

原

ReLU激活函数：简单之美

2016年11月13日 17:46:24

对半独白

阅读数：35226

标签：

ReLU激活函数

神经网络

机器学习

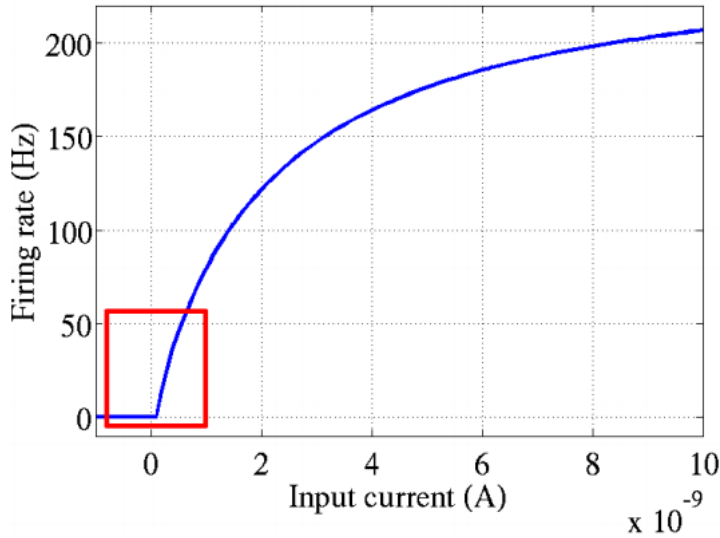
更多

版权声明：本文为博主原创文章，未经博主允许不得转载。<https://blog.csdn.net/cherryvlei/article/details/53149381>

导语

在深度神经网络中，通常使用一种叫**修正线性单元(Rectified linear unit, ReLU)**作为神经元的激活函数。ReLU起源于神经科学研究：2001年，D

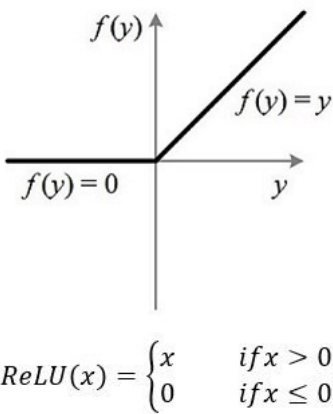
从生物学角度模拟出了脑神经元接受信号更精确的激活模型，如下图：



其中横轴是时间(ms)，纵轴是神经元的放电速率(Firing Rate)。同年，Attwell等神经科学家通过研究大脑的能量消耗过程，推测神经元的工作方式具有稀疏性；2003年Lennie等神经科学家估测大脑同时被激活的神经元只有1~4%，这进一步表明了神经元的工作稀疏性。而对于ReLU函数而言，类似表现是其相比于其他线性函数(如purlin)和非线性函数(如sigmoid、双曲正切)又有何优势？下面请各位看官容我慢慢道来。

简单之美

首先，我们来看一下ReLU激活函数的形式，如下图：



从上图不难看出，ReLU函数其实是分段线性函数，把所有的负值都变为0，而正值不变，这种操作被成为**单侧抑制**。可别小看这个简单的操作，正因为有了这单侧抑制，才的神经元也具有了稀疏激活性。尤其体现在深度神经网络模型(如CNN)中，当**模型增加N层之后，理论上ReLU神经元的激活率将降低2的N次方倍**。这里或许有童鞋会问：R

那么问题来了：这种稀疏性有何作用？换句话说，我们为什么需要让神经元稀疏？不妨举栗子来说明。当看名伶演唱时，我们以根据故事情节来理解，这时用到的是我们的大脑左半球；而当看蒙面唱将时，我们可以跟着歌手一起哼唱，这时用到的则是我们的大脑右半球。左半球主要负责逻辑思维，而右半球负责感性思维。也就是说，当我们在进行运算或者欣赏时，都会有一部分神经元处于激活或是抑制状态，可以说是各司其职。再比如，生病了去医院看病，检查指标，但跟病情相关的通常只有那么几个。与之类似，当训练一个深度分类模型的时候，和目标相关的特征往往也就那么几个，因此通过ReLU实现稀疏性，能够更好地挖掘相关特征，拟合训练数据。

此外，相比于其它激活函数来说，ReLU有以下优势：**对于线性函数而言，ReLU的表达力更强**，尤其体现在深度网络中；**对于非线性函数而言，ReLU的梯度为常数**，因此**不存在梯度消失问题(Vanishing Gradient Problem)**，使得模型的收敛速度维持在一个稳定状态。这跟描述一下什么是梯度消失问题：当梯度小于1时，预测值与真实值之间的误差每传播一层会衰减一次，如果在深层模型中使用sigmoid作为激活函数，这种衰减尤为明显，将导致模型无法收敛。

结语

以上便是我对ReLU激活函数的一些理解，在此感谢各位看官的耐心阅读，不足之处希望多多指教。后续内容将会不定期奉上，欢迎大家关注小斗公众号。



转型人工智能 可以吗？一个小测试就让你知道怎么学

人工智能技术向前发展，也必然会出现一些岗位被人工智能取代，但我们相信，随着人工智能的发展，会有更多的新的、属于未来的工作岗位出现，是社会发展的必然产物，也许只能是与时俱进了

想对作者说点什么？

我来说两句

- qq_39597462

简单明了，棒

(3个月前)

#3楼
- yangzixuan_0608

赞

(4个月前)

#2楼
- wangsanlantg

很好，讲的非常清晰！！

(10个月前)

#1楼

查看回复(1)

深度学习：激活函数的比较和优缺点，sigmoid，tanh，relu

1万

1、什么是激活函数2、为什么要用3、都有什么激活函数4、sigmoid，Relu，softmax1. 什么是激活函数如下图，在神经元中，输入的 inputs 通过加权，...

Deep learning系列（七）激活函数

2.5万

介绍了三种神经网络激活函数。

从0到1学好区块链开发，一年编程经验学完月薪可达40K+

立即申请试学，成为时代颠覆者

神经网络为什么要有激活函数，为什么relu能够防止梯度消失

1.4万

第一个问题：为什么引入非线性激励函数？如果不用激励函数（其实相当于激励函数是f(x) = x），在这种情况下你每一层输出都是上层输入的线性函数，...

人工神经网络之激活函数-RELU函数

973

一句话概括：不用sigmoid和tanh作为激活函数，而用ReLU作为激活函数的原因是：加速收敛。因为sigmoid和tanh都是饱和(saturating)的。何为饱和？个...

深度学习——**激活函数Sigmoid/Tanh/ReLU**

一般情况下，使用ReLU会比较好 1、使用 ReLU，就要注意设置 learning rate，不要让网络训练过程中出现很多“dead” 神经元； 2、如果“dead”无法解决...

吴恩达deeplearning之CNN—卷积神经网络入门

1.边界检测示例假如你有一张如下的图像，你想让计算机搞清楚图像上有什么物体，你可以做的事情是检测图像的垂直边缘和水平边缘。如下是

远离假传奇！这游戏爆率+999倍，有充值有VIP算我输！

贪玩游戏 · 顶新

一次面试引发的**问题：激活函数Relu，sigmoid，tanh的角逐**

记得之前去某东面试，面试官问我CNN的激活函数有哪些？我说：Relu/sigmoid /tanh 问：有区别吗？我心想我只用过relu，至于区别我想应该

ReLU激活函数的缺点

训练的时候很“脆弱”，很容易就“die”了，训练过程该函数不适应较大梯度输入，因为在参数更新以后，ReLU的神经元不会再有激活的功能，导致梯度永远...

RELU 激活函数及其他相关的函数

本博客仅为作者记录笔记之用，不免有很多细节不对之处。 还望各位看官能够见谅，欢迎批评指正。 更多相关博客请猛戳：http://blog.csdn.net/cyh_24 ...

博主推荐



悟乙己

关注

295篇文章



Quincuntial

关注

626篇文章



Alice熹爱学习

关注

278篇文章



对半独白

关注

原创89

粉丝77

喜欢27

评论51

等级：博客4

访问：11万+

积分：1840

排名：3万+

勋章：恒



图像识别



最新文章

几何分布和二项分布有什么区别？

什么是内网穿透？

认识生活中的泊松分布

什么是极大似然估计？

为什么样本方差计算是除以n-1？

个人分类

随笔1篇

论文记录3篇

<https://blog.csdn.net/cherryivlei/article/details/53149381>

3/5

集群学习笔记	5篇
小窍门	1篇
面试经验	1篇
展开	

归档	
2018年9月	2篇
2018年8月	3篇
2018年7月	2篇
2018年6月	3篇
2018年5月	2篇
展开	

热门文章

ReLU激活函数：简单之美

阅读量：34895

如何直观理解AUC评价指标？

阅读量：24744

如何直观理解交叉熵及其优势？

阅读量：5818

论文实验笔记之二：EViews6跑ARIMA

阅读量：3945

AlphaGo技术剖析：揭开围棋大脑的神秘面纱

阅读量：3017

最新评论

机器学习方法篇(7)-----L...

C12H22O11_H2O：[reply]cherrylvlei[/reply] 谢谢回复，还有一点疑惑，在Input Gat...

机器学习方法篇(7)-----L...

cherrylvlei：[reply]C12H22O11_H2O[/reply] Cell的状态

机器学习方法篇(7)-----L...

C12H22O11_H2O：请问Scf代表的含义是什么呢？

机器学习方法篇(6)-----朴...

qq_23934649：真的写得太好了，明白了为什么simple rnn忘事。也知道它是怎么做bp的了。真的要10086个赞

机器学习方法篇(6)-----朴...

qq_23934649：真的写得太好了，明白了为什么simple rnn忘事。也知道它是怎么做bp的了。真的要10086个赞

十大笔记本排名



联系我们

4





扫码联系客服



扫码下载APP

关于 招聘 广告服务 网站地图

京ICP证09002463号 百度提供站内搜索

©2018 CSDN版权所有

kefu@csdn.net

400-660-0108

QQ客服

客服论坛

经营性网站备案信息

网络110报警服务

中国互联网举报中心

北京互联网违法和不良信息举报中心



官方公众号



区块链大本营

4

