



1点点加盟要求



联系我们



请扫描二维码联系客服

webmaster@csdn.net

400-660-0108

QQ客服 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN版权所有

京ICP证09002463号

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



工业键盘



相关资料与课程推荐



文章搜索

文章分类

31

Deep Learning论文笔记之（一）K-means特征学习

标签：Deep Learning

2013年08月15日 14:14:09

56918人阅读 评论(19) 收藏

分类：机器学习（45） 计算机视觉（72） Deep Learning（17）

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/zouxy09/article/details/9982495>

Deep Learning论文笔记之（一）K-means特征学习

zouxy09@qq.com

<http://blog.csdn.net/zouxy09>

自己平时看了一些论文，但老感觉看完过后就会慢慢的淡忘，某一天重新拾起来的时候又有看过一样。所以想习惯地把一些感觉有用的论文中的知识点总结整理一下，一方面在整理过自己的理解也会更深，另一方面也方便未来自己的勘察。更好的还可以放到博客上面与大家交为基础有限，所以对论文的一些理解可能不太正确，还望大家不吝指正交流，谢谢。

本文的论文来自：

[Learning Feature Representations with K-means](#), Adam Coates and Andrew Y. Ng. In Neural Network of the Trade, Reloaded, Springer LNCS, 2012

下面是自己对其中的一些知识点的理解：

《Learning Feature Representations with K-means》

自从Deep Learning之风盛起之时到现在，江湖上诞生了很多都可以从无标签数据中学习到分级的特征的算法。大部分情况，这些算法都涉及到一个多层网络，而训练和调整这个网络需tricks。最近，我们发现K-means聚类算法也可以被作为一个非常快的训练方法。它的优点是快实现！当然了，K-means也不是万能神丹，它也存在自身的局限性。在本文中，我们就关注K-n方方面面。总结了最近的K-means算法的效果和介绍使用k-means来有效地学习图像的特征的巧。

一、概述

非监督学习的一般流程是：先从一组无标签数据中学习特征，然后用学习到的特征提取选取有标签数据特征，然后再进行分类器的训练和分类。之前说到，一般的非监督学习算法都有hyper-parameters需要调整。而，最近我们发现对于上面同样的非监督学习流程中，用K-means法来实现特征学习，也可以达到非常好的效果，有时候还能达到state-of-the-art的效果。亮瞎了俗眼。

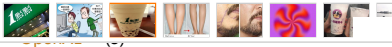
托“bag of features”的福，K-means其实在特征学习领域也已经略有名气。今天我们就不要迷失在其往日的光芒中了。在这里，我们只关注，如果要K-means算法在一个特征学习系统良好的性能需要考虑哪些因素。这里的特征学习系统和其他的Deep Learning算法一样：直接从输入（像素灰度值）中学习并构建多层的分级的特征。另外，我们还分析了K-means算法与江他知名的特征学习算法的千丝万缕的联系（天下武功出少林，哈哈）。

<https://blog.csdn.net/zouxy09/article/details/9982495>

1/9



1点点加盟要求



联系我们



请扫描二维码联系客服

webmaster@csdn.net

400-660-0108

QQ客服 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN 版权所有

京ICP证09002463号

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

展开

阅读排行

Deep Learning (深度学习) ...	(797957)
Deep Learning (深度学习) ...	(537975)
Deep Learning (深度学习) ...	(532331)
Deep Learning 论文笔记之 (...	(386399)
机器学习中的范数规则化之 (...	(383488)
Deep Learning (深度学习) ...	(327936)
从最大似然到EM算法浅解	(288785)
Deep Learning (深度学习) ...	(276971)
Deep Learning (深度学习) ...	(264855)
Deep Learning (深度学习) ...	(219139)

最新评论

- Deep Learning 论文笔记...
weixin_38356015 : [reply]qq295456059[reply] 你好，我是CNN刚入门，本科毕设就是做这个的...
- Deep Learning 论文笔记...
jinhuan_hit : [reply]atlantisbuaa[reply] 请问data、util这两个文件夹从哪里获...
- 机器学习中的范数规则化之（一）L0...
BIT_666 : 学习了，谢谢博主！
- 语音信号处理之（二）基音周期估计（...
raymus : 楼主，可以发一份源码给我吗，505211818@qq.com。感激不尽！
- 语音信号处理之（三）矢量量化（Ve...
Joel_Wang : 红字那应该是码矢吧，博主是不是笔误啊
- 光流Optical Flow介绍与...
tkingcer : flow的可视化，源于 Middlebury col or code ，链接见 http://vi...
- 从最大似然到EM算法浅解
sjx19871109 : 讲得很好，生动形象，所给的案例能够较好地阐明EM算法的背景。
- 从最大似然到EM算法浅解
Nicco_773 : jensen不等式那里是不是写错了啊？不应该是 $E[x] \geq E[f(x)]$ 吗？
- 从最大似然到EM算法浅解
世界已黑白 : 讲的很好，对于使用EM算法进行编程设计来说，很受用，不然只是一大堆的公式证明 而不清楚应用背景...

经典的K-means聚类算法通过最小化数据点和最近邻中心的距离来寻找各个类中心。江湖个别名，叫“矢量量化vector quantization”（[这个在我的博客上也有提到](#)）。我们可以把K-n成是在构建一个字典 $D \in \mathbb{R}^{n \times k}$ ，通过最小化重构误差，一个数据样本 $x^{(i)} \in \mathbb{R}^n$ 可以通过这个字典映射个k维的码矢量。所以K-means实际上就是寻找D的一个过程：

$$\begin{aligned} & \underset{D, s}{\text{minimize}} \sum_i \|D s^{(i)} - x^{(i)}\|_2^2 \\ & \text{subject to } \|s^{(i)}\|_0 \leq 1, \forall i \\ & \text{and } \|D^{(j)}\|_2 = 1, \forall j \end{aligned}$$

这里， $s^{(i)}$ 就是一个与输入 $x^{(i)}$ 对应的码矢量。 $D^{(j)}$ 是字典D的第j列。K-means毕生的目标就是足上面这些条件的一个字典D和每个样本 $x^{(i)}$ 对应的码矢量 $s^{(i)}$ 。我们一起来分析下这些条件。首先字典D和码矢量 $s^{(i)}$ ，我们需要能很好的重构原始的输入 $x^{(i)}$ 。数学的表达是最小化 $x^{(i)}$ 和它的 $s^{(i)}$ 。这个目标函数的优化需要满足两个约束。首先， $\|s^{(i)}\|_0 \leq 1$ ，意味着每个码矢量 $s^{(i)}$ 被约束只有一个非零元素。所以我们寻找一个 $x^{(i)}$ 对应的新的表达，这个新的表达不仅需要更好的保留息，还需要尽可能的简单。第二个约束要求字典的每列都是单位长度，防止字典中的元素或者得任意大或者任意小。否则，我们就可以随意的放缩 $D^{(j)}$ 和对应的码矢量，这样一点用都没有。

这个算法从精神层面与其他学习有效编码的算法很相似，例如sparse coding：

$$\begin{aligned} & \underset{D, s}{\text{minimize}} \sum_i \|D s^{(i)} - x^{(i)}\|_2^2 + \lambda \|s^{(i)}\|_1 \\ & \text{subject to } \|D^{(j)}\|_2 = 1, \forall j \end{aligned}$$

Sparse coding也是优化同样类型的重构。但对于编码复杂度的约束是通过在代价函数中增惩罚项 $\lambda \|s^{(i)}\|_1$ ，以限制 $s^{(i)}$ 是稀疏的。这个约束和K-means的差不多，但它允许多于一个非在保证 $s^{(i)}$ 简单的基础上，可以更准确的描述 $x^{(i)}$ 。

虽然Sparse coding比K-means性能要好，但是Sparse coding需要对每个 $s^{(i)}$ 重复的求解一个凸题，当拓展到大规模数据的时候，这个优化问题是非常昂贵的。但对于K-means来说，对 $s^{(i)}$ 的解就简单很多了：

$$s_j^{(i)} = \begin{cases} D^{(j)\top} x^{(i)} & \text{if } j == \arg \max_l |D^{(l)\top} x^{(i)}| \\ 0 & \text{otherwise.} \end{cases}$$

这个求解是很快的，而且给定s求解D也很容易，所以我们可以通过交互的优化D和s（[可以我博客上的EM算法的博文](#)）来快速的训练一个非常大的字典。另外，K-means还有一个让人青方是，它只有一个参数需要调整，也就是要聚类的中心的个数k。

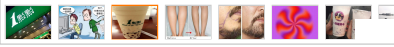
二、数据、预处理和初始化

这里我们采用的是包含很多小的图像patches的数据集，每个patch是16x16的灰度图，对每样本我们将其拉成一个256维的列向量。这些patches可以在无标签图像中随机的裁剪得到。为一个“完备complete”的字典（至少有256个类中心的字典），我们需要保证有足够的用以patches，这样每个聚类才会包含一定合理数量的输入样本。对于16x16的灰度patch来说，m=100,000个是足够的。实际上，训练一个k-means字典比其他的算法（例如sparse coding）需训练样本个数要多，因为在k-means中，一个样本数据只会贡献于一个聚类的中心，换句话说一只能属于一个类，与其他类就毫无瓜葛了，该样本的信息全部倾注给了它的归宿（对应类的中

2.1、预处理 Pre-processing



1点点加盟要求



联系我们



请扫描二维码联系客服

webmaster@csdn.net

400-660-0108

QQ客服 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN 版权所有

京ICP证09002463号

经营性网站备案信息

网络110报警服务

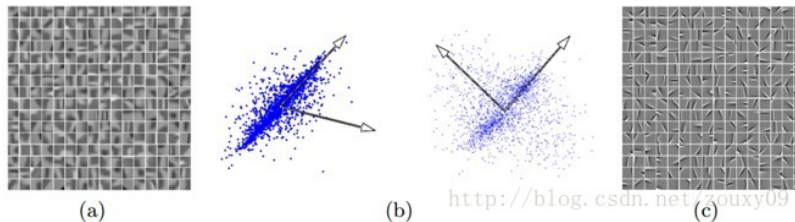
中国互联网举报中心

北京互联网违法和不良信息举报中心

在训练之前，我们需要对所有的训练样本patches先进行亮度和对比度的归一化。具体做法每个样本，我们减去灰度的均值和除以标准差。另外，在除以标准差的时候，为了避免分母制噪声，我们给标准差增加一个小的常数。对于[0, 255]范围的灰度图，给方差加10一般是ok的

$$x^{(i)} = \frac{\tilde{x}^{(i)} - \text{mean}(\tilde{x}^{(i)})}{\sqrt{\text{var}(\tilde{x}^{(i)}) + 10}}$$

对训练数据进行归一化后，我们就可以在上面运行k-means算法以获取对应的聚类中心了（图a展示了没有经过白化处理的自然图像中学习到的聚类中心，图b展示了没有和的效果。左边是没有经过白化的，因为数据存在相关性，所以聚类中心会跑偏。右边是经过了白化的，可以看到聚类中心是更加正交的，这样学习到的特征才会不一样。图c展示的是从经过白化图像patches学习到的聚类中心。



图a是由k-means从没有经过白化处理的自然图像中学习到的聚类中心。图b展示了没有和的效果。左边是没有经过白化的，因为数据存在相关性，所以聚类中心会跑偏。右边是经过了白化的，可以看到聚类中心是更加正交的，这样学习到的特征才会不一样。图c展示的是从经过白化图像patches学习到的聚类中心。

实现whitening白化一个比较简单的方法是ZCA白化（可以参考UFLDL）。我们先对数据点协方差矩阵进行特征值分解 $\text{cov}(x) = V D V^T$ 。然后白化后的点可以表示为：

$$V(D + \epsilon_{zca} I)^{-1/2} V^T x$$

ϵ_{zca} 是一个很小的常数。对于对比度归一化后的数据，对16x16的patch，可以设 $\epsilon_{zca}=0.01$ ，对于patch，可以设 $\epsilon_{zca}=0.1$ 。需要注意的一点是，这个常数不能太小，如果太小会加强数据的高频会使特征学习更加困难。另外，因为数据的旋转对K-means没有影响，所以可以使用其他的白化方法，例如PCA白化（与ZCA不同只在于其旋转了一个角度）。

在白化后的数据中运行k-means可以得到清晰的边缘特征。这些特征和sparse coding啊，L1等方法学到的初级特征差不多。如图c所示。另外，对于新的数据，可能需要调整归一化的参数化的参数 ϵ_{zca} ，以满足好的效果（例如，图像patches具有高对比度，低噪声和少低频波动）。

2.2、初始化 Initialization

一般的K-means聚类算法存在一个比较常见的问题，就是会出现空的聚类。通俗的讲，聚类后的一个类里面居然没有样本（接近于没有）。那么很明显，这个类就一点意义都没有，甚至还危害人间。这我们当然得做些努力来避免这种情况的发生了。那就得找原因了吧。其实这情况下可以认为是中心的初始化的不恰当导致的。常用的中心初始化方法就是随机地从样本中挑出来作为k个初始的聚类中心。但这不是个明智的选择。它有可能导致图像趋于稠密聚集某些因为如果随机选择的patches，也就是训练样本本身就在某个区域分布非常密，那么我们随机去类中心的时候，就会出现就在这个数据分布密集的地方被选出了很多的聚类中心（因为原数据聚集，所以被选中的概率会大）。这是不好的，因为本身这个密集的区域，一个聚类中心才是好硬搞出几个聚类中心，好好的队伍就会被强拆了，还搞得其他比较分散的数据点找不到归宿，类中心离它好遥远，大部分的聚类中心都让富二代给占有了，社会资源分配不当啊，哈哈。这出现一种情况，就是如果样本分布密集的话，它们都几乎从属于一个聚类中心，那么其他聚类几乎没有样本属于它。这样也不好。所以，一个比较好的初始化方式就是从正态分布中随机化聚类中心，然后归一化他们到单位长度。另外，因为经过了白化的阶段，我们期望我们的数据成分已经在一定程度上被组织为球面spherical分布了，而在一个球面上随机地选取一些向量显得温柔多了。



1点点加盟要求



联系我们



请扫描二维码联系客服

webmaster@csdn.net

400-660-0108

QQ客服 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN版权所有

京ICP证09002463号

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

另外，还存在一些启发性的方法去改善k-means的这个问题。例如启发性得重新初始化空中心。实践中证明，这种方法对图像数据的效果比较好。当空类出现的时候，我们再随机挑选替代这个空类中心，一般都是足够有效的。但好像一般也没啥必要。事实上，对于一个足够规模来说，我们会训练很多的聚类中心，如果一个聚类中心包含的数据点太少了，我们直接抛弃了，不要这个聚类中心了，因为俺们的聚类中心够多，多你一个不多，少你一个不少的。

还有一种改善的方法叫做聚类中心的抑制更新damped updates。它 31 次的迭代中，都会下方式更新聚类中心：

$$\begin{aligned} \mathcal{D}_{\text{new}} &:= \arg \min_{\mathcal{D}} \|\mathcal{D}S - X\|_2^2 + \|\mathcal{D} - \mathcal{D}_{\text{old}}\|_2^2 \\ &= (SS^T + I)^{-1}(XS^T + \mathcal{D}_{\text{old}}) \\ &\propto XS^T + \mathcal{D}_{\text{old}} \\ \mathcal{D}_{\text{new}} &:= \text{normalize}(\mathcal{D}_{\text{new}}). \end{aligned}$$

需要注意的是，这种形式的抑制并不会对“大”的聚类有很大的影响（因为 XS^T 的第j列会占第j列要大得多），它只能防止小的聚类在一次的迭代过程中被拉得太远（就像滤波的方法一样值不仅受当前值的影响，还受历史值的影响，这样就会防止这个值出现突变。例如 $a_{\text{old}}=1$ ，新5，那么 $a_{\text{new}}=0.5*5+0.5*a_{\text{old}}=3$ ，也就是 a_{new} 会跳到3而不是一下子跳到5，当然还取决于前值0.5，看我们更信任新值还是历史值。但如果新值很大，那就没有特别大的效果了。例如新值那么 $a_{\text{new}}=50$ ，它还是比 $a_{\text{old}}=1$ 要大多了）。我们将这个k-means训练算法概括如下：

1. Normalize inputs:

$$x^{(i)} := \frac{x^{(i)} - \text{mean}(x^{(i)})}{\sqrt{\text{var}(x^{(i)}) + \epsilon_{\text{norm}}}}, \forall i$$

2. Whiten inputs:

$$\begin{aligned} [V, D] &:= \text{eig}(\text{cov}(x)); \quad // \text{ So } VDV^T = \text{cov}(x) \\ x^{(i)} &:= V(D + \epsilon_{\text{zca}}I)^{-1/2}V^Tx^{(i)}, \forall i \end{aligned}$$

3. Loop until convergence (typically 10 iterations is enough):

$$\begin{aligned} s_j^{(i)} &:= \begin{cases} \mathcal{D}^{(j)T}x^{(i)} & \text{if } j == \arg \max_l |\mathcal{D}^{(l)T}x^{(i)}| \\ 0 & \text{otherwise.} \end{cases} \quad \forall j, i \\ \mathcal{D} &:= XS^T + \mathcal{D} \\ \mathcal{D}^{(j)} &:= \mathcal{D}^{(j)} / \|\mathcal{D}^{(j)}\|_2 \forall j \end{aligned}$$

http://blog.csdn.net/zouxy09

三、与稀疏特征学习的比较

上面我们提到，K-means能像ICA，sparse coding一样学习到方向性的边缘特征。那我们会这是偶然吗（因为边缘太常见了，所以一般的学习模型都可以发现它们）？！还是说这k-means本身就具有和ICA相似的稀疏分解的能力。因为每个 $s^{(i)}$ 只能包含一个非零元素（相应的心），所以k-means会尝试去学习一个可以很好的描述整个输入图像的聚类中心。因此，它也证一定可以学到和ICA或者sparse coding学到的滤波器一样。这些算法可以学习到真正的“分布”，因为一张图像可以通过一个字典的几列来联合描述，而不是只通过一列。尽管如此，k-n天时地利人和的时候依然能发现数据的稀疏投影，虽然作为一个聚类方法它好像没有明确的告诉我们它还能做这个。所以我们看到的结果也不是幻觉或者偶然。

虽然k-means和ICA和sparse coding有着经验性的相似，但它有一个很大的缺点：它发现数据投影方向的能力很大程度上取决于输入数据的维数和数量。如果数据的维数增加，我们就需要量的样本，才能得到好的结果。这点就比较难伺候了。例如，在实验中，对64x64的patch，我500,000个样本来训练k-means，结果还是很烂。我们只学习到了很少的边缘。在这种情况下，在很多的接近空数据点的聚类，所以需要非常多的数据才能让k-means发挥光与热。所以，在一个折衷了，是否要采用k-means，就要看我们的数据有多少维，我们能提供多少数据（一



1点点加盟要求



联系我们



请扫描二维码联系客服

✉ webmaster@csdn.net

☎ 400-660-0108

👤 QQ客服 🗨 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN版权所有

京ICP证09002463号

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

到相同的结果，比sparse coding要多很多）。对于适当的维度（例如100维），使用k-means的度比其他算法快那么多，那么我情愿花点力气去提供更多的样本给它。但对于很高维的数据，他算法例如sparse coding会比k-means性能要好，甚至速度也要更快。

4、在图像识别上的应用

在这里没什么好说的，k-means就作为一种学习特征的方法，取代一般的非监督学习过程特征学习那块就好了。可以参考UFLDL中的“卷积特征提取”和“池化”另外，文中也介绍选择参数的方法，很多论文也可以看到，这里也不说了。这里比较有趣一点是，文中说到，测的结果，如果我们的识别系统设计有大量有效的有标签训练数据的话，那么选择一个简单的前向编码器（特征映射函数）的话，效果都可以挺好。但如果我们的有训练数据不多，那一个复杂的编码函数会更好。例如，在大规模情况下，采用软阈值非线性函数：

$$f(x; \mathcal{D}, \alpha) = \max\{0, \mathcal{D}^T x - \alpha\}$$

就可以工作的比较好。其中， α 是个可调的常数。相反，复杂的sigmoid非线性函数效果就没了：

$$f(x; \mathcal{D}, b) = (1 + \exp(-\mathcal{D}^T x + b))^{-1}$$

5、构建深度网络

在上面一个章节中，我们是先从一些无标签数据中随机提取一些patches，然后用k-means其中的特征，得到一个字典。然后对于一个大的图像，我们用这个字典（特征提取器）去卷积每个局部感受野，从而提取整幅图像的特征。然后我们再通过一个pooling步骤去对提取到的特征降维。当然了，如果我们可以通过上面这些提取到的特征去学习更高层的特征，那就更好了。的方法就是，对无标签数据集 X 中，我们通过上面的方法得到pooling后的特征后，然后得到新集 Z ，我们再在这个数据库 Z 中用同样的特征学习过程去学习新的特征。从而得到第二层的特征有个问题，就是这个 Z 一般都是非常高维的，所以需要非常大的特征字典来学习（例如 $k=100$ 更多）。但这样，对k-means的要求也高了，我们需要提供更多的样本来训练。

该文提出了一种pair-wise “dependency test”的方法来改善这个问题。我们通过一个“关”来定义数据集 Z 中两个特征 z_j 和 z_k 的依赖性。我们需要在 Z 的特征表达之上学习更高级的特定“dependency test”，我们可以以一种相对简单的方式来选择合适的感受野：我们挑了一个 z_0 然后使用“dependency test”来寻找和 z_0 具有很强依赖性的 R 特征。然后只用这 R 特征作为k-means输入。如果我们选取的 R 足够小（例如100或者200），那么归一化和白化过后，再用k-means来一般都可以达到好的效果。因为输入的维数很小，所以我们只需要训练一百个左右的聚类中心，用少量的训练样本就可以运行k-means了。

详细的算法介绍和效果见原文。


六、总结

本文中，我们讨论了用k-means来搭建特征学习系统的一些关键点，现在总结如下：

- 1、需要先对数据进行均值和对比度的归一化。
- 2、使用白化来消去数据的相关性。注意白化参数的选择。
- 3、通过高斯噪声和归一化来初始化k-means的聚类中心。
- 4、使用抑制更新来避免空聚类和增加稳定性。
- 5、注意数据的维数和稀疏的影响。K-means通过寻找数据分布的“heavy-tailed”方向来找到输入稀疏投影。但如果数据没有经过适当的白化，或者数据维数太高了，或者存在无效的数据，那往往是不好的。
- 6、对于高维数据，k-means需要增加大量的样本来能得到好的效果。



1点点加盟要求



联系我们



请扫描二维码联系客服

✉ webmaster@csdn.net

☎ 400-660-0108

💬 QQ客服 💬 客服论坛

关于 招聘 广告服务 百度

©1999-2018 CSDN版权所有

京ICP证09002463号

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

Deep Learning 论文笔记之（一）K-means特征学习 - CSDN博客

7、一些外在的参数（例如pooling，编码方法的选择等等）对于系统性能的影响比学习算法大。所以最好先化时间和成本在通过更多的交叉检验的方法来挑选好的参数比花更多的成本在法上要明智的多。

8、对于图像识别来说，更多的聚类中心往往是有帮助的。当然，前提是我们要有足够的训练样

9、当有标签数据多的时候，采用一个简单的编码器。如果标签数据有门，例如每类只有百来样本），那么采用一个复杂的编码器会更好。

10、尽可能地使用局部感受野。K-means是否成功，其瓶颈在于输入数据维数。越低越好。法手动的选择好的局部感受野，那么尝试去通过一个自动的依赖性测试，dependency test来帮助中挑选出一组低维的数据。这对于深度网络来说是有必要的。

•


上一篇

运动检测（前景检测）之（二）混合高斯模型GMM

•


下一篇

Deep Learning论文笔记之（二）Sparse Filtering稀疏滤波



FS 飞速

光传输方案
轻松扩容 远距传输



您还没有登录,请[\[登录\]](#)或[\[注册\]](#)

查看评论



大木elina

11楼 2016-10-12 08:28发表

我的导师要求我学习elm，完全是小白，应该从哪部分学起呢



科研菜鸟的成长之路

Re: 2017-10-31 11:08发表

回复大木elina：ELM那么简单的算法。。



bluekv525

10楼 2016-04-21 21:06发表

[查看更多评论](#)

卷积神经网络CNN入门——代码

 s151506879 2016年07月27日 14:27 

Deep Learning论文笔记之（五）CNN卷积神经网络代码理解 zouxy09@qq.com <http://blog.csdn.net/zouxy09>

Deep Learning论文笔记之（六）Multi-Stage多级架构分析

Deep Learning论文笔记之（六）Multi-Stage多级架构分析zouxy09@qq.com<http://blog.csdn.net/zouxy09> 自了一些论文...

 zouxy09 2013年08月16日 17:48  32575

程序员不会英语怎么行？

北大猛男教你：不背单词和语法，一个公式学好英语
广告

Deep Learning（深度学习）学习笔记整理系列之（三）

Deep Learning（深度学习）学习笔记整理系列 zouxy09@qq.com <http://blog.csdn.net/zouxy09> 作者：Zouxy versi 01...

<https://blog.csdn.net/zouxy09/article/details/9982495>

6/9



1点点加盟要求



联系我们



请扫描二维码联系客服

✉ webmaster@csdn.net

☎ 400-660-0108

👤 QQ客服 💬 客服论坛

关于 招聘 广告服务 🌐 百度

©1999-2018 CSDN版权所有

京ICP证09002463号

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

Deep Learning 论文笔记之（一）K-means特征学习 - CSDN博客

 zouxu09 2013年04月09日 00:37 📖 532796

深度学习（二十四）矩阵分解之基于k-means的特征表达学习

本篇博文主要讲解2012年大神吴恩达所发表的一篇paper: 《Learning Feature Representations with K-means》。可能于k均值聚类，这个算法之前已经非常...

 hjimce 2015年12月29日 21:04 📖 5118

K-means聚类算法

 rex_huang61

额...想了很久不知道写什么好。最后还是觉得写K-means聚类算法好了。K-means聚类算法类 (prototype-based clustering) 中一种非常古老的...

加盟一只酸奶牛

一只酸奶牛加盟费多少

百度广告

在K均值聚类的时候发现在某一次迭代过程中，有一类变为了空类，如何解释这个现象

导致这种情况的原因有很多，下面是几个可能的方面: 1.数据本身的问题 比如虽然有一组体育类的新闻，但是这组数据实际娱乐类的距离比较近，那么很有可能导致结果体育类的分类个数为0。 ...

 lvjun93 2013年01月15日 22:16 📖 2203

Kmeans、Kmeans++和KNN算法比较

 loadstar_kun 2014年09月21日 12:38 📖


K-Means介绍 K-means算法是聚类分析中使用最广泛的算法之一。它把n个对象根据他们的属性分为k个聚类以便使的聚类满足：同一聚类中的对象相似度较高；而不同聚类中的对象相...

基本Kmeans算法介绍及其实现

 qll125596718 2012年11月30日 06:26 📖 1

1.基本Kmeans算法[1] 选择K个点作为初始质心 repeat 将每个点指派到最近的质心，形成K个簇 重新计算每个簇的质心 1. 发生变化或达到最大迭代次数时间复杂度：O(tKm...

数据预处理之数据归一化

 u012751110 2016年04月09日 11:20 📖

转载来自ufldl.stanford.edu/wiki/index.php/数据预处理 一、简单缩放 分为：最大值缩放和均值缩放 在简单缩放中，目的是通过对数据的每一个维度的值进行重新调...

特征选择(feature selection)

 u012328159 2016年12月31日 15:00 📖

特征选择 feature selection终于有时间把好久之前就想写的关于特征选择的基本介绍补上来了，主要想从以下几个方面介绍特征选择的动机-为什么要特征选择 2. 常见的特征...

程序员不会英语怎么办？

老司机教你一个数学公式秒懂天下英语

Deep Learning笔记之一：K-means特征聚类算法

聚类是数据挖掘中用来发现数据分布和隐含模式的一项重要技术，聚类分析是指事先不了解一批样品中的每个样品的先验知识，而唯一的分类依据是样品的特征，利用某种相似性度量的方法，把特征相同的或相近...

https://blog.csdn.net/zouxu09/article/details/9982495

7/9



1点点加盟要求



联系我们



请扫描二维码联系客服

✉ webmaster@csdn.net

☎ 400-660-0108

👤 QQ客服 🗣 客服论坛

关于 招聘 广告服务 🌐 百度

©1999-2018 CSDN版权所有

京ICP证09002463号

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

Deep Learning 论文笔记之（一）K-means特征学习 - CSDN博客

 gdut2015go 2015年06月03日 15:01 📄 8255

k-means多维数据聚类


2014年06月09日 22:33 7KB [下载](#)

使用稀疏化（字典学习）工具spams做分类

 qiao1245 2016年11月03日 15:40 🗉

spams简介SPAMS (SPArse Modeling Software)是一个为解决各种稀疏估计问题的开源工具箱，其主页为http://svl.gforge.inria.fr...

特征选择

 a1223031949 2016年10月06日 16:57

转载：http://blog.csdn.net/shanglianlm/article/details/49464445 1. 前言 1-1 问题：从 D 维特征中选择 dD 个特...

软件测试工程师的尴尬

软件测试工程师


百度广告

kmeans特征提取原理，详细代码图解分析

 whiteinblue 2014年05月28日 22:02 🗉


一、数据、数据预处理 2.1数据：下载的CIFAR-10数据库；其中的每个data_batch都是10000x3072大小的，即有1w个片，每个图片都是32*32且rgb三通道的，这里的...

Kmeans++及字典学习和图像分割

 llcchh012 2015年06月05日 19:45 🗉

1. Kmeans++Kmeans 中对聚类中心的初始化比较敏感，不同的初始值会带来不同的聚类结果，这是因为 Kmeans 仅仅函数求近似最优解，不能保证得到全局最优解。在常规的 Kmeans...

机器学习（十三）k-svd字典学习

 hjimce 2016年03月05日 18:44 🗉

给定训练数据Y，Y的每一列表示一个样本，我们的目标是求解字典D的每一列(原子)。K-svd算法，个人感觉跟k-means是k-means的一种扩展，字典D的每一列就相当于k-means的聚类中...

k-means优化 & k-means距离的选择 &k-medoids对比

一、k-means：在大数据的条件下，会耗费大量的时间和内存。优化k-means的建议：1、减少聚类的数目K。因为，每要跟类中心计算距离。2、减少样本的特征维度。比如说，通过PCA等进行...

 lujiandong1 2016年05月05日 11:24 🗉 4878

技术外文文献看不懂？教你一个公式秒懂英语

不背单词和语法，一个公式学好英语

【机器学习-斯坦福】学习笔记10 K-means聚类算法 很全面

K-means也是聚类算法中最简单的一种了，但是里面包含的思想却是不一般。最早我使用并实现这个算法是在学习韩爷爷数据挖掘的书中，那本书比较注重应用。看了Andrew Ng的这个讲义后才有些明白K...

 zkl99999 2015年06月11日 22:09 🗉 2651

【OpenCV学习笔记 020】K-Means聚类算法介绍及实现

https://blog.csdn.net/zouxy09/article/details/9982495

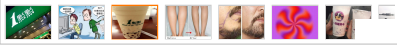
8/9

Opencv Kmeans聚类算法介绍及实现

 DaveBobo 2016年12月01日 17:03  6705




1点点加盟要求



联系我们



请扫描二维码联系客服
 webmaster@csdn.net
 400-660-0108
 QQ客服  客服论坛

关于 招聘 广告服务  百度
©1999-2018 CSDN版权所有
京ICP证09002463号

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


31



