

# 矩阵分解之基于k-means的特征表达学习

:**= = 6** 

[...]

评论(2)

5118人阅i

登录 注册

₩ 摘要视图

收藏

原文地址: http://blog.csdn.net/hjimce/article/details/50429317

作者: hjimce

# 一、相关理论

因为最近搞的项目,要么没有数据、要么训练数据很少,精度老是提高不上去,另一方面 到网上一篇推荐文章:《Convolutional Clustering for Unsupervised Learning》,于是就操起家 习CNN无监督学习方面的文章,看能不能把精度提高上去。

本篇博文主要讲解2012年大神吴恩达他们曾发表的一篇 《Learning Feature Representations with K-means》。可能大家对于k均值聚类,这个算法之 非常熟悉了,觉得这个算法是一个很简单的算法,估计会觉得没什么可以学的东西,因此对其 趣。然而其实很多人只是表层理解的k-means算法,对于k-means怎么用于图片分类无监督特 学习、k-means与矩阵分解的关系、与压缩感知相关理论之间的关系,却极少有人去深究理解 的我,看到这篇文献题目的时候,曾心想:"这种文献肯定很水,连k-means这种算法也好意思 发表文章"。好吧,这也许就是所谓的:坐井观天,惭愧啊。人生只有意识到自己是菜鸟的时候 去努力,才会进步。

本篇博文我将从k-means与矩阵分解开始讲起,矩阵分解推导部分主要参看文献:《k-Clustering Is Matrix Factorization》,对公式的推导过程不感兴趣的可以跳过,个人感觉推导 燥乏味。矩阵分解部分是为了对k-means有进一步的了解。因为后面要讲的特征表达学习,文 的公式也是一个矩阵分解的公式。第三部分,我们将详细讲解文 《Learning Feature Representations with K-means》,深入学习如何利用k-means算法进行 征学习,并提高图片分类的精度,领略k-means的牛逼之处。

## 二、k-means与矩阵分解

这几天刚好看到这篇paper:《k-Means Clustering Is Matrix Factorization》,因此就用这篇文 思路讲解。最近突然领悟到矩阵分解威力好强大,好多机器学习的算法都根源于矩阵分解,如 NMF、ICA、LDA、稀疏编码等,所以就决定对矩阵分解的奥义做一个总结,决定写一个系列 做一个总结,不然算法一多,就会乱掉。下面的的讲解,主要翻译自文献:《k-Means Clust Matrix Factorization ».

#### 1、k-means算法回顾

k均值算法可以说是机器学习中,最常用的一个聚类算法,给定数据集X∈Rm,我们的目的前 们分成C1,C2.....CK,其中k就是我们要聚类的个数。我们知道k-means算法的原理就是最小 的函数:

$$\sum_{i=1}^k \sum_{j=1}^n \langle z_{ij} \| x_{j} - \mu_i \|^2$$

其中,xj是每一个数据点,ui是聚类中心,当xj聚类为Ci时,那么zij的值就是1,如果不是那么 0,即:



关于 招聘 广告服务

©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

▲ QQ客服 ● 客服论坛

\* 百度

个人简介

声明:博文的编写,主要参考网上资料,并结合个 人见解,仅供学习、交流使用,如有侵权,请联系 博主删除,原创文章转载请注明出处。博主qq:1 393852684。内推阿里,欢迎大牛砸简历.

#### 博客专栏



深度学习 文章:68篇

阅读:942030

#### 文章分类

图像处理 (19)

机器学习 (20)

(87) 深度学习

数据挖掘 (4) 基础知识 (29)

图形处理 (14)

算法移植优化 (12)

数据结构 (18)

#### 文章存档

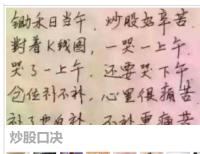
2018年3月 (6)

2018年2月 (12)

2018年1月

2017年10月 (4)

2017年9月 (1)















请扫描二维码联系客服

webmaster@csdn.net

**2**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

深度学习(七十三)pytorch学... SyGoing: 我也在学不过的确感觉这个框架比ten sorflow方便 这么多框架 已经学不过来了

深度学习(十七)基于改进Coars... CM\_Kun :您好,在校大学生一枚,能不能提供 一下源码,以便学习 2858591047@qq.com;

深度学习(七十-- ) 3D CNN时容

SyGoing : 师兄 我现在就在做行为识别这一块 C3D貌似就是想图像分类,对视频片段分类,但是 真实场景中,动作开...

深度学习(二十二)Dropout浅... b52\_bomber : [reply]weixin\_40717708[/reply] 请看这里:https://www.j...

深度学习(二十二)Dropout浅... b52\_bomber : [reply]weixin\_40717708[/reply] 不需要搞的这么麻烦,上万次测试你不可...

深度学习 (二十七)可视化理解卷积神... dta0502 : 博主您好,我想问下为什么反激活还 是Relu,那样原来小于0的那部分不是丢失了吗?

х	早	ij.	麦	ž	į

$$z_{ij} = \begin{cases} 1, & \text{if } x_j \in C_i \\ 0, & \text{otherwise.} \end{cases}$$

## 2、证明

## (1)证明题目

凸

开始前我们先明确一下我们的目标,我们的目标是证明k-means的目标函数可以被分解成如下

$$\sum_{i=1}^{k} \sum_{j=1}^{n} z_{ij} \left\| x_{j} - \mu_{i} \right\|^{2} = \left\| X - MZ \right\|^{2}$$

其中,X是数据样本矩阵,每一列表示一个样本。M是聚类中心矩阵,它的每一列就是我们的 小vui, Z表示一个编码矢量:

$$z_{ij} = \begin{cases} 1, & \text{if } x_j \in C_i \\ 0, & \text{otherwise.} \end{cases}$$

就像上初中的时候,考试考证明题一样,我们的目标就是要证明上面的等式成立。也就是k-m 过程就相当于矩阵分解的过程。如果对证明过程不感兴趣的可以跳过,因为证明过程确实是 的。而且这也不是本篇博文的重点,这只是为了让大家先对k-means的求解其实是可一个归结 矩阵分解问题的。

#### (2)定义

为了后面推导方便起见,我们先对一些代号,做一下定义。我们定义:数据样本矩阵X,X的每 示一个样本xi,我们把矩阵X的L2范数定义为所有样本向量的长度平方和,即:

$$\left\|X\right\|^2 = \sum_{l,j} \left.x_{lj}^2 = \sum_{j} \left\|x_j\right\|_{\text{nttp:}}^2 = \sum_{l,j} \left.x_j^T x_j = \sum_{\text{net}_j} \left(X^T X\right)_{jj} = \text{tr}\left[X^T X\right]_{\text{net}_j}^T = \sum_{l,j} \left(X^T X\right)_{jj} = \text{tr}\left[X^T X\right]_{\text{net}_j}^T = \sum_{l,j} \left(X^T X\right)_{lj}^T = \sum_{l,j} \left(X$$

OK,啰嗦了这么多,我们要大干一场了,下面开始证明等式:

$$\sum_{i=1}^{k} \sum_{j=1}^{n} z_{ij} \left\| x_j - \mu_i \right\|^2 = \left\| X - MZ \right\|^2$$

成立。

#### (3)证明成立

我们先把等式左边展开:

$$\begin{split} \sum_{i,j} z_{ij} \left\| x_j - \mu_i \right\|^2 &= \sum_{i,j} z_{ij} \left( x_j^T x_j - 2 x_j^T \mu_i + \mu_i^T \mu_i \right) \\ &= \underbrace{\sum_{i,j} z_{ij} | x_j^T x_j / + 2}_{T_1} \underbrace{\sum_{i,j} z_{ij} | x_j^T \mu_i}_{T_2} + \underbrace{\sum_{i,j} z_{ij} | \mu_i^T \mu_i}_{T_2}. \end{split}$$

然后上面的T1、T2、T3项我们可以进一步化简为:

$$T_1 = \sum_{i,j} z_{ij} \, x_j^T x_j = \sum_{i,j} z_{ij} \, ig\| x_j ig\|^2$$
 
$$= \sum_j ig\| x_j ig\|_{ ext{tp://blog.csdn.net/}}^2$$
 
$$= ext{tr} ig[ X^T X ig]$$





请扫描二维码联系客服

webmaster@csdn.net

**2**400-660-0108

■ QQ客服 ● 客服论坛

关于 招聘 广告服务 <sup>™</sup> 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

$$T_{2} = \sum_{i,j} z_{ij} x_{j}^{T} \mu_{i} = \sum_{i,j} z_{ij} \sum_{l} x_{lj} \mu_{li}$$

$$= \sum_{j,l} x_{lj} \sum_{i} \mu_{li} z_{ij}$$

$$= \sum_{j,l} x_{lj} (MZ)_{lj}$$

$$= \sum_{j} \sum_{l} (X^{T})_{jl} (MZ)_{lj}$$

$$= \sum_{j} (X^{T}MZ)_{jj}$$

$$= \operatorname{tr}[X^{T}MZ]$$

$$T_{3} = \sum_{i,j} z_{ij} \mu_{i}^{T} \mu_{i} = \sum_{i,j} z_{ij} \|\mu_{i}\|^{2}$$

$$= \sum_{i} \|\mu_{i}\|^{\frac{1}{2}p://b\log. \operatorname{csdn. net/}}$$

我们把**等式右边也展开**:

$$\begin{aligned} \left\| \boldsymbol{X} - \boldsymbol{M} \boldsymbol{Z} \right\|^2 &= \operatorname{tr} \left[ \left( \boldsymbol{X} - \boldsymbol{M} \boldsymbol{Z} \right)^T \left( \boldsymbol{X} - \boldsymbol{M} \boldsymbol{Z} \right) \right] \\ &= \underbrace{\operatorname{tr} \left[ \boldsymbol{X}^T \boldsymbol{X} \right]}_{T_4} - 2 \underbrace{\operatorname{tr} \left[ \boldsymbol{X}^T \boldsymbol{M} \boldsymbol{Z} \right]}_{T_5} + \underbrace{\operatorname{tr} \left[ \boldsymbol{Z}^T \boldsymbol{M}^T \boldsymbol{M} \boldsymbol{Z} \right]}_{T_6} \end{aligned}$$

这个时候,我们可以看到T4=T1,T2=T5。因此接着我们只需要证明:T3=T6,证明过程如下:

T6项等价于:

$$\mathrm{tr}ig[oldsymbol{Z}^Toldsymbol{M}^Toldsymbol{M}oldsymbol{Z}oldsymbol{Z}^Tig]_{\mathrm{blog.\,csdn.\,net}/}$$

据此,我们可以进一步推导:

$$\begin{split} \operatorname{tr} \big[ \boldsymbol{M}^T \boldsymbol{M} \boldsymbol{Z} \boldsymbol{Z}^T \big] &= \sum_i \left( \boldsymbol{M}^T \boldsymbol{M} \boldsymbol{Z} \boldsymbol{Z}^T \right)_{ii} \\ &= \sum_i \sum_l \left( \boldsymbol{M}^T \boldsymbol{M} \right)_{il} (\boldsymbol{Z} \boldsymbol{Z}^T)_{li} \\ &= \sum_i \left( \boldsymbol{M}^T \boldsymbol{M} \right)_{ii} (\boldsymbol{Z} \boldsymbol{Z}^T)_{ii}^{\operatorname{sdn. net}/} \\ &= \sum_i \left\| \boldsymbol{\mu}_i \right\|^2 n_i \end{split}$$

其中 $Z*Z^T$ 是一个对角矩阵。至此T3=T6证明成立,因此我们最后所要证明的等式也就成立毕。

啰嗦了这么多,无非就是想要说一句话:k-means也是一个矩阵分解的求解过程。

## 三、特征表达学习

OK,言归正传,前面啰嗦了一大堆,仅仅只是为了说明一个问题:k-means的求解过程是阵分解的过程。本部分要讲的利用k-means做特征学习,才是我们要关注的东西。接着我们要西就是文献:《Learning Feature Representations with K-means》到底是怎么实现无监督习。本部分同时也参考了前辈zouxy的博文:http://blog.csdn.net/zouxy09/article/details/9982感谢前辈。

开始前先要讲一下spherical K-means算法,这个算法与我们平时学到的k-means算法稍微同,是k—means算法的一个进化版。所以需要先熟悉一下这个算法。spherical K-means故名意思应该是我们的数据点集全部分布在一个球面上,然后进行聚类。因此在采用这个算法的时





请扫描二维码联系客服

webmaster@csdn.net

**2**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

要对每个数据点归一化处理,使得我们要聚类的数据,分布在一个球面上,然后我的个人理解球面聚类,应该是用了余弦相似度度量替代了欧式距离度量(具体没有细看,只是个人浅薄理能有误)。

## 1、字典与矢量量化

首先文献采用的k—means算法是:spherical K-means,其满足如下条件:

minimize 
$$\sum_{i} ||\mathcal{D}s^{(i)} - x^{(i)}||^{\frac{|\mathcal{D}|^{2^{2}}}{\square}}$$
subject to 
$$||s^{(i)}||_{0} \leq 1, \forall i$$
and 
$$||\mathcal{D}^{(j)}||_{2} = 1, \forall j$$

其中,D我们由称之为字典,字典这个词把我难住了,我不是信息专业的,从没听过字典,好的名字。然而其实说白了D就是一个变换矩阵。D(j)就是表示D的第一句,D的每一列就是1值的聚类中心向量,因为我们的聚类是在单位球面上,所以聚类中心的模长等于1。

S又称之为特征向量编码,当D的第j列(聚类中心)与x(i)最近的时候,s(i)对应的元素不为0,部为0。这个就是对x(i)的一个编号罢了,比如我们对X进行聚类,聚类为5个类,最后如果xi被第3个类别,那么编码矢量si=(0,0,1,0,0)。编码这个词听起来好像也有点牛逼的样子,其实可以把s理解为特征。而k-means特征提取的过程,说白了就是计算s。

K-mean无监督学习的过程,说白了就是为了学习字典D,也就是每个聚类中心。然后当我们个新数据x的时候,我们就用D去计算S,计算公式如下:

$$s_j^{(i)} = \begin{cases} \mathcal{D}^{(j)\top} x^{(i)} & \text{if } j == \underset{l}{\operatorname{arg\,max}} |\mathcal{D}^{(l)\top} x^{(i)}| \\ 0 & \text{http://blag.csdn.net/} \end{cases}$$

简单理解一下上面的那个符号:argmax。因为我们是在一个球面上,两点间如果最近,那么该是最大的(点积、余弦)。

# 2、算法流程

至今为止,我们都还没扯到图像,扯到图像的特征学习,接着我们就要开始结合实际应用,means的特征学习过程。

## (1)数据获取

对于图像,我们要如何进行k-means无监督特征学习?我们从一堆的图片中,随机的抽取出一图片块(16\*16大小的块),然后把这些16\*16的图片块(这个是灰度图),拉成一个一维的矿也就是256维的特征向量,这就是我们的训练数据了。文献抽取出了100000个16\*16的图片块字典学习。

#### (2)数据预处理

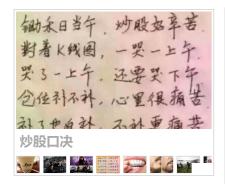
a、亮度和对比度归一化。采用如下公式对样本xi进行归一化:

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

这个简单,搞过机器学习的人都知道,上面公式分母加了一个数:10,是为了避免除数是0。

#### b、白化

这个我们可以采用PCA白化也可以采用ZCA白化,不过我们需要注意的是:白化的过程对于我果非常重要,所以千万不要忘了白化。这个文献对是否采用白化进行了对比:





请扫描二维码联系客服

webmaster@csdn.net

**2**400-660-0108

▲ QQ客服 ● 客服论坛

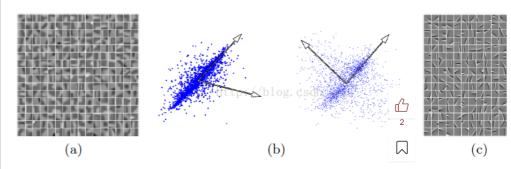
关于 招聘 广告服务 \*\*\* 百度 ©1999-2018 CSDN版权所有 京ICP证09002463号

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



左边是没用使用白化,聚类出来的中心。右边则是我们采用白化,学习100 约字典。

(3)最后就是spherical k-means的聚类过程了。

最后算法总流程如下:

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:

$$[V, D] := \operatorname{eig}(\operatorname{cov}(x)); \ // \text{ So } VDV^{\top} = \operatorname{cov}(x)$$

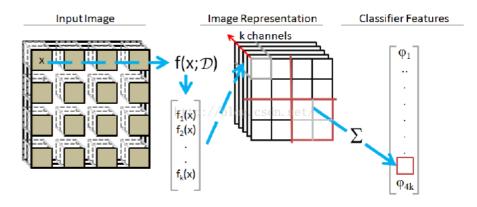
$$x^{(i)} := V(D + \epsilon_{\operatorname{zca}}I)^{-1/2}V^{\top}x^{(i)}, \forall i$$
http://blog.csdn.net/

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

$$\begin{split} s_j^{(i)} &:= \begin{cases} \mathcal{D}^{(j)\top} x^{(i)} & \text{if } j == \underset{l}{\arg\max} |\mathcal{D}^{(l)\top} x^{(i)}| \\ 0 & \text{otherwise.} \end{cases} \\ \mathcal{D} &:= X S^\top + \mathcal{D} \\ \mathcal{D}^{(j)} &:= \mathcal{D}^{(j)} / ||\mathcal{D}^{(j)}||_2 \forall j \end{split}$$

## 3、图片识别应用

前面我们介绍了,给定一堆无标签图片,我们通过K-means算法,进行学习字典D。接着我们道它具体是怎么用于图片识别分类的。下面是算法的总流程:



(1)通过无监督学习,学习到一系列的字典D。

(2)接着我们利用学习到的字典,把一张输入图片,进行函数映射:

$$f: \mathcal{R}^n \to \mathcal{R}^k$$

我们可以把这个过程看成是卷积,比较容易理解。这个映射过程不一定是非重叠的。

(3)采用pooling进行降维。

(4)如果是构建多层网络,那么就重复上面的过程。如果单层网络,那么我们就后面接一个分类如svm。

后面就不再细讲了,个人感觉:http://blog.csdn.net/zouxy09/article/details/9982495 讲的很好自己再怎么细讲,也没有前辈讲的很细,自叹不如啊,所以本篇博文到此结束。





请扫描二维码联系客服

webmaster@csdn.net

▲ QQ客服 ● 客服论坛

**2**400-660-0108

关于 招聘 广告服务 酒 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

### 参考文献:

- 1、 《k-Means Clustering Is Matrix Factorization》
- 2. 《Learning Feature Representations with K-means》
- 3, http://blog.csdn.net/zouxy09/article/details/9982495
- 《An Analysis of Single-Layer Networks in Unsupervised Feature L
- 《Convolutional Clustering for Unsupervised Learning》

\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*\*作者:hjimce 时间:2015.12.12 联系QQ:13938<sup>-ეeo</sup>4 原创文章,转 

- 深度学习(二十三)Maxout网络学习
- 深度学习 (二十五)基于Mutil-Scale CNN的图片语义分割、法向量估计

#### 安全工程师证书2018年取消认证



消息显示,2018年注册安全工程师证,将取消认证,专家分析这是注安18年的大变革

您还没有登录,请[登录]或[注册]

#### 查看评论



ajason6208

1楼 2016-06-10 15:25发表

你好博主,我想請問一下「把一张输入图片,进行函数映射」這段我不太能夠理 解,希望可以給的具體的解釋,謝謝博主



回复ajason6208:这步算法说的简单一点就是:先把输入图片进行分 块,求取每个块的字典稀疏表示;也就是把一张输入图片映射成字典 稀疏表示

Re: 2016-06-10 18:12发表

## 【机器学习】K-Means 聚类是特殊的矩阵分解问题



**cyh24** 2015年12月26日 15:59 Q

【机器学习】K-Means 聚类是特殊的矩阵分解问题。 本博客是该论文《k-Means Clustering Is Matrix Factorization》 记。论文证明了传统的K-Means...

# NMF 非负矩阵分解 -- 原理与应用



🔞 gg 26225295 2016年04月21日 18:13 👊

1.原理发现写关于非负矩阵的博文还是蛮多的,还是以自己的角度总结一下自己的最近看的若干东西以及对非负矩阵分解 些资料链接。NMF,全称为non-negative matrix factoriza...

### 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

广告



# PCA, Kmeans, NMF和谱聚类之间的联系



本文对这几种算法进行统一分析和归纳。1. PCA与Kmeans降维-主成分分析(PCA)-第4.1 🏶 採类与Kmeans聚类-第5节kmeans假设数据服从高斯分布,所以对于非高斯分布的...





#### 利用k-means实现图像识别(车型识别)



夏 zxs1129 2018年01月16日 16:49

锄承日当午,炒股办车苦 對着 K线图,一哭一上午 哭3一上午,还要哭下年 仓佐补不补,心里很痛 引了中的补, 乙科重病共 炒股口决 And the second s

# 联系我们



请扫描二维码联系客服

webmaster@csdn.net **2**400-660-0108

▲ QQ客服 ● 客服论坛

关于 招聘 广告服务 酒 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

前段时间做了一个车型识别的小项目,思路是利用k-means算法以及词袋模型来做的。 这边只记录一下我那个项目的思路,核心思想是k-means算法和词汇树。

近年来图像识别的方法非常3

# 【机器学习】数据处理中白化Whitening的作用图解分析

之前在看斯坦福教程中whiteining这一章时,由于原始图像相邻像素值具有高度相关性,所以介数据信息冗余,对于白 的描述主要有两个方面:1,减少特征之间的相关性;2,特征具有相同的方差(协方...

🚫 LG1259156776 2016年04月12日 12:24 🔘 4184

# 软件测试工程师

软件测试工程师是做什么的

百度广告



 $\overline{\odot}$ 

简单易学的机器学习算法——K-Means算法 ❷ google19890102 2014年05月18日 15:25 □

一. 聚类算法的简介

# Coursera-吴恩达-机器学习-(编程练习7)K均值和PCA(对应第8周课程)

此系列为 Coursera 网站Andrew Ng机器学习课程个人学习笔记(仅供参考)课程网址:https://www.coursera.org/le hine-learning ...



🥡 malele4th 2018年01月16日 17:07 🕮 462

# coursera斯坦福机器学习公开课作业7(K-means聚类, PCA降维)



2015年12月17日 15:36 10.95MB 下载

#### 深入理解K-Means聚类算法



🍘 taoyangi8932 2016年12月18日 20:50 🕮

概述什么是聚类分析聚类分析是在数据中发现数据对象之间的关系,将数据进行分组,组内的相似性越大,组间的差别越 类效果越好。不同的簇类型聚类旨在发现有用的对象簇,在现实中我们用到很多的簇的类型,使用...

## k-means++和k-means||



一、k-means算法缺点: 1、需要提前指定k 2、k-means算法对初始化非常敏感 k-means++主要用来解决第二个缺点 二 ns++算法原理分析 k-means++算法...

## 程序员不会英语怎么行?

老司机教你一个数学公式秒懂天下英语



# 机器学习算法与Python实践之(五)k均值聚类(k-means)

机器学习算法与Python实践之(五)k均值聚类(k-means)zouxy09@qq.comhttp://blog.csdn.net/zouxy09 法与Python实践这个...

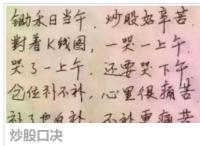
机

#### K-means分类



( klliu\_control 2016年03月22日 09:58

K-means聚类算法 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 是在学习韩爷爷那本数据挖掘的书中,那本书比较注重应用。看了...















请扫描二维码联系客服

webmaster@csdn.net **2**400-660-0108

▲ QQ客服 ● 客服论坛

关于 招聘 广告服务 酒 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

# K-means算法原理以及应用(R)



csqazwsxedc 2016年08月16日 23:58 🔘

K-means是一种无监督学习算法,是聚类算法中最简单的一种了。不同与一些分类的监督学习算法,比如逻辑回归、SVM 森林等,k-means聚类无需给定Y变量,只有特征X。下面是k-means算法原...

#### **K-Means**





**少 u010248552** 年11月08日 11:43 Q

K-Means版权声明:本文由callback发布于https://www.zybuluo.com/callback/note/942 🛴 本文可以被全部的转载: 使用,但请注明出处,如果有问...



## 高端网站设计

素马专业网站设计制作与开发机构

百度广告



# Kmeans和kmeans++



聚类算法分类划分聚类 Partitioning Methods: K-means 层次聚类 Hierachical Methods: Bottum-Up, Top-Down, BIR E,...

## Scikit-learn机器学习实战之Kmeans



mingtian715 2017年01月04日 09:27 □

Scikit-learn中Kmeans实例

# K-means算法原理与R语言实例



🌃 baimafujinji 2016年01月23日 17:57 🕮

聚类是将相似对象归到同一个簇中的方法,这有点像全自动分类。簇内的对象越相似,聚类的效果越好。支持向量机、神 讨论的分类问题都是有监督的学习方式,现在我们所介绍的聚类则是无监督的。其中,K均值(K...

### k-means+python | scikit-learn中的KMeans聚类实现( + MiniBatchKMeans)

之前一直用R,现在开始学python之后就来尝试用Python来实现Kmeans。 之前用R来实现kmeans的博客:笔记 | 多种 模型以及分群质量评估(聚类注意事项、使用技巧) ...



**Ѿ** sinat\_**26917383** 2017年04月19日 15:07 ♀ 26933

# 恭喜:一个公式教你秒懂天下英语

老司机教你一个数学公式秒懂天下英语



### 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。 发生变化或达到最大迭代次数时间复杂度:O(tKm...

#### k-means和bp神经网络要点总结



😨 u012073449 2016年07月27日 14:48 🚨

机器学习k-means算法和bp神经网络