



博客 (<http://blog.csdn.net/?ref=toolbar>) 学院 (<http://edu.csdn.net/?ref=toolbar>)

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

GitChat (<http://gitbook.cn/?ref=csdn>) 更多 ▾



6

scikit-learn工具学习 - random,mgrid,np.r_,np.c_, scatter, axis, pcolormesh, contour, decision_function



2015年10月28日 17:18:52

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

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

7322

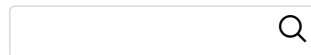
scikit-learn 练习题

题目: Try classifying classes 1 and 2 from the iris dataset with SVMs, with the 2 first features. Leave out 10% of each class and test prediction performance on these observations. (链接: http://scikit-learn.org/stable/tutorial/statistical_inference/supervised_learning.html)

官方提供的答案如文末代码段

通过这段源代码, 我们主要可以学习到如下几个常用函数的使用:

numpy 库



weixin_3506...

(<http://my.csdn.net/>)

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



crossky_jing (

+ 关注



(http://blog.csdn.net/crossky_jing)



雅马哈钢琴价格

3d立体画

码云

未开通

原创

粉丝

喜欢

(https://github.com/crossky_jing)

16

14

8

他的最新文章

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

scikit-learn工具学习 - cross_validation, 一条语句里用for循环 (http://blog.csdn.net/crossky_jing/article/details/49472657)

POJ1015 (http://blog.csdn.net/crossky_jing/article/details/28127351)



内容举报

poj 3253---Fence Repair (http://blog.csdn.net/crossky_jing/article/details/22729743)



返回顶部

POJ3349--Snowflake Snowflake s (http://blog.csdn.net/crossky_jing/arti)

```
import numpy as np
```

1、random

用法：产生伪随机数

样例：

```
np.random.seed(0) //产生以0为种子的伪随机数生成器
order_arr = np.random.permutation(100) //返回100个伪随机数，返回值是一个array
```

2、mgrid

用法：返回多维结构，常见的如2D图形，3D图形。对比np.meshgrid，在处理大数据时速度更快，且能处理多维（np.meshgrid只能处理2维）

用法：np.mgrid[第1维，第2维，第3维，...]

返回多值，以多个矩阵的形式返回，第1返回值为第1维数据在最终结构中的分布，第2返回值为第2维数据在最终结构中的分布，以此类推。（分布以矩阵形式呈现）

例如np.mgrid[X, Y]

样本 (i, j) 的坐标为 (X[i, j], Y[i, j])，X代表第1维，Y代表第2维，在此例中分别为横纵坐标。

例如1D结构（array），如下：

```
1 >>> pp = np.mgrid[-5:5]
2 >>> pp
3 array([-5., -2.5, 0., 2.5, 5.] )
```

例如2D结构（2D矩阵），如下：

cle/details/21934629)

POJ1804 (http://blog.csdn.net/crossky_jing/article/details/212899)



雅马哈钢琴价格



3d立体画



清理化粪池



他的热门文章

scikit-learn工具学习 - random,mgrid,np.r_ ,np.c_ , scatter, axis, pcolormesh, contour, decision_function - CSDN博客
7310

scikit-learn工具学习 - cross_validation , 一条语句里用for循环 (http://blog.csdn.net/crossky_jing/article/details/49472657)

6647

内容举报

poj1753解题报告 (http://blog.csdn.net/crossky_jing/article/details/9704397)

844

返回顶部

实现有大小限制的并发阻塞队列 (http://blog.csdn.net/crossky_jing/article/details/17784493)

```
1 >>> pp = np.mgrid[-1:1:2j,-2:2:3j]
2 >>> x , y = pp
3 >>> x
4 array([[ -1., -1., -1.],
5        [  1.,  1.,  1.]])
6 >>> y
7 array([[ -2.,  0.,  2.],
8        [ -2.,  0.,  2.]])
```



6

例如3D结构 (3D立方体), 如下:



768

poj1328--Radar Installation
dn.net/crossky_jing/article/
7)

497



雅马哈钢琴价格



平面设计学习

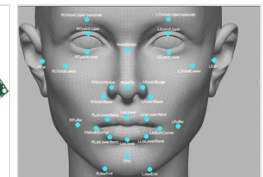


3d立体画

广告



回收内存条



人脸识别



内容举报



返回顶部

```
1 >>> pp = np.mgrid[-1:1:2j,-2:2:3j,-3:3:5j]
2 >>> print pp
3 [[[-1. -1. -1. -1. -1.]
4  [-1. -1. -1. -1. -1.]
5  [-1. -1. -1. -1. -1.]]
6
7  [[ 1.  1.  1.  1.  1.]
8  [ 1.  1.  1.  1.  1.]
9  [ 1.  1.  1.  1.  1.]]]
10
11
12 [[[-2. -2. -2. -2. -2.]
13  [ 0.  0.  0.  0.  0.]
14  [ 2.  2.  2.  2.  2.]]
15
16 [[-2. -2. -2. -2. -2.]
17  [ 0.  0.  0.  0.  0.]
18  [ 2.  2.  2.  2.  2.]]]
19
20
21 [[[-3. -1.5 0.  1.5 3.]
22  [-3. -1.5 0.  1.5 3.]
23  [-3. -1.5 0.  1.5 3.]]
24
25 [[-3. -1.5 0.  1.5 3.]
26  [-3. -1.5 0.  1.5 3.]
27  [-3. -1.5 0.  1.5 3.]]]]
```



6



3、 np.r_ , np.c_



内容举报



返回顶部

用法：concatenation function

np.r_按row来组合array，

np.c_按column来组合array

```

1  >>> a = np.array([1,2,3])
2  >>> b = np.array([5,2,5])
3  >>> //测试 np.r_
4  >>> np.r_[a,b]
5  array([1, 2, 3, 5, 2, 5])
6  >>>
7  >>> //测试 np.c_
8  >>> np.c_[a,b]
9  array([[1, 5],
10         [2, 2],
11         [3, 5]])
12 >>> np.c_[a,[0,0,0],b]
13 array([[1, 0, 5],
14         [2, 0, 2],
15         [3, 0, 5]])

```

matplotlib.pyplot 库

```
import matplotlib.pyplot as plt
```

1、scatter

用来画散点图的，对样本点着色。如下：X为一个n*2的矩阵，代表n个2维样本点，且每个样本点对应一个label y，用y来对颜色变量c赋值来区分颜色，按照cmap来布局。

```
plt.scatter(X[:, 0], X[:, 1], c=y, zorder=10, cmap=plt.cm.Paired)
```



内容举报



返回顶部

2、axis

用法：设置布局策略

例如：plt.axis('tight')，表明采用紧致方案，需要将样本的边缘作为画布的边缘。

3、pcolormesh

用法：类似np.pcolor，是对坐标点着色。

np.pcolormesh(X, Y, C, **kwargs)

例如有样本点 $(X[i, j], Y[i, j])$ ，对样本周围（包括样本所在坐标）的四个坐标点进行着色，C代表着色方案，kwargs里可以设置着色配置。



- 1 $(X[i, j], Y[i, j])$,
- 2 $(X[i, j+1], Y[i, j+1])$,
- 3 $(X[i+1, j], Y[i+1, j])$,
- 4 $(X[i+1, j+1], Y[i+1, j+1])$.

样例：plt.pcolormesh(XX, YY, Z>0, cmap=plt.cm.Paired)

4、contour

用法：画轮廓

样例：plt.contour(XX, YY, Z, colors=['k', 'k', 'k'], linestyle=['-', '-', '-'], levels=[-0.5, 0, 0.5])

svm 库

from sklearn import svm

1、decision_function



内容举报



返回顶部

用法：Distance of the samples X to the separating hyperplane. 即样本点到超平面的距离。

样例：

```

1 x_min = X[:, 0].min()
2 x_max = X[:, 0].max()
3 y_min = X[:, 1].min()
4 y_max = X[:, 1].max()
5
6
7 XX, YY = np.mgrid[x_min:x_max:200j, y_min:y_max:200j] //分别得到样本第1维和第2维的分布：
8 Z = clf.decision_function(np.c_[XX.ravel(), YY.ravel()]) //用np.c_()将XX, YY拉平后的两个array按照列合并(此时是n*2的举证，

```

附录（完整代码）：

http://scikit-learn.org/stable/_downloads/plot_iris_exercise.py (http://scikit-learn.org/stable/_downloads/plot_iris_exercise.py)



内容举报



返回顶部

```

1  """
2  =====
3  SVM Exercise
4  =====
5
6  A tutorial exercise for using different SVM kernels.
7
8  This exercise is used in the :ref: `using_kernels_tut` part of the
9  :ref: `supervised_learning_tut` section of the :ref: `stat_learn_tut_index`.
10 """
11 print(__doc__)
12
13 import numpy as np
14 import matplotlib.pyplot as plt
15 from sklearn import datasets, svm
16
17 iris = datasets.load_iris()
18 X = iris.data
19 y = iris.target
20
21 X = X[y != 0, :2]
22 y = y[y != 0]
23
24 n_sample = len(X)
25
26 np.random.seed(0)
27 order = np.random.permutation(n_sample)
28 X = X[order]
29 y = y[order].astype(np.float)
30
31 X_train = X[:.9 * n_sample]
32 y_train = y[:.9 * n_sample]

```



内容举报



返回顶部


```
33 X_test = X[.9 * n_sample:]
34 y_test = y[.9 * n_sample:]
35
36 # fit the model
37 for fig_num, kernel in enumerate(('linear', 'rbf', 'poly')):
38     clf = svm.SVC(kernel=kernel, gamma=10)
39     clf.fit(X_train, y_train)
40
41     plt.figure(fig_num)
42     plt.clf()
43     plt.scatter(X[:, 0], X[:, 1], c=y, zorder=10, cmap=plt.cm.Paired)
44
45     # Circle out the test data
46     plt.scatter(X_test[:, 0], X_test[:, 1], s=80, facecolors='none', zorder=10)
47
48     plt.axis('tight')
49     x_min = X[:, 0].min()
50     x_max = X[:, 0].max()
51     y_min = X[:, 1].min()
52     y_max = X[:, 1].max()
53
54     XX, YY = np.mgrid[x_min:x_max:200j, y_min:y_max:200j]
55     Z = clf.decision_function(np.c_[XX.ravel(), YY.ravel()])
56
57     # Put the result into a color plot
58     Z = Z.reshape(XX.shape)
59     plt.pcolormesh(XX, YY, Z > 0, cmap=plt.cm.Paired)
60     plt.contour(XX, YY, Z, colors=['k', 'k', 'k'], linestyles=['--', '-', '--'],
61                 levels=[-.5, 0, .5])
62
63     plt.title(kernel)
64     plt.show()
```



内容举报



返回顶部



发表你的评论



(http://my.csdn.net/weixin_35068028)

6



sherryxiaoyu (/sherryxiaoyu) 2017-08-08 16:36

1楼

(/sherryxiaoyu) 数据有4个特征，为什么只取前两个呢，是为了作图还是SVM只能做二维数据呢



回复

**numpy.c_和numpy.squeeze的用法 (http://blog.csdn.net/zhuzuwei/article/details/78022629)**

numpy.c_ = 将切片对象沿第二个轴（按列）转换为连接。例子：np.c_[np.array([1,2,3]), np.array([4,5,6])] Out[96]: array([...



zhuzuwei (http://blog.csdn.net/zhuzuwei) 2017年09月18日 19:13 5662

**numpy.c/numpy.r (http://blog.csdn.net/hjhjhkkk/article/details/78044538)**

1、numpy.c 官网介绍为https://docs.scipy.org/doc/numpy-1.8.1/reference/generated/numpy.c_.html#numpy.c_ ...



hjhjhkkk (http://blog.csdn.net/hjhjhkkk) 2017年09月20日 19:32 271

**Python薪酬到底多高？**

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？



内容举报



返回顶部

(http://www.baidu.com/cb.php?c=lgF_pyfqHmknjnvPjc0lZ0qnfK9ujYzP1nYPH0k0Aw-5Hc3rHnYnHb0TAq15HfLPWRznjb0T1YvPyc3nhmsnARLrjfdmHub0AwY5HDdnHf3njmsn1R0lgF_5y9YIZ0lQzq-uZR8mLPbUB48ugfEIAqspynETZ-YpAq8nWqdlAdxTvqdThP-5yF_UvTkn0KzujYk0AFV5H00TZcqn0KdpyfqHRLPjnvnfKEpyfqHc4rj6kP0KWpyfqP1cvrHnz0AqLUWYs0ZK45HcsP6KWThnqPjnzP1f)

用Python开始机器学习（4：KNN分类算法）(http://blog.csdn.net/lsidd/article/details/4135...)



KNN分类算法（K-Nearest-Neighbors Classification）是



lsidd (http://blog.csdn.net/lsidd) 2014年11月23日 17:24 40526

matplotlib中使用Pcolor绘制二维图 (http://blog.csdn.net/rumswell/article/details/6581845)

很容易的,下面的代码试一下就行了 #使用Pcolor绘制二维图

```
from pylab import *
a=arange(-2.0,2.001,0.5)
b=arange(-2.0,2.001,0.5)
x,...
```



rumswell (http://blog.csdn.net/rumswell) 2011年07月03日 14:52 8091



numpy.r_ (http://blog.csdn.net/huhu0769/article/details/52742395)



numpy.r_是将一系列的序列合并到一个数组中，调用是要用中括号[],而不是()。它有两种类型的参数，放在最前面的控制参数，已经后面逗号分开的要合并的序列。控制参数其他的情况都很好理解，本文...



huhu0769 (http://blog.csdn.net/huhu0769) 2016年10月06日 09:44 2004




内容举报



返回顶部


Python基础-numpy使用 (<http://blog.csdn.net/lishuandao/article/details/52438852>)

2 NumPy-快速处理数据 标准安装的Python中用列表(list)保存一组值,可以用来当作数组使用,不过由于列表的元素可以是任何对象,因此列表中所保存的是对象的指针。这样为了保存一个简单的[1...

 lishuandao (<http://blog.csdn.net/lishuandao>) 2016年09月05日 11:37 2326


numpy.zeros (np.zeros) 使用方法--python学习笔记31 (http://blog.csdn.net/qq_26948675...)

翻译: 用法: zeros(shape, dtype=float, order='C') 返回: 返回来一个给定形状和类型的用0填充的数组; 参数: shape:形状
dtype:数...

 qq_26948675 (http://blog.csdn.net/qq_26948675) 2017年01月10日 23:29 38788

Python科学绘图实例附代码 (<http://blog.csdn.net/u012176591/article/details/44700551>)

 Python绘图精简实例附代码作者: 金良 (golden1314521@gmail.com) csdn博客: <http://blog.csdn.net/u012176591Python>
绘图精简实例附代码 ...

 u012176591 (<http://blog.csdn.net/u012176591>) 2015年04月05日 10:47 9899

mgrid与ogrid (http://blog.csdn.net/Exupery_/article/details/76572546)

numpy.mgrid numpy.ogrid numpy.mgrid.numpy.lib.index_tricks.nd_grid的替代。在给定范围内返回一个ndarray,是网格上的点。import ...

 Exupery_ (http://blog.csdn.net/Exupery_) 2017年08月02日 10:52 968


Python闲谈 (一) mgrid慢放 (<http://blog.csdn.net/u013534498/article/details/51399035>)





内容举报


返回顶部

不论是利用Mayavi还是matplotlib绘制三维图表，里面都用到了numpy中的一个函数叫mgrid。本次博客我简单地讲一下mgrid是干什么用的，以及一个三维曲面是如何绘制出来的。...


 u013534498 (<http://blog.csdn.net/u013534498>) 2016年05月13日 21:53 5661


matplotlib中使用imshow绘制二维图 (<http://blog.csdn.net/rumswell/article/details/6581134>)

 这里所指的二维图,是二维矩阵数据的平面色彩显示# -*- coding:utf-8 -*-from matplotlib import mplimport matplotlib.pyplot as p...⁶


 rumswell (<http://blog.csdn.net/rumswell>) 2011年07月02日 23:42 49476

python之numpy的用法 (<http://blog.csdn.net/Cyuhong/article/details/50876364>)

 随机抽样 (numpy.random) 简单的随机数据 rand(d0, d1, ..., dn) 随机值 >>> np.random.rand(3,2) array([[0....

 Cyuhong (<http://blog.csdn.net/Cyuhong>) 2016年03月13日 11:01 1733


Python与数据分析1 (<http://blog.csdn.net/shengshengwang/article/details/52860491>)

 NumPy提供了两种基本的对象：ndarray (N-dimensional array object) 和 ufunc (universal function object) 。ndarray(下文统一称之...

 shengshengwang (<http://blog.csdn.net/shengshengwang>) 2016年10月19日 16:56 1318

scikit-learn : matplotlib.pyplot常用画图功能总结 (1) (<http://blog.csdn.net/mmc2015/artic...>)

参考： http://matplotlib.org/api/pyplot_api.html 1、 matplotlib.pyplot.plot(*args, **kwargs) , 最简单的沿坐标轴划线函...

 mmc2015 (<http://blog.csdn.net/mmc2015>) 2015年08月18日 09:17 14237





内容举报


返回顶部


python给scatter设置颜色渐变条colorbar (http://blog.csdn.net/yefengzhichen/article/detail...

python设置scatter颜色渐变 import matplotlib.pyplot as plt cm = plt.cm.get_cmap('RdYlBu') xy = range(20) ...

 yefengzhichen (http://blog.csdn.net/yefengzhichen) 2016年10月09日 14:30 111439


python sklearn decision_function、predict_proba、predict (http://blog.csdn.net/o11015...

import matplotlib.pyplot as plt import numpy as np from sklearn.svm import SVC X = np.array([[[-1,-1]...

 o1101574955 (http://blog.csdn.net/o1101574955) 2017年04月17日 15:37 3763

Matplotlib 图例中文显示（大小修改） (http://blog.csdn.net/qq_28119401/article/details/509...

legend学习代码如下：#-- coding: utf-8 -- from pylab import * myfont =matplotlib.font_manager.FontProper...

 qq_28119401 (http://blog.csdn.net/qq_28119401) 2016年03月17日 22:12 9489


机器学习实战python版决策树以及Matplotlib注解绘制决策树 (http://blog.csdn.net/XD_Senio...

这一章代码比较难懂，主要是matplotlib的函数调用参数多，调用灵活，让初学者费解。 import matplotlib.pyplot as plt decisio
nNode = dict(b...

 XD_Senior (http://blog.csdn.net/XD_Senior) 2015年11月26日 14:18 4849

使用scikit-learn进行机器学习的简介(教程1) (http://blog.csdn.net/prom1201/article/details/5...

一、机器学习：问题设定 通常，一个学习问题是通过分析一些数据样本来尝试预测未知数据的属性。如果每一个样本不仅仅是一个单独的数字，比如一个多维的实例（multivariate data），也就是说...

 prom1201 (http://blog.csdn.net/prom1201) 2016年06月07日 15:49 4129



内容举报

返回顶部