

DOI: 10.11830/ISSN.1000-5013.201707010



# 采用 BP 算法和深度 SAE 网络的 学生综合能力评价方法

张永梅<sup>1</sup>, 付昊天<sup>1</sup>, 张睿<sup>2</sup>

(1. 北方工业大学 计算机学院, 北京 100144;  
2. 太原科技大学 计算机科学与技术学院, 山西 太原 030024)

**摘要:** 针对现有评价方法需人工提取特征且评价准确率低的问题, 提出基于反向传播(BP)算法的深度堆栈编码器(SAE)网络的学生综合能力评价方法. 通过 SAE 网络对输入的学生各项指标成绩进行无监督训练, 将 SAE 学习到的特征结合相应的样本标签, 利用柔性最大值分类器(Softmax)进行有监督式分类. 采用 BP 算法进行反向传播调整隐层权重, 优化整个模型, 以避免过拟合现象的发生. 结果表明: 该评价方法有利于解决需对传统神经网络进行人工提取和分析特征的问题, 可提高评价结果的准确率.

**关键词:** 反向传播算法; 深度神经网络; 堆栈式自编码器; 综合能力评价

中图分类号: TP 183      文献标志码: A      文章编号: 1000-5013(2018)05-774-07

## Evaluation Method for Students' Comprehensive Abilities Using Deep SAE Networks and BP Algorithm

ZHANG Yongmei<sup>1</sup>, FU Haotian<sup>1</sup>, ZHANG Rui<sup>2</sup>

(1. College of Computer Science and Technology, North China University of Technology, Beijing 100144, China;  
2. School of Computer Science and Technology, Taiyuan University of Science and Technology, Taiyuan 030024, China)

**Abstract:** Since the existing evaluation methods need to extract features manually, and the evaluation accuracy is lower. The paper proposes an evaluation method for students' comprehensive abilities based on deep SAE Networks and back propagation (BP) algorithm. The method adopts stacked autoencoder (SAE) network for students' various scores with unsupervised training, and utilizes the features extracted by SAE network and the corresponding sample labels to train the network via supervised Softmax classifier. The method uses BP algorithm to adjust the weights for hidden layers and optimize the entire model, and avoids the occurrence of over-fitting. The method can help to solve the problem of manual extraction and analysis features for traditional neural networks, and improve the evaluation accuracy.

**Keywords:** back propagation algorithm; deep neural network; stacked autoencoder; comprehensive ability evaluation

在学生综合能力评价领域,很少有学者采用深度学习法进行评价.而如何客观地评价计算机专业学生的综合能力并建立有效的评价体系,是当前亟需解决的问题.目前,评价方法主要包括层次分析法、K最近邻方法和神经网络评价方法<sup>[1-3]</sup>.深度神经网络采用模拟人脑的多层结构,对输入数据从底层到高

收稿日期: 2017-07-03  
通信作者: 张永梅(1967-),女,教授,博士,主要从事图像处理的研究. E-mail: zhangym@ncut.edu.cn.  
基金项目: 国家自然科学基金资助项目(61371143); 北方工业大学优势学科科研基金资助项目(XN044); 太原科技大学博士科研启动项目(20162036); 北方工业大学教育教学改革和课程建设研究项目(XN093-002)

层进行逐级提取特征, 最终形成适合模式分类且较为理想的特征, 避免传统神经网络等需要进行人工特征选取与分析的问题. 目前, 很少有学者将深度神经网络应用到评价中的原因有: 1) 由于训练数据样本不够多, 模型较为复杂, 表达能力更强的深度神经网络会过多地拟合当前的模型, 导致过拟合; 2) 由于很多深度神经网络模型在训练时, 需要大量的标签数据才能进行训练, 而在实际的评价中, 找到很多带有标签(评价结果)的数据往往较难. 深度堆栈编码器(SAE)是深度学习中一个重要的模型, 它能很好地抽象评价数据的稀疏性特征, 并避免因维度过高导致的过拟合问题. 因此, 本文提出一种基于反向传播(BP)算法的深度 SAE 网络模型的学生综合能力评价方法.

# 1 基于深度 SAE 网络的学生综合能力评价方法

## 1.1 学生综合能力评价的 SAE 网络模型

SAE 网络模型是一种无监督学习方法, 可以提取到一些非线性特征, 主要包括单层的 SAE 单元和用于进行分类(给出评价结果)的分类器. 与监督式学习算法对比, 无监督学习方法不是“教会”网络如何学习特征, 而是让网络自己发现特征. 由于需要给出学生综合能力的多个评价结果, SVM 分类器的时间复杂度较高, 约为  $O(n^3)$ , 而且 SVM 分类器的分类效果在一定程度上依赖于核函数的选取.

柔性最大值分类器(Softmax)是一个线性分类器, 它可解决多分类问题, 其时间复杂度约为  $O(n)$ . 采用 Softmax 分类器对 SAE 学习到的特征进行分类<sup>[4]</sup>, SAE 与 Softmax 二者结合构建的 SAE 模型, 结合了无监督学习与有监督学习的优点, 能够对学生综合能力进行有效评价.

单层的 SAE 单元由编码器和解码器 2 部分组成. 高维度的数据, 即输入每个学生综合能力的各项分数, 输入形式是一个或多个向量  $I_n=[X_1, X_2, \dots, X_n]$ . 这些输入向量通过编码器形成低维度的编码, 并通过解码器将编码结果进行相应解码, 以便更完整地恢复输入的原始数据, 每一个隐层的特征编码都是从处于较低层的 SAE 单元传向更高层的单元<sup>[5]</sup>. 通过对输入数据进行逐层编码, 可以从训练数据通过无监督式提取得到评价体系权重特征. Softmax 分类器连接 SAE 神经网络模型的最高层, 即 SAE 网络的输出层, 通过计算输入特征的后验概率对测试数据进行分类. 在训练网络时, 加入 BP 算法, 通过有监督微调, 可进一步确定较优的网络模型; 同时, 不断调整网络隐层单元的权重和训练参数, 可达到更好的分类效果, 并避免过拟合问题的发生. 文中采用的深度 SAE 网络模型, 如图 1 所示.

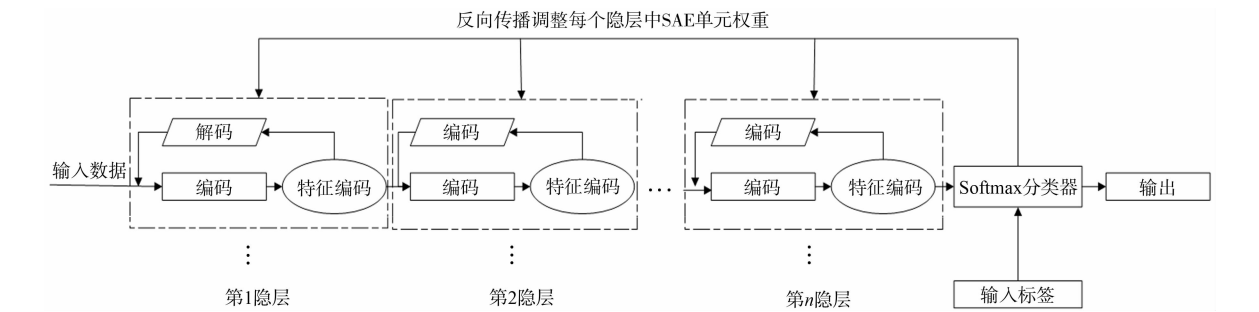


图 1 深度 SAE 网络模型  
Fig. 1 Deep SAE network model

输入数据是从网络的输入层节点输入到网络层, 输入的是一个或多个学生综合能力的各指标成绩组成的输入矩阵  $M=(l_1, l_2, l_3, \dots, l_n)$ . 由于 SAE 组成的神经网络采用无监督学习算法, 因此, 不需要输入每个评价客体的标签, 即评价结果; 之后, 进入每个 SAE 单元, 在单元中进行编码和解码, 并自动提取到适应于本模型的特征向量, 将提取到的特征输入到网络的下一层, 执行编码、解码, 并得到特征. 以此类推, 当网络最高层 SAE 单元提取到最终特征时, 将提取到的特征输入到 Softmax 分类器中, Softmax 分类器计算输入矩阵  $M$  提取特征的分类结果的后验概率, 并通过损失函数计算评价结果与专家评价的结果  $R=(R_1, R_2, R_3, \dots, R_n)$  之间的损失值. 通过对损失函数求导, 计算造成损失值的各个参数, 并通过反向传播算法更新这些参数的值, 以便得到更好的分类模型, 提高模型的拟合能力<sup>[6]</sup>.

## 1.2 单层 SAE 单元模型

SAE 单元模型又称为迭代自动编码器模型, 是一种编码器和解码器对等结构的神经网络, 从无标

签的测试集数据学习特征. 单层的 SAE 单元模型, 如图 2 所示. 它通过编码、解码使输入数据近似于输出数据. 编码可以看作是单元从数据中以无监督学习方式学习到的特征. 神经网络的输入, 即每个学生综合能力的各项得分表示为一个向量, 利用 S 形 (Sigmoid) 函数对输入的向量进行编码<sup>[7]</sup>.

在带有稀疏性的训练目标下, 网络训练结果大部分是隐含层节点处于抑制状态, 并将隐含层节点输出值限定在所设定的目标附近. 稀疏性被认为是一种广泛存在的现象, 通过在网络中引入稀疏性, 能够增强网络的鲁棒性, 为网络提供较低的输入耦合, 从而增强网络的泛化能力<sup>[8]</sup>. 稀疏神经网络的表现更加接近生物的神经网络, 即使在连接度很低的情况下, 也能保证一定的输出精度.

1.3 基于 BP 算法的深度 SAE 网络的学生综合评价方法模型训练

上述深度 SAE 网络的神经网络参数传递方式是单向的, 即从输入到 SAE 网络的低层 SAE 单元, 再到高层 SAE 单元, 最终将参数传递到分类器. 单向传播容易导致过拟合问题, 即训练好的模型过度适合于当前的输入数据, 对当前输入的分类效果极好, 而对于不参与当前网络训练的测试数据, 则识别效果较差<sup>[9]</sup>.

反向传播训练可通过有监督式学习将高层得到的参数误差反向传递到低层次网络, 通过更新低层网络的权值, 达到在训练 SAE 网络正向传播时, 网络能通过各层 SAE 单元的权值提取学生成绩特征, 并通过 Softmax 分类器给出评价结果. 如果评价结果与标签 (专家评价结果) 之间存在误差, 则需要通过损失函数计算误差, 并通过损失函数对各层 SAE 单元权值求偏导, 计算该节点需要更新的权值并更新. 经过多次训练, 网络的权值不断优化, 更能贴近专家给出的评价结果. 如果不进行 BP 训练, SAE 只能利用随机初始化的权值进行评价, 评价的正确率将保持不变, 而不会随着训练提高评价的正确率. 因此, 需要将 SAE 网络结合 BP 训练算法进行评价<sup>[10]</sup>.

基于 BP 算法的深度 SAE 网络的学生综合能力评价方法, 如图 3 所示. 在深度 SAE 网络的每一层编码和解码操作时, 输入数据的特征及训练参数得以不断传递, 而反向传播算法可以计算预设输出和实际输出之间的误差, 并在下次训练时对参数进行调整.

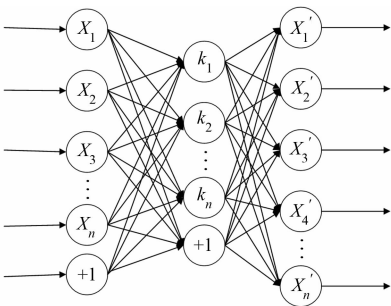


图 2 单层 SAE 网络模型  
Fig. 2 Single layer SAE network model

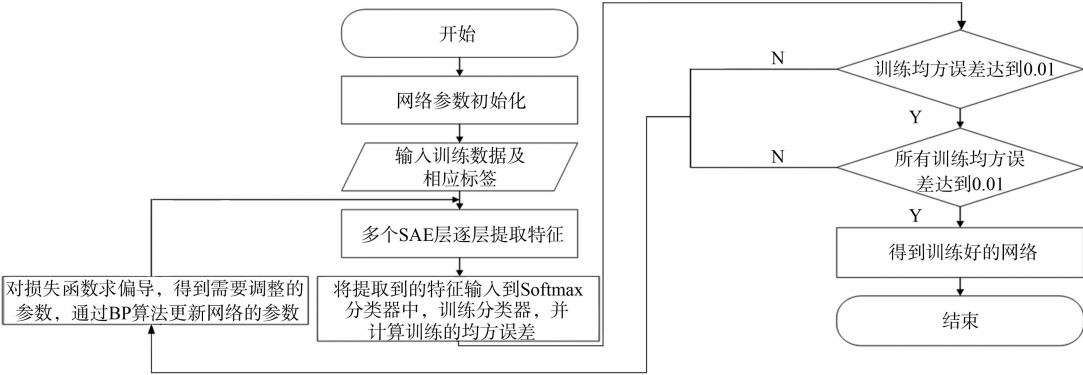


图 3 网络训练算法流程图  
Fig. 3 Algorithm flowchart for training network

首先, 对学习率、动量、每个单层 SAE 的权重等网络参数进行初始化, 输入数据包括每个学生的各项指标得分及训练标签, 即专家对该生各科成绩的评价结果. 然后, 通过前向传播, 将学生成绩输入到第一层 SAE, 进行编码和解码. 最后, 将解码结果依次传入到下一个 SAE 层, 最后一层 SAE 将输出传入到 Softmax 分类器, 并进行训练. 这里对 Softmax 分类器训练主要是使其能够具有足够的泛化能力, 即对训练集以外的数据能够给出正确的评价<sup>[11]</sup>.

对于训练分类器来说, 可以将其看作是一个具有很多参数的假设函数. 它通过不断计算当前网络输出的标签与正确标签 (专家给出的评价结果) 之间的差, 即代价函数的值, 并计算各个参数的偏导数, 调

整造成网络输出与实际输出偏差的参数值,使假设函数对输入数据拟合程度更高,从而对新的输出也具有拟合能力.假设函数的定义为

$$h_{\theta}(x^{(i)}) = \begin{bmatrix} p(y^{(i)} = 1 | x^{(i)}; \theta) \\ p(y^{(i)} = 2 | x^{(i)}; \theta) \\ \vdots \\ p(y^{(i)} = k | x^{(i)}; \theta) \end{bmatrix} = \frac{1}{\sum_{j=1}^k \exp(\theta_j^T x^{(i)})} \begin{bmatrix} \exp(\theta_1^T x^{(i)}) \\ \exp(\theta_2^T x^{(i)}) \\ \vdots \\ \exp(\theta_k^T x^{(i)}) \end{bmatrix}. \tag{1}$$

式(1)中: $x^{(i)}$ 为输入训练集,是每组学生的成绩; $y^{(i)}$ 为正确标签,即专家的评价结果;由于需要进行 6 个等级的评估,即将输入数据分成 6 类,取  $k=6$ ;  $p(y^{(i)} = k | x^{(i)}; \theta)$  为当前网络对于输入数据属于第  $k$  类的概率; $\theta$  为输入到分类器中的网络参数; $h_{\theta}(x^{(i)})$  的输出为当前数据分别属于第  $k$  类的概率,概率最大的输出就认为是网络进行分类的结果<sup>[12]</sup>.

由于分类器的输出与专家的评价结果很有可能不同,因此,需要定义损失函数,即

$$J(\theta) = -\frac{1}{m} \left[ \sum_{i=1}^m \sum_{j=1}^k 1\{y^{(i)} = j\} \lg \frac{\exp(\theta_j) x^{(i)}}{\sum_{l=1}^k \exp(\theta_l) x^{(i)}} \right]. \tag{2}$$

式(2)中:假设函数是衡量当前网络的输出,即假设函数的输出与实际输出的偏差有多大; $1\{\cdot\}$ 是一个示性函数,表示分类器预测结果与标签是否一样,若二者相同,输出 1,否则,输出 0. 训练的最终目标是使损失函数最小化<sup>[13]</sup>.

为防止发生过拟合,加入一种惩罚范式,即权重衰落项. 权值衰减是以某个小因子降低损失函数的参数值,进而减少过大的参数值对梯度下降优化的影响,避免过拟合现象的出现. 惩罚范式的损失函数  $J^*(\theta)$  定义为

$$J^*(\theta) = J(\theta) + \frac{\lambda}{2} \sum_{i=1}^k \sum_{j=0}^n \theta_{i,j}^2. \tag{3}$$

式(3)中: $J(\theta)$ 为损失函数; $\frac{\lambda}{2} \sum_{i=1}^k \sum_{j=0}^n \theta_{i,j}^2$  为加入的惩罚范式. 于是,新的损失函数  $J^*(\theta)$  为严格的凸函数,可采用随机梯度下降优化方法来保证  $J^*(\theta)$  的全局最小值. 优化  $J^*(\theta)$ ,需对  $\theta_j$  进行偏导,即

$$\nabla_{\theta_j} J(\theta) = -\frac{1}{m} \sum_{i=1}^n (x^i (1\{y^{(i)} = j\} - p(y^{(i)} | x^{(i)}; \theta))) + \lambda \theta_j. \tag{4}$$

通过式(4)的多次迭代,可得  $J^*(\theta)$  的全局最小值点. 为求得式(3)的最小值,采用 BP 算法对参数  $\theta$  进行负梯度更新,有

$$\theta_j' = \theta_j - \nabla_{\theta_j} J(\theta). \tag{5}$$

至此,完成 Softmax 分类器训练过程,并得到一套参数. 对于网络测试,即输入数据为训练集之外的数据,没有标签,需要网络通过之前,将学习到的参数对输入的若干组学生成绩进行评定<sup>[14]</sup>. 此处,网络仅进行前向传播,通过 SAE 逐层提取输入成绩的特征,输入到分类器中. 由于 Softmax 分类器已经训练完成,即式(1)中的假设函数已经对数据拟合出较好的效果,可通过式(1)计算当前输入所属 6 个级别的概率. 概率最大的类别就是分类结果,即基于 BP 算法深度 SAE 网络对于当前学生各项成绩综合给出的评价结果,完成网络泛化能力测试.

2 结果与分析

以大学计算机专业本科生的数据为例(如果其他专业有足够的样本数,也可进行网络构建、训练与学生评价,但网络结构可能需要根据数据形式稍作调整),综合各大高校计算机专业的本科生课程、实践等多方面能力,进行进一步细化,生成 48 个 3 级指标,可较为科学、合理地评价学生的综合能力. 指标包括主观、客观评价指标. 例如,课程成绩、学生参加竞赛情况、学生分析问题能力、掌握编程语言的熟练程度、逻辑思维能力等指标. 软件环境为 Windows 10 64 bit, MATLAB 2016a;硬件环境为 i7-3770 CPU, 8 GB RAM. 部分训练集数据,如表 1 所示. 为便于数据的处理,已将数据归一化到[0,1]区间内.

部分学生综合能力的评价结果,如表 2 所示. 表 2 中:专家根据 48 个指标的得分情况,对学生综合

能力的评价结果分为 A,B,C,D,E,F 6 个等级,A 为最好,F 为最差. 每行的“1”所在列代表该行对应学生的评价结果,该评价结果将用于 BP 神经网络的训练和正确率的计算.

表 1 部分训练集数据

Tab.1 Part of training set data

数据结构	C 程序设计	离散数学	操作系统	...	论文发表情况
0.99	0.72	0.91	0.76	...	0.81
0.83	0.91	0.77	0.75	...	0.96
0.78	0.87	0.77	0.65	...	0.71
0.89	0.90	0.83	0.78	...	0.93
0.79	0.87	0.94	0.85	...	0.89
⋮	⋮	⋮	⋮	⋮	⋮
0.79	0.81	0.71	0.90	...	0.94

表 2 部分学生编程能力的评价结果

Tab.2 Evaluation results of some students' programming abilities

A	B	C	D	E	F
0	0	1	0	0	0
0	1	0	0	0	0
0	0	1	0	0	0
0	0	1	0	0	0
0	0	0	0	0	1
⋮	⋮	⋮	⋮	⋮	⋮
0	0	0	0	1	0

结合文献[15]推荐的参数和大量实验数据,得到较优参数. 3 种深度网络的参数,如表 3 所示. 实验发现,学习率为 0.01,0.03,0.10,0.30 时,结果较好;学习率为 0.30 时,结果最好. 通过去除一些梯度下降过程中的噪声和振动,加入动量使网络可更快地突破识别错误率的“停滞期”. 迭代次数是决定训练和识别效果的重要参数之一,在一定范围内,训练次数与效果呈正比关系. 但过多的训练并不能继续提高识别效果,且每次实验会消耗大量时间.

表 3 3 种深度网络的参数

Tab.3 Parameters of three deep networks

置信网络		BP 神经网络		SAE 神经网络	
参数	取值	参数	取值	参数	取值
神经网络的隐层数	3	神经网络的隐层数	3	SAE 的隐层数	3
每个隐层的节点数	{48,20,30}	每个隐层的节点数	{48,30,20}	每个隐层的节点数	{48,30,20}
激活函数	Sigmoid	激活函数	Sigmoid	激活函数	Sigmoid
学习率	0.3	学习率	0.3	学习率	0.3
动量	0.5	动量	0.5	动量	0.5
迭代次数	100	迭代次数	100	迭代次数	100
每个 batch 处理数据的个数	20	每个 batch 处理数据个数	20	稀疏目标值	0.05
				加入“0”值百分比	0.01
				每个 batch 处理数据个数	20

为兼顾实验效果和时效性,在 30 次的实验中,每次训练的迭代次数设为 100 次. 对于每个 batch 的处理次数,当选取越小次数(最小值为 1)时,其训练和识别效果越好. 但当训练次数较小时,神经网络每次(每个 batch)的训练时间会非常长,其耗时甚至是选择 20 个处理数量的数十倍,且识别效果又没有比每个 batch 处理 20 个数量的效果好. 为了使 SAE 网络具有稀疏性,且防止网络训练出现过拟合现象,加入“0”值.

为验证文中所提评价方法的有效性,实验选用 20 000 名学生的指标作为深度神经网络的训练数据;另外,选取 15 000 名学生的指标作为测试数据,并将其分为 30 组,用于测试网络的泛化能力,即评价结果的准确性. 500 名学生为一组的分组情况与一般大学一个年级的计算机专业的学生数量类似,其评价结果的正确率与运行时间更具有参考价值. 这些成绩数据由北方工业大学计算机学院和教学处提供,共 36 位副教授或教授参与评价.

每次随机选取 3 位专家组成一个评价小组,对每位学生进行综合能力评价. 当 3 位专家评价结果相同时,直接给出其中任意一位专家的评价成绩作为该生综合能力的最终评价结果;当只有 2 位专家评价结果相同时,采取少数服从多数的方法,在 2 位评价结果相同的专家中,选取其中任一评价成绩作为该生的最终评价结果;当 3 位专家的评价结果都不相同时,将 3 位专家的评价结果量化为 1~6 分,即评价 A 得 6 分,依次递减,评价 F 得 1 分,综合 3 位专家评价的结果,取其平均值的四舍五入值作为最后得分结果,将最后的得分结果转化为 A~F,即得到学生的最终评价结果. 评价时,删除学生的个人信息,如名字、学号等,且每个评价小组的专家独立评价,彼此不进行商讨.

在 30 次的实验中,每次的正确率、平均正确率及平均运行时间的统计结果,如表 4 所示.由表 4 可知:与深度置信网络评价方法、深度 BP 神经网络评价方法、深度 SAE 网络评价方法、K 最近邻评价方法相比较,文中方法的平均正确率最高,运行时间最长.

表 4 不同评价方法的统计结果  
Tab. 4 Statistical results of different evaluation methods

评价方法	正确率/%					平均 正确率/%	平均运行 时间/s
	1 #	2 #	3 #	...	30 #		
深度置信网络评价方法	82.9	82.2	81.8	...	83.1	82.4	274
深度 BP 神经网络评价方法	79.9	80.0	81.7	...	81.1	80.9	185
文中方法	86.8	84.5	85.6	...	86.4	85.8	287
深度 SAE 网络评价方法	83.7	81.1	82.5	...	84.2	82.7	263
K 最近邻评价方法	68.5	67.2	68.6	...	68.2	68.6	62

由表 4 还可知:文中方法在对不同组的学生的综合素质评价时,其正确率存在一些波动.产生波动的原因可能是:训练集数据是从所有收集到的数据中随机选取的,而数据集中每条数据的某一项分数之间的方差较小,提取 SAE 网络特征又不够充分,导致在对测试集数据进行评价时,由于输入数据的特性与训练集存在差异,使网络对于输入数据的泛化能力有所差异,从而导致正确率出现波动.在时间消耗上:1) 由于网络规模较大,神经网络需要多次迭代训练才能达到较为理想的结果;2) 由于在网络测试时,每次输入的数据量一样,因此,每次进行测试时,所消耗的时间基本相同,最多有±1 s 的差别,这可能是由一些程序外部的原因造成的,如 CPU 轮询、操作系统的调度等.

虽然,文中方法比基于深度置信网络的评价方法的平均运行时间多约 13 s,但平均正确率却提升了 3.4%.由此可见,该方法在评价平均正确率上的性价比高于深度置信网络.由于基于 BP 算法深度 SAE 神经网络的学生综合能力评价方法需要对输入的特征数据进行编码和解码,因此,消耗的时间比深度 BP 神经网络更多,但是平均正确率却比深层 BP 神经网络提高约 5%.在对测试时间不是很敏感的情况下,可通过多花费一些时间来提升分类结果.K 最近邻评价法是一种懒惰的学习方法,不需要训练集,因此,其消耗的时间主要用在计算测试集的距离上,在训练上消耗的时间几乎为 0.因此,K 最近邻评价法总的时间消耗最少.

目前,深度神经网络主要包括深度卷积神经网络、深度循环神经网络和深度置信网络.深度卷积神经网络常用于图像识别,在卷积核提取特征时,保留图像像素之间的二维空间位置关系.由于对学生综合能力评价时,采用的数据不存在数据间的空间位置关系,且由于卷积神经网络的计算量很大,会产生额外的时间浪费<sup>[16]</sup>,因此,卷积神经网络不可取.循环神经网络是一种时间结构参数共享的神经网络,网络的输出会保留在系统中.因此,可提取一些当前输入数据的上下文信息,如句子中、单词之间的语义上下文关系.但由于文中的输入数据,即一组学生成绩的各项得分之间的上下文关系并不是很密切,使用循环神经网络不但不能提取到上下文特征,反而会造成额外的计算开销,性价比较低.深度置信网络是限制玻尔兹曼机串联形成的深度神经网络,通过每个限制玻尔兹曼机自动提取特征,并可通过 BP 算法进行反向传播进行网络训练.深度置信网络的输入是一个一维数组,不需要考虑各组数据之间的空间关系及上下文关系,不会造成额外的时间开销,可较好地用于学生综合能力的评价.

3 结束语

提出基于 BP 算法的深度 SAE 网络的学生综合能力评价方法.SAE 采用无监督学习方式,Softmax 为监督式学习算法,构建的 SAE 模型结合二者的优点.利用 BP 算法,进行有监督微调,可进一步确定较优的网络模型.实验结果表明:该评价方法可以对学生的综合能力进行有效地评价.

参考文献:

[1] SAMUEL O W, ASOGBON G M, SANGAIAH A K, *et al.* An integrated decision support system based on ANN and Fuzzy\_AHP for heart failure risk prediction[J]. Expert Systems with Applications, 2017, 68(2017):163-172.

- DOI:10.1016/j.eswa.2016.10.020.
- [2] DENG Zhenyun,ZHU Xiaoshu,CHENG Debo,*et al.* Efficient  $k$ NN classification algorithm for big data[J]. *Neurocomputing*,2016,195(C):143-148. DOI:10.1016/j.neucom.2015.08.112.
- [3] 杨玮,岳婷,曹薇,等.采用BP神经网络的猪肉冷链物流预警模型与仿真[J]. *华侨大学学报(自然科学版)*,2015,36(5):511-516. DOI:10.11830/ISSN.1000-5013.2015.05.0511.
- [4] FARIAS G,DORMIDO-CANTO S,VEGA J,*et al.* Automatic feature extraction in large fusion databases by using deep learning approach[J]. *Fusion Engineering and Design*,2016,112:979-983. DOI:10.1016/j.fusengdes.2016.06.016.
- [5] WANG Yasi,YAO Hongxun,ZHAO Sicheng. Auto-encoder based dimensionality reduction[J]. *Neurocomputing*,2016,184(C):232-242. DOI:10.1016/j.neucom.2015.08.104.
- [6] JIANG Wenhao,GAO Hongchang,CHUNG Fulai,*et al.* The L2,1-norm stacked robust autoencoders for domain adaptation[C]//Thirtieth AAAI Conference on Artificial Intelligence. Phoenix; AAAI,2016:1723-1729.
- [7] CAO Lele,HUANG Wenbing,SUN Fuchun. Building feature space of extreme learning machine with sparse denoising stacked-autoencoder[J]. *Neurocomputing*,2016,174(Part A):60-71. DOI:10.1016/j.neucom.2015.02.096.
- [8] 尚敬文,王朝坤,辛欣,等.基于深度稀疏自动编码器的社区发现算法[J]. *软件学报*,2017,28(3):648-662. DOI:10.13328/j.cnki.jos.005165.
- [9] 李阳辉,谢明,易阳.基于降噪自动编码器及其改进模型的微博情感分析[J]. *计算机应用研究*,2017,34(2):373-377. DOI:10.3969/j.issn.1001-3695.2017.02.012.
- [10] OUYANG Yuanxin,LIU Wenqi,RONG Wenge,*et al.* Autoencoder-based collaborative filtering[C]//International Conference on Neural Information Processing. Kuching; Springer,2014:284-291. DOI:10.1007/978-3-319-12643-2\_35.
- [11] LABBACI H,MEDJAHED B,BINZAGR F,*et al.* A deep learning approach for web service interactions[C]//IEEE/WIC/ACM International Conference on Web Intelligence. Leipzig; IEEE Press,2017:848-854. DOI:10.1145/3106426.3106492.
- [12] SUK H I,LEE S W,SHEN Dinggang,*et al.* Latent feature representation with stacked auto-encoder for AD/MCI diagnosis[J]. *Brain Structure and Function*,2015,220(2):841-859. DOI:10.1007/s00429-013-0687-3.
- [13] LÜ Yisheng,DUAN Yanjie,KANG Wenwen,*et al.* Traffic flow prediction with big data: A deep learning approach[J]. *IEEE Transactions on Intelligent Transportation Systems*,2015,16(2):865-873. DOI:10.1109/TITS.2014.2345663.
- [14] KAMPER H,ELSNER M,JANSEN A,*et al.* Unsupervised neural network based feature extraction using weak top-down constraints[C]//IEEE International Conference on Acoustics, Speech and Signal Processing. South Brisbane; IEEE Press,2015:5818-5822. DOI:10.1109/ICASSP.2015.7179087.
- [15] MONTAVON G,ORR G,MLLER K R. Neural networks: Tricks of the trade[M]. Berlin; Springer,2012:437-478.
- [16] 吴琼,陈锻生.多尺度卷积循环神经网络的情感分类技术[J]. *华侨大学学报(自然科学版)*,2017,36(6):875-879. DOI:10.11830/ISSN.1000-5013.201606077.

(编辑:李宝川 责任编辑:黄晓楠 英文审校:吴逢铁)