

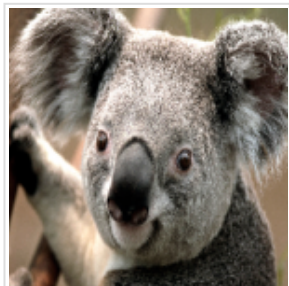
lujiandong1的专栏

目录视图

摘要视图

RSS 订阅

个人资料



BYR_jiandong



访问：132446次

积分：3525

等级：BLOG > 5

排名：第8044名

原创：220篇 转载：41篇

译文：0篇 评论：22条

【有奖投票】玩转Dragonboard 410c 的正确姿势 CSDN日报20170406 —— 《代码很烂，所以离职。》 Python数据分析与机器学习 博客搬家，有礼相送

机器学习中的内核方法 李政轩的视频笔记

标签：机器学习 核方法

2015-06-04 08:18

1873人阅读

评论(0)

收藏

举报

分类：机器学习 (32) ▼

版权声明：本文为博主原创文章，未经博主允许不得转载。

目录(?)

[+]

课程链接：<http://www.powercam.cc/slide/6552>。

Kernel Method 的基本思想：

关闭



文章搜索

文章分类

ACM DP问题 (7)

ACM其他文章 (2)

STL (5)

人生感悟 (1)

C++基础 (20)

C# (1)

ACM 贪心算法 (2)

ACM 哈希 (3)

二分法 (3)

STL 空间配置器 (1)

poj Trie树 (1)

poj 调试经验 (2)

POJ KMP (2)

poj 暴力法 (1)

C++调试 (3)

More Effective C++读书笔记 (3)

Effective C++读书笔记 (25)

机器学习 (33)

北邮人论坛 (2)

杂项 (6)

数学 (2)

opencv (3)

操作系统基本概念 (10)

Linux基础学习 (10)

程序优化 (3)

深入理解计算机系统 (5)

The Basic Idea

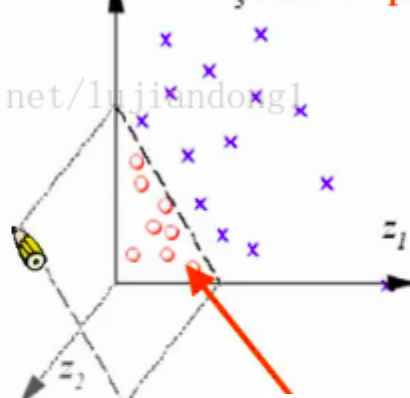


feature mapping $\phi: R^2 \rightarrow R^3$

$$(x_1, x_2) \mapsto (z_1, z_2, z_3) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$$

original space X

z_3 feature space H



$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1 \Rightarrow$$

Kernel的基本思想是，将低维空间不可分数据映射到高纬度的空间，比如说左是：

关闭

百度云

云计算新用户
注册送520元代金券

立即领取

[Python 基础 \(25\)](#)[SVM \(1\)](#)[推荐系统 \(1\)](#)[机器学习讲座笔记 \(1\)](#)[Kaggle学习笔记 \(10\)](#)[特征工程 \(2\)](#)[caffe教程及遇到问题的解决方案 \(2\)](#)[leetcode \(7\)](#)[自然语言处理 \(10\)](#)[哈工大SCIR 神经网络和深度学习转载 \(11\)](#)[tensorflow调研 \(37\)](#)

文章存档

[2017年03月 \(5\)](#)[2017年02月 \(9\)](#)[2017年01月 \(2\)](#)[2016年12月 \(8\)](#)[2016年11月 \(31\)](#)[展开](#)

阅读排行

[安装scikit-learn , win7 64 \(8805\)](#)[SVM的两个参数 C 和 gamma \(6803\)](#)[Python 列表的清空 \(3361\)](#)[import sys sys.path.append \(3062\)](#)[error LNK2019: 无法解析 \(3018\)](#)[连续特征离散化达到更好 \(2567\)](#)[配置caffe的python接口及 \(2194\)](#)

$$\frac{x_1^2}{a^2} + \frac{x_2^2}{b^2} = 1$$

, 将数据映射到三维空间, 就可以得到线性的分类面,

$$\left| \frac{1}{a^2} x_1^2 + \frac{1}{b^2} x_2^2 = 1 \Rightarrow \frac{1}{a^2} \cdot z_1 + 0 \cdot z_2 + \frac{1}{b^2} \cdot z_3 = 1 \Rightarrow (x_1, x_2) \Rightarrow (z_1, z_2, z_3) = (x_1^2, 0, x_2^2) \right|$$

总结：在低维空间线性不可分，映射到高维空间，线性可分的概率会增大，比如说，数据在一维空间线性可分的难度比二维空间线性可分的难度大，二维空间线性可分的难度比在三维空间线性可分的难度大。以此类推。

Kernal的另一个关键点事Kernal Function:

[关闭](#)

caffe 教程 Fine-tuning C: (1938)

机器学习中的内核方法 (1873)

DBN的训练过程 (1790)

评论排行

连续特征离散化达到更好 (4)

machine learning week6 (3)

error LNK2019: 无法解析 (2)

深入分析C++引用 (1)

人为什么会浮躁 (1)

Python 列表的清空 (1)

XGBoost Stopping to Av (1)

Andrew Ng的 Machine L (1)

配置caffe的python接口及 (1)

caffe 教程 Fine-tuning C: (1)

推荐文章

* Android安全防护之旅---带你把Apk混淆成中文语言代码

* TensorFlow文本摘要生成 - 基于注意力的序列到序列模型

* 创建后台任务的两种代码模式

* 一个屌丝程序员的人生 (六十)

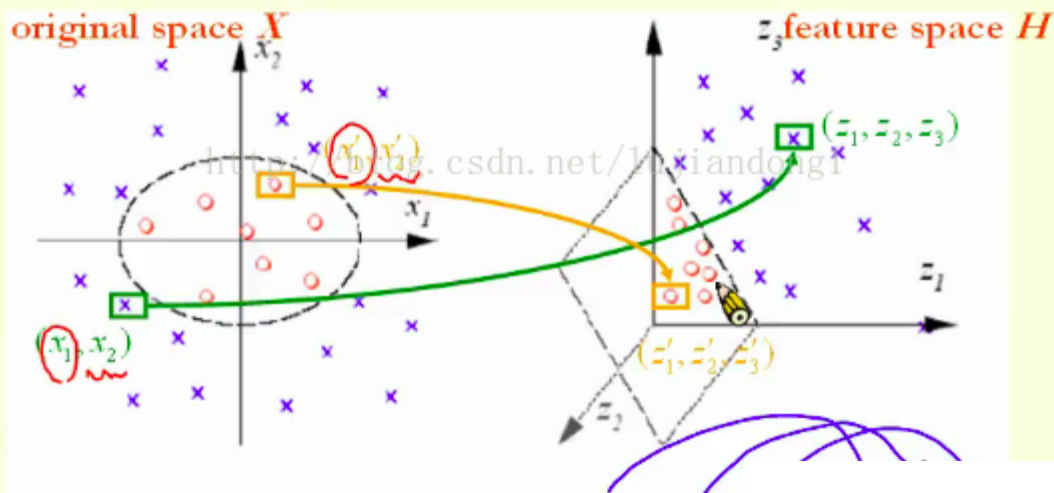
* WKWebView与js交互之完美解决方案

* 年轻人,“砖砖瓦瓦”不应该成为你的梦想!

The Basic Idea

feature mapping $\phi: R^2 \rightarrow R^3$

$$(x_1, x_2) \mapsto (z_1, z_2, z_3) = (x_1^2, \sqrt{2}x_1x_2, x_2^2)$$



$$\begin{aligned} \langle \phi(x_1, x_2), \phi(x'_1, x'_2) \rangle &= \langle (z_1, z_2, z_3), (z'_1, z'_2, z'_3) \rangle = \langle (x_1^2, \sqrt{2}x_1x_2, x_2^2), (x_1'^2, \sqrt{2}x_1'x_2', x_2'^2) \rangle \\ &= x_1^2x_1'^2 + 2x_1x_2x_1'x_2' + x_2^2x_2'^2 = (x_1x_1' + x_2x_2')^2 = (\langle x, x' \rangle)^2 := k(x, x') \end{aligned}$$

2维空间映射到3维空间的结果是：

$$(x_1, x_2) \rightarrow (z_1, z_2, z_3); (x'_1, x'_2, x'_3) \rightarrow (z'_1, z'_2, z'_3)$$

积为：

关闭

百度云

云计算新用户
注册送520元代金券

立即领取

最新评论

Andrew Ng的 Machine Learning fup1303: 写的挺好的, 可惜只有2和4, 有其他的课程笔记吗?

tensorflow MNIST数据集上简单: 倾城一少: 博主, MLP网络的全称是什么?

tensorflow CNN for mnist xjbada: 我运行这个代码为什么会出现这个错误呢*** TypeError: __init__() got an...

tensorflow中关于队列使用的实验 yuehanliushuang: very good

tesnsorflow 使用LSTM进行分类 qq_27590277: 为什么说我出错 TypeError: __init__() got an unexpected ...

error LNK2019: 无法解析的外部: m0_37640107: 多谢博主! 我的opencv用vs生成的时候也一直报这个错误, 上网其他方案都没用, 直到看了你的才发现...

machine learning week6 诊断机: Starry5cm: rand_seq=round(rand(1,i)*(m-1))+1;%生成i个随机序列 0~m这里改...

machine learning week6 诊断机: Starry5cm: rand_seq=round(rand(1,i)*(m-1))+1;%生成i个随机序列 0~m这里改...

人为什么会浮躁 annipiao: 相当有见地的分析, 受教了

tensorflow中dropout的用法, 防止 Wxlong: 博主你好, 你在文中说“train的时候才是dropout起作用的时候, train和test的时候不应...

$$\langle \phi(x_1, x_2), \phi(x'_1, x'_2) \rangle = K(x, x')$$

。也就是说, 我们的内核函数 $K()$, 计算的

是新空间的数据点之间的相似度。

内核方法总结: 1、将数据从第维映射到更高维。2, 映射到高维后, 数据可以线性可分。3、kernel function可以计

算高纬度空间的几何性质, 角度, 距离, 并不需要得到映射函数 ϕ 。

现在推导, 可以利用kernel function计算高纬度空间的距离, 角度。

The Basic Idea

Recall that

$$\begin{aligned} \langle \phi(x_1, x_2), \phi(x'_1, x'_2) \rangle &= \langle (z_1, z_2, z_3), (z'_1, z'_2, z'_3) \rangle = \langle (x_1^2, \sqrt{2}x_1x_2, x_2^2), (x_1'^2, \sqrt{2}x_1'x_2', x_2'^2) \rangle \\ &= x_1^2x_1'^2 + 2x_1x_2x_1'x_2' + x_2^2x_2'^2 = (x_1x_1' + x_2x_2')^2 = (\langle x, x' \rangle)^2 := \kappa(x, x') \end{aligned}$$

Distance in the Feature Space

$$\begin{aligned} \|\phi(x) - \phi(x')\|^2 &= (\phi(x) - \phi(x'))^T (\phi(x) - \phi(x')) \\ &= \phi(x)^T \phi(x) - 2\phi(x)^T \phi(x') + \phi(x')^T \phi(x') \\ &= \langle \phi(x), \phi(x) \rangle - 2\langle \phi(x), \phi(x') \rangle + \langle \phi(x'), \phi(x') \rangle \\ &= \kappa(x, x) - 2\kappa(x, x') + \kappa(x', x') \end{aligned}$$

Angle in the Feature Space

$$\begin{aligned} \langle \phi(x), \phi(x') \rangle &= \|\phi(x)\| \cdot \|\phi(x')\| \cos \theta \\ \Rightarrow \cos \theta &= \frac{\langle \phi(x), \phi(x') \rangle}{\|\phi(x)\| \cdot \|\phi(x')\|} = \frac{\langle \phi(x), \phi(x') \rangle}{\sqrt{\langle \phi(x), \phi(x) \rangle} \sqrt{\langle \phi(x'), \phi(x') \rangle}} \end{aligned}$$

关闭



云计算新用户
注册送520元代金券



立即领取

在国家 +0086 中国内地(China)

机号码 157

校验码 27978

用户名 速达移动

置密码 *****

认密码 *****

短信验证码接口

总结：在高维空间中要计算数据向量之间的角度，距离，只需要知道kernel function即可，并不需要知道 ϕ 。

现在引出内积矩阵，也叫做Gram Matrix ,Kernel Matrix。Kernel Matrix其实就是把所有的样本点映射到高维空间，然后在高维空间中计算内积，形成内积矩阵。内积矩阵的计算，也只需要，kernel function,不需要映射函数

$$K = \begin{bmatrix} \langle \phi(x_1), \phi(x_1) \rangle & \cdots & \langle \phi(x_1), \phi(x_N) \rangle \\ \vdots & \ddots & \vdots \\ \langle \phi(x_N), \phi(x_1) \rangle & \cdots & \langle \phi(x_N), \phi(x_N) \rangle \end{bmatrix}$$

$$= \begin{bmatrix} K(x_1, x_1) & \cdots & K(x_1, x_N) \\ \vdots & \ddots & \vdots \\ K(x_N, x_1) & \cdots & K(x_N, x_N) \end{bmatrix}$$

现在，举个小例子，讲解kernel function 的使用。

关闭



在国家 +0086 中国内地(China)

手机号码 157

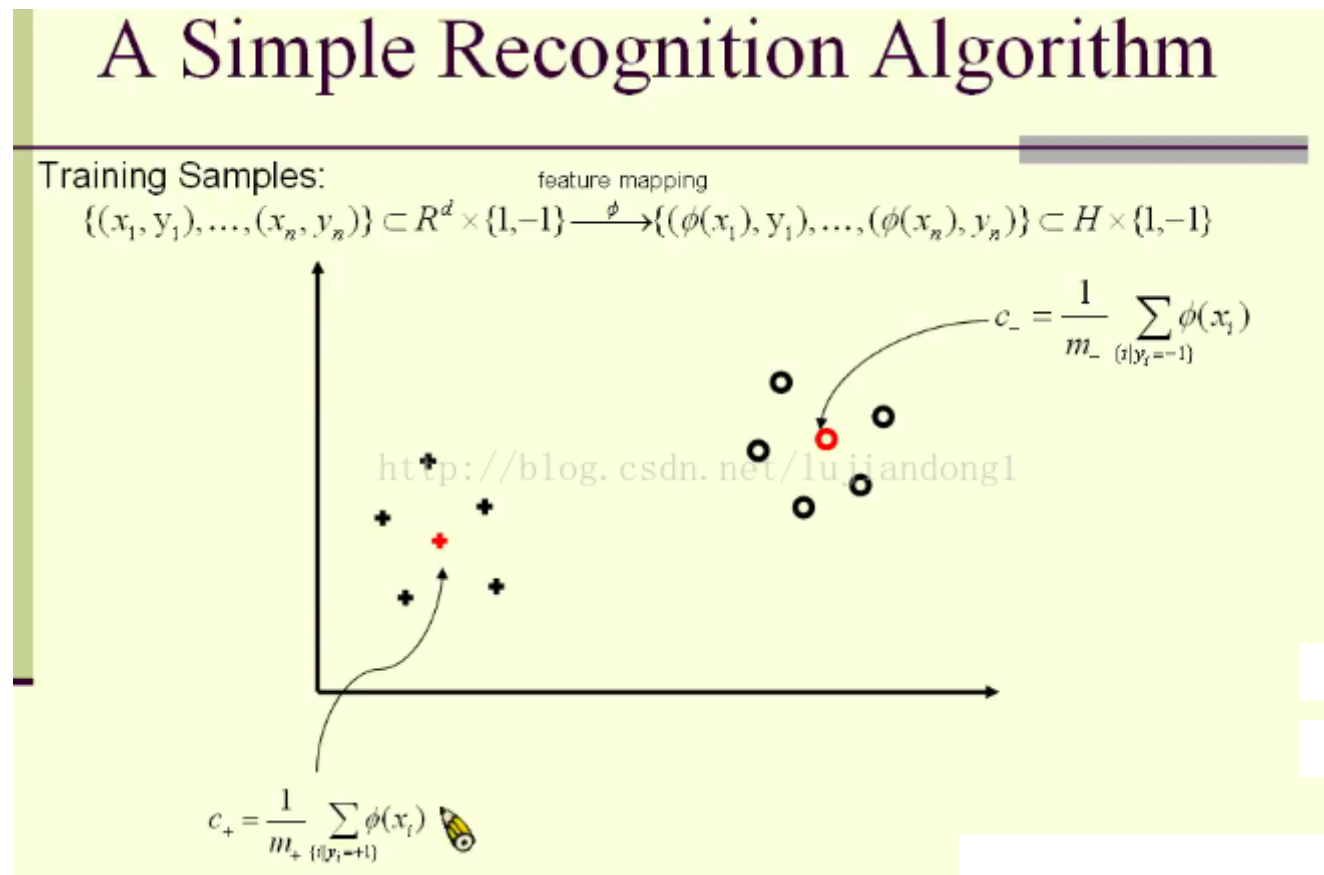
校验码 27978

用户名 速达移动

置密码

认密码

短信验证码接口



在地位空间中的数据点: (x_n, y_n) , 映射到高维空间得到 $(\phi(x_n), y_n)$

高维空间中的数据分布, 显而易见, 在高维的空间中, 数据已经是线性可分了

策函数呢。

关闭

百度云

云计算新用户

注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

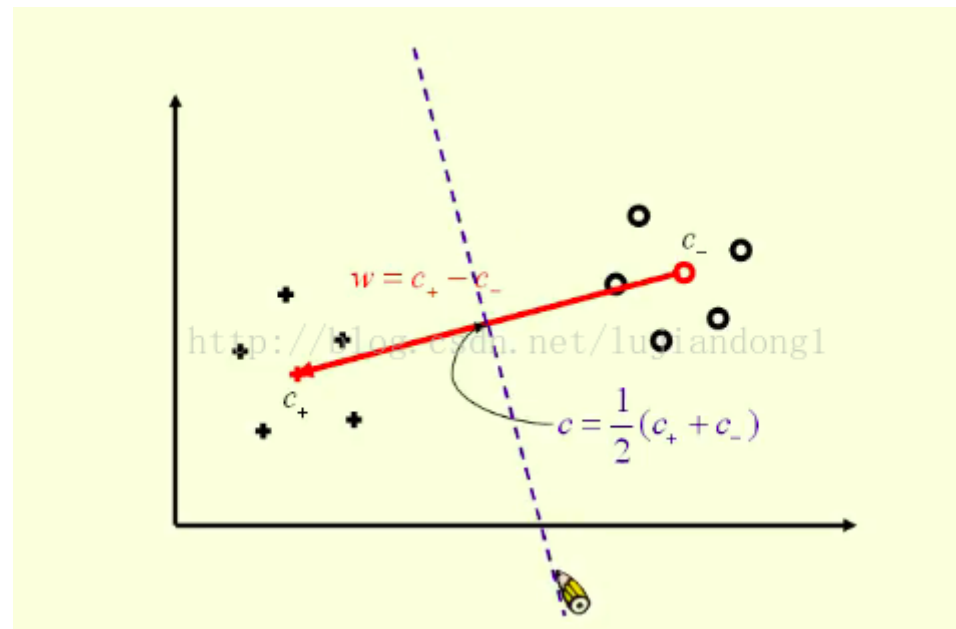
校验码 27978

用户名 速达移动

置密码 *****

认密码 *****

短信验证码接口



$$c_+ = \frac{1}{m_+} \sum_{\{i|y_i=+1\}} \phi(x_i)$$

其中， c_+ 是正类别（标签值为+1的数据的中心点）

$$c_- = \frac{1}{m_-} \sum_{\{i|y_i=-1\}} \phi(x_i)$$

c_- 是负类别（标签值为-1的数据的中心点）

1 → →、

关闭

。那么中间那条虚线就是对应于我们的分类平面，虚线和红线的交点坐标为，
那样就可以得到决策函数如下：

百度云

云计算新用户
注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

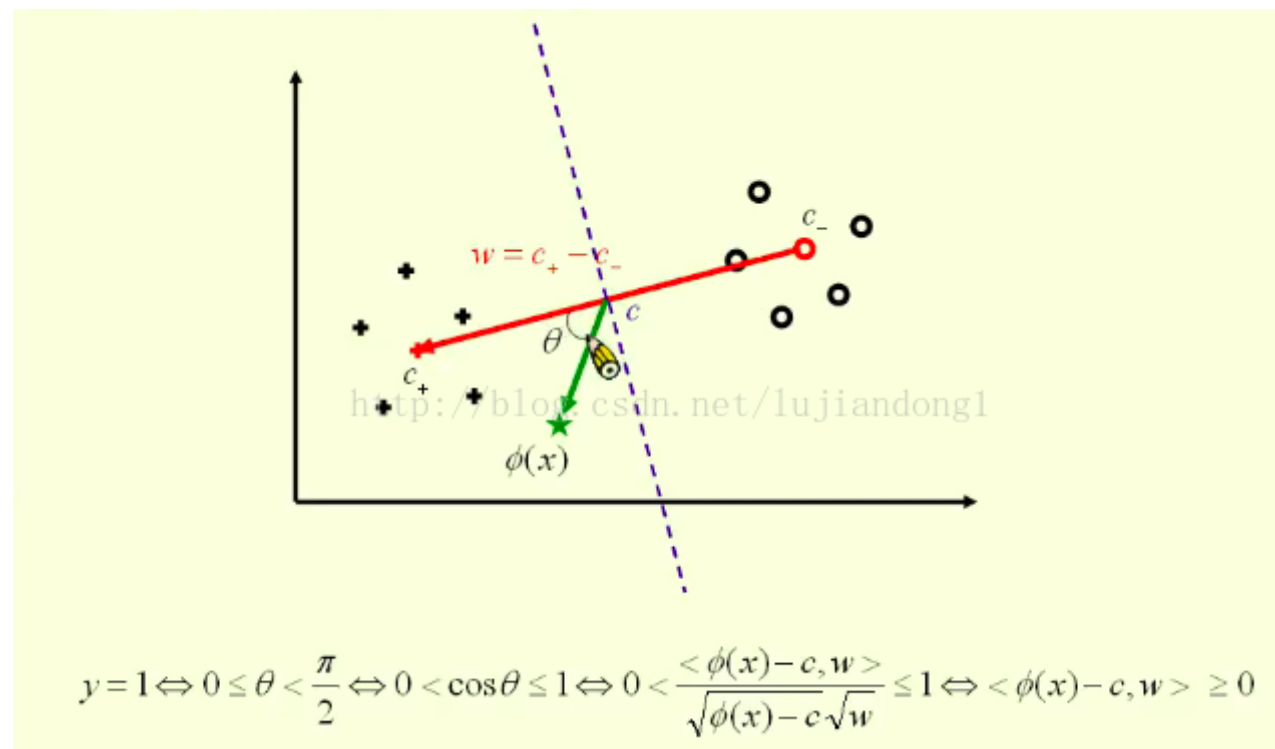
校验码 27978 重新发送(F)

用户名 速达移动

置密码 *****

认密码 *****

短信验证码接口



关闭

百度云

云计算新用户

注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

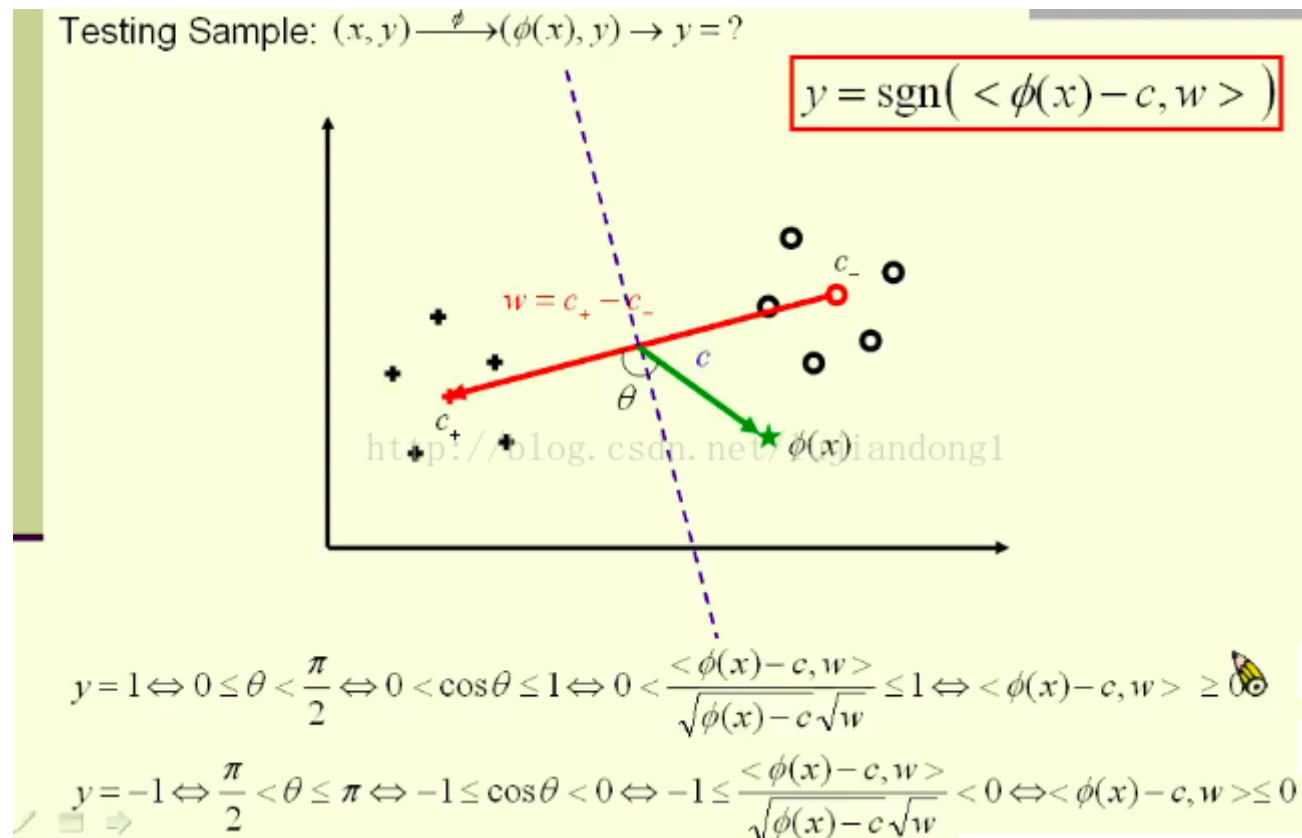
校验码 27978 重新发送(F)

用户名 速达移动

置密码 *****

认密码 *****

短信验证码接口



所有，要判断在高维的空间中，未知的点对应的表是啥，关键是计算出

$$y = \text{sgn}(\langle \phi(x) - c, w \rangle)$$

。那么如何用核函数计算

百度云

云计算新用户
注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

验证码 27978 重新发送(F)

用户名 速达移动

密码 *****

验证码 *****

短信验证码接口

$$\begin{aligned}
 \langle \phi(x) - c, w \rangle &= w^T (\phi(x) - c) = w^T \phi(x) - w^T c = (c_+ - c_-)^T \phi(x) - \frac{1}{2} (c_+ - c_-)^T (c_+ + c_-) \\
 &= \left(\frac{1}{m_+} \sum_{\{i|y_i=1\}} \phi(x_i) - \frac{1}{m_-} \sum_{\{i|y_i=-1\}} \phi(x_i) \right)^T \phi(x) - \left(\frac{1}{2} c_+^T c_+ - c_+^T c_- + \frac{1}{2} c_-^T c_- \right) \\
 &= \left(\frac{1}{m_+} \sum_{\{i|y_i=1\}} \phi(x_i)^T \phi(x) - \frac{1}{m_-} \sum_{\{i|y_i=-1\}} \phi(x_i)^T \phi(x) \right) - b \\
 &= \left(\frac{1}{m_+} \sum_{\{i|y_i=1\}} \kappa(x, x_i) - \frac{1}{m_-} \sum_{\{i|y_i=-1\}} \kappa(x, x_i) \right) - b
 \end{aligned}$$

其中，b也是可以用kernel function函数解决的，所以，只要我们知道kernel function 我们是可以用k来计算出y的。

现在进行总结：对于kernel function 映射函数是不一定需要的，只要有kernel function即可，但是不是随便定义一个kernel function 就可以找到相应的映射,其kernel matrix 是半正定矩阵。

关闭



在国家 +0086 中国内地(China)

手机号码 157

校验码 27978 重新发送(F)

用户名 速达移动

置密码 *****

认密码 *****

短信验证码接口

A function

$$\kappa: X \times X \rightarrow \mathbb{R}$$

satisfies the **finitely positive semi-definite** property if it is a symmetric function for which the matrices formed by restriction to any finite subset of the space X are positive semi-definite.

$$\begin{matrix} \{x_1, \dots, x_n\} \\ \uparrow \\ \text{Any} \end{matrix} \rightarrow K = \begin{bmatrix} \kappa(x_1, x_1) & \dots & \kappa(x_1, x_n) \\ \vdots & & \vdots \\ \kappa(x_n, x_1) & \dots & \kappa(x_n, x_n) \end{bmatrix} \geq 0$$

现证明 $k(x, z) = \langle x, z \rangle$ 对应的 kernel function 是正半定矩阵。

关闭

百度云

云计算新用户
注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

校验码 27978 重新发送(F)

用户名 速达移动

置密码 *****

认密码 *****

短信验证码接口

Finetely Positive Semi-definite Functions

$$(AB)^T = B^T A^T$$

■ Example: Show that

$$K(x, z) = \langle x, z \rangle$$

is a finetely positive semi-definite function.

Given any $\{x_1, \dots, x_n\}$.

$$\Rightarrow K = \begin{bmatrix} K(x_1, x_1) & \dots & K(x_1, x_n) \\ \vdots & & \vdots \\ K(x_n, x_1) & \dots & K(x_n, x_n) \end{bmatrix} = \begin{bmatrix} \langle x_1, x_1 \rangle & \dots & \langle x_1, x_n \rangle \\ \vdots & & \vdots \\ \langle x_n, x_1 \rangle & \dots & \langle x_n, x_n \rangle \end{bmatrix}$$

$$= \begin{bmatrix} x_1^T x_1 & \dots & x_1^T x_n \\ \vdots & & \vdots \\ x_n^T x_1 & \dots & x_n^T x_n \end{bmatrix} = \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} [x_1 \dots x_n]$$

$$\Rightarrow y^T K y = y^T \begin{bmatrix} x_1^T \\ \vdots \\ x_n^T \end{bmatrix} [x_1 \dots x_n] y = ([x_1 \dots x_n] y)^T ([x_1 \dots x_n] y) = \| [x_1 \dots x_n] y \|^2$$

Dual Representation:

在读关于核映射的paper时，我们经常看到如下式子：

$$f(x) = w^T \phi(x)$$

关闭

百度云

云计算新用户

注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

校验码 27978

用户名 速达移动

置密码 *****

认密码 *****

短信验证码接口

如果是预测问题，那么 $f(x)$ 就是预测值，如果是分类问题，那么 $f(x)$ 就是对应的类别。举个例子，对于刚才我们的分类问题，

$$y = \text{sgn}(\langle \phi(x) - c, w \rangle)$$

$$\Leftrightarrow y = \text{sgn}(f(x)), \quad f(x) = w^T \phi(x) - w^T c,$$

$$w^T = \frac{1}{m_+} \sum_{\{i|y_i=+1\}} \phi(x_i) - \frac{1}{m_-} \sum_{\{i|y_i=-1\}} \phi(x_i)$$

可以看出， w 是 $\phi(x_i)$ 的线性组合，所以，判决平面完全由训练样本所控制的。在统计上面有个定理，因为我们所掌握的只是训练数据，所以，我们计算处理的判决平面，投影方向等等，都是由训练样本所决定的。

$$w = \sum_{i=1}^N \phi(x_i)$$

关闭

百度云

云计算新用户
注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

校验码 27978 重新发送(F)

用户名 速达移动

重置密码

验证码

短信验证码接口

Dual Representation

$$w = \sum_{i=1}^N \alpha_i \phi(x_i)$$

$$f(\phi(x)) = w^T \phi(x) + b$$

$$= \left(\left(\sum_{i=1}^N \alpha_i \phi(x_i) \right)^T \phi(x) \right) + b$$

$$= \left(\sum_{i=1}^N \alpha_i \underbrace{\phi(x_i)^T \phi(x)}_{\text{inner product}} \right) + b$$

$$= \left(\sum_{i=1}^N \alpha_i \underbrace{K(x_i, x)}_{\text{inner product}} \right) + b$$

$\phi \rightarrow \text{dim} = 100$

顶

0

踩

0

关闭

百度云

云计算新用户
注册送520元代金券

立即领取

在国家 +0086 中国内地(China)

手机号码 157

验证码 27978 重新发送(7)

用户名 速达移动

密码 *****

确认码 *****

短信验证码接口



上一篇 [安装scikit-learn, win7 64位 \(装了一下午, 终于搞定了\) 解决ImportError: DLL load failed](#)

下一篇 [SVM的两个参数 C 和 gamma](#)

我的同类文章

机器学习 (32)

- LSTM的完整推导过程, 附上.. 2016-11-16 阅读 239
- 机器学习常见的六大错误 2016-11-02 阅读 111
- pandas使用get_dummies进... 2016-10-17 阅读 1721
- XGBoost Stopping to Avoid ... 2016-10-10 阅读 417
- k-means优化 & k-means距... 2016-05-05 阅读 769
- 机器学习经典博客链接 2015-06-15 阅读 409
- 随机森林的几个重要参数 2016-11-05 阅读 292
- 安装jpye出现_jpye错误的... 2016-10-18 阅读 194
- xgboost cross_validation&自.. 2016-10-11 阅读 896
- AUC的理解与应用场景 2016-07-23 阅读 167
- vowpal_wabbit是在单机上性... 2015-12-30 阅读 514

更多文章



云服务器性价比之王

猜你在找

[统计机器学习入门——线性模型选择与正则化2](#)

[python数据分析与机器学习实战](#)

[统计机器学习入门——线性模型选择与正则化1](#)

[统计机器学习入门——分类2](#)

[统计机器学习入门——重抽样方法](#)

[R语言与机器学习中的回归方法](#)

[转R语言与机器学习中的回归方](#)

[R语言与机器学习中的回归方法](#)

[机器学习实践系列之9 - 视频结](#)

[机器学习基础林軒田笔记之七](#)



云计算新用户 注册送520元代金券



立即领取

关闭

在国家 +0086 中国内地(China) 手机号 157 验证码 27978 用户名 验证码 验证码

短信验证码接口

新东方
XDF.CN

北京总部

美国夏令营
全真课堂体验 提升留学背景

广告

马上报名

查看评论

暂无评论

您还没有登录,请[登录](#)或[注册](#)

* 以上用户言论只代表其个人观点,不代表CSDN网站的观点或立场

核心技术类目

全部主题 Hadoop AWS 移动游戏 Java Android iOS Swift 智能硬件 Docker OpenStack
VPN Spark ERP IE10 Eclipse CRM JavaScript 数据库 Ubuntu NFC WAP jQuery
BI HTML5 Spring Apache .NET API HTML SDK IIS Fedora XML LBS Unity
Splashtop UML components Windows Mobile Rails QEMU KDE Cassandra CloudStack
coremail OPhone CouchBase 云计算 iOS6 Rackspace Web App SpringSide Maemo
Compuware 大数据 aptech Perl Tornado Ruby Hibernate ThinkPHP HBase Pure S
Angular Cloud Foundry Redis Scala Django Bootstrap

关闭

[公司简介](#) | [招贤纳士](#) | [广告服务](#) | [联系方式](#) | [版权声明](#) | [法律顾问](#) | [问题报告](#) | [合作伙伴](#) | [论坛反馈](#)[网站客服](#) [杂志客服](#) [微博客服](#) webmaster@csdn.net 400-600-2320 | 北京创新乐知信息技术有限公司 版权所有 | 江苏知

京 ICP 证 09002463 号 | Copyright © 1999-2016, CSDN.NET, All Rights Reserved

百度云

云计算新用户
注册送520元代金券

立即领取