文章编号: 1005-2542(2007) 05-0563-05

考虑问题随机性的神经网络

[1], [1], [2]
(1. 五邑大学, 江门 529020; 2. 佛山科学技术学院, 佛山 528000)

【摘要】 。
。 ,
, 。 , 。
, , ,
, 。 , , ,
。

关键词: ; ;
中图分类号: TU 4 文献标识码:

Neural Network Study on Random Problem

PAN Hua[1], *LI Ben-qiang*[1], *LEI Yuan-xin*[2]
(1. Wuyi Univ., Jiangmen 529020, China; 2. Foshan Univ., Foshan 528000, China)

【Abstract】 The study on artificial neural network dealing with random samples is a challenging problem. For an artificial neural network model, the choice of the error of training and the evaluation of the predicted results is concerned. To our knowledge, the error of training is mostly chosen by trial-and-error. The predicted results were often evaluated by using relative error method, and it is often regarded that the less the relative error is, the better the predicting precision is. ut these methods are no longer fitted for random training samples. Therefore, combined with the design of concrete mi proportion, by taking into account the stochastic characteristic of training samples, a new method for determining training error and evaluating predicted results of a neural network model was studied. Several formulas for random training samples were provided. n the end, a practical e ample was given. perimental results ustify that the method studied in this paper is useful, and its conclusion is close to engineering actualities. t has certain practical significance.

Key words: neural network; prediction; randomness

人工神经网络的研究是从人脑的生理结构出发来研究人的智能行为, 模拟人脑信息处理的功能。它具有许多优良性能, 如允许大规模并行处理、高度的非线性映射能力和泛化能力等, 适合于在特定的输入和输出之间建立复杂的映射关系, 这对于那些有一定的实际经验却又难于建立有效数学力学模型的工程领域具有重要的现实意义, 这方面的研究工作非常丰富, 如: Shi [1] 用累积分布函数法处理神经网络的样本数据, 与线性变换方法相比, 大大降低了网络模型预测的错误率。文献 [2] 中提出了多变量神经自适应控制器通用结构。文献 [3, 4] 中分别给出了网络结构每层隐节点数的起始值和上限值的计

收稿日期: 2006-11-14 修订日期: 2007-03-28
作者简介: 潘 华(1963-), 女, 硕士。研究方向为土木工程及系统工程在土木工程中的应用。lb ph 126. com

混凝土梁 神经网络学习和预测结果的

本数与网

文献[7]中

神经网络

神经网

法 文

本 资

经 络

中

文

混凝土

标参数

配合比

的 网

完成 对

文献

模型。文

质量等级

土 性能

通 和城

及 应用领

多 部分是

关系, 且都

上 随机不确定

土木工程中的

配

神

的

方法

但

般是

性。

时间

练

测结果

为训

实测

后,

混凝土性能指标的随机 析

是土木工程中的重要建筑材料, 它

混凝土的方向发展。由于组成高性能

复杂性, 生产中常常受到各种随机因 的

凝土性能指标(如强度等) 往往表现出一定

性。对于同一配合比的同一盘混凝土 不同

测得的性能指标值各不相同, 也就 训练

随机的, 其分布一般服从正态分布

高性能混凝土配合比设计问题的

凝土的各组份(如水泥用量、砂 量 用量

用量等) 与混凝土性能指标 之间 射关

上分析知, 组份和性能指标之 随机 系

于性能指标为随机变量, 因 选择 计特

为某性能指标的代表值 确 网络

输入为混凝土各组份的 , 输出 混凝

标的平均值。

设 Y_i 为混凝 某一性 指标

落度等), 则 Y_i (μ_i, σ_i^2), 其中

值, σ_i^2 为方 是

$$\frac{1}{n} y_{i1}$$

$$Y_i \quad N\left\{\mu_i, \sigma_i^2\right.$$

给定置信概率

$$P \quad \alpha/2$$

$$P\left\{\left|\frac{Y_i-\mu_i}{\sigma_i/\sqrt{n}}\right| \quad \alpha/2\right\} = 1-\alpha$$

得平均值 μ_i 的置信区间为

(2)

在工程

凝土取 3 种试件

值, 即上面的 n 3。

由以上分析可以看

计算必须要给出一个确定值,

到的性能指标值的

μ_i

结果的评价等问题，都将与输入输出变量为确定性关系时有所不同。

2 网络训练误差的选取

众所周知， 学习算法中的关键是 在学习过程中以尽可能快地减小训练误差的方式进行，通常是梯度方向。但减小到什么程度学习结束呢 即如何选取训练误差控制值，这是在实际应用中经常碰到的问题，一般是经过反复试算调整确定，带有一定的盲目性。一般地，神经网络模型的输入输出变量按其性质的不同可分为确定性变量和非确定性变量（本文涉及的是随机变量）。对前者相对容易些，可以结合实际问题对误差的要求计算确定 而对后者为非确定性变量就更为复杂。下面结合混凝土性能指标的随机特性，研究神经网络模型的输出变量为随机变量时确定训练误差控制值的方法。

由第 1 节中分析知，作为训练样本的 Y_i 不一定是混凝土性能指标 μ_i 的真实反映，训练误差的大小反映的是网络预测计算值 μ_i 与实测值 Y_i（而非 μ_i）之间的差值大小，这样训练误差 E 的大小并不一定能反映训练结果的好坏。但通过式(2)可以得到 μ_i 的置信区间，因此，若通过训练使网络预测计算值落在该置信区间内，就可以认为此时的样本学习反映了 μ_i 的情况，学习就可以结束。因此，设某一性能指标 μ_i 的置信区间的 1 2 为 Δ_i，

$$\mu_i - Y_i \leqslant \Delta_i \tag{3}$$

则全网络平均训练误差

$$E = \frac{1}{2}\sum_{i=1}(\mu_i - Y_i)^2 \leqslant \frac{1}{2}\sum_{i=1}\Delta_i^2$$

即

$$E \leqslant \frac{1}{2n}\ {}^2_{\alpha 2}\sum_{i=1}\sigma_i^2 \tag{4}$$

式中 n 为计算 Y_i 时所采用试件的个数 为网络输出变量总数。

由于激励函数 Sigmoid 的函数值输出范围在接近 0, 1 时，曲线较平缓，变化速度非常慢，为了减少网络的学习时间，将输入输出数据归一化变换为 0 1, 0 9 之间[6]，这样 Sigmoid 函数在该区间内变化梯度比较大，网络收敛时间大大缩短，改善了网络性能，其变换方法为

$$= 0\ 1\quad \frac{0\ 8(\ -\ _{\min})}{_{\mathrm{ma}} - \ _{\min}},\qquad \in\ 0\ 1, 0\ 9 \tag{5}$$

式中 为样本原始值 $_{\mathrm{ma}}$、$_{\min}$是样本原始值中的最大、最小值 为归一化后的值。

由式(5)可得 Y_i 经归一化处理后的均方差为

$$\sigma_i = \frac{0\ 8}{Y_{i,\mathrm{ma}} - Y_{i,\min}}\sigma_i \tag{6}$$

将式(6)代入式(4)，得

$$E \leqslant \frac{0\ 32}{n}\ {}^2_{\alpha 2}\sum_{i=1}\left(\frac{\sigma_i}{Y_{i,\mathrm{ma}} - Y_{i,\min}}\right)^2 \tag{7}$$

因此，全网络平均初始训练误差取为

$$\varepsilon = \frac{0\ 32}{n}\ {}^2_{\alpha 2}\sum_{i=1}\left(\frac{\sigma_i}{Y_{i,\mathrm{ma}} - Y_{i,\min}}\right)^2 \tag{8}$$

3 预测结果的评价方法

对模型预测的结果，一般是用平均相对误差值（如 5 ）来评价[12, 19]。由于混凝土的性能指标具有随机性，即用来训练的样本本身是随机变量，显然此时相对误差值的大小不能反映预测结果的好坏。因此，下面将针对混凝土性能指标的随机性这一特点，研究合适的评价方法。

由第 1 节的分析知，作为训练样本的 Y_i 不一定是混凝土性能指标 μ_i 的真实反映，因此，预测值与实测的 Y_i 的接近程度不一定能反映预测的优劣。实际上，对预测结果的评价面临着两方面的问题 ①首先要判断出每次预测的正确性 ②每次预测正确的概率如何。现分析如下

由式(2)可得平均数 μ_i 的置信区间，即

$$\left(Y_i - \ {}_{\alpha 2}\frac{\sigma_i}{\sqrt{n}},\ Y_i\ \ {}_{\alpha 2}\frac{\sigma_i}{\sqrt{n}}\right)$$

因此可以认为，对于神经网络的每次预测，若神经网络预测的结果落在该置信区间内（或预测值 μ_i 与 Y_i 之差的绝对值 $\leqslant\ {}_{\alpha 2}\frac{\sigma_i}{\sqrt{n}}$），则接受该预测结果，预测正确 否则，拒绝它，预测失败。其可信度为 $1-\alpha$（如 α 取 0 05，则可信度为 95 ）。

因为一次预测正确与否并不能正确反映模型的预测性能，必需考查多次预测的情况。因此，定义如下预测正确率

$$\text{正确率}(\ ) = \frac{\text{测试正确的次数}(n)}{\text{所有的测试次数}(N)} \tag{9}$$

显然正确率的大小能够反映模型的预测性能。然而由于混凝土性能指标的随机性，在有限的测试次数 N 的条件下，该正确率的可信度如何呢 对此，可以先通过式(9)计算预测正确概率的估计值，然后再分析出预测正确的概率的置信区间。

由于每次的预测，要么正确，要么失败。因此，预测的结果可以用(0－1)分布的随机变量来描述。设随机变量 表示模型一次预测的结果，每次预测

正确的概率为 , 则 的分布为

$$P(\quad = \quad) = \quad (1- \quad)^{1-} , \quad = 0, 1 \quad (10)$$

由于 未知, 为估计 , 选取充分大容量的检测样本 $x_1, x_2, \cdots, x_N$, 这样由中心极限定理

$$= \frac{x_1 \quad x_2 \quad \cdots \quad x_N - N}{\sqrt{N} \sqrt{\quad (1- \quad)}} =$$

$$\frac{Nx - N}{\sqrt{N \quad (1- \quad)}}$$

5 结 语

本文在分析了 … 的随
特性的基础上，应 … 了网络
络误差初始值的设 …
计价方法，从而 …
高了网络训练效 …
自己确定一 …
结果正 … 结果
性。应用 … 价
的工程 …
的依据 …
值得注 … ，对 …
差，会造成 … 训练

参考文献：

… pr…
… neural netw…
… engineering, …
[2] … 结合模糊 …
… 与 … 94, … 22
[3] … 神经 … 及其 … 中的应用
… 出版社 …
[4] … 的线性 … 络
… 计算机学报 …
[5] … maximum heat … a
… for the algo… 0, 1)
金属 … 径 …
4. … 敷 …