



论文检测报告

报告编号：78AF9FF1800A4A6EBECF527385943A9C

送检文档：BP神经网络课设报告

论文作者：郭立程

文档字数：12430

检测时间：2019-02-14 15:50:01

检测范围：互联网，中文期刊库（涵盖中国期刊论文网络数据库、中文科技期刊数据库、中文重要学术期刊库、中国重要社科期刊库、中国重要文科期刊库、中国中文报刊报纸数据库等），学位论文库（涵盖中国学位论文数据库、中国优秀硕博论文数据库、部分高校特色论文库、重要外文期刊数据库如Emerald、HeinOnline、JSTOR等）。

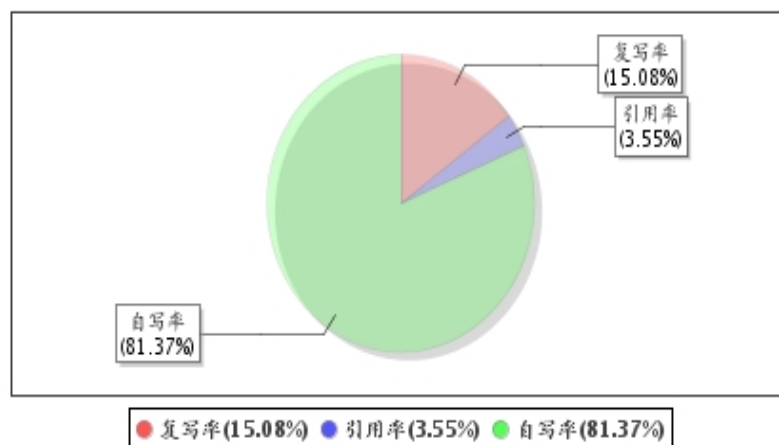
一、检测结果：

总相似比：18.63% [即复写率与引用率之和]

检测指标：自写率 81.37% 复写率 15.08% 引用率 3.55%

相似比：互联网 12.91% 学术期刊 0.95% 学位论文 1.22%

其他指标：表格 0 个 脚注 0 个 尾注 0 个



章节抄袭比

0.00% 课程编号：0521172B课程性质：必修

12.50% 本设计题为BP神经网络的实现，主要任务是

0.00% 主要有这些：

12.50% 设定好学习样本个数，取样本个数的训练数据

50.00% 开发工具：Visual Studio 2

0.00% 输入初始参数：

0.00% 五、总结：



12.50% 首先是对隐藏层节点数目的考虑，过多或者过

二、相似文献汇总：

序号	标题	文献来源	作者	出处	发表时间
1	最全解析:大数据和机器学习有什么区别-51CTO.COM	互联网		互联网	
2	BP算法在IDS中应用与实现的研究	学位论文	甘刚	硕博学位论文	2005
3	C++实现的简单BP神经网络-未命名专栏-CSDN博客	互联网		互联网	
4	基于BP神经网络的电力系统短期负荷预测-手机知网	互联网		互联网	
5	请问神经网络如何更改输入层到隐藏层的权值-CSDN问答	互联网		互联网	
6	百度快照	互联网		互联网	
7	基于BP神经网络的变压器故障诊断	学术期刊	朱浪 王蕾 潘丰	江南大学学报 ：自然科学版	2012
8	BP原理与实现-CSDN博客	互联网		互联网	

三、全文相似详情：

（红色字体为相似片段、浅蓝色字体为引用片段、深蓝色字体为可能遗漏的但被系统识别到与参考文献列表对应的引用片段、黑色字体为自写片段）



课程编号：0521172B课程性质：必修

数据结构课程设计报告

2019年1月20日至2019年2月15日

一、实验概述：1.1课程设计题目：

本设计题为BP神经网络的实现，主要任务是利用C++语言实现BP神经网络，并利用BP神经网络解决蠕虫分类问题。1.2课程设计目的：

主要有这些：

- 1、熟悉自己所学的多种数据结构.
 - 2、理解BP神经网络的工作原理
 - 3、利用C++实现BP神经网络
 - 4、利用BP神经网络实现蠕虫分类
- 1.3系统主要内容与功能：1.3.1设计内容：

设定好学习样本个数，取样本个数的训练数据存储于文本文件中，在C++ main函数中设定学习效率和步长，读取训练输入和对应的输出进行训练，通过训练得到对应的权值，然后进行数据测试。

1.3.2设计功能：

本课程设计由3个主要功能实现，它们分别为初始化数据，训练和测试蠕虫分类是否准确。

初始化数据：

```
srand((unsigned)time(NULL));  
for(inti=0;i  HN;i++)  
{  
for(intj=0;j  INnum;j++)  
W[i][j]=double(rand()&nbsp;%100)/100;//初始化输入层到隐层的权值  
}  
for(intii=0;ii  ONnum;ii++)
```



```
for(intjj=0;jj  HN;jj++)
```

```
V[ii][jj]=double(rand()&nbsp;%100)/100;//初始化隐层到输出层的权值
```

这段代码是为了初始化输入层到隐层的权值，以及隐层到输出层的权值。

input_sample(int m)：这段代码是为了初始化训练数据的输入值。

output_sample(int m)：这个函数是为了初始化数据对应的输出值

训练：

本功能主要是逐步训练，直至收敛。

calculate_HIDDEN_OUTPUT() :本函数主要用于求隐层输出。

calculate_OUTPUT_OUTPUT() : 本函数主要用于求输出层输出。

inaccuracy_Output_Hidden(int m);本函数主要用于求解输出层到隐层的一般化误差

inaccuracy_Hidden_Input() : 本函数主要用于求解隐层到输入层的一般化误差。

update_Output_Hidden：本函数主要是用于调整输出层至隐层的权值。

update_Hidden_Input：本函数主要是用于调整隐层至输入层的权值。saveWeight() : 本函数主要是用于保存计算出来的权值。

测试分类是否准确：

数据结构:

```
struct
```

```
doubleinput[INnum];
```

```
doubleteach[ONnum];
```

```
}Study_Data[SampleCount];
```

代表了训练样本；

```
double Weight_Input[NUM_HIDDEN][mInputNode];
```

代表了输入层至隐藏层权值；



```
double Weight_Output[mOutPutNode][NUM_HIDDEN];  
代表了隐层至输出层权值；  
double Sample_Input[mInputNode];  
代表了单个样本输入数据；  
double Sample_Output[mOutPutNode];  
代表了单个样本期望输出值；  
double ORIGIN_W[NUM_HIDDEN][mInputNode];  
代表了隐藏层到输入层的旧权数组；  
double ORIGIN_V[mOutPutNode][NUM_HIDDEN];  
代表了输出层到隐藏层的旧权数组；  
double HIDDEN_INPUT[NUM_HIDDEN];  
代表了隐层的输入；  
double OUTPUT_INPUT[mOutPutNode];  
代表了输出层的输入；  
double hiddenLayerOutput[NUM_HIDDEN];  
代表了隐层的输出；  
double OUTPUT_OUTPUT[mOutPutNode];  
代表了输出层的输出；  
double inaccuracy_sample[SampleCount];  
代表了第m个样本的总误差。1.3.3实验环境与工具：操作系统：Windows企业版2016长期服务版；  
开发工具：Visual Studio 2017社区版；  
开发语言：C++。
```



二、实验原理：

2.1结构构成：

BP神经网络是一种根据误差逆向传播的多层前馈网络。其由输入层，隐藏层和输出层构成的。它模拟了人脑的神经网络的结构，众所周知，人大脑传递信息的基本单位是神经元，人脑中有大量的神经元，每个神经元与多个神经元相连接，就是这样，构成了人的大脑。人的大脑的神经网络传播是按照一定的规律，不是神经元会对所有传过来的刺激进行反应。首先是积累刺激，而BP神经网络，类似于上述，是一种简化的生物模型。每层神经网络都是由神经元构成的，单独的每个神经元相当于一个感知器。输入层是单层结构的，输出层也是单层结构的，而隐藏层可以有多层，也可以是单层的。输入层、隐藏层、输出层之间的神经元都是相互连接的，为全连接。总得来说，BP神经网络结构就是，输入层得到刺激后，会把他传给隐藏层，至于隐藏层，则会根据神经元相互联系的权重并根据规则把这个刺激传给输出层，输出层对比结果，如果不对，则返回进行调整神经元相互联系的权值。这样就可以进行训练，最终学会，这就是BP神经网络模型。

2.2算法流程：

先把统一处理过的训练样本输入到BP神经网络。

设定步长，学习效率，最大学习次数，生成初始权值，阈值。

计算从输入层开始到输出层结束的各层的输入输出值。

计算输出层的误差，如果误差小于设定误差，便结束。如果大于设定误差，便调整权值和阈值，执行步骤。

2.3数学原理：

图1 三层BP神经网络示意图

在三层BP神经网络中，

输入向量,也就是输入层神经元为：

(1)

隐藏层输入向量，也就是隐藏层神经元：

(2)

输出层输出向量，也就是输出层神经元：

(3)



期望输出向量可以表示为：

(4)

输入层到隐藏层之间的权值用数学向量可以表示为

(5)

这里的列向量为隐藏层第j个神经元对应的权重；隐藏层到输出层之间的权值用数学向量可以表示为：

(6)

式(6)中的列向量为输出层第k个神经元对应的权重。

输入层，隐藏层，输出层之间的数学关系如下所示：

就输出层而言

就隐藏层而言

在上述表达式中，激活函数 $f(x)$ 必须为单调函数，我们选取sigmoid函数：(11)

由于 $f(x)$ 是连续函数，是可导函数，且

这些就构成了BP神经网络的输入层，输出层和隐藏层。

如果BP网络的输出结果和期望输出结果不一致时，那么便有了输出误差E，表达式如下：

(12)

展开到隐藏层，则有如下表达式：

(11)

展开到输入层，则有以下表达式：

(13)

根据这些，我们可以明白，BP神经网络的输入误差产生是由于输出层到隐藏层的权值和隐藏层到输入层的权值决定，因此调整输出层到隐藏层的权值和隐藏层到输入层的权值可以改变误差E。

显而易见的，调整输出层到隐藏层的权值和隐藏层到输入层的权值的思想是让这个误差不断缩小这样才能满足我们的要求，因此我们需要让这些权值修改量与误差E的



负梯度下降量成正比，也就是如下表达式：

式（14）、（15）中常数反映了学习效率。

综上所述，通过数学原理可以看出，在BP学习算法中，输入层，输出层和隐藏层权值调整方式都是一样的。由三个条件决定，它们分别是：学习效率，输入层，输出层和隐藏层的各层误差信号和各层输入信号 X/Y 。其中最为重要的是输出层误差信号，它直接意味着和实际期望结果的差异，代表着与我们预期结果的差距，而前面的各层的误差都是从后往前传递计算的误差。

三、实验结果：

输入初始参数:

图2 初始化参数

正在训练:

图3 程序训练运行时

训练成功：

图4 程序训练完成

测试数据：

图5 测试算法训练结果

值得一提的是，本次课程设计仍然存在一定的不足。本次课程设计中采用的学习效率是一个定值，没有发生过改变。这样的话，在某些地方用函数逼近的时候出现了一些误差。所以，如果能够将这个学习效率能够自动调整，在一开始是一个数值比较大的学习效率，这样可以更快速地接近理想函数，等到快接近理想函数的时候，这个时候的话采用较小的学习效率，这样可以使最后的函数更加准确，这样的话训练时间会缩短，并且训练的效果会更佳。

实验一开始，采取的参数不是很合适，采取了较大的学习效率，这样导致训练失败了，一开始采取的训练次数也较低，最后的结果也是训练失败，因此，学习效率和步长都不可以太大，如果太大，波动也大了，这样的话，发散的概率提高了太多。学习效率和步长也不可以过小，如果过小，训练时间会很长，影响性能。然后预定误差对实验的进行影响也比较大，如果预定误差值过小，则实验最后的效果会出现较大的偏差，如果预定误差值较大，则训练次数和时间会过大。

所以在本次实验中，一定要把握好参数的设定，否则结果会很差。

四、算法分析：



五、总结：

首先是对隐藏层节点数目的考虑，过多或者过少都不能科学地反映出样本的规律性。如果数量太少，则不能够将训练数据的规律性概括出来。如果数量过多，也不是很好，因为可能会把一些干扰性的并不是规律的内容给记录下来，这样会使得过于准确导致某些数据还是判断失误，而且会使得训练时间变长很多。比较妥当的一个做法就是根据公式计算出一个初始的实验值。

至于究竟是选取每输入一个样本就进行权值调整,然后计算误差,或者是输入完所有样本数据之后再进行权值调整，根据本次实验，我得出的结论是如果样本数量多的话，输入完所有样本数据之后再进行权值调整的话能更快得进行收敛。

这次课程设在VC++6.0上完成了BP神经网络实现的编程，明白了BP神经网络计算过程分两步，分别是正向计算，反向计算。对于正向传播而言，计算从输入层开始到输出层结束的各层的输入输出值，对于反向传播，便是如果正向计算的输出值不满足预期，便从后往前调整权值，直至最后的数据符合预期要求。

通过本次实验，对BP神经网络理解变得非常透彻，熟悉了BP神经网络的结构，对人工智能的应用有了一定的认识，也对数据结构的学习印象更深刻，提高了我学习相关知识的兴趣，期间遇到了很多困难，而且编程的实现对我而言比较繁琐，遇到了很多BUG，但是都一一细心解决了。

通过熟悉其数学原理，我深刻理解了BP神经网络已经非常成熟。它的优点是可以解决非线性映射问题。它的网络结构也非常灵活。输入层，输出层和隐藏层的神经元数量可以根据具体的需要而设定。但是它的缺点也存在，比如说它的学习速度比较慢，就本课程设计而言，学了较大次数才能收敛，而且对于初始参数的设定而言，得不到很多的理论指导。

附录（源代码）：

头文件necessary.h：

四、指标说明：

1. 总相似比即类似于重合率。总相似比即送检论文中与检测范围所有文献相似的部分（包括参考引用部分）占整个送检论文的比重，总相似比=复写率+引用率。
2. 引用率即送检论文中被系统识别为引用的部分占整个送检论文的比重（引用部分一般指正确标示引用的部分）。
3. 自写率即送检论文中剔除雷同片段和引用片段后占整个送检论文的比重，一般可用于论文的原创性和新颖性评价，自写率=1-复写率-引用率。



4. 复写率即送检论文中与检测范围所有文献相似的部分（不包括参考引用部分）占整个送检论文的比重。

5. 红色字体代表相似片段；浅蓝色字体代表引用片段、深蓝色字体代表可能遗漏的但被系统识别到与参考文献列表对应的引用片段；黑色字体代表自写片段。

五、免责声明：

鉴于论文检测技术的局限性以及论文检测样本库的局限性，格子免费检测系统不保证检测报告的绝对准确，相关结论仅供参考，不做法律依据。

格子免费检测系统服务中使用的论文样本，除特别声明者外，其著作权归各自权利人享有。根据中华人民共和国著作权法相关规定，格子免费检测系统为学习研究、介绍、评论、教学、科研等目的引用其论文片段属于合理使用。除非经原作者许可，请勿超出合理使用范围使用其内容和本网提供的检测报告。