

I'm a DMer & Mler & NlPer & IRer

博文有原创，有转载。均为分享学习之用。P.S. 部分转载可能未注明出处。若有冒昧，请与我联系:(ILoveDataMining AT gmail DOT com)

导航

博客园

首页

订阅 XML

管理

统计信息

随笔 - 186

文章 - 0

评论 - 54

Trackbacks - 0

NEWS

昵称: wentingtu

园龄: 4年5个月

粉丝: 322

关注: 1

+ 加关注

搜索

随笔分类

A-memo

Big-Data Mining (Hadoop)(7)

Book Reading Notes

Business Intelligence(2)

Code of My toolkit

Computational Advertising(7)

Data Analysis(5)

Data Mining(33)

Dataset, Dictionary,

CorpusLink(1)

English(2)

Experiments

Financial

HKU DB seminar

Idea Depository

Jobs Preparation

Learning to Rank(1)

Life(4)

Love, Friendship and more

Machine Learning(33)

Marketing

Mobile Internet(1)

News/Event Detection

Optimization

Original(1)

Paper/Slides Reading Notes(4)

Programming(Python,Java)(14)

Programming(R,SAS,MATLAB)

(41)

Project

Recommendation System(30)

Semantic web

Social network(1)

SPARK

StuDY(15)

Text Mining & NLP(15)

Working Diary

随笔档案

2015年6月 (1)

2014年12月 (2)

2014年6月 (2)

BP神经网络模型与学习算法

一，什么是BP

"BP (Back Propagation) 网络是1986年由Rumelhart和McClelland为首的科学家小组提出，是一种按误差逆传播算法训练的多层前馈网络，是目前应用最广泛的神经网络模型之一。BP网络能学习和存贮大量的输入-输出模式映射关系，而无需事前揭示描述这种映射关系的数学方程。它的学习规则是使用最速下降法，通过反向传播来不断调整网络的权值和阈值，使网络的误差平方和最小。BP神经网络模型拓扑结构包括输入层(input)、隐层(hidden layer)和输出层(output layer)。"

我们现在来分析下这些话：

- "是一种按误差逆传播算法训练的多层前馈网络"

BP是后向传播的英文缩写，那么传播对象是什么？传播的目的是什么？传播的方式是后向，可这又是什么意思呢。

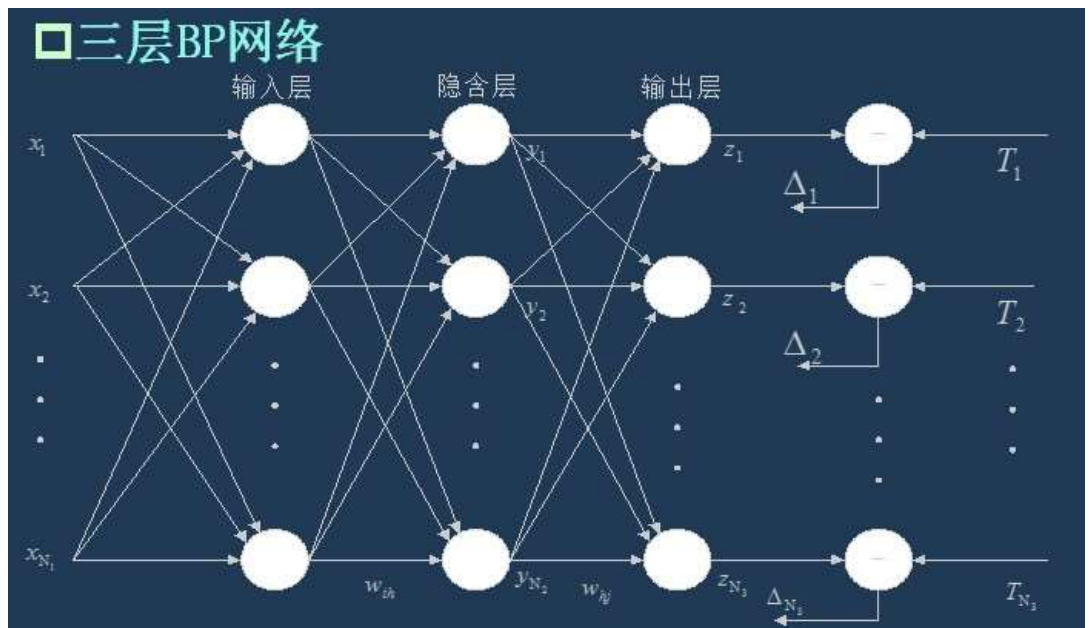
传播的对象是误差，传播的目的是得到所有层的估计误差，后向是说由后层误差推导前层误差：

即BP的思想可以总结为

利用输出后的误差来估计输出层的直接前导层的误差，再用这个误差估计更前一层的误差，如此一层一层的反传下去，就获得了所有其他各层的误差估计。

- "BP神经网络模型拓扑结构包括输入层(input)、隐层(hidden layer)和输出层(output layer)"

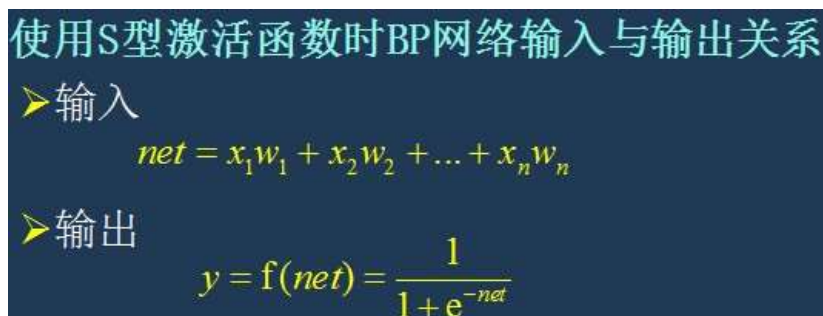
我们来看一个最简单的三层BP：



- "BP网络能学习和存贮大量的输入-输出模式映射关系，而无需事前揭示描述这种映射关系的数学方程。"

BP利用一种称为激活函数来描述层与层输出之间的关系，从而模拟各层神经元之间的交互反应。

激活函数必须满足处处可导的条件。那么比较常用的是一种称为S型函数的激活函数：



那么上面的函数为什么称为是S型函数呢：

2014年5月 (1)
 2014年3月 (2)
 2014年1月 (1)
 2013年7月 (1)
 2013年6月 (1)
 2013年5月 (1)
 2013年2月 (1)
 2013年1月 (5)
 2012年12月 (3)
 2012年11月 (2)
 2012年9月 (2)
 2012年7月 (1)
 2012年6月 (11)
 2012年5月 (21)
 2012年4月 (19)
 2012年3月 (33)
 2012年2月 (7)
 2012年1月 (5)
 2011年12月 (61)
 2011年10月 (1)
 2011年9月 (2)

ABOUT ME

Homepage
 Micro-Blog (Sina)

BLOG

A Link Collection

STUDY

Course
 朝花夕拾

最新评论

1. Re:非常好的协同过滤入门文章
 好文章

--xx ee

2. Re:R语言基础入门
 关注

--junjunang

3. Re:决策树模型组合之（在线）随机森林与GBDT
 源码微盘失效了。。。同求下载！

--Lhfcws

4. Re:推荐中的GRAPH MODEL
 不错，介绍精辟，到位

--samjustin

5. Re:BP神经网络模型与学习算法
 @gystld请问你有正版的PEW软件吗。急需，谢谢！
 1377069951@qq.com...

--叶孤傲

6. Re:机器学习算法复习-随机森林求大神指导！

--sfqy

7. Re:BP神经网络模型与学习算法
 神经网络源码：

--cofftech

8. Re:决策树模型组合之（在线）随机森林与GBDT
 不错，可是微盘分享的资源不在了呢？

--DebugMyself

9. Re:Learning to Rank入门小结 + 漫谈
 部分图片不显示

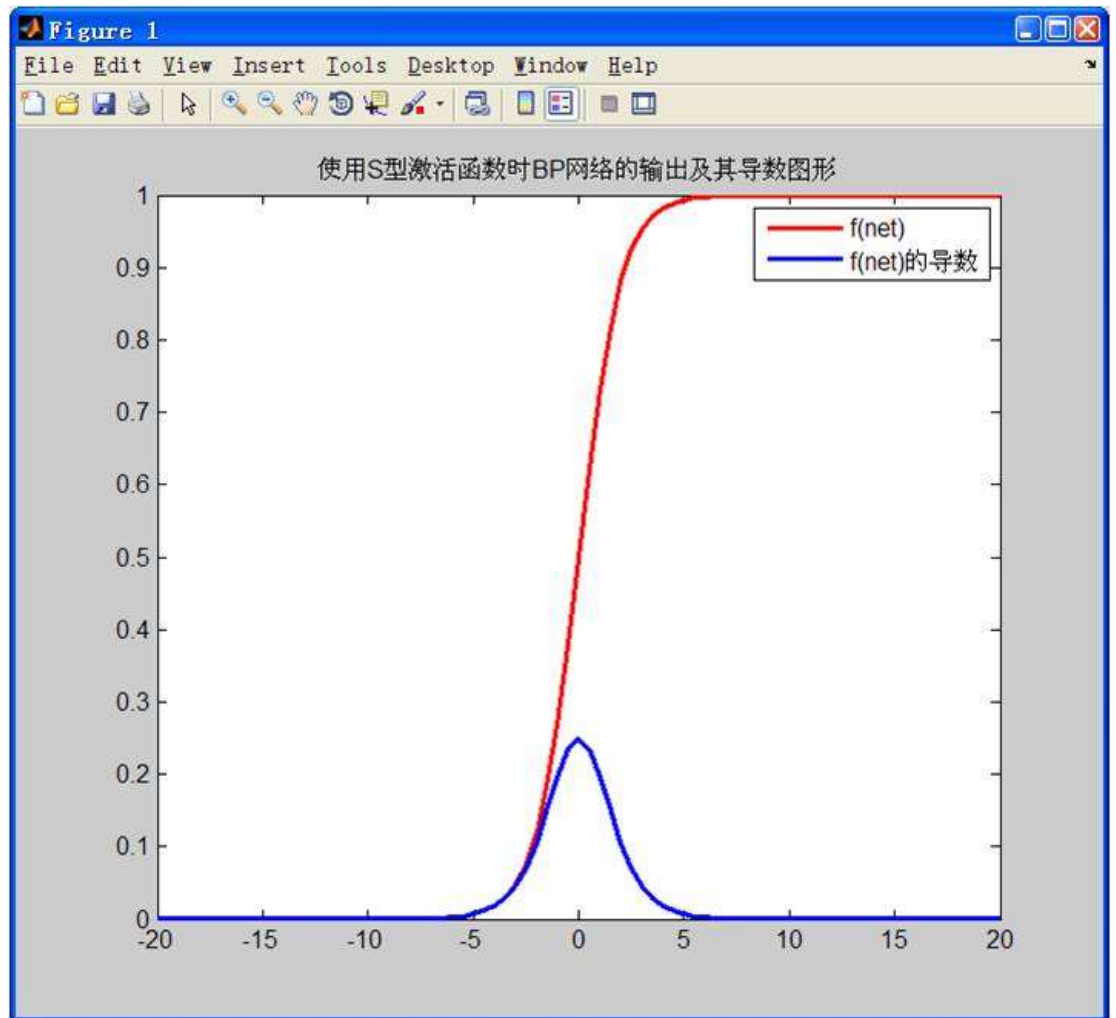
--Rice1202

10. Re:BP神经网络模型与学习算法
 推荐一款最易用中文excel神经网络软

我们来看它的形态和它导数的形态：

p.s. S型函数的导数：

$$f'(net) = \frac{1}{1+e^{-net}} - \frac{1}{(1+e^{-net})^2} = y(1-y)$$



神经网络的学习目的：

希望能够学习到一个模型，能够对输入输出一个我们期望的输出。

学习的方式：

在外界输入样本的刺激下不断改变网络的连接权值

学习的本质：

对各连接权值的动态调整

学习的核心：

权值调整规则，即在学习过程中网络中各神经元的连接权变化所依据的一定的调整规则。

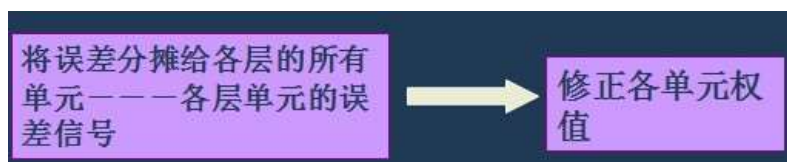
二，有监督的BP模型训练过程

1. 思想

有监督的BP模型训练表示我们有一个训练集，它包括了：input X 和它被期望拥有的输出 output Y

所以对于当前的一个BP模型，我们能够获得它针对于训练集的误差

所以BP的核心思想就是：将输出误差以某种形式通过隐层向输入层逐层反传，这里的某种形式其实就是：



也就是一种 "信号的正向传播 ----> 误差的反向传播" 的过程：

件包，有BP、RBF、PNN、GRNN、Elman、SOM、LVQ七种神经网络功能。一键操作，易用易懂。。...

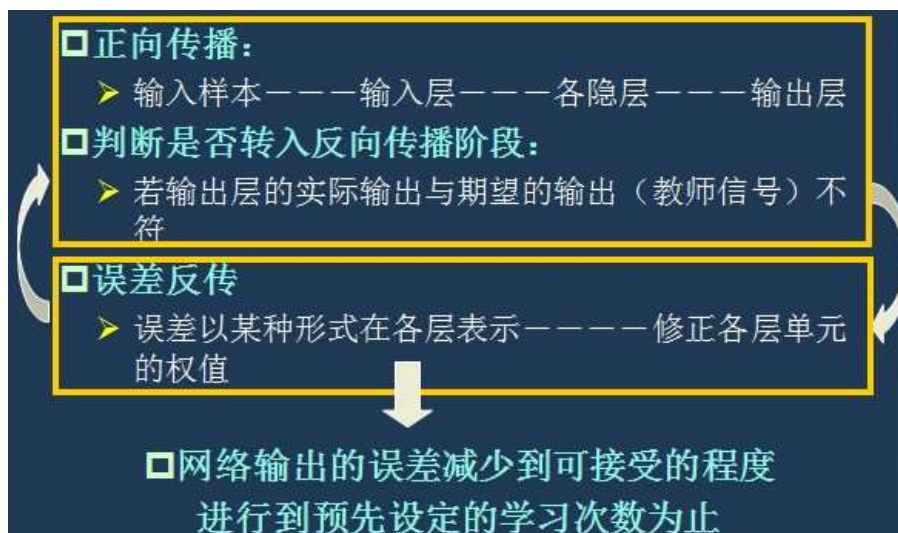
-gystld

阅读排行榜

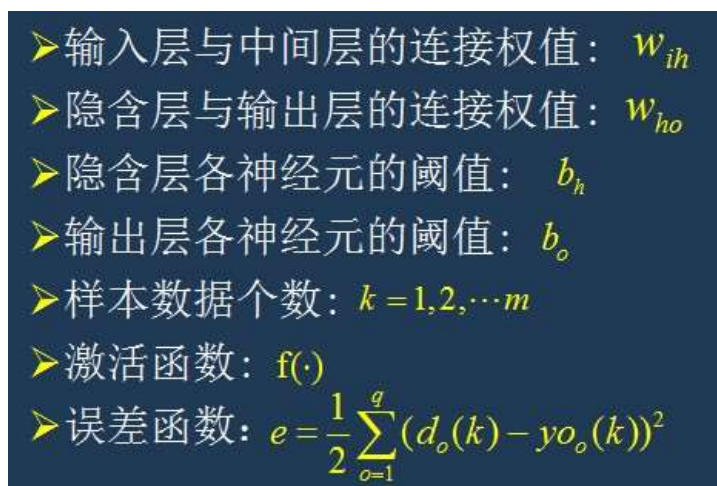
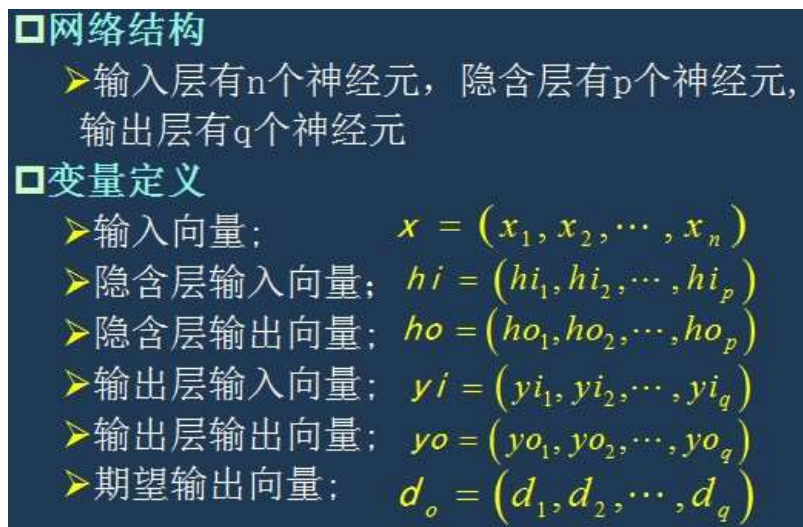
1. BP神经网络模型与学习算法 (61588)
2. R语言基础入门(41434)
3. 机器学习算法复习-随机森林 (32165)
4. python数据挖掘领域工具包 (16466)
5. R与矩阵运算总结(13256)
6. R语言多元分析系列(12805)
7. Learning to Rank入门小结 + 漫谈(11483)
8. RapidMiner数据挖掘入门(7959)
9. 决策树模型组合之（在线）随机森林与GBDT(7618)
10. R语言编程入门(7416)

推荐排行榜

1. 非常好的协同过滤入门文章(7)
2. BP神经网络模型与学习算法(6)
3. 推荐系统常用数据集(5)
4. python数据挖掘领域工具包(3)
5. R语言基础入门(3)
6. ICML 2012 推荐系统部分文章小结及下载(3)
7. 总结一下网上可以得到的R语言资源(2)
8. 机器学习算法复习-随机森林(2)
9. R语言多元分析系列(2)
10. 简单的社交网络分析（基于R）(2)



2.具体



□ 第一步，网络初始化

➤ 给各连接权值分别赋一个区间 $(-1, 1)$ 内的随机数，设定误差函数 e ，给定计算精度值 ε 和最大学习次数 M 。

□ 第二步，随机选取第 k 个输入样本及对应期望输出

$$x(k) = (x_1(k), x_2(k), \dots, x_n(k))$$

$$d_o(k) = (d_1(k), d_2(k), \dots, d_q(k))$$

□ 第三步，计算隐含层各神经元的输入和输出

$$hi_h(k) = \sum_{i=1}^n w_{ih} x_i(k) - b_h \quad h = 1, 2, \dots, p$$

$$ho_h(k) = f(hi_h(k)) \quad h = 1, 2, \dots, p$$

$$yi_o(k) = \sum_{h=1}^p w_{ho} ho_h(k) - b_o \quad o = 1, 2, \dots, q$$

$$yo_o(k) = f(yi_o(k)) \quad o = 1, 2, \dots, q$$

□ 第四步，利用网络期望输出和实际输出，计算误差函数对输出层的各神经元的偏导数 $\delta_o(k)$

$$\frac{\partial e}{\partial w_{ho}} = \frac{\partial e}{\partial yi_o} \frac{\partial yi_o}{\partial w_{ho}} \quad \frac{\partial yi_o(k)}{\partial w_{ho}} = \frac{\partial (\sum_h w_{ho} ho_h(k) - b_o)}{\partial w_{ho}} = ho_h(k)$$

$$\frac{\partial e}{\partial yi_o} = \frac{\partial (\frac{1}{2} \sum_{o=1}^q (d_o(k) - yo_o(k))^2)}{\partial yi_o} = -(d_o(k) - yo_o(k)) yo'_o(k)$$

$$= -(d_o(k) - yo_o(k)) f'(yi_o(k)) \square -\delta_o(k)$$

$$\begin{aligned} \frac{\partial e}{\partial hi_h(k)} &= \frac{\partial (\frac{1}{2} \sum_{o=1}^q (d_o(k) - yo_o(k))^2)}{\partial ho_h(k)} \frac{\partial ho_h(k)}{\partial hi_h(k)} \\ &= \frac{\partial (\frac{1}{2} \sum_{o=1}^q (d_o(k) - f(\sum_{h=1}^p w_{ho} ho_h(k) - b_o))^2)}{\partial ho_h(k)} \frac{\partial ho_h(k)}{\partial hi_h(k)} \\ &= \frac{\partial (\frac{1}{2} \sum_{o=1}^q ((d_o(k) - f(\sum_{h=1}^p w_{ho} ho_h(k) - b_o))^2)}{\partial ho_h(k)} \frac{\partial ho_h(k)}{\partial hi_h(k)} \\ &= -\sum_{o=1}^q (d_o(k) - yo_o(k)) f'(yi_o(k)) w_{ho} \frac{\partial ho_h(k)}{\partial hi_h(k)} \\ &= -(\sum_{o=1}^q \delta_o(k) w_{ho}) f'(hi_h(k)) \square -\delta_h(k) \end{aligned}$$

□第六步，利用输出层各神经元的 $\delta_o(k)$ 和隐含层各神经元的输出来修正连接权值 $w_{ho}(k)$ 。

$$\Delta w_{ho}(k) = -\mu \frac{\partial e}{\partial w_{ho}} = \mu \delta_o(k) h o_h(k)$$

$$w_{ho}^{N+1} = w_{ho}^N + \eta \delta_o(k) h o_h(k)$$

□第七步，利用隐含层各神经元的 $\delta_h(k)$ 和输入层各神经元的输入修正连接权。

$$\Delta w_{ih}(k) = -\mu \frac{\partial e}{\partial w_{ih}} = -\mu \frac{\partial e}{\partial i_h(k)} \frac{\partial i_h(k)}{\partial w_{ih}} = \delta_h(k) x_i(k)$$

$$w_{ih}^{N+1} = w_{ih}^N + \eta \delta_h(k) x_i(k)$$

□第八步，计算全局误差

$$E = \frac{1}{2m} \sum_{k=1}^m \sum_{o=1}^q (d_o(k) - y_o(k))^2$$

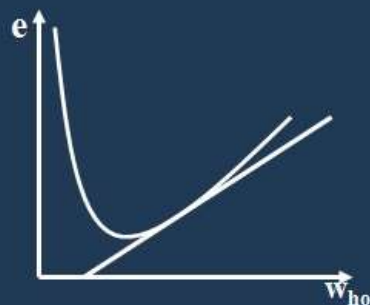
□第九步，判断网络误差是否满足要求。当误差达到预设精度或学习次数大于设定的最大次数，则结束算法。否则，选取下一个学习样本及对应的期望输出，返回到第三步，进入下一轮学习。

这里解释下根据误差对权值的偏导数来修订权值：

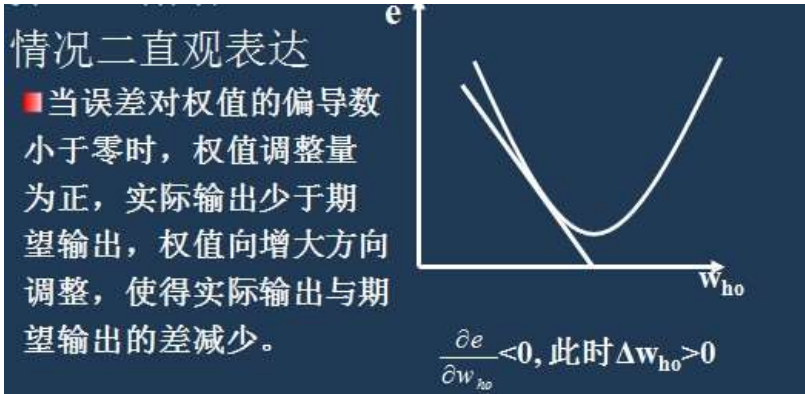
BP算法直观解释

➤情况一直观表达

■ 当误差对权值的偏导数大于零时，权值调整量为负，实际输出大于期望输出，权值向减少方向调整，使得实际输出与期望输出的差减少。



$$\frac{\partial e}{\partial w_{ho}} > 0, \text{ 此时 } \Delta w_{ho} < 0$$



标签: neural networks

好文要顶

关注我

收藏该文







wentingtu

关注 - 1

粉丝 - 322

+ 加关注

60

(请您对文章做出评价)

« 上一篇: 每天工作4小时的程序员
» 下一篇: 特征抽取之 χ^2 统计量

posted on 2012-06-05 11:43 wentingtu 阅读(61588) 评论(9) 编辑 收藏

Feedback

#1楼 2014-01-05 14:52 Shirley 杨

对已经归一化的数据进行网络训练的时候，newff函数对输入和输出矩阵有什么具体的要求吗？

支持(o) 反对(o)

#2楼 2014-07-08 09:19 Tsien

good!

支持(o) 反对(o)

#3楼 2014-08-21 10:25 gogooooo7

第4步有个地方显示有问题

支持(o) 反对(o)

#4楼 2014-11-03 11:09 wangysa

Good!!

支持(o) 反对(1)

#5楼 2015-05-03 16:28 大爷就是我

第五步呢，

支持(o) 反对(o)

#6楼 2015-05-03 16:29 大爷就是我

是漏掉了还是你写错了步骤顺序，我没看懂

支持(o) 反对(o)

#7楼 2015-05-10 10:28 gystld

推荐一款最易用中文excel神经网络软件包，有BP、RBF、PNN、GRNN、Elman、SOM、LVQ七种神经网络功能。一键操作，易用易懂。<http://www.plsexcelword.com/>。

支持(1) 反对(0)

#8楼 2015-06-18 08:39 cofftech

神经网络源码:<http://www.eyesourcecode.com/f/AI/1>

支持(0) 反对(1)

#9楼 2015-07-29 21:58 叶孤傲

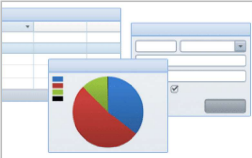
@gystld
请问你有正版的PEW软件吗。急需，谢谢！1377069951@qq.com

支持(0) 反对(0)

[刷新评论](#) [刷新页面](#) [返回顶部](#)

注册用户登录后才能发表评论，请 [登录](#) 或 [注册](#)，[访问](#) 网站首页。

【推荐】50万行VC++源码:大型组态工控、电力仿真CAD与GIS源码库
【阿里云SSD云盘】速度行业领先



jQuery MiniUI

快速开发WEB软件，支持
Java、.Net

- 最新IT新闻：
- 马云：我和我的阿里巴巴是幸运的
 - 当年网景的首席技术专家 现在成谷歌搜索新掌门
 - 报告称32%iPhone用户仍使用4英寸设备
 - 马斯克下个梦想：建垂直起飞电动飞机
 - Visa推出多款开发者工具 发力移动支付
- » 更多新闻...

- 最新知识库文章：
- 消息队列（Message Queue）基本概念
 - 编程每一天（Write Code Every Day）
 - Docker简介
 - Docker简明教程
 - Git协作流程
- » 更多知识库文章...