CSDN

博客 (http://b//b/y/ws/dos/det//?ee?*tebelloalibar)

学院 (http://edu.csdn.net?ref=toolbar)

下载 (http://download.csdn.net?ref=toolbar)

GitChat (http://gitbook.cn/?ref=csdn)

全 发Chat

____登录 (https://passport.csdn.net/account/mobileregister?ref=toolbar&action=mobileRegister) (http://mp.blognesw/gitchat

[深度学习Deep Learning (05) :: Batch Normalization(BN)批标准化

2017年01月09日 21:37:00

/DeepLearning_Python)

标签: github (http://so.csdn.net/so/search/s.do?q=github&t=blog) /

深度学习 (http://so.csdn.net/so/search/s.do?q=深度学习&t=blog) /

BN (http://so.csdn.net/so/search/s.do?q=BN&t=blog) /

批标准化 (http://so.csdn.net/so/search/s.do?q=批标准化&t=blog)

∩ 2707

(http://blog.csdn.net

89

/u013082989)

码云

莫失莫忘Lawlite (http://...

未开诵 (https://gite

OLIALCOVW

Unable to Conn

The Proxy was unable to connect to the remote site. responding to requests. If you feel you have reached please submit a ticket via the link provided below.

URL: http://pos.baidu.com/s?hei=250&wid=300&di=u %2Fblog.csdn.net%2Fu013082989%2Farticle%2Fdet

四、Batch Normalization(BN)批标准化

1、说明

• 参考论文: http://jmlr.org/proceedings/papers/v37/ioffe15.pdf (http://jmlr.org/proceedings /papers/v37/ioffe15.pdf)

• github地址: https://github.com/lawlite19/DeepLearning_Python (https://github.com/lawlite19

• 或者查看这里 (https://github.com/lawlite19/DeepLearning Python/blob/master/paper/%EF%BC %88BN%EF%BC

%89Batch%20Normalization%20Accelerating%20Deep%20Network%20Training%20by%20Reducing%20 我放在github上了: https://github.com/lawlite19/DeepLearning_Python/blob/master/paper/%EF %BC%88BN%EF%BC

%89Batch%20Normalization%20Accelerating%20Deep%20Network%20Training%20by%20Reducing%20 (https://github.com/lawlite19/DeepLearning Python/blob/master/paper/%EF%BC%88BN%EF %BC%89Batch%20Normalization%20Accelerating%20Deep%20Network%20Training%20by%20Reducin

2、论文概述

- 2015年Google提出的Batch Normalization
- 训练深层的神经网络很复杂,因为训练时每一层输入的分布在变化,导致训练过程中的饱和,称这种 现象为: internal covariate shift
- 需要降低学习率Learning Rate和注意参数的初始化
- 论文中提出的方法是对于每一个小的训练batch都进行标准化(正态化)
 - 允许使用较大的学习率
 - 不必太关心初始化的问题

加入CSDN. 享回財産監例子安不需要使用。Dropper方法避免过拟合

● 此方法在 ImageNet classification 比赛中获得 4.82% top-5 的测试错误率

●他的最新文章

更多文章 (http://blog.csdn.net/u013082989)

阿里云GPU服务器上Torch安装与测试 (h ttp://blog.csdn.net/u013082989/article/det ails/78897235)

论文记录_MobileNets Efficient Convoluti onal Neural Networks for Mobile Vision A pplication (http://blog.csdn.net/u0130829 89/article/details/77970196)

论文记录-Pruning Filters For Efficient Co nvNets (http://blog.csdn.net/u013082989/ article/details/77943240)

论文记录-Deep Compression:Compressi ng DeepNeural Networks With Pruning, T rained Quantization And Huffman (http:// blog.csdn.net/u013082989/article/details/ 77915375)

深度学习(08)_RNN-LSTM循环神经网 络-03-Tensorflow进阶实现 (http://blog.cs 返回顶部 dn.net/u013082989/article/details/736933 92)

他的热门文章

登录 注册 服务器被入侵(minerd挖矿程序)(http:

3、BN 思路

- 如果输入数据是白化的(whitened), 网络会更快的收敛
 - 白化目的是降低数据的冗余性和特征的相关性,例如通过线性变换使数据为0均值和单位方差

● 并非直接标准化每一层那么简单,如果不考虑归一化的影响,可能会降低梯度下降的影响

● 标准化与某个样本和所有样本都有关系

● 解决上面的问题,我们希望对于任何参数值,都要满足想要的分布;

 Equation Quota Exceeded The World's #1 Online Equation Editor

• 对于反向传播,需要计算:



Equation Quota Exceeded

The World's #1 Online Equation Editor The World's #1 Online Equation Editor

● 这样做的计算代价是非常大的,因为需要计算x的协方差矩阵

• 然后白化操作: Equation Quota Exceeded The World's #1 Online Equation Editor

- 上面两种都不行或是不好,进而得到了BN的方法
- 既然白化每一层的输入代价非常大,我们可以进行简化
- 简化1
 - 标准化特征的每一个维度而不是去标准化所有的特征,这样就不用求协方差矩阵了



• 标准化操作:



■ 需要注意的是标准化操作可能会降低数据的表达能力,例如我们之前提到的Sigmoid函数:

//blog.csdn.net/u013082989/article/detail s/51971121)

18889

scrapy爬虫框架将数据保存Mysql数据库 中 (http://blog.csdn.net/u013082989/articl e/details/52589791)

□ 18015

[Python]使用Scrapy爬虫框架简单爬取图 片并保存本地 (http://blog.csdn.net/u0130 82989/article/details/52507601)

11047

GitHub合并 (merge) 代码时冲突解决 (http://blog.csdn.net/u013082989/article/d etails/51099269)

3 8769

递归求数组的最大值, 主要说下递归执 行的过程 (http://blog.csdn.net/u0130829 89/article/details/50784385)

\$5607





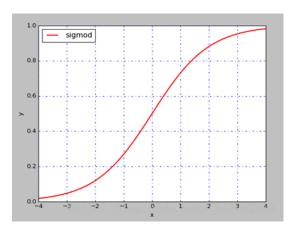
加入CSDN,享受更精准的内容推荐,与500万程序员共同成长!

登录

注册

第2页 共9页

2018/1/4 上午11:38



● 标准化之后均值为0,方差为1,数据就会落在近似线性的函数区域内,这样激活函数的意义就 不明显

● 所以对于每个 , 对应一对参数:



, 然后令:

The World's #1 Online Equation Editor



● 从式子来看就是对标准化的数据进行缩放和平移,不至于使数据落在线性区域内,增加数据的



恢复到原来的值了)

● 但是这里还是使用的全部的数据集,但是如果使用随机梯度下降,可以选取一个batch进行训练

● 简化2

- 第二种简化就是使用 mini-batch 进行 随机梯度下降
- 注意这里使用 mini-batch 也是标准化每一个维度上的特征,而不是所有的特征一起,因为若果 mini-batch 中的数据量小于特征的维度时,会产生奇异协方差矩阵 ,对应的行列式的值为0,非 满秩
- 假设mini-batch 大小为 m 的 B



● 作者给出的批标准化的算法如下: 加入CSDN,享受更精准的内容推荐,与500万程序员共同成长!

 \triangle 内容举报

TOP 返回顶部

登录

Input: Values of x over a mini-batch: $\mathcal{B} = \{x_{1...m}\}$; Parameters to be learned: γ , β

Output: $\{y_i = BN_{\gamma,\beta}(x_i)\}$

$$\mu_{\mathcal{B}} \leftarrow \frac{1}{m} \sum_{i=1}^{m} x_i \qquad \qquad \text{// mini-batch mean}$$

$$\sigma_{\mathcal{B}}^2 \leftarrow \frac{1}{m} \sum_{i=1}^m (x_i - \mu_{\mathcal{B}})^2$$
 // mini-batch variance

$$\widehat{x}_i \leftarrow \frac{x_i - \mu_{\mathcal{B}}}{\sqrt{\sigma_{\mathcal{B}}^2 + \epsilon}}$$
 // normalize

$$y_i \leftarrow \gamma \widehat{x}_i + \beta \equiv \mathrm{BN}_{\gamma,\beta}(x_i)$$
 // scale and shift

Algorithm 1: Batch Normalizing Transform, applied to activation x over a mini-batch $\log \cosh n$ and n = 100

P

算法中的 ε 是一个常量, 为了保证数值的稳定性

● 反向传播求梯度:

● 因为:

powered by
CODECOGS

Please obtain a licence from www.CodeCogs.com
Equation Quota Exceeded

0 , .

The World's #1 Online Equation Editor



Equation Quota Exceeded

The World's #1 Online Equation Editor



The World's #1 Online Equation Editor





The World's #1 Online Equation Editor

⚠
内容举报

命 返回顶部

加入CSDN,享受更精准的内容推荐,与500万程序员共同成长!

登录

注册



● 对于BN变换是可微分的,随着网络的训练,网络层可以持续学到输入的分布。

4、BN 网络的训练和推断

- 按照BN方法,输入数据 x 会经过变化得到 BN (x) ,然后可以通过随机梯度下降进行训练,标准化是 在mini-batch上所以是非常高效的。
- 但是对于推断我们希望输出只取决于输入,而对于输入只有一个实例数据,无法得到 mini-batch 的其 他实例,就无法求对应的均值和方差了。
- 可以通过从所有训练实例中获得的统计量来**代替**mini-batch中m个训练实例获得统计量均值和方
- 我们对每个 mini-batch 做标准化,可以对记住每个 mini-batch 的B,然后得到全局统计量



• 所以推断采用 BN 的方式为:

 \triangle 内容举报

TOP 返回顶部

加入CSDN,享受更精准的内容推荐,与500万程序员共同成长!

登录





作者给出的完整算法:

Input: Network N with trainable parameters Θ ; subset of activations $\{x^{(k)}\}_{k=1}^K$

Output: Batch-normalized network for inference, $N_{\rm BN}^{\rm inf}$

- 1: $N_{\rm BN}^{\rm tr} \leftarrow N$ // Training BN network
- 2: **for** k = 1 ... K **do**
- Add transformation $y^{(k)} = BN_{\gamma^{(k)},\beta^{(k)}}(x^{(k)})$ to $N_{\rm BN}^{\rm tr}$ (Alg. 1)
- Modify each layer in $N_{\rm BN}^{\rm tr}$ with input $x^{(k)}$ to take $y^{(k)}$ instead
- 5: end for
- 6: Train $N_{\mathrm{BN}}^{\mathrm{tr}}$ to $\Theta \cup \{\gamma^{(k)}, \beta^{(k)}\}_{k=1}^K$ optimize parameters
- 7: $N_{\rm BN}^{\rm inf} \leftarrow N_{\rm BN}^{\rm tr}$ // Inference BN network with frozen // parameters
- 8: **for** k = 1 ... K **do**
- // For clarity, $x \equiv x^{(k)}$, $\gamma \equiv \gamma^{(k)}$, $\mu_{\mathcal{B}} \equiv \mu_{\mathcal{B}}^{(k)}$, etc. Process multiple training mini-batches \mathcal{B} , each of 10: size m, and average over them:

$$E[x] \leftarrow E_{\mathcal{B}}[\mu_{\mathcal{B}}]$$

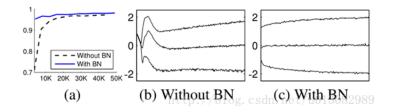
$$Var[x] \leftarrow \frac{m}{m-1} E_{\mathcal{B}}[\sigma_{\mathcal{B}}^2]$$

In $N_{\mathrm{BN}}^{\mathrm{inf}}$, replace the transform $y = \mathrm{BN}_{\gamma,\beta}(x)$ with $y = \frac{\gamma}{\sqrt{\mathrm{Var}[x] + \epsilon}} \cdot x + \left(\beta - \frac{\gamma \, \mathrm{E}[x]}{\sqrt{\mathrm{Var}[x] + \epsilon}}\right)$

Algorithm 2: Training a Batch-Normalized Network

5、实验

● 最后给出的实验可以看出使用BN的方式训练精准度很高而且很稳定。



 \triangle 内容举报

> TOP 返回顶部

版权声明:本文为博主原创文章,未经博主允许不得转载。

加入CSDN,享受更精准的内容推荐,与500万程序员共同成长!

登录

Д

深度学习优化策略-1>Batch Normalization(BN)

今天给大家带来深度学习的优化策略篇的第一篇Batch Normalization(BN)。BN可以看做对输入样本的一种约束,最大 作用是加速收敛,减少模型对dropout, careful weigh...



(http://blog.csdn.net/lqfarmer/article/details/71155733)

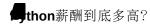
matlab标准化和反标准化——zscore



🥵 zb1165048017 2016年03月25日 09:27 🕮 5138

分 先来看自带函数zscore的使用 >> A=[1 2 3;4 5 6]; >> [B,A_mean,A_std]=zscore(A) B = -0.7071 -0.7071 -0.7...

p://blog.csdn.net/zb1165048017/article/details/50976858)



Python2017年表现异常突出,这次我们拿出10个招聘网站的Python数据,对比后发现....你猜到了 吗?



(http://www.baidu.com/cb.php?c=lgF_pyfqnHmknjnvPjc0IZ0qnfK9ujYzP1nYPH0k0Aw-

uZR8mLPbUB48ugfElAqspynETZ-YpAq8nWqdlAdxTvqdThP-

5yF UvTkn0KzujY4rHb0mhYqn0KsTWYs0ZNGujYkPHTYn1mk0AqGujYknWb3rjDY0APGujYLnWm4n1c0ULl85H00TZbqnW0v0APzm1Ykn16kP0)

数据标准化/归一化normalization



🖏 pipisorry 2016年08月19日 09:42 🕮 42206

http://blog.csdn.net/pipisorry/article/details/52247379数据的标准化(normalization)和归一化 数据的标准化(normaliza.

(http://blog.csdn.net/pipisorry/article/details/52247379)

<深度学习优化策略-1>Batch Normalization(BN)

今天给大家带来深度学习的优化策略篇的第一篇Batch Normalization (BN) 。BN可以看做对输入样本的一种约束,最大 作用是加速收敛,减少模型对dropout, careful weigh...



🍘 lqfarmer 2017年05月03日 21:47 🕮 1281

(http://blog.csdn.net/lgfarmer/article/details/71155733)

【深度学习】论文导读:google的批正则方法(Batch Normalization: Accelerat...

google2015年的论文,首次提出批正则方法,优化深度神经元网络的学习摘要在深度网络的训练中,每一层网络的输入 都会因为前一层网络参数的变化导致其分布发生改变,这就要求我们必须使用一个很小的学...



🥵 mao_xiao_feng 2016年11月13日 17:07 🕮 829

(http://blog.csdn.net/mao xiao feng/article/details/53150037)

Æ 内容举报

TOP

返回顶部



加入CSDN,享受更精准的内容推荐,与500万程序员共同成长! Unable to Connect to Site.

登录

注册

第7页 共9页

基于Theano的深度学习(Deep Learning)框架Keras学习随笔-05-模型

原地址: http://blog.csdn.net/niuwei22007/article/details/49207187可以查看更多文章介绍完了优化器和目标函数,那么剩 下的就是训练模型了。这...

(dlaicxf 2016年10月18日 11:02 Q 734

http://blog.csdn.net/dlaicxf/article/details/52847051)

[**TensorFlow** 学习笔记-05]批标准化(Bacth Normalization, BN)

[版权说明] TensorFlow 学习笔记参考: 李嘉璇 著 TensorFlow技术解析与实战黄文坚 唐源 著 TensorFlow实战郑泽宇 顾 医字著 TensorFlow实战Googl...

(http://blog.csdn.net/caicaiatnbu/article/details/72742293)

★ 度学习论文笔记][Weight Initialization] Batch Normalization: Accelerating De...

loffe, Sergey, and Christian Szegedy. "Batch normalization: Accelerating deep network training by re...

Hao_Zhang_Vision 2016年09月20日 13:54 🕮 706

(http://blog.csdn.net/Hao_Zhang_Vision/article/details/52595249)

Neural Networks and Deep Learning-神经网络与深度学习-zh.zip

2017年12月26日 17:42 3.06MB 下载



Deep Learning Book 深度学习 [带标签PDF完全版]

2017年12月09日 08:14 15.91MB 下载



基于深度学习的图像分类Image classification with deep learning常用模型

本文中, 我会根据下大家image classification常用的cnn模型, 针对cifar10 (for 物体识别), mnist (for 字符识别) & Ima geNet(for 物体识别)做一个mo...

🚇 GarfieldEr007 2016年03月31日 12:52 🕮 5714

(http://blog.csdn.net/GarfieldEr007/article/details/51025323)

MIT Deep Learning book 深度学习 中文+英文版本

2017年12月06日 17:16 28.88MB 下载



深度学习 英文原版 Deep Learning(Ian Goodfellow Yoshua Be...

2017年11月28日 14:38 24.15MB 下载



【deep learning】深度学习用于图片的分类和检测总结

【deep learning】深度学习用于图片的分类和检测总结

🧌 qq_30214939 2017年06月11日 10:10 🗘 427

加入6066011610夏爱更精准的内容推荐39/每6606万程源5为共同成战)!

登录

注册

Æ

内容举报

TOP

返回顶部

第8页 共9页

(最新python版深度学习) Deep Learning with python

2017年12月06日 08:53 5.74MB 下载



深度学习方法及应用 英文版 Deep Learning Methods and Appli...

/L 2017年12月05日 12:28 7.6MB 下载



二牛的《深度学习》笔记,Deep Learning速成教程

實锋网(搜索"雷锋网"公众号关注)按:本文由Zouxy责编,全面介绍了深度学习的发展历史及其在各个领域的应用,并解释了深度学习的基本思想,深度与浅度学习的区别和深度学习与神经网络之间的关系。深度学...

w huangting1 2016年08月23日 00:05 🕮 5743

p://blog.csdn.net/huangting1/article/details/52281889)

●ural Network and Deep Learning(神经网络和深度学习)中...





deep learning(深度学习)中文版

2017年11月05日 00:26 26.5MB 下载



Multimodal Deep Learning(多模态深度学习)未完待续

提出一种在深度网络上的新应用,用深度网络学习多模态。特别的是,我们证明了跨模态特征学习——如果在 特征学习过程中多模态出现了,对于一个模态而言,更好的特征可以被学习。此外,我们展示了如...

📦 s2010241013 2016年06月22日 08:53 🕮 7464

(http://blog.csdn.net/s2010241013/article/details/51731657)

<u>/</u>ì\ 内容举报

TOP 返回顶部