证候的诊断用一个神经网络实现。

方式 2 具有相同或相似症状的多个证候的诊断用一个神

经网络实现。

对于方式 1, 神经网络与证候是一一对一的关系, 它有基于可信度因子和基于可信度区间两种模型, 如图 1、图 2 所示。图 1 是基于可信度因子的模型, 此时每一症状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证候的可信度因子; 图 2 是基于可信度区间的模型, 此时每一症状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证候的可信度区间的下限和上限。

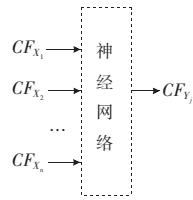


图 1 基于可信度因子的辨证模型(1)

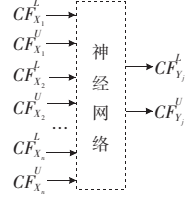


图 2 基于可信度区间的辨证模型(1)

对于方式 2, 神经网络与证候是一对多的关系, 它也有基于可信度因子和基于可信度区间两种模型, 如图 3、图 4 所示。图 3 是基于可信度因子的模型, 此时每一症状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证候的可信度因子; 图 4 是基于可信度区间的模型, 此时每一症状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证候的可信度区间的下限和上限。

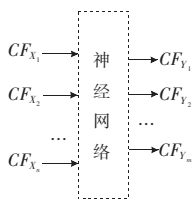


图 3 基于可信度因子的辨证模型(2)

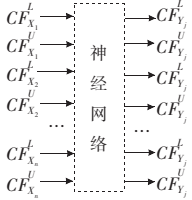


图 4 基于可信度区间的辨证模型(2)

(6) 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。

(7) 测试神经网络

以测试样本测试神经网络。

5 仿真示例

利用 MATLAB7.0 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, 基于可信度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

以 Rule_1 作为示例规则。

(2) 形式化编码

将规则 Rule_1 中的症状“舌黑”、“舌干燥”和证候“肾阴虚”进行形式化编码如表 1 所示。

(3) 获取典型的可信度组合

假设根据专家经验获得的症状与证候的可信度区间组合

表 1 症状与证候的可信度组合

编号	症状				证候	
	舌黑 X_1		舌干燥 X_2		肾阴虚 Y_1	
	$CF_{X_1}^t$	$CF_{X_1}^f$	$CF_{X_2}^t$	$CF_{X_2}^f$	$CF_{Y_1}^t$	$CF_{Y_1}^f$
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

如表 1 所示,从表 1 中可以很容易人为地总结出症状与证候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”到症状与证候的可信度组合的规律,并能够将此规律推广应用。

(4) 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 $test_target=[0.3 \ 0.4; 0.6 \ 0.7]^T$

(5) 设计神经网络

这里仍然采用三层 BP 网络结构。根据图 3 所示的基于可信度区间的辨证模型(1),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。

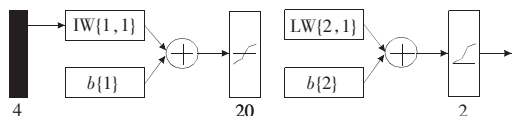


图 5 最终确定的神经网络结构

(6) 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。

(7) 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

$[0.199 \ 8 \ 0.299 \ 8; 0.500 \ 1 \ 0.600 \ 1; 0.700 \ 1 \ 0.800 \ 1; 0.799 \ 7 \ 0.899 \ 7]^T$

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

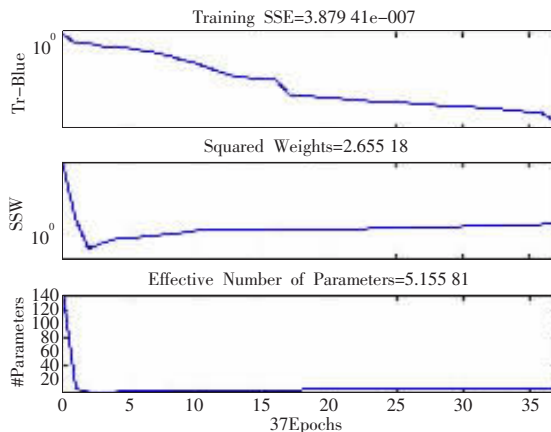


图 6 神经网络训练过程曲线

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

$[0.299 \ 8 \ 0.399 \ 8; 0.600 \ 2 \ 0.700 \ 2]^T$

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。

6 讨论

文中提出的基于可信度因子和可信度区间的中医辨证不确定性推理模型,反映了中医辨证的思维方式。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,故采用改进的 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中医辨证不确定性推理,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症状与证候的可信度组合,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,例如:症状的不确定性用可信度因子度量,证候的不确定性用可信度区间度量。用神经网络来实现不确定推理的关键是需要能够正确反映症状与证候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至更一般的诊断情形中。(收稿日期:2006 年 12 月)

参考文献:

- [1] 朱文锋.中医(辅助)诊疗系统的研究[J].中国中医基础医学杂志, 2003,9(1):8-11.
- [2] 陈五零,王存冉,郭荣江.神经网络模型及其在中医诊断方面的应用[J].中华医学杂志,1991,71(2):111-113.
- [3] 周志坚,王宗源,邓兆智.神经网络在类风湿性关节炎病情分级中的应用初探[J].生物医学工程学杂志,1999,16(4):479-482.
- [4] 边沁,何裕民,施小成,等.基于 MFBP 算法的中医证型的神经网络模型初探[J].中国中医基础医学杂志,2001,7(5):66-69.
- [5] 王璐,吴南健,温殿忠.人工神经网络在孕妇脉象判别中的应用[J].黑龙江大学自然科学学报,2003,20(4):66-72.