

# 一种基于BP神经网络的属性重要性计算方法

潘庆先<sup>1,2</sup>,董红斌<sup>1</sup>,韩启龙<sup>1</sup>,王莹洁<sup>2</sup>,丁蕊<sup>1</sup>

(1. 哈尔滨工程大学计算机科学与技术学院,黑龙江哈尔滨 150001; 2. 烟台大学计算机与控制工程学院,山东烟台 264005)

**摘要:** 人工神经网络学习作为机器学习的重要方式,在人工智能、模式识别、图像处理等领域已成功应用;BP网络作为神经网络学习的精华,它利用误差反传的方式不断修正权重以达到最佳拟合。多属性决策问题是决策理论研究领域的热点,当研究的问题涉及多个属性时,需要分析各属性的重要程度,即属性的权重。针对多分类输出结果的多输入属性相关性和重要性问题,提出了利用BP神经网络计算复杂输入属性的重要性方法;并对神经网络的节点数量、网络层数、学习策略、学习因子等进行研究,建立了适合属性重要性计算的BP神经网络模型;以烟台大学学生评教数据作为具体实例,利用k-fold方法验证其可行性和有效性。

**关键词:** BP神经网络;属性重要性;多分类输出;学生评教

**中图分类号:** TP183      **文献标识码:** A      doi:10.3969/j.issn.0253-2778.2017.01.003

**引用格式:** 潘庆先,董红斌,韩启龙. 一种基于BP神经网络的属性重要性计算方法[J]. 中国科学技术大学学报, 2017,47(1):18-25.

PAN Qingxian, DONG Hongbin, HAN Qilong, et al. A computing method for attribute importance based on BP neural network[J]. Journal of University of Science and Technology of China, 2017,47(1):18-25.

## A computing method for attribute importance based on BP neural network

PAN Qingxian<sup>1,2</sup>, DONG Hongbin<sup>1</sup>, HAN Qilong<sup>1</sup>, WANG Yingjie<sup>2</sup>, DING Rui<sup>1</sup>

(1. College of Computer Science and Technology, Harbin Engineering University, Harbin 150001, China;

2. College of Computer and Control Engineering, Yantai University, Yantai 264005, China)

**Abstract:** As an important method for machine learning, artificial neural network has been applied successfully in artificial intelligence, pattern recognition, image processing and other fields. As the essence of neural network learning, BP network utilizes the error back propagation to correct weights continually in order to achieve the best-fit. The multi-attribute decision-making problem is a hotspot in decision theory. When involving multiple attributes, it needs to analyze the importance degrees for different attributes, i. e., weights of attributes. According to the correlation and importance problems of multiple input attributes for multi-classification output results, an importance method for calculating complex input attributes based on BP neural network was proposed. In addition, the BP neural network model for calculating the importance degrees of attributes was established through researching the number of nodes,

**收稿日期:** 2016-03-01; **修回日期:** 2016-09-17

**基金项目:** 国家自然科学基金(61472095, 61502140, 61572418)资助。

**作者简介:** 潘庆先,男,1979年生,博士生/副教授,研究方向:人工智能、机器学习等。E-mail: ytu\_pqx@163.com

**通讯作者:** 董红斌,男,博士/教授。E-mail: donghongbin@hrbeu.edu.cn

the layers of network, learning strategies and learning factors in neural networks. The data of teaching evaluation of Yantai University is utilized to verify the feasibility and validity of the proposed method through applying  $k$ -fold approach.

**Key words:** BP neural network; importance of attributes; multi-classification output; teaching evaluation

## 0 引言

人工神经网络 (artificial neural network, ANN)<sup>[1-2]</sup> 作为机器学习的重要研究内容之一, 是人们在人脑神经系统研究的基础上, 利用数学理论对人脑神经网络进一步抽象, 从而建立的简化网络模型. BP(back propagation) 网络<sup>[3-4]</sup> 是人工神经网络中应用最广、最具代表性的机器学习方法, 已经成功用于图像处理、模式识别等领域.

多属性决策<sup>[5-6]</sup> 问题是决策理论研究领域的热点, 当研究的决策问题涉及多个属性时, 我们需考虑各属性的重要程度, 即各属性的权重分配问题. 目前确定属性权重的方法大致分为两类<sup>[7]</sup>: 第一类为主观赋权法, 比如专家会议法、最小平方和法和 Delphi 法等; 主观赋权法是以政策的导向为目标, 设计者以主观经验设计属性并设置其权重, 这种方法的有效性过度依赖设计者的经验, 缺乏对现实数据之间关系的分析, 存在主观臆断的缺点. 第二类为客观赋权法, 如王国胤等提出的基于粗糙集理论的属性约简及规则提取方法<sup>[8]</sup>; 陈圣兵等提出一种基于相似关系的不完备信息熵, 并证明该信息熵的若干性质, 给出一个不完备数据特征选择算法<sup>[9]</sup>; 潘庆先等提出基于关联规则数据挖掘方法来计算属性权重, 并应用于学生评教系统中<sup>[10]</sup>; 李亮对权系数理论进行了研究, 提出了权系数理论, 给出了权系数的计算流程<sup>[11]</sup>.

上述方法主要在未知的复杂的特征提取、抽取和属性约简等方面进行了研究, 在特定的时间段、特定的应用领域取得了一定的成功, 主要对输出结果为二分类问题的多输入属性的重要程度进行了研究, 对于多分类输出结果的多相关属性的重要程度及其属性的相关性研究较少, 尤其是利用神经网络的方法来研究属性重要性问题, 近年的文献中还没有被发现. 本文提出了一种基于 BP 神经网络的属性重要性计算方法, 利用  $k$ -fold 方法对真实数据进行训练和验证, 对复杂的多属性进行相关性及其重要程度进行分析, 通过误差反传的方式不断修正属性权值以达到最佳拟合.

本文的主要贡献:

(I) 针对多分类输出结果, 提出了基于 BP 神经网络属性重要性的计算方法;

(II) 对神经网络的节点数量、网络层数、学习策略、学习因子等进行研究, 建立了适合属性重要性计算的神经网络模型.

## 1 相关工作

### 1.1 人工神经网络

长期以来, 研究者们对人脑的工作机理和思维本质进行研究, 试图构造具有人类智能的系统. 当前关于人工神经网络的定义尚不统一, 但可以简单地表述为<sup>[12]</sup>: 人工神经网络是一种旨在模仿人脑结构及其功能的信息处理系统.

#### 1.1.1 人工神经网络的数学模型

人们根据生物神经网络的机理, 将其抽象为数学模型, 人工神经元模型如图 1 所示. 其中  $x_1, x_2, \dots, x_n$  为输入信号,  $o_j$  为输出信号,  $w_{ij}$  为相应输入信号的权重,  $T$  为阈值,  $f$  为激活函数.

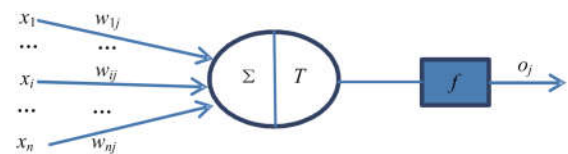


图 1 人工神经元模型

Fig. 1 Artificial neuron model

令  $x_i(t)$  为第  $j$  个神经元在  $t$  时刻接收的来自第  $i$  个神经元的输入信息,  $o_j(t+1)$  为在  $t+1$  时刻第  $j$  个神经元的输出信息, 那么神经元  $j$  的状态可用如下数学式表示为

$$o_j(t+1) = f\left\{\left[\sum_{i=1}^n w_{ij}x_i(t)\right] - T_j\right\} \quad (1)$$

定义净输入为

$$\text{net}'_j(t) = \sum_{i=1}^n w_{ij}x_i(t) \quad (2)$$

则可用矩阵表示为

$$\text{net}'_j = \mathbf{W}_j^T \mathbf{X} \quad (3)$$

式中,

$$\mathbf{W}_j = (w_{1j}, w_{2j}, \dots, w_{nj})^T,$$

$$\mathbf{X} = (x_1, x_2, \dots, x_n)^T.$$

令  $x_0 = -1, \omega_{0j} = T_j$  (或  $x_0 = 1, \omega_0 = -T_j$ )

则有

$$\left. \begin{aligned} -T_j &= x_0 \omega_{0j} \\ \text{net}'_j - T_j &= \text{net}_j = \sum_{i=0}^n \omega_{ij} x_i = \mathbf{W}_j^T \mathbf{X} \\ o_j &= f(\text{net}_j) = f(\mathbf{W}_j^T \mathbf{X}) \end{aligned} \right\} \quad (4)$$

1.1.2 BP 神经网络

BP 网络<sup>[13]</sup>是一种多层前馈网络,其误差按反向传播的方式逐层进行调整,学习过程分为信号的正向传播和误差反向传播两个过程。

BP 的学习规则采用梯度下降法,通过将误差逆向传播的方式调整权值和阈值以达到最佳拟合。BP 网络拓扑结构如图 2 所示。

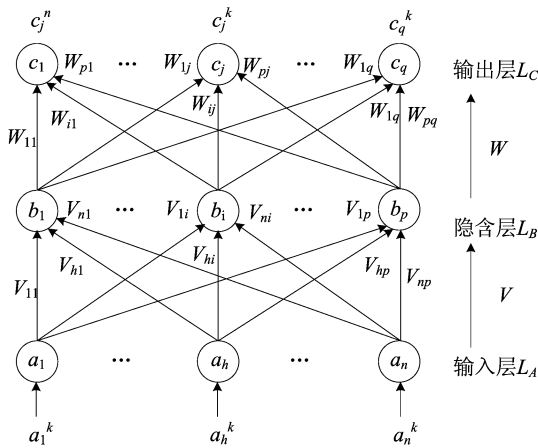


图 2 BP 网络的拓扑结构

Fig. 2 BP network topology

(I) BP 网络的特点

一般 BP 网络由多层构成,层与层之间采用全连接,同层之间无连接,网络的传递函数是可微的,一般采用 Sigmoid 函数 (Log-Sigmoid、Tan-Sigmoid) 作为激活函数, BP 网络属于有导师的监督学习。

(II) BP 算法

BP 网络在学习训练阶段信号为正向传播,误差为反向传播。其中在反向传播阶段,首先用输出层的误差调整上一层的权值和阈值,再根据得到的误差来调整更上一层的权值和阈值,按这样的方式逐层逆向传播下去,最终得到所有层的误差,根据误差逐步调整输入的权值与阈值。其流程如图 3 所示。

BP 网络只是在训练过程中反向传播误差,在测试阶段只需正向传播。衡量假设相对于训练样例的训练误差公式如下:

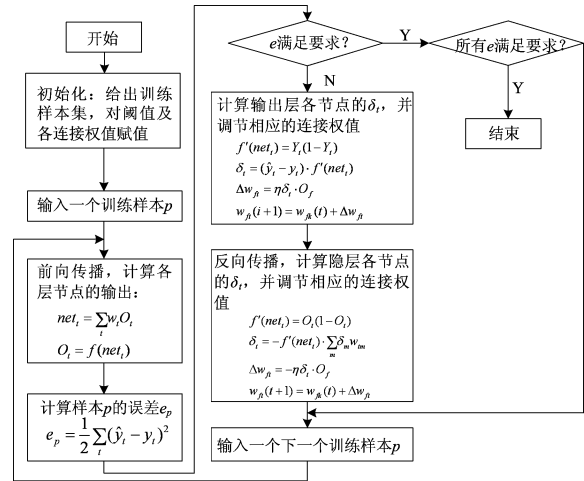


图 3 BP 算法流程图

Fig. 3 BP algorithm flow chart

$$E(\vec{w}) = \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2 \quad (5)$$

BP 算法采用非线性规划方法中的最速下降法修正网络权值,使误差达到最小二乘意义下最小。梯度下降是从任意的初始权向量开始,沿误差曲面下降最快的方向进行搜索,直达到达全局的最小误差点。根据数学中导数的几何意义,通过梯度来得到最速下降的方向,  $E$  对于  $\vec{w}$  的梯度记作  $\nabla E(\vec{w})$ 。梯度训练法则为  $\vec{w} \leftarrow \vec{w} + \Delta \vec{w}$ 。由

$$E(\vec{w}) = \frac{1}{2} \sum_{d \in D} (t_d - o_d)^2 = \frac{1}{2} \sum (t_d^2 - 2 * t_d * o_d + o_d^2) \quad (6)$$

得到梯度的分量为

$$\frac{\partial E}{\partial w_i} = \sum_{d \in D} (t_d - o_d) (-x_{id}) \quad (7)$$

式中,  $\Delta \vec{w} = -\eta \nabla E(\vec{w}) \quad \Delta w_i = \eta \sum_{d \in D} (t_d - o_d) x_{id}$ 。

在解决实际的问题中,误差曲面是高度复杂的,这也使网络的学习成为 NP 难问题,所以 BP 算法容易出现收敛速度慢、易陷入局部极小点等问题。为了解决算法存在的缺陷,研究人员提出了很多措施,比如通过调整学习率来提高收敛的速度;通过加入动量项抑制振荡等;还探索使用其他的梯度优化算法来解决收敛速度问题,比如牛顿法、拟牛顿法、共轭梯度法等;还可以通过选择合理的激励函数和目标函数来降低误差曲面的复杂性;同时在训练前,对样本数据进行合理的整理,如样本筛选、特征压缩、数据归一化等来加快学习的速度。

## 1.2 属性重要性

属性的重要性指各属性对输出结果的重要程度或贡献程度,一般用权重进行衡量.传统的属性权重定义方法多采用专家会议法,通过领域专家根据自己的经验对案例的特征属性进行赋权,这种方法过度依赖于专家的主观判断和经验,存在误差,往往会遭到被评价者的反对.另外的方法是客观赋权法,常用的方法是采用粗糙集理论、信息熵等方法对案例属性的权重进行赋值;2010年,潘庆先等利用关联规则数据挖掘技术,通过计算各指标与输出结果的关联程度对各指标进行赋权,当时只考虑了输入指标为“优秀”,输出评价等级为“优秀”的关联程度,缺乏对全局数据的考量.

针对多值输入和多分类输出问题,根据 BP 神经网络反向修正权值的特点,本文提出了利用神经网络方法计算属性的重要度.

## 2 基于 BP 神经网络的属性重要性计算模型

### 2.1 网络模型的构建

本文选用 BP 网络作为基本网络模型.网络结构中隐层可以设置多层,输入节点的个数与特征属性的数量一致,输出层节点个数与分类的种类数量相关,一般采用 1-of- $n$  方式.

网络的初始权值、阈值、学习算法、数值修改频度、节点变换函数及参数、学习率等参数的设置对网络的收敛速度和效果都有直接的影响.初始权值、阈值采用随机设置的方式.数值修改频度、节点变换函数、学习率等参数,在工程实践中主要靠经验和试凑.对网络规模、学习参数的科学化研究也是神经网络需要研究的课题.

本文实验根据学生评教的具体应用构造网络模型,网络由输入层、隐层 1、隐层 2、输出层组成,输入层节点个数与特征属性的个数一致,输出层与结果的类别数一致,为得到输入层的权值隐层 1 设置一个节点,为提供收敛速度隐层 2 设置了 10 个节点.

### 2.2 学习策略的选择

传统的 BP 神经网络学习具有收敛速度慢、易于陷入局部最优的缺陷.为了改进学习策略,本文采用了自适应的动量梯度下降法,以提高收敛速度并避免振荡.

自适应动量梯度下降法是在传统的梯度下降法

中引入动量因子  $\eta$  和学习率  $\alpha$ .  $\eta$  相当于阻尼系数,用于避免陷入局部最小值,用学习率  $\alpha$  来避免学习过程的震荡性.

$$w(k+1) = w(k) + \alpha[(1-\eta)\Delta\tau w(k) + \eta\Delta\tau w(k-1)] \quad (8)$$

式中,  $\Delta\tau w(k)$  为  $k$  时刻的负梯度,学习率  $\alpha > 0$ , 动量因子  $\eta \in [0, 1]$ .

动量法实质是将最后一次权值变化的影响通过动量因子来传递,本文对动量因子的取值进行了验证,当  $\eta = 0.95$  时,网络的训练避免了陷入局部极小值,落入全局最小值.训练结果如图 4 所示.

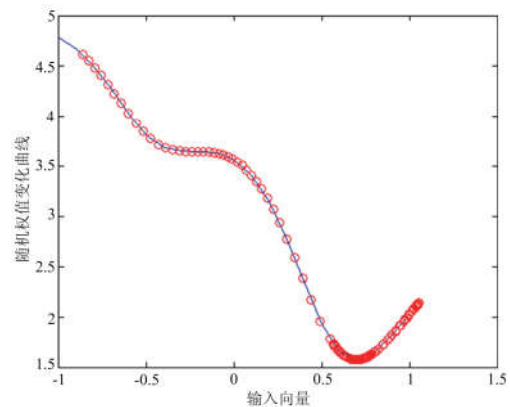


图 4  $\eta=0.95$  时的训练结果

Fig. 4 The training results of  $\eta=0.95$

学习率  $\alpha$  在传统的 BP 算法中为常数,由于误差曲面的平坦程度不同,现实中很难找到这样一个常数作为最佳学习率,本实验采用了自适应调节学习率的方法,公式表示为

$$\alpha(k+1) = \begin{cases} 1.05\alpha(k), & \text{SSEY}k + 1Y < \text{SSE}(k) \\ 0.7\alpha(k), & \text{SSE}(k+1) > \text{SSE}(k) \\ \alpha(k), & \text{其他} \end{cases} \quad (9)$$

通过实验证实,采用自适应调节学习率的方法,网络训练的次数远小于固定学习率的训练次数,所以自适应学习率的网络训练非常有效.

### 2.3 k-fold 交叉方法

$k$ -fold 交叉方法<sup>[14]</sup>是把训练数据集平均分成  $k$  份,在训练的过程中用其中的  $k-1$  份作为训练数据,把剩下的一份作为验证集,按这样的方式进行  $k$  次实验,每次实验都使用其中不同的一份作为验证集.

使用交叉验证方法进行  $k$  次实验,得到每次最佳性能的迭代次数  $i_j$  ( $j \in [1, k]$ ),然后通过

$$n = \frac{1}{k} \sum_{j=1}^k i_j \quad (10)$$

计算得到均值  $n$ 。最后,运行一次反向传播算法,训练所有  $m$  个实例并迭代  $n$  次。

$k$ -fold 交叉验证法中  $k$  的取值对泛化的误差是有影响的,在训练样本足够多的情况下, $k$  的取值越大,评估结果越可靠。本文实验  $k$  值取为 10,对数据集进行训练与验证。

表 1 学生评教指标(理论课)  
Tab. 1 Teacher evaluation index (theory course)

序号	评价指标	评价等级(分优良中差)
1	为人师表,对学生热情,严格要求	
2	对教学工作有热情,讲课认真,遵守教学纪律	
3	认真辅导答疑,重视作业检查并及时解决问题	
4	讲课思路清晰,阐述准确,用普通话授课,语言生动	
5	注意启发引导,师生相互沟通,学生积极性高	
6	理论联系实际,注重学生创新能力的培养	
7	板书清楚、有条理或简洁、实用;有效运用各种教学手段	
8	内容充实,讲授训练,突出重点,讲清难点	
9	教材选用合适,注重内容更新	
总评	对该教师的总体评价	

学生评教的原始数据如表 2 所示。其中 BPR 为教师编号, PGR 为学生学号, PGNR 为课程编号。

表 2 学生评教原始数据  
Tab. 2 Raw data table of student evaluation

BPR	PGR	PGNR	$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_7$	$i_8$	$i_9$	zp	ZGPI
25012	201421505130	252000011	良	优	优	优	优	良	优	优	优	优	老师给我很深的印象
25016	201325505131	252000035	优	优	优	优	优	优	优	优	优	优	内容丰富,有趣
25012	201325605108	252000011	优	优	良	良	优	优	优	优	优	优	普通话需加强
25032	201325501101	252000278	优	中	优	优	优	优	优	优	优	良	
21021	201321501132	212000218	优	优	优	优	优	优	良	良	良	优	
22047	201321502191	222003245	优	优	优	优	优	优	优	良	优	优	讲得挺好的,希望生动些

把表 1 中的序号 1-9 所对应的指标作为网络的输入,对应输入层的 9 个节点,指标 10“对应教师的总体评价”作为输出。对输入数据进行数值化,用数值 1 代替等级“优”,0.8 代替“良”,0.5 代替“中”,0.1 代替“差”,输出采用 1-of- $n$  的方式,在神经网络中用 4 个节点作为输出。为使结果快速收敛,设置两个隐层。为了便于表示指标的权值,第一层隐层设置一个节点。神经网络的 Matlab 模型如图 5 所示。

在训练数据的选取与预处理阶段,根据以往多

### 3 实验与仿真

本文以 2014—2015 烟台大学 2 学期的学生评教数据为实验数据,以学生评教指标(理论课)为实验对象,利用神经网络方法计算各指标的权重。学生评教(理论课)指标体系首先利用专家法、问卷调查法选取主要指标,如表 1 所示。

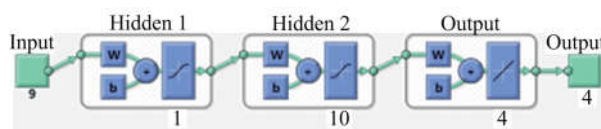


图 5 神经网络 Matlab 模型

Fig. 5 The Matlab model of neural network

年的分析数据,选取多年评价在前 10% 和后 10% 的评价成绩,去除奇异性数据后,选取了 200 条数据作为训练数据,并对数据进行标注,处理后的数据如表

3 所示. 其中  $i_1-i_9$  为输入,  $o_1-o_4$  作为输出, 分别代表 “优、良、中、差” 四类. 训练采用  $k$ -fold 交叉验证方

表 3 训练数据样例表

Tab. 2 Sample of training data

$i_1$	$i_2$	$i_3$	$i_4$	$i_5$	$i_6$	$i_7$	$i_8$	$i_9$	$o_1$	$o_2$	$o_3$	$o_4$
1	0.8	1	1	0.5	1	0.1	0.5	0.1	1	0	0	0
0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	0.1	1	0	0	0
0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0.5	0	1	0	0
0.5	0.8	0.5	0.5	0.1	0.5	0.5	0.5	0.5	0	1	0	0
0.5	0.8	0.5	0.8	0.5	0.5	0.5	0.8	0.5	0	1	0	0
0.8	0.8	0.8	0.8	1	1	1	1	1	0	0	1	0
0.8	1	0.8	1	0.8	0.8	1	1	1	0	0	1	0
1	1	1	1	1	1	1	1	1	0	0	0	1
1	0.8	0.8	1	1	0.8	1	1	0.8	0	0	0	1
1	1	1	1	1	1	1	1	1	0	0	0	1

本文采用 Matlab 工具进行实验与仿真, 学习策略选用自适应  $lr$  动量梯度下降法, 动量因子取 0.95, 学习速率的调整采用公式(9), 调节因子分别取 1.05 和 0.7. 隐层的激活函数采用 TanSig, 输出层采用 PureLin. 实验运行结果如图 6 所示. 输出结果如表 4 所示, 其中  $o_1-o_4$  为输出分别代表优、良、中、差四个等级, 从表中可以看出, 对于多分类输出

的数值可以将其看作其输出的置信度, 按数值的大小可以判断输出的类别, 比如样例 1,  $o_1$  的输出数值为 0.932 255 远远高于其他输出值, 可以判断此输入为“差”等级, 而且置信度非常高. 以此作为训练模型, 对 100 个测试用例进行测试, 准确率达到 96%, 说明此模型是可靠的.

表 4 多分类输出结果样例表

Tab. 4 The output results sample of multi classification

	$o_1$	$o_2$	$o_3$	$o_4$
样例 1	0.932 255	-0.032 96	0.224 629	0.064 393
样例 2	1.102 405	0.140 075	-0.334 66	-0.147 03
样例 3	-0.011 17	0.875 413	0.228 272	0.043 651
样例 4	-0.045 52	0.825 099	0.094 356	-0.090 66
样例 5	-0.043 2	0.744 52	0.088 791	-0.001 87
样例 6	0.070 381	0.188 918	0.786 767	0.412 175
样例 7	0.070 381	0.188 918	0.786 767	0.412 175
样例 8	-0.003 1	-0.001 48	0.163 48	0.792 728
样例 9	-0.013 22	0.070 819	0.108 454	0.728 914
样例 10	-0.007 95	0.018 481	0.132 535	0.779 466

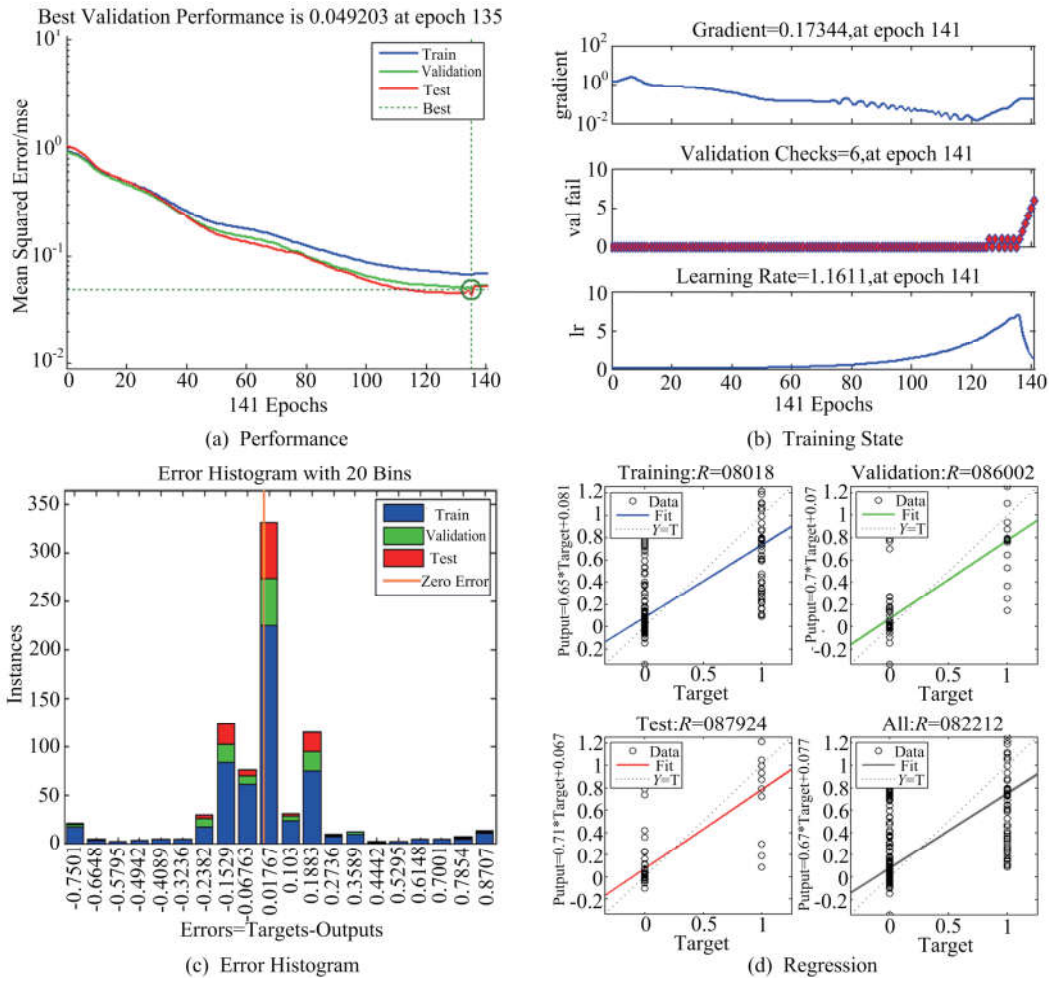


图 6 实验运行结果

Fig. 6 Experimental results

设  $\omega_1 - \omega_9$  为输入层 9 个节点向第一隐层节点传播的权重, 因输入层 9 个节点与学生评教中的 9 个指标相对应, 其权重的大小反映了学生对该指标的重视程度. 实验权值及阈值如表 5 所示, 从表中可

以看出各指标的权重之间的差异, 其中  $\omega_7$  数值最大, 表明在学生更看重指标 7;  $\omega_3$  最小, 说明“认真辅导答疑, 重视作业检查并及时解决问题”不太受到学生的重视.

表 5 输入权值表

Tab. 5 The weight of input

$\omega_1$	$\omega_2$	$\omega_3$	$\omega_4$	$\omega_5$	$\omega_6$	$\omega_7$	$\omega_8$	$\omega_9$	阈值 b
-0.04607	0.139918	-0.51065	0.024413	-0.25193	0.183732	0.986376	0.049169	0.677114	0.154449

## 4 结论

特征工程是重要但又是费力的, 特征工程是用目标问题所在的特定领域知识或者自动化的方法来生成、提取、删减或者组合变化得到特征. 特征的选择及属性的重要性将直接影响分类的结果, 本文提出了用神经网络的方法来计算属性的重要程度, 并以学生评教为具体实例, 验证了其有效性, 对相关领

域的研究具有较好的借鉴意义.

同时由于评价者自身松弛程度的不同, 比如同样评价“良”, 比较严格的用户表示很好, 比较宽松的用户表示一般, 这对评价结果的分类也将产生一定的影响, 在以后的研究中应加入用户的特征参数<sup>[15]</sup>; 随着机器学习研究的不断深入, 利用深度学习<sup>[16]</sup>、集成学习<sup>[17-19]</sup>的方法来研究复杂问题, 包括学生主观留言数据的情感分析; 通过建立可信模型, 来验证评价数据的可信性.

## 参考文献(References)

- [1] XU Libin, LI Yang, XU Ning. Soy sauce classification by geographic region and fermentation based on artificial neural network and genetic algorithm [J]. *Journal of Agricultural and Food Chemistry*, 2014, 62(51): 12294-12298.
- [2] YAİCI W, ENTCHEV E. Performance prediction of a solar thermal energy system using artificial neural networks [J]. *Applied Thermal Engineering*, 2014, 73(1): 1348-1359.
- [3] HAN X H, XIONG X Y, DUAN F. A new method for image segmentation based on BP neural network and gravitational search algorithm enhanced by cat chaotic mapping [J]. *Applied Intelligence*, 2015, 43(4): 855-873.
- [4] LIU Y K, XIE F, XIE C L, et al. Prediction of time series of NPP operating parameters using dynamic model based on BP neural network [J]. *Annals of Nuclear Energy*, 2015, 85: 566-575.
- [5] VAN VALKENHOEF G, TERVONEN T, POSTMUS D. Notes on 'Hit-And-Run enables efficient weight generation for simulation-based multiple criteria decision analysis' [J]. *European Journal of Operational Research*, 2014, 239(3): 865-867.
- [6] ZHANG X L, XU Z S. Extension of TOPSIS to Multiple Criteria decision making with Pythagorean fuzzy sets [J]. *International Journal of Intelligent Systems*, 2014, 29(12): 1061-1078.
- [7] 刘文军, 谷云东. 属性依赖性及其重要性度量 [J]. *数学的实践与认识*, 2009, 39(7): 148-156.
- [8] 王国胤, 姚一豫, 于洪. 粗糙集理论与应用研究综述 [J]. *计算机学报*, 2009, 32(7): 1229-1246.
- WANG G Y, YAO Y Y, YU H. A survey on rough set theory and applications [J]. *Chinese Journal of Computers*, 2009, 32(7): 1229-1246.
- [9] 陈圣兵, 王晓峰. 基于信息熵的不完备数据特征选择算法 [J]. *模式识别与人工智能*, 2014, 27(12): 1131-1137.
- CHEN S B, WANG X F. Feature Selection Algorithm for incomplete data based on information entropy [J]. *Pattern Recognition and Artificial Intelligence*, 2014, 27(12): 1131-1137.
- [10] 潘庆先, 于萍, 娄兰芳. 关联规则算法的研究及其在教学评价中的应用 [J]. *烟台大学学报(自然科学与工程版)*, 2010, 23(2): 127-131.
- [11] 李亮. 评价中权系数理论与方法比较 [D]. 上海: 上海交通大学, 2009.
- [12] 韩立群. 人工神经网络理论、设计及应用 [M]. 北京: 化学工业出版社, 2007.
- [13] 刘智斌, 曾晓勤, 刘惠义, 等. 基于 BP 神经网络的双层启发式强化学习方法 [J]. *计算机研究与发展*, 2015, 52(3): 579-587.
- LIU Z B, ZENG X Q, LIU H Y, et al. A heuristic two-layer reinforcement learning algorithm based on bp neural networks [J]. *Journal of Computer Research and Development*, 2015, 52(3): 579-587.
- [14] ESMAEELZADEH S R, ADIB A, ALAHDIN S. Long-term streamflow forecasts by adaptive neuro-fuzzy inference system using satellite images and K-fold cross-validation (Case study: Dez, Iran) [J]. *KSCIE Journal of Civil Engineering*, 2015, 19(7): 2298-2306.
- [15] TANG D Y. User modeling with neural network for review rating prediction [C] // *Proceedings of the 24th International Conference on Artificial Intelligence*. Buenos Aires, Argentina: AAAI Press, 2015: 1340-1346.
- [16] 尹宝才, 王文童, 王立春. 深度学习研究综述 [J]. *北京工业大学学报*, 2015, (1): 48-59.
- [17] 周志华, 陈世福. 神经网络集成 [J]. *计算机学报*, 2002, 25(1): 1-8.
- ZHOU Z H, CHEN S F. Neural network ensemble [J]. *Chinese Journal of Computers*, 2002, 25(1): 1-8.
- [18] TANG S, ZHANG Y D, XU Z X. An efficient concept detection system via sparse ensemble learning [J]. *Neurocomputing*, 2015, 169: 124-133.
- [19] MONGE D A, HOLEC M, ŽELEZNÝ F, et al. Ensemble learning of runtime prediction models for gene-expression analysis workflows [J]. *Cluster Computing*, 2015, 18(4): 1317-1329.