

# The Improved Neural Network Model about English College Examination Anxiety

Ma-hong NiuYaqing Li Yamin

Agriculture University of Hebei, Baoding, China

**Abstract**—A predicted grade method based on the relationship between precompetitive state anxiety and the predicted grade is presented. Then Genetic algorithm is combined in order to improve traditional model. On the basis of the qualitative and quantitative study, the future achievements have been predicted.

**Key words**—neural network, Test anxiety, English test

## 高校英语考试焦虑与成绩 BP 模型构建及分析

马宏 牛亚卿 李亚民

河北农业大学，河北保定，中国，071000

**摘 要** 将改进的 BP 神经网络应用于考试焦虑与英语成绩之间的关系描述。结合传统神经网络的不足，引入遗传算法对其进行优化，在此基础上预测受试者的英语水平考试成绩。

**关键词** 神经网络，考试焦虑，英语考试

### 1. 引言

大量研究证实，高校学生英语考试表现和两个方面有关，一是受试者的个人特性，二是考试试题的难度和范畴，但是考生个人性格特质方面的研究并不多见。本研究聚焦于高校英语考生的语水平考试焦虑，有文献其定义为 EPTA。EPTA 概念为新近提出，围绕此概念已经定义出一些模型，但这些模型均为线性模型，难以体现出英语水平与考试焦虑之间的“非线性”关系。本研究将改进的 BP 神经网络应用于 EPTA 与英语水平考试成绩之间的关系描述。结合传统神经网络收敛过慢以及易陷入局部最小点等不足，引入遗传算法对其进行优化，使其学习速率显著提升，网络收敛性能明显改善。在此基础上预测受试者的英语水平考试成绩。

### 2. 神经网络预测模型

#### 2.1 数据的采集

随机抽取 20 名高校非英语专业大三学生。其中男生 10 名、女生 10 名。在 30min 内，被试填写完 EPTAS 量表。归一化后的数据如下表所示：

表 1：归一化后的学习样本

样本	EPTAS 听力	EPTAS 口语	EPTAS 阅读	EPTAS 写作	CET-4 成绩
1	0.0025	0.0023	0.0031	0.0012	0.83
2	0.0024	0.0022	0.0031	0.0012	0.92
3	0.00111	0.0028	0.0038	0.0032	0.96
4	0.0035	0.004	0.0035	0.0029	0.99
5	0.00111	0.00111	0.0011	0.00111	0.79
6	0.0026	0.0022	0.0027	0.00111	0.93
7	0.0028	0.003	0.0031	0.0033	0.91
8	0.0047	0.005	0.0047	0.0052	1.01
9	0.003	0.0027	0.0022	0.0026	0.84
10	0.0028	0.0022	0.0025	0.0026	0.86

#### 2.2 神经网络优化

遗传算法对传统的神经网络进行优化，使之能够适应对学生成绩的预测。这种优化算法通过遗传变异来使神经网络初始权重得到改善，算法的实质通过两个步骤对设计网络进行训练：第一步是引入遗传算法，对 BP 网络的权重进行优化；第二步则结合 BP 神经网络实现网络训练的目标。

具体的步骤为:

(1)引入合理的编码模式对系统的阈值编码,通过该编码能够产生随机的分布,而这些分布则和BP网络的阈值一一对应。

(2)结合选定的训练样本集,对样本的误差值进行估算,定义系统的适应度为结合误差所求出的平方和。如果误差足够小,可以认为所得到的适应度较大,否则便可判定适应度比较小。通过适应度的大小来预测阈值是否合适。

(3)从样本集中选取适应度足够大的个体,将这些个体遗传到下一代个体中。

(4)以交叉和变异运算来处理此群体,从而使之能够构建下一代的群体。

(5)对以上的第二步至第四步进行重复,便能够得到一组阈值,通过不断的进化,一直到得到了训练目标符合本研究的条件,算法结束。

通过以上的流程,便能够产生一组合理的编码信息,而基于遗传算法的原理则可知这组编码信息即为所求的最优解。再基于遗传算法对其进行解码操作,最终把这组初始全值进行转换,使之成为由权重构成的集合,最终成为神经网络的阈值,注入到神经网络。

## 2.3 神经网络预测

### (1)神经网络参数的确定

#### ①网络隐含层数的选取

结合神经网络的特性,本研究将网络层次定为3层,将对学生成绩造成影响的因素数目当做神经网络神经元个数,因此神经网络输入层为4个节点;神经网络的输出是学生成绩的成绩值,因为学生成绩仅包括一列,因此神经网络输出层为1个节点。

#### ②确定隐含层的神经元个数

隐含层的神经元个数是十分重要的,本层所含有的神经元的个数能够对神经网络非线性产生比较大的影响,并和需要解决问题的复杂度相关,而一个问题其具体的解决复杂度是难以进行量化的,所以目前并不存在一个标准的算法得出最准确的隐含层所含单元数,如果设置了比较少的隐层节点,则神经网络难以获取足够的信息数据,不能实现网络训练,如果设置了比较多的隐层节点,则神经网络有可能变得过拟合,影响性能。

本研究通过“黄金分割”确定神经网络的隐含层的具体节点数。此模式和传统的“对分法”比较,能够提升搜索速度,缩短搜索时间,算法复杂度较低。当前已经有不少研究能够证实黄金分割法在寻优方面的先进性。

此处设BP网络输入层所拥有的节点数量是I个,输出层为J个。

神经网络的输入层,选取5个因子,输出层则为学生成绩的具体成绩,所以有 $I=4, J=1$ 。可知隐含层节点区间是介于2到15之间。

#### 试验步骤1:

a. 寻找首个试验点位置。在[2, 15]上取 $g_1=0.618 \times (15-2)+2=10$ ,同时标记首个点的具体结果,以 $E(g_1)$ 表示均方误差MSE的实际值;BP网络结构为:4-10-1,记为BP1。在95此循环操作中后出现收敛,因此得到 $E(g_1)$ 的值为0.00262220。

b. 标记第2个点的具体结果,以 $E(g_2)$ 表示均方误差MSE的实际值;BP网络结构为:4-7-1,记为BP2。在63次循环操作中后出现收敛,因此得到 $E(g_2)$ 的值为 $E(g_2)=0.00239930$ 。

由于,所以在2个模型之间,显然BP2性能更好。所以将新的实验区间改为[2, 10]。

#### 试验步骤2:

a. 结合所确定的新区间[2, 10],寻找试验点3,有,由此可推知神经网络是4-7-1结构。网络循环之后收敛。

b. 确定第四个试验点的位置,推知神经网络是4-5-1结构,网络循环之后收敛。

因为,在对这两个试验点比较之后可知BP3性能更佳。所以保留区间[a,  $g_3$ ],将实验区间改为[2, 7]。

#### 试验步骤3:

a. 结合所确定的新区间[2, 7]寻找试验点5,有, BP网络结构为:4-5-1,  $E(g_5)=0.0024002$ 。

b. 确定第六个试验点的位置,有,推知神经网络是4-4-1结构,网络循环之后收敛。

因为,在对这两个试验点比较之后可知BP3性能更佳。所以保留区间[ $g_6$ , b],将实验区间改为[4, 7]。

#### 试验步骤4:

a. 结合所确定的新区间[4, 7]寻找试验点7,有,推知神经网络是4-6-1结构,网络循环之后收敛。

b. 推知神经网络是4-5-1结构,因为,在对这两个试验点比较之后可知BP3性能更佳。所以保留区间[ $g_7$ , b],将实验区间改为[4, 6]。

结合以上的分析,可知 $E(g_5)$ 均方差最低。

在寻优性能上,表现最好的是BP3模型,BP3收敛前的循环次数是88次,所得到的 $MSE=0.0025001$ ,故将本研究的学生成绩神经网络结构的隐含层节点定位于五节点结构,即形成4-5-1的三层神经网络结构。

### ③确定学习速率

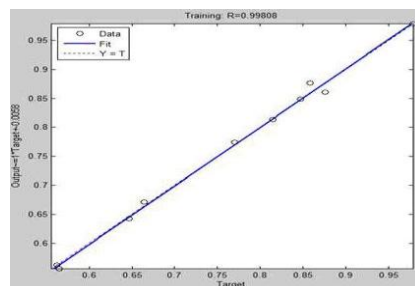
为了保证神经网络不出现振荡,学习速率的选取应寻找较小的值,一般情况下学习速率的区间是[0.01, 0.9]。为降低学习速率确定所需的训练开销,本研究最终选取了基于自适应调整的速率,因此BP网络能够结合训练的具体情况来自动配置学习速率。

归一化后的测试样本如下表:

表 2.归一化后的测试样本

样本	EPTAS 听力	EPTAS 口语	EPTAS 阅读	EPTAS 写作	CET-4 成绩
1	0.0035	0.0032	0.0033	0.0034	0.93
2	0.0021	0.0023	0.0025	0.0029	0.92
3	0.0022	0.0025	0.00111	0.00111	0.88
4	0.003	0.0031	0.0028	0.0026	0.99
5	0.0038	0.0039	0.0036	0.0031	0.81
6	0.0032	0.0036	0.0036	0.0037	0.88
7	0.0048	0.0037	0.004	0.0033	0.96
8	0.0011	0.003	0.00111	0.00111	0.99
9	0.00111	0.00111	0.00111	0.00135	0.89
10	0.0035	0.0032	0.0033	0.0034	0.93

在神经网络训练完成的基础上，对其拟合度进行检验，方法是结合训练样本目标向量和 BP 网络之间的拟合值，样点的分布如果趋近于 1:1，则可以证实模型的估算值与实地测量值接近，说明预测效果较为满意。在 1:1 线的附近集中着绝大多数样本。学生成绩实测数据和拟合数据之间的相关值是 0.953，均方根误差是 4.24。可知本研究所构建的改进神经网络一方面能够实现较好的建模精度，另一方面也拥有满意的检验精度，能够实现较为准确的预测。测试数据拟合曲线图如下：



图：测试数据拟合曲线图

可知预测数据能精确的逼近实际值。

### 3. 结束语

本研究把改进的神经网络模型应用在高校学生英语水平考试成绩的预测上。应该强调的是，EPTA 仅仅是预测高校学生英语考试成绩的参数之一，除了 EPTA，还有一些重要的指标也能够对英语考试成绩造成显著影响，例如语言能力、自尊、成就动机等，在本研究的基础上，可以将以上参数作为 BP 网络更多的输入，从而使预测的精度更加准确。

### 参考文献(References)

- [1] Yanlai Li; Kuanquan Wang; Hua Zhang. Step acceleration based training algorithm for feedforward neural networks. Pattern Recognition. 2012. Proceedings 16th International Conference on. 2012. Page(s):84 -87 Vol.2.
- [2] Riedmiller M., Braun H. A direct adaptive method for faster backpropagation learning: the RPROP algorithm. Neural Networks. 1993. IEEE International Conference on. 2013. Page (s): 586 -591 vol.1