

使用不同置信级训练样本的神经网络学习方法

高学星^{*①} 孙华刚^② 侯保林^①

^①(南京理工大学机械工程学院 南京 210094)

^②(军械技术研究所 石家庄 050000)

摘要: 针对含不同置信级样本的模型拟合问题, 该文提出了一种基于神经网络的二次学习方法。文中指出真实模型是实验模型的一种变异, 提出逼近真实模型期望值的神经网络, 是融合先验样本和真实样本的最佳网络。首先, 以先验样本为训练样本进行第1次神经网络学习, 并计算取决于硬点信息的软点误差容量区间; 然后, 同时将先验样本和真实样本作为训练样本, 利用软点误差容量区间和硬点误差敏感系数, 对神经网络训练过程中输入/目标对的误差进行修改, 通过第2次学习得到既能精确拟合真实样本, 又能最大化利用先验样本信息的综合网络。与基于知识的神经网络(KBNN)相比, 该方法更加简单, 可控性更强并具有更加明确的逻辑意义。

关键词: 神经网络; 模型拟合; 基于知识的神经网络(KBNN); 先验知识

中图分类号: TP183

文献标识码: A

文章编号: 1009-5896(2014)06-1307-05

DOI: 10.3724/SP.J.1146.2013.01099

A Neural Network Learning Method Using Samples with Different Confidence Levels

Gao Xue-xing^① Sun Hua-gang^② Hou Bao-lin^①

^①(School of Mechanical Engineering, Nanjing University of Science and Technology, Nanjing 210094, China)

^②(Ordnance Technical Institute, Shijiazhuang 050000, China)

Abstract: To solve the model-fitting problem with different confidence levels of samples, a Neural-Network (NN)-based twice learning method is proposed. It is pointed out that the real model is a variation of experimental model. The neural network approximation to the mathematical expectation of real model, is believed to be the best network fusing the information of prior samples and real samples. In the first learning, neural network is trained using the prior samples only, and the error capacity intervals of the soft points, which are determined by the information of hard points, are calculated. Then, both prior samples and real samples are included in the training samples. The import-objective errors in the process of NN training are modified, using soft point error capacity intervals and hard point error-sensitivity coefficients. The expected network is generated by the second learning, with accurate fitting to the real samples and efficacious utilization of the prior samples. In contrast with Knowledge-Based Neural Networks (KBNN), this method is simpler and more amenable to manipulation with definite logical significance.

Key words: Neural Networks (NN); Model-fitting; Knowledge-Based Neural Networks (KBNN); Prior knowledge

1 引言

机器学习算法在很多应用领域被证明很有使用价值, 可以有效地解决数据挖掘等需要动态适应条件变化, 或者超出人类当前认知范围的某些困难问题^[1]。机器学习的效果严重依赖于训练样本的数量和质量^[2], 通常, 来自真实使用过程的真实样本比较少而且分布不全面, 此时真实样本不适合直接用来建模^[3]。1990年, 文献[4]提出了联合符号化领域理

论和神经网络的基于知识的神经网络方法(Knowledge-Based Neural Networks, KBNN)。与传统神经网络相比, KBNN提供了一种使用特定问题知识(problem-specific knowledge)来补偿样本缺少的学习途径^[5,6]。KBNN在非线形动态建模^[7]、控制^[8]、监测诊断^[9,10]、预测^[11,12]和知识库的补充^[13]等方面得到了应用。但总体来说, 其应用并不广泛^[14], 主要有两方面原因: 对于复杂系统, 先验领域知识不易获得, 不准确的领域知识甚至会对算法造成误导^[15]; 另外, 尚无证据表明神经网络的结构因素变化具有明确的意义^[16]。

2013-07-25 收到, 2013-12-09 改回

国家自然科学基金(51175266/E050604)资助课题

*通信作者: 高学星 hss1737@gmail.com

实际上, 真实样本稀少时, 不仅可以利用公式化的领域知识进行补充, 实验样本作为数据形式的先验知识, 同样可以与真实样本结合进行训练学习。可以将实验样本学习到的神经网络作为另一种形式的领域知识, 应用 KBNN 对真实样本进行学习。KBNN 会通过初始网络权值进行一定范围内精化修正的方法实现对真实样本的进一步学习。

本文基于不同置信级样本(真实样本与实验样本)间的误差分析, 针对含先验样本的拟合问题, 提出了一种不同于 KBNN 的神经网络 2 次学习方法, 通过修改标准神经网络算法中输入/目标对误差进行第 2 次训练, 以获得融合先验样本和真实样本的综合网络。

2 基本思路

虽然实验或者虚拟实验更容易获得大量可规划的样本, 但大多数情况下的实验模型相对真实模型都有一定的误差, 该误差可能是因为实验条件的不完全真实性, 或者样本的获取途径差异, 所以真实样本和实验样本是具有不同置信度级别的。不妨称实验样本为软点样本, 真实使用中获取的样本为硬点样本。若只简单地将二者合在一起进行神经网络训练, 对于偏离软点的硬点, 要么会导致过拟合, 要么会成为被忽略的“噪点”。需要一种改进的训练算法, 使其既能对硬点精确拟合, 又能最大限度地利用软点提供的信息。

可以认为实验模型是真实模型的一种变异; 同样也可认为真实模型是实验模型的一种变异, 即: 真实值 $h(x)$ = 实验值 $s(x)$ + 实验模型误差 $\varepsilon(x)$ 。它们的期望值为

$$E(h(x)) = s(x) + E(\varepsilon(x)) \quad (1)$$

我们所需要的是, 模型在综合所有已知信息下的期望 $E(h(x))$ 的神经网络近似 $neth(x)$, 来作为真实模型 $h(x)$ 的代理模型, 此网络是涵盖当前已知信息的最佳网络。

假设通过先验软点样本学习到的神经网络是 $nets(x)$ 。如果网络 $neth(x)$ 和 $nets(x)$ 都是训练良好的, 应有: $neth(x) \approx E(h(x))$, $nets(x) \approx s(x)$ 。代入式(1), 得

$$neth(x) \approx nets(x) + E(\varepsilon(x)) \quad (2)$$

由于实际模型是实验模型的一种变异, 已知实验模型而不知其它信息(误差来源不清楚, 没有硬点样本数据)的情况下, 真实模型的期望就是实验模型, $h(x)$ 是 $s(x)$ 的无偏估计, 即 $E(h(x)) = s(x)$, $E(\varepsilon(x)) = 0$, 对的 $s(x)$ 采样点即是 $E(h(x))$ 的训练样本。当存在硬点样本时, $E(\varepsilon(x))$ 将与硬点样本数据有关。

3 针对单硬点样本的拟合算法

3.1 单硬点样本的模型误差期望值分析

若已知某个硬点 (x_0, y_0) , 以及由一系列软点训练得到的先验网络 $nets(x)$, 假设 $neth(x_0) = y_0$, 则

$$E(\varepsilon(x_0)) \approx y_0 - nets(x_0) \quad (3)$$

对于 x_0 点之外, 考虑两种特殊情形:

(1) 若 $\varepsilon(x)$ 是相互独立, 不具有自相关性的随机值, 即

$$E(\varepsilon(x)) = \begin{cases} 0, & x \neq x_0 \\ y_0 - nets(x_0), & x = x_0 \end{cases}$$

此时误差函数是离散的, 没有“刚性”的。

(2) 若 $\varepsilon(x)$ 是常数值, $E(\varepsilon(x)) = E(\varepsilon(x_0)) = y_0 - nets(x_0)$, $neth(x)$ 是 $nets(x)$ 的沿 y 方向的平移, 此时误差函数是完全“刚性”的。

实际中, 实验模型的误差来源是复杂的, $\varepsilon(x)$ 可能既包含确定性成分, 又包含具有一定自相关性的不确定性成分, 若存在 x_0 使得 $E(\varepsilon(x_0)) \neq 0$, 可能在部分区域 $E(\varepsilon(x))$ 偏离零值, 除了在 x_0 点, $E(\varepsilon(x))$ 仍是未知的。

由于实验样本具有一定的置信度, 不考虑硬点存在的情况下, 根据中心极限定理, 假设 $\varepsilon(x)$ 服从正态分布 $N(0, \sigma_1^2)$, 其概率密度函数为

$$f_1(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{\varepsilon^2}{2\sigma_1^2}} \quad (4)$$

如果不考虑实验样本的置信度限制, 假设过硬点 (x_0, ε_0) 的函数 $\varepsilon(x)$ 在另一点 x_1 的值 $\varepsilon(x_1)$ 的概率分布为 $N(\varepsilon_0, \sigma_2^2)$ 。其概率密度函数为

$$f_2(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_2} e^{-\frac{(\varepsilon-\varepsilon_0)^2}{2\sigma_2^2}} \quad (5)$$

在实验样本的置信度限制下, 有且只有一个硬点 (x_0, ε_0) 时, 根据贝叶斯公式, $\varepsilon(x_1)$ 的期望值为

$$E(\varepsilon(x_1)) = \frac{\int_{-\infty}^{\infty} \varepsilon \cdot f_1(\varepsilon) \cdot f_2(\varepsilon) d\varepsilon}{\int_{-\infty}^{\infty} f_1(\varepsilon) \cdot f_2(\varepsilon) d\varepsilon} \quad (6)$$

由式(4)和式(5), $f_1(\varepsilon) \cdot f_2(\varepsilon) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{\varepsilon^2}{2\sigma_1^2}} \cdot \frac{1}{\sqrt{2\pi}\sigma_2}$

$$\cdot e^{-\frac{(\varepsilon-\varepsilon_0)^2}{2\sigma_2^2}}。$$

对上式进行整理得

$$f_1(\varepsilon) \cdot f_2(\varepsilon) = \frac{1}{2\pi\sigma_1\sigma_2} \cdot e^A \cdot e^B \quad (7)$$

式中 $A = \frac{1}{2\sigma_1^2\sigma_2^2} \cdot \left(\frac{\sigma_1^4\varepsilon_0^2}{\sigma_1^2 + \sigma_2^2} - \sigma_1^2\varepsilon_0^2 \right)$, $B = -\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2}$

$\cdot \left(\varepsilon - \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2} \right)^2$, 由于 A 中不含 ε 项, 故

$$\int_{-\infty}^{\infty} f_1(\varepsilon) \cdot f_2(\varepsilon) d\varepsilon = \frac{1}{2\pi\sigma_1\sigma_2} \cdot e^A \cdot \int_{-\infty}^{\infty} e^B d\varepsilon \quad (8)$$

$$\int_{-\infty}^{\infty} \varepsilon \cdot f_1(\varepsilon) \cdot f_2(\varepsilon) d\varepsilon = \frac{1}{2\pi\sigma_1\sigma_2} \cdot e^A \cdot \int_{-\infty}^{\infty} \varepsilon \cdot e^B d\varepsilon \quad (9)$$

积分时，以 $\varepsilon + \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2}$ 代替 ε ，有

$$\int_{-\infty}^{\infty} e^B d\varepsilon = \int_{-\infty}^{\infty} e^{-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \varepsilon^2} d\varepsilon \quad (10)$$

$$\int_{-\infty}^{\infty} \varepsilon \cdot e^B d\varepsilon = \int_{-\infty}^{\infty} \left(\varepsilon + \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2} \right) \cdot e^{-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \varepsilon^2} d\varepsilon \quad (11)$$

$\varepsilon \cdot e^{-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \varepsilon^2}$ 是关于 ε 的奇函数，故

$$\int_{-\infty}^{\infty} \varepsilon \cdot e^{-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \varepsilon^2} d\varepsilon = 0$$

从而

$$\int_{-\infty}^{\infty} \varepsilon \cdot e^B d\varepsilon = \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2} \cdot \int_{-\infty}^{\infty} e^{-\frac{\sigma_1^2 + \sigma_2^2}{2\sigma_1^2\sigma_2^2} \varepsilon^2} d\varepsilon \quad (12)$$

将式(8)和式(9)代入式(6)，约去相同项，可得

$$\begin{aligned} E(\varepsilon(x_1)) &= \frac{\int_{-\infty}^{\infty} \varepsilon \cdot f_1(\varepsilon) \cdot f_2(\varepsilon) d\varepsilon}{\int_{-\infty}^{\infty} f_1(\varepsilon) \cdot f_2(\varepsilon) d\varepsilon} \\ &= \frac{\int_{-\infty}^{\infty} \varepsilon \cdot e^B d\varepsilon}{\int_{-\infty}^{\infty} e^B d\varepsilon} = \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2} \end{aligned} \quad (13)$$

特别地， $\varepsilon(x)$ 不具有关于 x 的自相关性时， σ_2 为无穷大， $\varepsilon_1 = E(\varepsilon(x_1)) = \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2} = 0$ ； $\varepsilon(x)$ 是确定的常数值时， $\sigma_2 = 0, \varepsilon_1 = \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2} = \varepsilon_0$ 。

在对 $\varepsilon(x)$ 不加限制的情况下， σ_2 可能为零或无穷大；如果限制 $\varepsilon(x)$ 的散布程度，显然 σ_2 和 x_1, x_0 的距离 $d = |x_1 - x_0|$ 有关。本文考虑 $\sigma_2/d \geq c \geq 0$ 的情况，即认为 $\varepsilon(x)$ 具有一定的随机性，其概率分布具有一定的散布程度，不会过度集中，其“刚度”不会过大。 $\sigma_2/d = c$ 时，

$$\varepsilon_1 = \frac{\sigma_1^2\varepsilon_0}{\sigma_1^2 + \sigma_2^2} = \frac{\varepsilon_0}{1 + (c/\sigma_1)^2 d^2} \quad (14)$$

若取 $\varepsilon_1 = 0.1\varepsilon_0$ ，可求得 $d = 3\sigma_1/c$ ，不妨称 $r = 3\sigma_1/c$ 为该硬点的影响范围。则超出该范围后，误差函数期望值受该硬点的影响将衰减到 10% 以下。注意到当 $d=0$ 时，式(14)同样成立。由 $c \leq \sigma_2/d \leq +\infty$ 得

$$0 \leq \frac{\sigma_1^2}{\sigma_1^2 + \sigma_2^2} \leq \frac{1}{1 + (c/\sigma_1)^2 \cdot d^2} = \frac{1}{1 + (3d/r)^2}$$

即 $E(\varepsilon(x))$ 应经过硬点 (x_0, ε_0) 且满足：

$$\left. \begin{aligned} 0 \leq E(\varepsilon(x)) &\leq \frac{1}{1 + \left\| \frac{3(x-x_0)}{r} \right\|^2} \varepsilon_0, \quad \varepsilon_0 > 0 \\ \frac{1}{1 + \left\| \frac{3(x-x_0)}{r} \right\|^2} \varepsilon_0 &\leq E(\varepsilon(x)) \leq 0, \quad \varepsilon_0 < 0 \\ E(\varepsilon(x)) &= 0, \quad \varepsilon_0 = 0 \end{aligned} \right\} \quad (15)$$

当输入变量是 n 维时，在 $\mathbf{X}(x^1, x^2, \dots, x^n)$ 的不同分量 x^k 上，通常需要限制不同的硬点影响范围 r^k ，此时应有 $E(\varepsilon(\mathbf{X}))$ 经过硬点 $(\mathbf{X}_0, \varepsilon_0)$ 且满足：

$$\left. \begin{aligned} 0 \leq E(\varepsilon(\mathbf{X})) &\leq \frac{1}{1 + \left\| 3\text{diag}\left(\frac{1}{r^1}, \frac{1}{r^2}, \dots, \frac{1}{r^n}\right)(\mathbf{X} - \mathbf{X}_0) \right\|^2} \varepsilon_0, \\ &\quad \varepsilon_0 > 0 \\ \frac{1}{1 + \left\| 3\text{diag}\left(\frac{1}{r^1}, \frac{1}{r^2}, \dots, \frac{1}{r^n}\right)(\mathbf{X} - \mathbf{X}_0) \right\|^2} \varepsilon_0 &\leq E(\varepsilon(\mathbf{X})) \leq 0, \\ &\quad \varepsilon_0 < 0 \\ E(\varepsilon(\mathbf{X})) &= 0, \quad \varepsilon_0 = 0 \end{aligned} \right\} \quad (16)$$

3.2 神经网络算法实现

只有一个硬点时，获取综合神经网络的算法步骤如下：

(1) 根据所有软点样本进行标准神经网络训练，获得先验网络 $\text{nets}(\mathbf{X})$ ，并求出模型误差在硬点处的期望值 $E(\varepsilon(x_0)) = y_0 - \text{nets}(\mathbf{X}_0)$ ；

(2) 根据主观经验(自变量的观察范围，对误差来源的主观判断等)，为每个输入方向 x^j 设定硬点影响范围 r^j ，将所有软点数据代入式(16)，对于每个样本点 \mathbf{X}_i ，可求出相应的误差容量区间 $[\text{ed}_i, \text{eu}_i]$ ， ed_i 和 eu_i 分别为根据式(16)求出的 $E(\varepsilon(\mathbf{X}_i))$ 的最小值和最大值；

(3) 以 BP 网络为例，在标准的网络训练过程中，需要不断计算输入/目标对的误差 $e_i = t_i - a_i$ ，来计算雅可比矩阵和性能函数(例如均方误差)，本文将此过程做如下修改：

对于软点，输入/目标对的误差改为如下：

$$e_i = \begin{cases} -(a_i - t_i - \text{ed}_i), & a_i - t_i < \text{ed}_i \\ 0, & \text{ed}_i \leq a_i - t_i \leq \text{eu}_i \\ -(a_i - t_i - \text{eu}_i), & a_i - t_i > \text{eu}_i \end{cases} \quad (17)$$

对于硬点，为使网络对硬点数据有更好的拟合，将硬点同时纳入到训练集、验证集和测试集中，且借鉴信息融合时常用的加权思想^[17]，将硬点的输入/

目标对误差进行放大:

$$e_0 = b(t_0 - a_0) \tag{18}$$

b 为硬点误差敏感系数, 可取一个较大的值;

(4)将步骤(3)计算得到的 e_i 代替标准 BP 训练算法中的 e_i 进行第 2 次网络训练, 得到融合了软硬点信息的综合网络 $neth(\mathbf{X})$ 。

由于神经网络的算法及选择合适的神经元数量限制, 所学习到的网络具有一定的光滑性, 一定程度上限制了 σ_2 的上限, 而 r 值的大小又限制了 σ_2 的下限, 此神经网络算法会在 $nets(\mathbf{X}) + ed$ 和 $nets(\mathbf{X}) + eu$ 之间寻找到一个相对合适的网络模型。下面用一个简单的算例来说明此算法的拟合效果。

3.3 算例 1

此算例以含单个硬点及多个软点的 1 维输入样本为对象, 图 1 描述了样本的分布情况, 并仅用软点训练得到了先验网络 $nets(x)$ (图 1 中的曲线), 图 2 为 $r=2$ 时每个软点的误差容量区间 $[ed, eu]$, 由上述算法训练得到的综合网络如图 3 所示。图中代表综合网络的曲线符合我们的主观判断。更改 $r=1.5$ 后训练得到的综合网络如图 4 所示, 对比图 3 和图 4 可以看出 r 值对最终网络的影响。

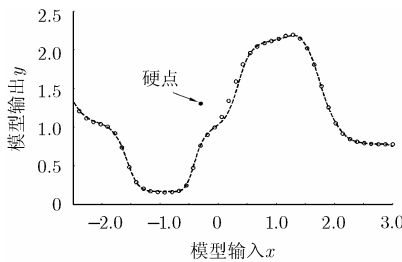


图 1 含 1 个硬点和多个软点的训练样本及先验网络

4 针对多硬点样本的拟合算法

当有多个硬点时, 可以参考式(16)求出软点 x_i 对应于每个硬点 x_j 的误差容量区间 $[ed_{ji}, eu_{ji}]$, 取

$$eu_i = \max_j \{eu_{ji}\} \tag{19}$$

$$ed_i = \min_j \{ed_{ji}\} \tag{20}$$

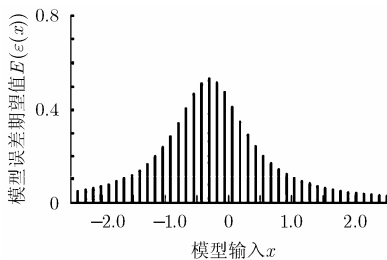


图 2 $r=2$ 时的误差容量区间 $[ed, eu]$

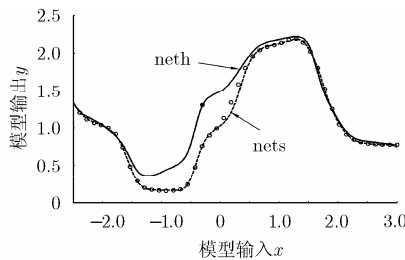


图 3 取 $r=2$ 时由本文算法获得的综合网络 $neth$ 与先验网络 $nets$ 对比

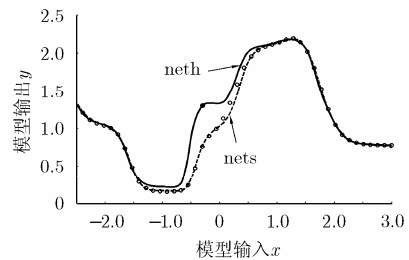


图 4 取 $r=1.5$ 时由本文算法获得的综合网络 $neth$ 与先验网络 $nets$ 对比

同样按照式(17)计算 e_i 。

同理, 对硬点误差进行放大

$$e_j = b(t_j - a_j) \tag{21}$$

式(21)中, b 为硬点误差敏感系数。算法的实现过程与本文 3.2 节相同。以下两个算例将说明此算法对多硬点样本的拟合效果。

4.1 算例 2

此算例以含 4 个硬点及多个软点的 1 维输入样本为对象, 图 5 中描述了样本的分布情况, 并仅用软点训练得到了先验网络 $nets(x)$, 有 3 个硬点偏离了先验网络。图 6 为 $r=2$ 时根据式(19)和式(20)计算得到的每个软点的误差容量区间 $[ed, eu]$, 由本文算法训练得到的综合网络如图 7 所示。

4.2 算例 3

此算例以含 3 个硬点及多个软点的 2 维输入样本为对象, 并仅用软点训练得到了先验网络 $nets$, 有 2 个硬点偏离了先验网络, 如图 8 所示, 图中 x^1, x^2 表示输入的两个维度。图 9 为本文算法训练得到的综合网络, 在保留了先验网络外形的同时对硬点具有更精确的拟合。图 10 中曲面为综合网络与先验网络的差值, 图中圆点和五角星分别为软点误差容量区间的上下限。可见 $neth-nets$ 形成于 eu 与 ed 之间的空间内。

5 结束语

本文针对包含大量低置信级样本(实验样本)和少量高置信级样本(真实样本)的神经网络训练问题, 提出涵盖所有已知信息下模型的期望 $E(h(x))$ 的神经网络近似 $neth(x)$, 是代理真实模型 $h(x)$ 的最佳网络。通过对样本间误差关系的分析, 提出此网络应等于先验网络与实验误差期望之和。引入了参数 r 对实验误差的散布程度进行约束, 通过计算软点误差容量区间 $[ed, eu]$ 和引入硬点误差敏感系数 b , 对标准神经网络训练过程中输入/目标对的误差进行了修改替换, 多种情形下的算例证明, 此方法在精确拟合高置信级样本的同时, 尽可能地利用了低置信级样本的信息。与 KBNN 相比, 该方法更加简

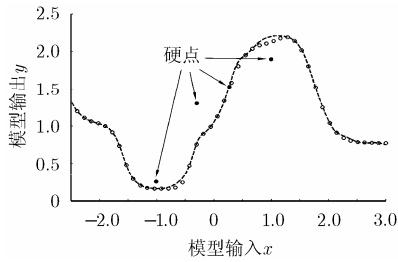


图5 含4个硬点和多个软点的训练样本及先验网络

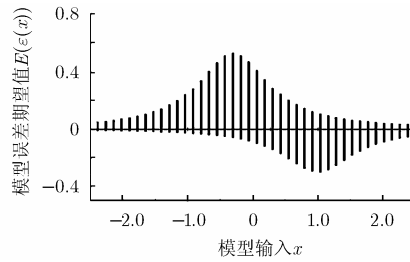


图6 $t=2$ 时的误差容量区间[ed,eu]

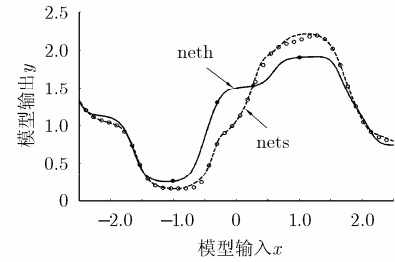


图7 本文算法获得的综合网络 neth 与先验网络 nets 对比

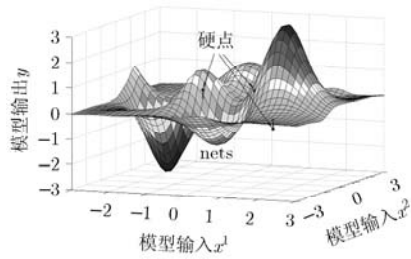


图8 含3个硬点的2维输入样本及先验网络

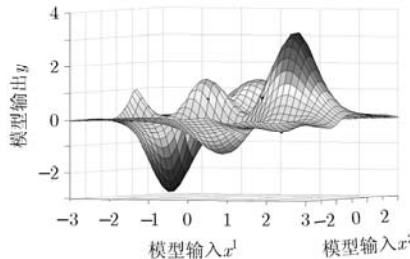


图9 本文算法获得的综合网络

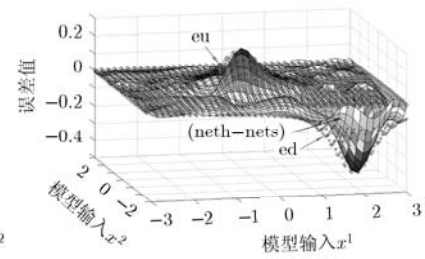


图10 综合网络与先验网络差值 neth-nets 及 eu, ed

单, 可操控性更强并具有更加明确的逻辑意义。本文算法可用于神经网络的离在线结合训练, 或系统变异后的重新学习。本文所做改进仅限于拟合问题, 分类问题需要进一步研究。

参考文献

- [1] Mitchell T M. Machine Learning[M]. New York: McGraw Hill Higher Education, 1997: 17.
- [2] 李露, 李斌, 杜艳玲, 等. 样本量对飞灰含碳量的神经网络预测精度的影响[J]. 电力与科技工程, 2012, 28(9): 38-41.
- [3] Milanic S, Strmcnik S, Sel D, et al. Incorporating prior knowledge into artificial neural networks-an industrial case study[J]. *Neurocomputing*, 2004, 62(1): 131-151.
- [4] Towell G G, Shavlik J W, and Noordewier M O. Refinement of approximate domain theories by knowledge-based neural networks[C]. Proceedings of the Association for the Advancement of Artificial Intelligence, Boston, 1990: 861-866.
- [5] Sahin S, Tolun M R, and Hassanpour R. Hybrid expert systems: a survey of current approaches and applications[J]. *Expert Systems with Applications*, 2012, 39(4): 4609-4617.
- [6] Simsek M, Zhang Q J, Kabir H, et al. The recent developments in knowledge based neural modeling[J]. *Procedia Computer Science*, 2012, 1(1): 1321-1330.
- [7] 姜海川, 苏宏业, 谢磊. 融合过程先验知识的递归神经网络模型及应用[J]. 化工学报, 2013, 64(5): 1665-1673.
- [8] 文刚. 用于步进电机细分控制的知识人工神经网络[J]. 现代电子技术, 2011, 34(7): 190-192.
- [9] Yu Jian-bo, Xi Li-feng, and Zhou Xiao-jun. Intelligent monitoring and diagnosis of manufacturing processes using an integrated approach of KBANN and GA[J]. *Computers in Industry*, 2008, 59(5): 489-501.
- [10] Serpen G, Tekkedil D K, and Orra M. A knowledge-based

artificial neural network classifier for pulmonary embolism diagnosis[J]. *Computers in Biology and Medicine*, 2008, 38(2): 204-220.

- [11] Akkala A, Bhatt D, Devabhaktuni V, et al. Knowledge-based neural network approaches for modeling and estimating radon concentrations[J]. *Environmental Progress & Sustainable Energy*, 2013, 32(2): 355-364.
- [12] 鲜于建川, 隽志才. 基于知识的神经网络在出行方式选择中的应用研究[J]. 计算机应用研究, 2008, 25(9): 2651-2654.
- [13] Gajzler M. The idea of knowledge supplementation and explanation using neural networks to support decisions in construction engineering[J]. *Procedia Engineering*, 2013, 57: 302-309.
- [14] Iqbal R A. Empirical learning aided by weak domain knowledge in the form of feature importance[C]. 2011 International Conference on Multimedia and Signal Processing, Guilin, 2011: 126-130.
- [15] Yoon H, Park C S, Kim J S, et al. Algorithm learning based neural network integrating feature selection and classification[J]. *Expert Systems with Applications*, 2013, 40(1): 231-241.
- [16] Jones J K, and Goel A K. Perceptually grounded self-diagnosis and self-repair of domain knowledge[J]. *Knowledge-Based Systems*, 2012, 27: 281-301.
- [17] 朱永生, 王成栋, 张优云. 基于证据加权调整方法的神经网络及其在故障诊断中的应用[J]. 机械工程学报, 2002, 38(6): 66-71.

高学星: 男, 1987年生, 博士生, 研究方向为参数不确定性与故障诊断。

孙华刚: 男, 1979年生, 高级工程师, 研究方向为机械系统故障诊断。

侯保林: 男, 1965年生, 教授, 博士生导师, 研究方向为火炮自动供输弹系统。