

几乎不包含什么语言学知识,其分析结果不利于进一步的利用和处理,尤其是不利于口语翻译中目标语的生成。因此,将符号主义与联结主义相结合的方法就受到了格外的青睐。Karlsruhe 大学的 Finn Dag Buo 等人将神经网络用于口语特征结构的分析,在进行组块分析时使用了多个含有一个隐藏层的前向网络,并采用 BP 算法和 PCL 算法进行网络参数的学习^[1]。汉堡大学的 SCREEN 系统^[2]等也使用了神经网络的方法。在将神经网络用于口语翻译时,需着重考虑两个问题,一是网络的拓扑结构,另一个是网络的学习算法。

结束语 口语翻译的研究极具理论和实践意义,对于我国计算语言学的研究者们来讲,汉外口语翻译的研究尤其是一个大有可为的领域。目前,我们已实现了一个面向会面安排领域的汉英口语语音翻译的原形系统^[3],取得了满意的效果,我们将继续在这方面进行深入研究。

参考文献

- [1] Hiroaki Kitano, Speech-to-Speech Translation: A Massively Parallel Memory-Based Approach, Kluwer Academic Publishers, Boston, 1994
- [2] 王海峰、高文、李生,面向受限领域的汉英语口语翻译,《语言工程》,第三届中国计算机智能接口与智能应用学术会议论文集,清华大学出版社,1997. 8
- [3] Alex Waibel, Interactive Translation of Conversational Speech, IEEE Computer, 29(7), 1996
- [4] Bill Z. Manaris, Brian M. Slator, Interactive Natural Language Processing: Building on Success, IEEE Computer, 29(7)1996
- [5] John. Hutchins, Latest Developments in MT Technology: Beginning a New Era in MT Research, MT Summit IV, Kobe, Japan, 1993. 7
- [6] 赵铁军、李生、高文,机器翻译研究的现状与发展方向,计算机科学,23(3)1996

(下转第55页)

(上接第76页)

(4)重新计算各子类的中心点: $C_i = \frac{1}{n_i} \sum_{x_j \in I} X_j$; n_i 为子类 i 中的样本总数, X_j 为所有属于子类 i 的样本。

(5)计算各子类的偏差值 b_i , 以子类中所有样本到中心点的距离的平均值来度量,即:

$$b_i^2 = \frac{1}{n_i} \sum_{x_j \in I} \|X_j - C_i\|^2$$

2.2 有监督学习过程

这一步通过学习得到输出层与隐层之间的权矢量,使网络的输出值与样本的理想输出之间的误差最小。定义误差函数 $E = \frac{1}{2} \|Y_{des} - Y\|^2$, Y_{des} 为样本理想输出矢量, Y 为网络实际输出矢量,由式(3)求得。采用梯度下降法求解,并使用动量因子 α 提高学习速度,在迭代过程中,学习速率 η 和动量因子 α 自动调整。权值学习的主要步骤如下:

(1)以小的随机数初始化权值矩阵;

(2)由公式(3)计算网络输出值 Y ;

(3)计算绝对误差

$$e = Y_{des} - Y;$$

(4)修正权值:

$$\Delta W_i(k+1) = \eta \cdot e \cdot R_i(X) + \alpha \cdot \Delta W_i(k),$$

$$W_i(k+1) = W_i(k) + \Delta W_i(k+1);$$

(5)计算系统误差,判断终止条件。若条件成立,退出循环;否则转(2)继续迭代。

小结 通过模拟实验,我们发现:采用 RBFNN 学习时,聚类粒度对规则总数和学习精度有影响,但测试精度基本稳定,而且精度较高。另外, RBFNN 学习速度快,一般只需几十次或几百次迭代就能得到稳定的输出。除聚类粒度外,与其它学习参数的初始值基本无关,因此学习精度容易控制。

参考文献

- [1] 刘有才、刘增良,模糊专家系统的原理与设计,北京:北京航空航天大学出版社,1995
- [2] Nikola K. Kasabov, Learning fuzzy rules and approximate reasoning in fuzzy neural networks and hybrid systems, Fuzzy Sets and Systems, V82, 1996
- [3] Y. J. Chen, Rule combination in a fuzzy neural network, Fuzzy Sets and Systems, Same to [2]
- [4] 赵群、保铮, RBFNN 的分类机理,通信学报, No. 2, 1996
- [5] 李凡,模糊专家系统,武汉:华中理工大学出版社, 1994

模糊系统 专家系统 神经网络 径向基函数

75-76,50

一种径向基函数神经网络生成模糊系统的方法研究

A Method for Generating Fuzzy Systems Using Radial Basis Function Neural Networks

王长琼 孙国正

TP18

(武汉交通科技大学港机 CAD/CAE 研究中心 武汉 430063)

摘要 In certain conditions, a RBFNN (Radial Basis Function Neural Networks) is equivalent to a fuzzy systems in function. So, the learning algorithm presented here, combining unsupervised and supervised learning, can automatically extract fuzzy rules and determine the number of rules from sample data.

关键词 Neural Networks, Fuzzy systems, Neuro-fuzzy system, Radial Basis Function

如何建立合适的模糊规则,是模糊系统设计的关键和难点。传统的方法是依靠统计分析或经验建立模糊规则库^[1],不仅难度大,而且建立的模糊系统缺乏适应能力。人工神经网络(ANN)技术的发展为模糊规则的自动获取提供了一条新途径。许多学者研究 ANN 与模糊系统的融合问题,其主要目的就是利用 ANN 的学习能力和自适应能力,从样本中提取模糊规则,形成具有自适应能力的模糊系统。尽管利用多层前馈网获取模糊规则^[2-3]具有较好的效果,但这种方法存在的问题是,模糊规则的产生过程解释困难,有时候还需初步估计规则数目,或研究其它方法对模糊规则进行精炼,即不能自动确定规则数目。

最近,径向基函数神经网络(RBFNN)引起许多研究者的兴趣。RBFNN 是一种使用局部作用场函数的神经网络,同样具有很强的非线性逼近能力,且

逼近速度快,在模式分类、预测及系统建模方面的应用前景非常广阔^[4]。

为了利用 RBFNN 实现模糊系统功能,首先必须弄清模糊系统与 RBFNN 的联系,分析用 RBFNN 实现模糊系统的可能性和特点;学习算法的好坏,将直接影响所生成的模糊系统性能,本文以下部分将按这一思路展开。

1 模糊系统与径向基函数神经网络

1.1 模糊系统

模糊系统实质上是利用模糊规则进行推理和决策的专家系统。如图1所示,一个模糊系统主要由模糊处理、推理机、模糊规则库和去模糊处理四部分组成。模糊处理和去模糊处理是为了提供模糊系统与环境之间的作用界面。各部分功能及系统工作过程如下:

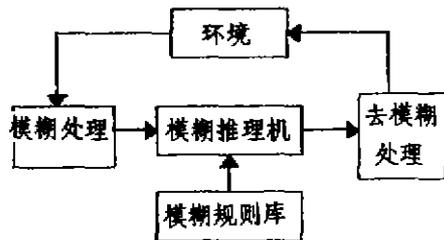


图1 模糊系统示意图

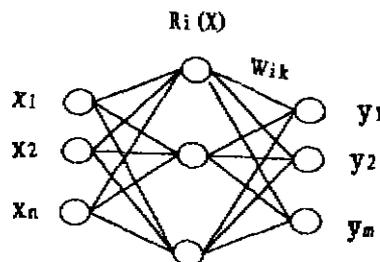


图2 RBFNN 拓扑结构示意图

王长琼 博士生,研究方向为神经网络、模糊专家系统和混合人工智能的理论及其在机械中的应用,孙国正 教授,博士生导师。

(a)模糊处理:将输入信息转换成便于系统处理的模糊数即隶属度,隶属度形式可由经验确定;

(b)模糊推理:将观测数据与模糊规则进行匹配,根据匹配度执行模糊规则;

(c)去模糊处理:由于模糊匹配会出现多条规则被执行的情况,必须将这种模糊的多值输出进行清晰化处理,得到便于理解的单一值,这就是去模糊处理。去模糊的方法有重心法、平均法、最大值法等。

模糊系统的工作能力取决于知识库中模糊规则的数量和质量,模糊知识获取方法的研究是模糊系统设计的一个非常重要的研究内容。

1.2 径向基函数神经网络 RBFNN

RBFNN 是一种使用局部作用场函数的神经网络,其拓扑结构如图2所示,图中 $X=(x_1, x_2, \dots, x_n)$ 为输入矢量;隐层激活函数可为任意形式的局部作用场函数。最常选用的是高斯函数,当选用高斯函数为隐层激活函数时,隐层第 i 个节点的输出为:

$$R_i(x) = \exp[-\|X - C_i\|^2 / 2b_i^2] \quad (1)$$

式中, C_i 为第 i 个节点高斯函数的中心点,与输入矢量同维数; b_i 是该节点高斯函数中的偏差值,输出层激活函数为线性函数,输出矢量为:

$$Y = (y_1, y_2, \dots, y_m)^T = f(x) = \sum_{i=1}^{i=L} W_i \cdot R_i(x) \quad (2)$$

这里, m, L 分别为输出层节点数和隐层节点数, $W_i = (w_{i1}, w_{i2}, \dots, w_{im})^T$ 为隐层第 i 个节点与输出层各节点之间的权值。归一化处理后的输出为:

$$Y = f(x) = \frac{\sum_{i=1}^{i=L} W_i \cdot R_i(X)}{\sum_{i=1}^{i=L} R_i(X)} \quad (3)$$

在 RBFNN 模型中,隐层节点数 L 、隐层各节点的 C_i 和 b_i 、以及隐层与输出层之间的权重 W 是需要通过学习得到的参数。

1.3 RBFNN 与模糊系统的关系

要想用 RBFNN 实现模糊系统功能,先必须探讨两系统之间的内在联系和实现条件。假设一个简单的模糊系统,其中第 i 条规则为:

IF x_1 IS A_{i1} and x_2 IS A_{i2} ... and x_n IS A_{in} ,

THEN y_1 IS w_{i1} and y_2 IS w_{i2} ... and y_m IS w_{im} ;

这里, $A_{i1}, A_{i2}, \dots, A_{in}$, 为输入变量的模糊取值,用其隶属度表示; w_{i1}, \dots, w_{im} 是结论分量。

如果用高斯函数 $R(X)$ 表示各前提的隶属度,对任一输入事实 X 而言,事实与规则的匹配度就是事实 X 与各子前提的匹配度,根据模糊数学理论^[5],就是 X 在各前提集中的隶属度 $R(X)$ 。假设该模糊系统共有 L 条规则,事实 X 与每条规则的匹配度分别为 $R_1(X), R_2(X), \dots, R_L(X)$; 每条规则当匹配

度等于1时,产生的第 k 个输出分量分别为 $w_{1k}, w_{2k}, \dots, w_{Lk}$ 。事实与每条规则的匹配度可看作是各条规则对结论的贡献程度(即权重)。如果选用重心法进行去模糊处理,则模糊系统中的第 k 个输出分量为:

$$y_k = \frac{\sum_{i=1}^{i=L} w_{ik} \cdot R_i(x)}{\sum_{i=1}^{i=L} R_i(x)} \quad (4)$$

由此可见,模糊系统的第 k 个输出分量与 RBFNN 的第 k 个输出分量具有相同的数学形式。因此,当一个模糊系统取高斯函数为其规则的隶属度、采用重心法进行去模糊处理、取模糊规则数等于隐层节点数时,可以用 RBFNN 实现模糊系统的功能。利用 RBFNN 的学习能力,就能从样本中提取模糊规则,这样,模糊系统的性能将直接取决于 RBFNN 的学习算法。

2 基于 RBFNN 的模糊系统的算法实现

一个具有自适应能力的模糊系统,应该具有随环境的变化而调整规则的能力,包括调整规则的数目和各条规则的参数。根据上一节的公式(4),基于 RBFNN 的模糊系统需学习的参数有:各条规则的前提隶属度 $R_i(x)$,即高斯函数的中心点 C_i 和偏差值 b_i 、各规则的结论分量 w_{ik} (即 RBFNN 的权矢量),模糊规则的数目即隐层节点数 L 。

本文研制的学习算法是一种将无监督学习与有监督学习结合起来的混合算法,该算法能自动形成隐层节点数,学习过程分为两步,第一步为无监督学习过程,通过竞争学习得到参数 L, C_i 和 b_i ; 第二步为有监督学习过程,学习得到输出层与隐层之间的权矢量。

2.1 无监督学习算法

这一竞争学习方法,对训练样本进行自动聚类,从而得到隐层诸参数。进行聚类时,采用欧氏距离作为聚类标准,以子类所允许的样本分布范围作为聚类粒度,用符号 r 表示。无监督学习过程如下:

(1)输入所有训练样本,选择聚类粒度 r 。

(2)用第一个训练样本对聚类中心点进行初始化,即 $C_1 = X_1 = (x_{11}, \dots, x_{1n})$,令隐层节点数 $L=1$ 。

(3)对所有训练样本,计算该样本到各子类中心点的欧氏距离。如果样本 X_i 到子类 i 的中心点 C_i 的距离最小,该样本有可能被划分到子类 i 中。令距离 $D_i = \min\|X_i - C_i\|$ 。若 $D_i \leq r$,则 X_i 确实属于子类 i ,转

(4)重新计算该子类中心点;否则,形成一个新子类,并以该样本为新子类的中心点。 (下转第50页)