

基于多层感知器网络的农作物疾病诊断系统

陈晓艳¹, 董朝轶², 李永亭¹, 刘月文¹

(1. 内蒙古工业大学 电力学院 电工基础教学中心, 内蒙古自治区 呼和浩特 010080;

2. 内蒙古工业大学 电力学院 自动化系, 内蒙古自治区 呼和浩特 010080)

摘 要: 农作物疾病的人工诊断效果常受到个人诊断经验和能力的限制, 无法达到最令人满意的诊断结果。将丰富的植物病理学诊断经验和知识编入专家系统, 利用模式识别算法, 对农作物常见疾病进行诊断, 可以大大提高诊断准确率, 有效地提高其产量和质量。主要研究了基于一种人工神经网络—多层感知器网络的模式识别技术在大豆疾病诊断中的应用。MLP神经网络通过模拟生物神经元细胞对外部刺激而产生的反应, 构成一种前向神经网络, 可以有效地解决非线性不可分问题。首先对大豆常见19种疾病症状进行了收集和整理, 构建试验样本集。然后, 利用反向传播算法对该网络进行训练和测试。测试结果表明, 该模型具有较高的农作物疾病诊断正确率和良好的泛化能力。

关键词: 反向传播算法; 多层感知器网络; 疾病诊断; 模式识别

中图分类号: TP183; S435.2

文献标识码: A

文章编号: 1673-629X(2011)11-0084-05

Study of Diagnosis for Diseases of Agricultural Crops Based on Multi-Layer Perceptron Networks

CHEN Xiao-yan¹, DONG Chao-yi², LI Yong-ting¹, LIU Yue-wen¹

(1. Center of Electrical and Electronic Teaching, College of Electric Power, Inner Mongolia University of

Technology, Huhhot 010080, China;

2. Department of Automatic Control, College of Electric Power, Inner Mongolia University of

Technology, Huhhot 010080, China)

Abstract: The manual diagnosis for the diseases of agricultural crops is often restricted by the individual ability and experiences so that one cannot obtain the precise results of diagnosis. To overcome this pitfall, the merge of expert systems with the rich pathological knowledge and the utilization of pattern recognition algorithm can significantly improve the precision of diagnosis. Therefore, it greatly increases the quantity and the quality of crop's production. A pattern recognition technique based on a multi-layer perceptron (MLP) neural network is applied to the diagnosis of soybean diseases. The MLP neural network, which is a novel and efficient feed-forward network, is based on the reflections of cortical neurons on the external stimulus and can be used to solve the problems of nonlinear inseparability. Firstly, nineteen typical symptoms of soybean diseases are collected, and then constructed to form an experimental sample set. Secondly, set up and train the MLP network model using back propagation algorithm. Finally, the test shows that the RBF model has high diagnosis precision and strong generalization ability.

Key words: back propagation algorithm; multi-layer perceptron neural network; disease diagnosis; pattern recognition.

0 引言

农作物病虫害诊断专家系统是一种计算机算法系统, 它依据存储于计算机中的专家知识和经验, 结合用户提供的信息、数据和事实, 模仿专家的思维方式采用计算机算法进行推理和判断, 对农作物疾病给予正确

的诊断, 从而达到有效治疗的目的。目前, 国内外农作物病虫害诊断专家系统已有极大的发展, 其范围涵盖了多种作物病虫害的防治、诊断与发生预测等。如, “农药处方与病虫害综合治理计算机咨询专家系统”及“蔬菜病虫害综合治理咨询系统”^[1,2]。这些系统不仅极大地丰富了农业作物病虫害诊断专家系统的研究内容, 推动了地区农业生产, 也为其他的开发研究提供了借鉴。

虽然已有的诊断系统在一定程度上可以对某些作物疾病做出有效诊断, 但是仍存在着诊断精度不高、泛化能力弱、容错能力差等缺点。文中采用一种人工神

收稿日期: 2011-03-23; 修回日期: 2011-07-08

基金项目: 2010年度内蒙古自治区教育自然科学基金重点项目(NJ10070); 内蒙古工业大学校基金(X200416; X200805)

作者简介: 陈晓艳(1975-), 女, 硕士, 讲师, 研究方向为模式识别、智能控制。

经元网络模式识别技术,即,利用多层感知器(Multi-layer Perceptron, MLP)网络来构造农作物疾病诊断系统^[3-5]。通过对大豆常见 19 种疾病症状的收集和整理,构造试验样本集,利用反向传播(Back Propagation, BP)算法^[6,7]对该网络进行训练和仿真,并对训练过程和诊断效果进行了分析。计算表明,该人工神经网络构造的大豆疾病诊断系统能够给出较为准确的诊断结果,并有可能作为一般方法进一步推广至更多农作物的疾病诊断,为农作物的增产增收做出贡献。

1 方法

1.1 多层感知器网络的基本介绍

多层感知器网络属于多层前向神经网络,能解决非线性不可分问题,由输入层、隐含层和输出层组成,其中隐含层可以为一层或多层。以含有两个隐层的 MLP 网络为例,其拓扑结构如图 1 所示。

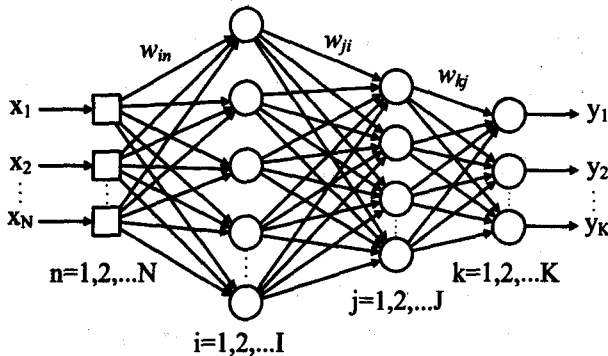


图 1 多层感知器模型(含有两个隐层 I 和 J)

输入层变量的个数相应于“样本”所具有的特征属性的维数,输出层变量的个数为预先规定的模式类别数,隐含层的个数以及隐节点的个数可由实际问题决定,亦可通过对训练后网络的诊断 MSE 残差寻优获得。MLP 网络一般具有可微的非线性激励函数,可将低维的非线性不可分问题映射到高维空间转换为非线性可分问题,最终通过输出层的网络进行正确分类。多层感知器的训练过程通过对输出层训练误差的反向传播实现,被称为 BP (Back Propagation) 算法,所以多层感知器也被称为 BP 神经网络。由于多层感知器所具有强大的计算和分类能力,已成为应用于模式分类领域最为广泛的一种人工神经网络^[8,9]。

1.2 BP 算法的推导

以图 1 所示网络为例具体推导 BP 算法,得到的结论可自然推广至其他 MLP 网络。设输入层变量数为 \$N\$,下标 \$n=1,2,\dots,N\$ 表示任意输入变量;隐层 1 具有 \$I\$ 个神经元,下标 \$i=1,2,\dots,I\$ 表示其中任意变量;隐层 2 包含 \$J=1,2,\dots,J\$ 个神经元,其中的任一神经元用下标 \$j\$ 表示,输出层对应 \$K\$ 个分类数量,下标 \$k=1,2,\dots,K\$ 表示其中任一类别输出,输出最大值者被判为模

式归属于该类。各层间突触连接的强度分别用 \$w_{in}\$、\$w_{ji}\$、\$w_{kj}\$ 表示(第一下标表示发出信号的神经元,第二下标表示接收信号的神经元)。

网络中所有神经元采用相同的 Log-Sigmoid 激励函数:

$$f(x) = \frac{1}{1 + \exp(-x)}, -\infty < x < \infty \quad (1)$$

使用上述的非线性激励函数形式的主要原因为,该函数其对 \$x\$ 的一阶导数可化为关于 \$f(x)\$ 的显式形式:

$$f'(x) = f(x)[1 - f(x)] \quad (2)$$

这使得 MLP 网络具有较为简化的训练权值更新表达式。

给定一个输入输出样本集 \$\{(X_p, Y_p)\}\$, \$p=1,2,\dots,P\$, 其中,对应任一训练样本 \$X_p = [x_{1p}, x_{2p}, \dots, x_{Np}]^T\$, 网络实际输出为 \$Y_p = [y_{1p}, y_{2p}, \dots, y_{Kp}]^T\$, 网络期望输出可表示为 \$d_p = [d_{1p}, d_{2p}, \dots, d_{Kp}]^T\$。这样,输入信号的正向传播过程可描述为:

$$u_i^1 = \sum_{n=1}^N w_{in} x_{np}, v_i^1 = f(u_i^1) = f\left(\sum_{n=1}^N w_{in} x_{np}\right), i=1,2,\dots,I$$

$$u_j^2 = \sum_{i=1}^I w_{ji} v_i^1, v_j^2 = f(u_j^2) = f\left(\sum_{i=1}^I w_{ji} v_i^1\right), j=1,2,\dots,J$$

$$u_k^3 = \sum_{j=1}^J w_{kj} v_j^2, v_k^3 = f(u_k^3) = f\left(\sum_{j=1}^J w_{kj} v_j^2\right), k=1,2,\dots,K$$

$$y_{kp} = v_k^3 = f(u_k^3) = f\left(\sum_{j=1}^J w_{kj} v_j^2\right), k=1,2,\dots,K$$

这里,使用符号 \$u\$ 表示神经元输入, \$v\$ 表示激励输出, \$u, v\$ 的上标标识当前层,下标指该层中的神经元。显然,迭代第 \$m\$ 次后,输出层第 \$k\$ 个神经元对于第 \$p\$ 个样本 \$(X_p, Y_p)\$ 的误差信号为 \$e_{kp}(m) = d_{kp}(m) - y_{kp}(m)\$。

对于全部 \$K\$ 个输出层的神经元误差总和为:

$$E_p(m) = \frac{1}{2} \sum_{k=1}^K e_{kp}^2(m) \quad (3)$$

考虑到全部 \$P\$ 个样本,则 \$m\$ 次训练后残差表示为 \$E(m) = \sum_{p=1}^P E_p(m)\$。网络训练的过程为发现合适的突触连接权值 \$w_{in}\$、\$w_{ji}\$、\$w_{kj}\$,使得 \$E_p(m)\$ 取最小。权值的更新过程如下所示:

$$w_{kj}(m+1) = w_{kj}(m) + \Delta w_{kj}(m), j=1,2,\dots,J; k=1,2,\dots,K$$

$$w_{ji}(m+1) = w_{ji}(m) + \Delta w_{ji}(m), i=1,2,\dots,I; j=1,2,\dots,J$$

$$w_{in}(m+1) = w_{in}(m) + \Delta w_{in}(m), n=1,2,\dots,N;$$

$i = 1, 2, \dots, I$ (4)

BP 学习算法可进一步描述为^[10]:

(1) 初始化网络参数, 将各突触连接权值 w_{in} 、 w_{ji} 、 w_{kj} 随机赋给非零初始值;

(2) 随机选择一个输入样本 X_p , 同时将迭代次数置零 ($m = 0$);

(3) 对该输入样本顺序逐层计算该网络中每层神经元的输入信号 u 和输出信号 v , 网络输出为 $y_{kp}(m) = v_k^k(m)$, $k = 1, 2, \dots, K$;

(4) 计算 $E_p(m)$, 判断该误差和是否满足预定的性能指标, 如果满足, 跳转至第 6 步 (输入下一训练样本); 若不满足, 则跳转至第 7 步 (需要进行下一轮 P 个样本的输入训练);

(5) 判断 $m + 1$ 是否大于规定的最大迭代次数, 是, 转至第 7 步, 否, 则对目前输入样本 X_p , 反向计算每层神经元的局部梯度值:

$$\delta_k^k(m) = y_{kp}(m)(1 - y_{kp}(m))(d_{kp}(m) - y_{kp}(m)),$$

$$k = 1, 2, \dots, K$$

$$\delta_j^j(m) = u_j^j(m)(1 - u_j^j(m)) \sum_{k=1}^K \delta_k^k(m) w_{kj}(m),$$

$$j = 1, 2, \dots, J$$

$$\delta_i^i(m) = u_i^i(m)(1 - u_i^i(m)) \sum_{j=1}^J \delta_j^j(m) w_{ji}(m),$$

$$i = 1, 2, \dots, I$$
 (5)

利用获得的局部梯度, 计算权值修正量 Δw :

$$\Delta w_{kj}(m) = \eta \delta_k^k(m) v_j^j(m)$$

$$\Delta w_{ji}(m) = \eta \delta_j^j(m) v_i^i(m)$$

$$\Delta w_{in}(m) = \eta \delta_i^i(m) x_{np}(m)$$
 (6)

η 为网络学习速率, 常为一大于零的较小正数, 选取过大则网络容易陷入振荡, 过小则训练过程过长, 不容易收敛。修正权值后, 令 $m = m + 1$, 返回第 3 步进行下一步权值修正;

(6) 对输入输出数据对 (X_p, Y_p) 训练结束后, 令 $p = p + 1$, 判断 p 是否等于 $P + 1$ 。是则调转至第 7 步, 否则返回第 2 步;

(7) 判断当前网络权值对于所有输入输出数据对 (X_p, Y_p) , $p = 1, 2, \dots, P$ 是否全部满足性能指标, 是则训练终结。否则置 $p = 1$ 返回第 1 步开始新一轮训练。

当 MLP 隐层和输出层神经元结构中包含常数偏置量 b_i 、 b_j 、 b_k 时, 可将各层的输入向量扩展一维, 新增元素为常数“1”; 同时, W_{in} 、 W_{IJ} 、 W_{KJ} 各增广一列, 按行依次添入 b_i 、 b_j 、 b_k , $i = 1, 2, \dots, I$, $j = 1, 2, \dots, J$, $k = 1, 2, \dots, K$ 。训练方法仍与步骤(1)~(7)相同。

2 仿真结果

农作物疾病诊断问题可以看作是由输入症状到输

出病症之间的非线性映射问题。采用 MLP 神经网络结构, 输入神经元对应于作物的各种属性, 输出层对应于疾病的类型。文中以大豆疾病诊断专家系统为例说明诊断的设计过程。大豆疾病诊断过程设计流程如图 2 所示。

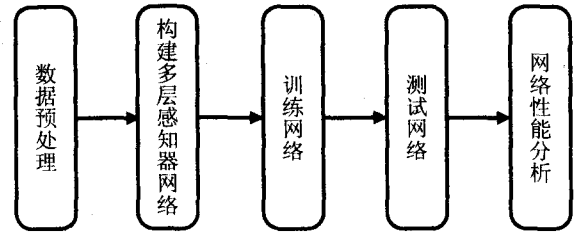


图 2 大豆疾病诊断过程设计流程

2.1 数据预处理

样本数据共包含 683 个获病大豆植株的 35 种特征属性 (包含降雨量、温度、植株生长季节、生长地海拔高度等气候地理条件; 出芽率、叶斑块大小、生长环境、茎部和根部形态等外观特征; 及有无菌丝体、麦角菌硬粒等病理条件等)^[11]。每种属性分别具有 2~7 种不同的取值, 可进一步量化为 -1 至 6 之间的整数形式。

这些大豆植株样本分别属于 19 种常见疾病, 包括:

1. diaporthe-stem-canker (间座壳茎腐病): 10
2. charcoal-rot (炭腐病): 10
3. rhizoctonia-root-rot (立枯丝核菌根腐病): 10
4. phytophthora-rot (疫霉根腐病): 40
5. brown-stem-rot (茎褐腐病): 20
6. powdery-mildew (白粉病): 10
7. downy-mildew (露菌病): 10
8. brown-spot (褐斑病): 40
9. bacterial-blight (细菌性斑点病): 10
10. bacterial-pustule (页烧病): 10
11. purple-seed-stain (紫斑病): 10
12. anthracnose (炭疽病): 20
13. phyllosticta-leaf-spot (叶点霉叶斑病): 10
14. alternaria-leaf-spot (交链孢霉叶斑病): 40
15. frog-eye-leaf-spot (灰斑病): 40
16. diaporthe-pod-&-stem-blight (间座壳荚枯病): 6
17. cyst-nematode (胞囊线虫病): 6
18. 2-4-d-injury (2,4-二氯苯氧乙酸损害): 1
19. herbicide-injury (除草剂药害): 4

以上疾病种类用 1~19 之间的整数进行量化表示。

样本的选取应注意以下两点原则:

- (1) 每类别选择的样本数量应可能相近, 以避免

网络训练陷入局部最优,保证训练后网络对各类别的样本具有程度相同的响应敏感度;

(2)在同一种类别内部,所选择的样本应尽量具有多样化的特征属性,以确保训练后的网络具有较高识别率。

样本的属性特征和类别量化完成以后,可以建立起对应的输入输出数据矩阵。输入数据矩阵为 (35×683) 阶,输出数据为 (19×683) 阶。

MLP 网络需要经过训练、测试等步骤才能形成具有模式识别能力的分类器,所以有必要对数据集进行划分,将原始数据矩阵进行随机地抽取,抽取出 3/4 的数据合并成集作为训练样本集(其中 1/3 为验证样本集,以防止网络过度训练),剩余 1/4 作为测试样本集。

首先利用数据训练集(512 个样本)对径向基网络进行训练。由于输入数据属性维数较大($N = 35$),需利用主分量分析法(Principal Component Analysis)来降低输入维数^[12]。主分量分析通过对输入数据施加一定的线性变换,将原有的高维输入数据投影为较低维的数据空间,并保留了原输入数据的主要信息,更方便于处理数据信息。主分量分析也可以看作是一种特征提取的过程,其主要目的是在大的输入空间中寻找合适的特征向量,并在所有的特征中提取主要特征,特征选择是指从数据空间变换到特征空间的过程。文中对输入数据特征提取后新构造的矩阵输入维数被压缩到 24,输入数据转换为 24×683 的新矩阵。

2.2 构建多层感知器网络及训练

事实上,研究已经表明,两层 MLP 网络在其隐层中使用 Log-Sigmoid 传输函数,在输出层使用线性传输函数,只要隐层中有足够的单元可用,就几乎可以以任意精度逼近任何 Borel 可积函数^[13]。这里使用单隐层加线性输出层的结构。这样权值更新矩阵简化为(注, J 层已经被消除):

$$w_{ki}(m+1) = w_{ki}(m) + \Delta w_{ki}(m), i = 1, 2, \dots, I; \\ k = 1, 2, \dots, K$$

$$w_{in}(m+1) = w_{in}(m) + \Delta w_{in}(m), n = 1, 2, \dots, N; \\ i = 1, 2, \dots, I;$$

$$\delta_k^K(m) = (d_{kp}(m) - y_{kp}(m)), k = 1, 2, \dots, K$$

$$\delta_i^I(m) = u_i^I(m) (1 - u_i^I(m)) \sum_{k=1}^K \delta_k^K(m) w_{ki}(m),$$

$$i = 1, 2, \dots, I$$

按下式计算权值修正量 Δw :

$$\Delta w_{ki}(m) = \eta \delta_k^K(m) v_i^I(m)$$

$$\Delta w_{in}(m) = \eta \delta_i^I(m) x_{np}(m)$$

$\eta > 0$ 为学习速率。

使其输入神经元个数等于输入向量维数,即($N = 24$)。隐层节点数 I 应选择为使验证数据集最小二乘

(MSE)残差不再减小时的值。 I 采用 1:1:50 共 50 个值分别进行构造,并分别进行网络训练,最后对于验证样本进行识别的最小二乘(MSE)残差进行计算,其结果如图 3 所示。对应最小残差 0.1468 的最优隐层节点数为 $I = 43$ 。所以选择 $I = 43$ 作为该 MLP 网路函数的隐层节点数。

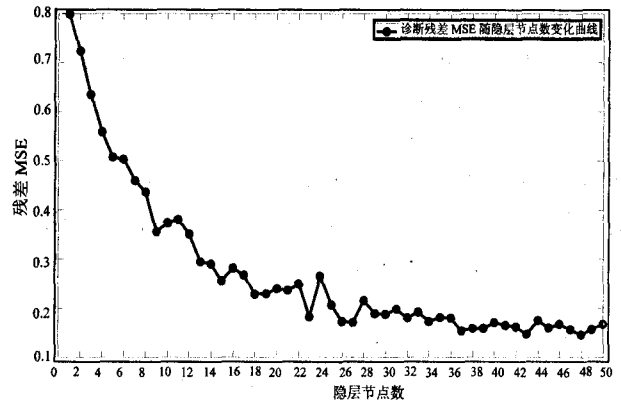


图 3 诊断 MSE 残差随隐层网络规模变化曲线

测试过程是用来检测训练完的网络的性能,将测试样本输入网络,利用上述构造的最优网络进行运算,测试样本的网络输出与实际输出之间的拟合程度如图 4 所示。

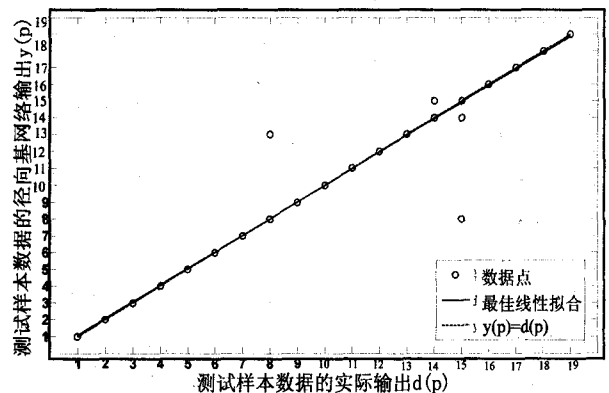


图 4 测试样本多层感知器网络输出与实际输出之间的拟合程度,线性相关系数 $R=0.9897$

测试样本的疾病类型用整数 1~19 表示,图中横坐标为每个测试样本的实际取值,纵坐标为 MLP 网络输出。理想情况为网络输出与实际取值相同,即表示诊断结果完全正确,图中所有数据点位于 45 度斜线上。实际诊断效果可以达到较高准确率,经过对 171 个测试样本进行测试,统计出诊断结果正确的测试样本为 162 个,错误的测试样本为 9 个,得出径向基网络识别正确率为 94.74%。

3 结束语

文中主要研究了基于一种人工神经网络(Artificial Neural Network)—多层感知器函数(Multi-layer Perceptron, MLP)网络的模式识别技术在大豆疾病诊

断中的应用。MLP 网络能够逼近任意的非线性映射,可以处理系统内在的难以解析的规律性,具有极快的学习收敛速度。MLP 网络对于相同的输入一定得出相同的输出;对于相近的输入则可得到相近的输出。这一特性可用于推理,专家知识作为训练样本,经过学习后被存储在网络中推理过程对输入模式做前向计算仅一次通过网络便可得出推理结论。这一特性使 MLP 网络具有抑制噪声的能力,可用于不确定性推理,并具有较强的网络泛化能力。

经过网络的训练以及测试,仿真结果显示所构建的 MLP 疾病诊断正确率较高,网络性能达到预定指标,并有可能进一步推广到其他农作物疾病的诊断过程,产生良好的经济和社会效益。

参考文献:

- [1] 唐乐尘,黄明真. 农药处方与病虫害综合治理计算机咨询专家系统[J]. 计算机与农业,1996(4):6-14.
- [2] 唐乐尘,郭 瑛. 蔬菜病虫害综合治理咨询系统[J]. 计算机与农业,1998(2):11-14.
- [3] 王俊杰,陈景武. BP 神经网络在疾病预测中的应用[J]. 数理医学杂志,2008,21(3):259-262.
- [4] 查青林,何羿婷. 基于神经网络分析方法探索类风湿关节炎证病信息对疗效的预测作用[J]. 中西医结合学报,2007,5(1):404-407.
- [5] 许林勇,白 毅. 基于人工神经网络模型的医院感染危险度预测[J]. 中南大学学报:医学版,2006,31(3):404-

(上接第 83 页)

的数学基础,又有直观的图形表示,既表示了每一个需求的功能,又表示了父子需求的耦合接口,并且将结构化程序设计的思想充分融入其中,自顶向下的方法能清晰地把握所分析问题的总体结构和层次结构,自底向上的方法则更加清楚地表示了子需求的精化描述,有利于系统分析人员和用户的理解、交流,为系统的可靠设计、开发提供了精确的形式化模型,减少需求表达的二义性,该需求模型在实际系统分析中得到了具体的应用,分析效果较好。

参考文献:

- [1] Wieggers K. Software Requirements[M]. 2nd ed. [s. l.]: Microsoft Press,2003.
- [2] 骆 斌,丁二玉. 需求工程—软件建模与分析[M]. 北京:高等教育出版社,2009.
- [3] Fu Chengcheng, Sun Youchao, Lu zhong. Hierarchical colored timed Petri nets for maintenance process modeling of civil aircraft[J]. International Journal of Plant Engineering and Management,2008,13(3):160-162.
- [4] Jensen K. An introduction to the theoretical aspect of colored

407.

- [6] Leung H, Haykin S. The complex back propagation algorithm[J]. IEEE Transactions on Signal Processing,2002,39(9):2101-2104.
- [7] Rumelhart D E, Hinton G E, Williams R J. Learning representations by back-propagating errors[J]. Nature,1986,323:533-536.
- [8] Rossi F, Conan-Guez B. Functional multi-layer perceptron: a non-linear tool for functional data analysis[J]. Neural Networks,2005,18:45-60.
- [9] Sultan A F, Swift G W, Fedirchuk D J. Detection of high impedance arcing faults using a multi-layer perceptron[J]. IEEE Transactions on Power Delivery,1992(7):1871-1877.
- [10] Leonard J, Kramer M A. Improvement of the back propagation algorithm for training neural networks[J]. Computers & Chemical Engineering,1990,14:337-341.
- [11] Michalski R S, Chilausky R L. Learning by being told and learning from examples: An experimental comparison of the two methods of knowledge acquisition in the context of developing an expert system for soybean disease diagnosis[J]. International Journal of Policy Analysis and Information Systems,1980(4):125-161.
- [12] Wood F, Esbensen K, Geladi P. Principal component analysis[J]. Chemometr. Intel. Lab. Syst,1987(2):37-52.
- [13] Maxwell K H, White H. Multilayer feed forward networks are universal approximators[J]. Neural Networks,1989(2):359-366.

Petri nets [C]//Lecture Notes in Computer Science. New York: Springer-Verlag,1994:245-248.

- [5] Jensen K. An introduction to the practical use of colored Petri nets [C]//Lectures on Petri nets II. LNCS 1492. Berlin: Heidelberg; New York: Springer-Verlag,1998:247-254.
- [6] 张国生. 分布式计算中的超接口体系结构模型[J]. 云南大学学报(自然科学版),2005,27(2):120-121.
- [7] 李景霞,闫春钢. 一种基于扩展颜色 Petri 网的 Web 服务组合验证机制[J]. 计算机科学,2009,36(10):146-147.
- [8] 李景霞,程久军. Web 服务组合的层次颜色 Petri 网描述模型[J]. 计算机工程,2009,35(24):39-40.
- [9] Bohm C, Jacopini G. Flow Diagram, Turning Machine and Language with Two Formation Rules[J]. CACM,1966,9(3):179-191.
- [10] Li Tong. An Approach to Modeling Software Evolution Processes[M]. [s. l.]: Springer,2008.
- [11] 李 彤,王黎霞,柳 青,等. 形式化软件功能分解的交互式规则[J]. 计算机工程与应用,2000(7):44-45.
- [12] 郝克刚,丁剑洁. 层次结构的 Petri 网[J]. 计算机科学与探索,2008,2(2):125-126.