

原

周志华《机器学习》课后习题解答系列（四）：Ch3.3 - 编程实现对率回归

2017年03月19日 13:53:42 Snoopy\_Yuan 阅读数：6978 标签：机器学习 对率回归 梯度下降法 python sklearn 更多

个人分类：机器学习

版权声明：本文为博主原创文章，转载时请注明来源。https://blog.csdn.net/Snoopy\_Yuan/article/details/63684219

这里采用Python-sklearn的方式，环境搭建可参考数据挖掘入门：Python开发环境搭建（eclipse-pydev模式）。

相关答案和源代码托管在我的Github上：PY131/Machine-Learning\_ZhouZhihua.

### 思路概要

编程实现对率回归：

- \* 采用sklearn逻辑斯蒂回归库函数实现，通过查看混淆矩阵，绘制决策区域来查看模型分类效果；
- \* 自己编程实现，从极大化似然函数出发，采用梯度下降法得到最优参数，然后尝试了随机梯度下降法来优化过程。

### 3.3 编程实现对率回归

3.3 编程实现对率回归, 并给出西瓜数据集 3.0α 上的结果.

所使用的数据集如下：

表 4.5 西瓜数据集 3.0α			
编号	密度	含糖率	好瓜
1	0.697	0.460	是
2	0.774	0.376	是
3	0.634	0.264	是
4	0.608	0.318	是
5	0.556	0.215	是
6	0.403	0.237	是
7	0.481	0.149	是
8	0.437	0.211	是
9	0.666	0.091	否
10	0.243	0.267	否
11	0.245	0.057	否
12	0.343	0.099	否
13	0.639	0.161	否
14	0.657	0.198	否
15	0.360	0.370	否
16	0.593	0.042	否
17	0.719	0.103	否

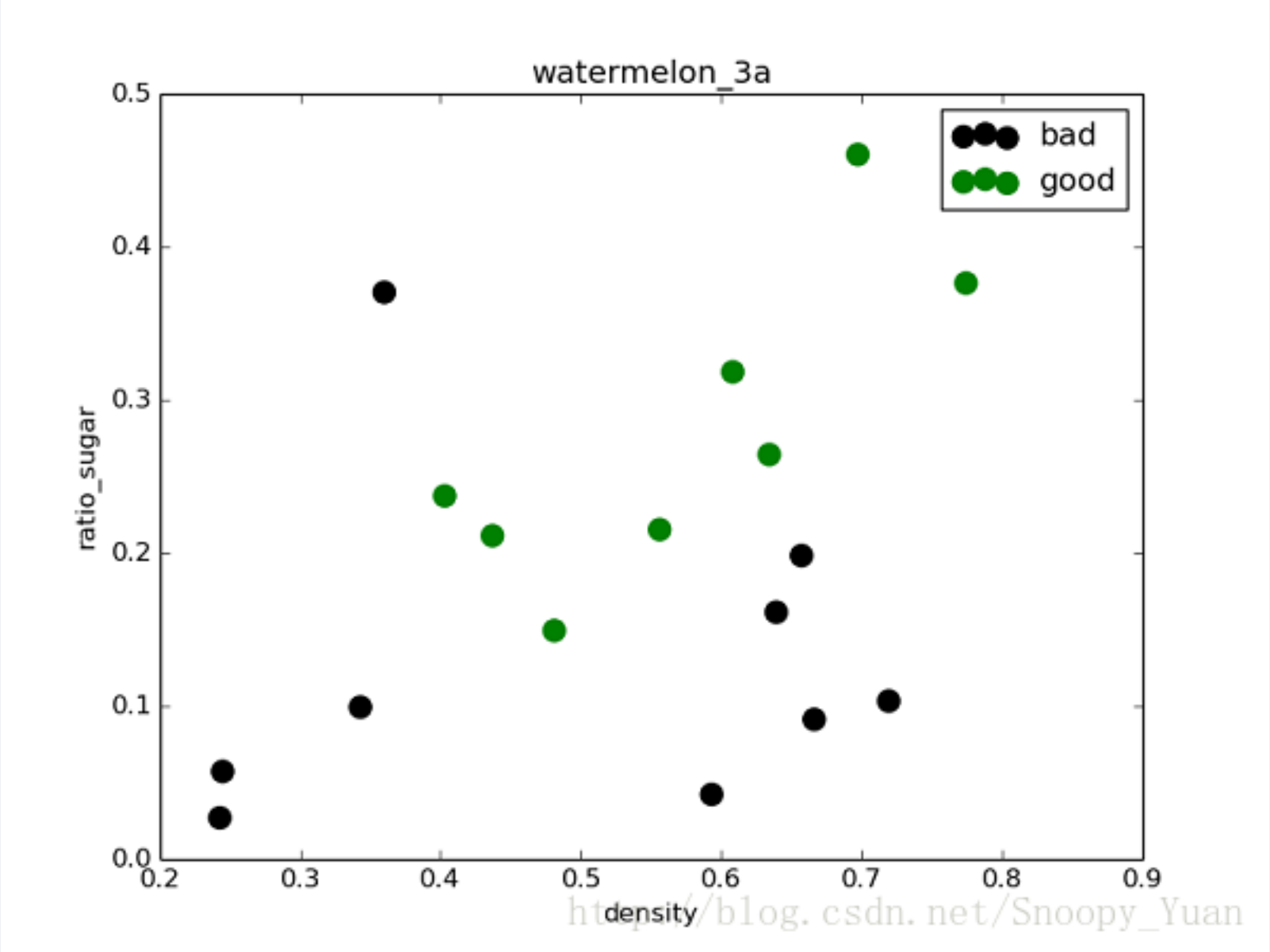
本题是本书的第一个编程练习，采用了自己编程实现和调用sklearn库函数两种不同的方式，详细解答和编码过程：（查看完整代码）：

#### 1.获取数据、查看数据、预处理：

观察数据可知，X包含（密度、含糖量）两个变量，y为西瓜是否好瓜分类（二分），由此生成.csv数据文件，在Python中用Numpy读取数据并采用matplotlib可视化数据：

```
1  '''
2  data importion
3  '''
4  import numpy as np
5  import matplotlib.pyplot as plt
6
7  # load the CSV file as a numpy matrix
8  dataset = np.loadtxt('../data/watermelon_3a.csv', delimiter=",")
9
10 # separate the data from the target attributes
11 X = dataset[:,1:3]
12 y = dataset[:,3]
13
14 # draw scatter diagram to show the raw data
15 f1 = plt.figure(1)
16 plt.title('watermelon_3a')
17 plt.xlabel('density')
18 plt.ylabel('ratio_sugar')
19 plt.scatter(X[y == 0,0], X[y == 0,1], marker = 'o', color = 'k', s=100, label = 'bad')
20 plt.scatter(X[y == 1,0], X[y == 1,1], marker = 'o', color = 'g', s=100, label = 'good')
21 plt.legend(loc = 'upper right')
22 plt.show()
```

数据散点图：



2.采用sklearn逻辑回归库函数直接拟合：

虽然样本量很少，这里还是先划分训练集和测试集，采用sklearn.model\_selection.train\_test\_split()实现，然后采用sklearn.linear\_model.LogisticRegression于训练集直接拟合出逻辑回归模型，然后在测试集上评估模型（查看混淆矩阵和F1值）。

样例代码：

```
1  '''
2  using sklearn lib for logistic regression
3  '''
4  from sklearn import model_selection
5  from sklearn.linear_model import LogisticRegression
6  from sklearn import metrics
7
8
9  # generalization of test and train set
10 X_train, X_test, y_train, y_test = model_selection.train_test_split(X, y, test_size=0.5, random_state=0)
11
12 # model training
```

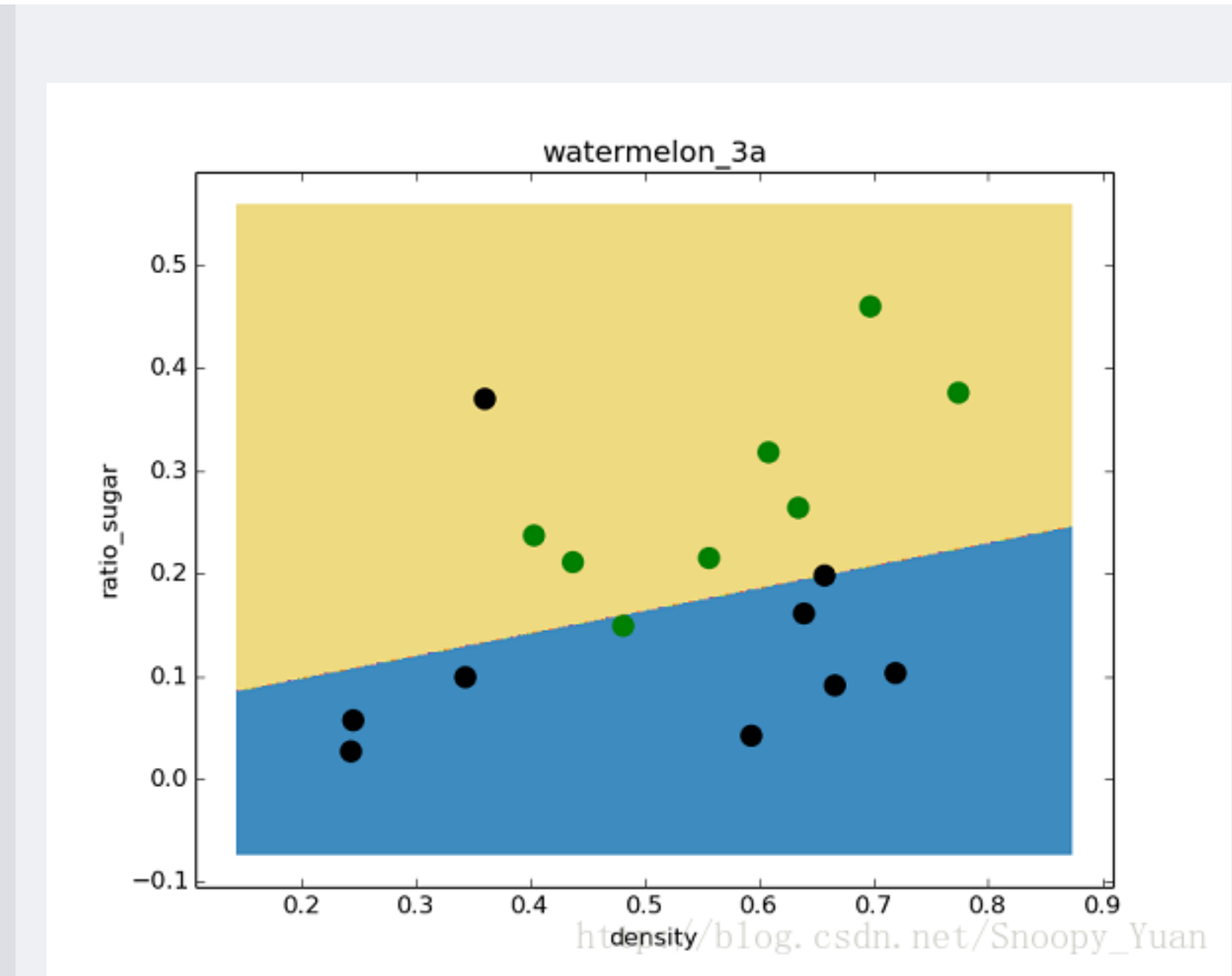
```
13 log_model = LogisticRegression()
14 log_model.fit(X_train, y_train)
15
16 # model testing
17 y_pred = log_model.predict(X_test)
18
19 # summarize the accuracy of fitting
20 print(metrics.confusion_matrix(y_test, y_pred))
21 print(metrics.classification_report(y_test, y_pred))
```

得出混淆矩阵和相关度量（查准率（准确率）、查全率（召回率），F1值）结果如下：

1	[[4 1]				
2	[1 3]]				
3		precision	recall	f1-score	support
4					
5		0.0	0.80	0.80	5
6		1.0	0.75	0.75	4
7					
8	avg / total	0.78	0.78	0.78	9

由混淆矩阵可以看到，由于样本本身数量较少，模型拟合效果一般，总体预测精度约为0.78。为提升精度，可以采用自助法进行重抽样扩充数据集，或是交叉验证选择最优模型。

下图是采用matplotlib.contourf绘制的决策区域和边界，可以看出对率回归分类器还是成功的分出了绝大多数类：



### 3.自己编程实现逻辑斯蒂回归

编程实现逻辑回归的主要工作是求取参数w和b（见书p59），最常用的参数估计方法是极大似然法，由于题3.1已经证得对数似然函数（见书3.27）是凸函数，因此存在唯一的最优解，这里考虑采用梯度下降法来迭代寻优。

回顾一下Sigmoid函数，即逻辑斯蒂回归分类器的基础模型：

$$y = \frac{1}{1 + e^{-(w^T x + b)}} \tag{3.18}$$

目的是基于数据集求出最优参数w和b，最常采用的是极大似然法，参数的似然函数为：

$$\ell(\boldsymbol{w}, b) = \sum_{i=1}^m \ln p(y_i | \boldsymbol{x}_i; \boldsymbol{w}, b), \quad (3.25)$$

根据书p59，最大化上式等价于最小化下式：

$$\ell(\boldsymbol{\beta}) = \sum_{i=1}^m \left( -y_i \boldsymbol{\beta}^T \hat{\boldsymbol{x}}_i + \ln \left( 1 + e^{\boldsymbol{\beta}^T \hat{\boldsymbol{x}}_i} \right) \right). \quad (3.27)$$

题3.2已证上式为凸函数，一定存在最小值，但按照导数为零的解析求解方式较为困难，于是考虑采用梯度下降法来求解上式最小值时对应的参数。

注：梯度下降法基本知识可参考书中附录p409页，也可直接采用书中p60式3.30偏导数公式。书中关于参数迭代改变式子如下：

$$\Delta \boldsymbol{x} = -\gamma \nabla f(\boldsymbol{x}), \quad (\text{B.17})$$

$$f(\boldsymbol{x} + \Delta \boldsymbol{x}) \simeq f(\boldsymbol{x}) + \Delta \boldsymbol{x}^T \nabla f(\boldsymbol{x}), \quad (\text{B.16})$$

对于迭代，可每次先根据(B.16)计算出梯度 $\nabla f(\boldsymbol{\beta})$ ，然后由(B.17)更新得出下一步的 $\Delta \boldsymbol{\beta}$ 。

接下来编程实现基本的梯度下降法：

(1)首先编程实现对象式3.27：

```
1 def likelihood_sub(x, y, beta):
2     '''
3     @param x: one sample variables
4     @param y: one sample label
5     @param beta: the parameter vector in 3.27
6     @return: the sub_log-likelihood of 3.27
7     '''
8     return -y * np.dot(beta, x.T) + np.math.log(1 + np.math.exp(np.dot(beta, x.T)))
9
10 def likelihood(X, y, beta):
11     '''
12     @param X: the sample variables matrix
13     @param y: the sample label matrix
14     @param beta: the parameter vector in 3.27
15     @return: the log-likelihood of 3.27
16     '''
17     sum = 0
18     m,n = np.shape(X)
19
20     for i in range(m):
21         sum += likelihood_sub(X[i], y[i], beta)
22
23     return sum
```

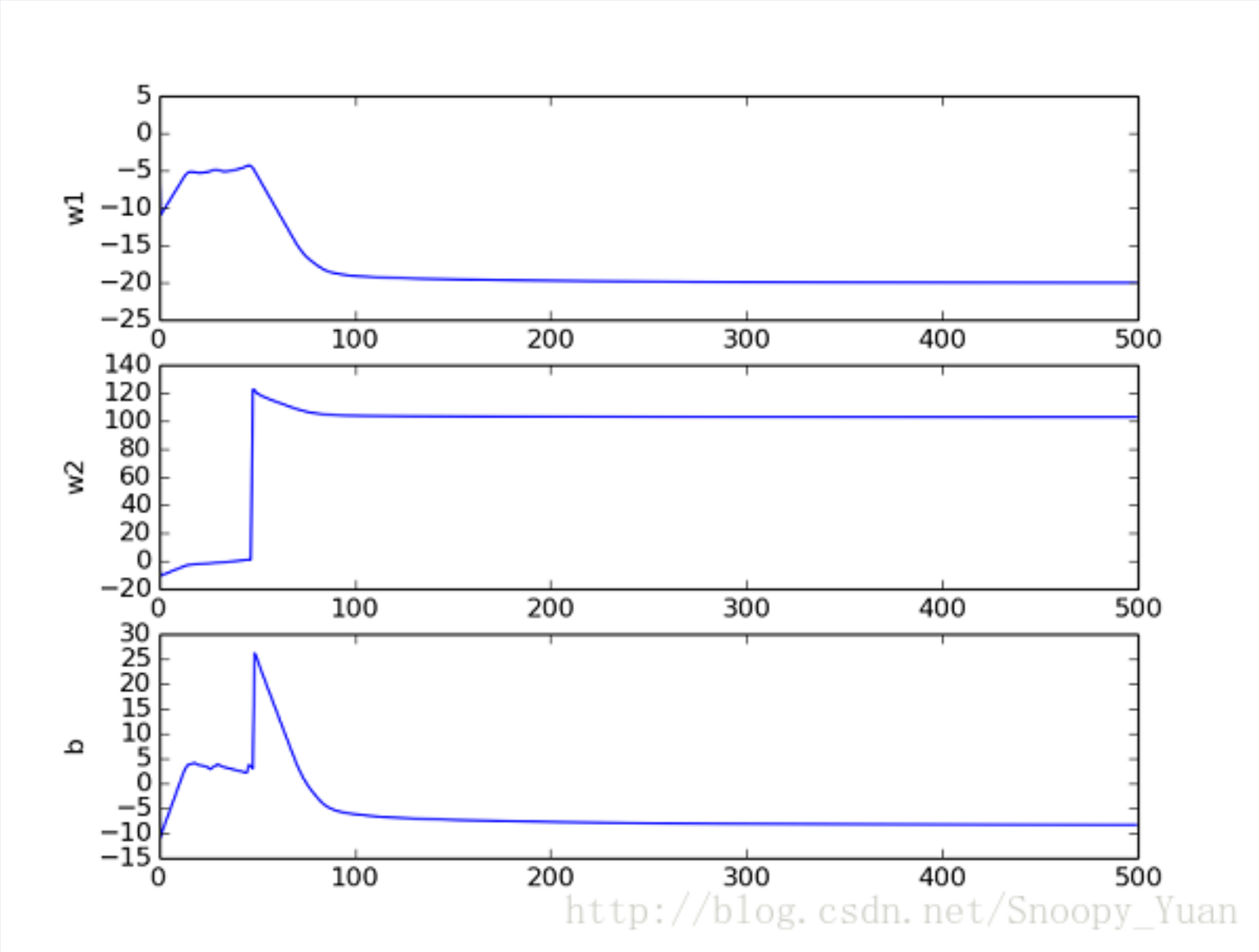
(2)然后基于训练集（注意 $\boldsymbol{x} \rightarrow [\boldsymbol{x}, 1]$ ），给出基于3.27似然函数的定步长梯度下降法，注意这里的偏梯度实现技巧：

```
1     '''
2     @param X: X is the variable matrix
3     @param y: y is the label array
4     @return: the best parameter estimate of 3.27
5     '''
6 def gradDscnt_1(X, y): #implementation of basic gradDscnt algorithms
7
8     h = 0.1 # step length of iteration
9     max_times= 500 # give the iterative times limit
```



```
10 m, n = np.shape(X)
11
12 beta = np.zeros(n) # parameter and initial to 0
13 delta_beta = np.ones(n)*h # delta parameter and initial to h
14 llh = 0
15 llh_temp = 0
16
17 for i in range(max_times):
18     beta_temp = beta.copy()
19
20     # for partial derivative
21     for j in range(n):
22         beta[j] += delta_beta[j]
23         llh_tmp = likelihood(X, y, beta)
24         delta_beta[j] = -h * (llh_tmp - llh) / delta_beta[j]
25         beta[j] = beta_temp[j]
26
27     beta += delta_beta
28     llh = likelihood(X, y, beta)
29
30 return beta
```

通过追踪参数，查看其收敛曲线，然后来调节相关参数（步长h，迭代次数max\_times）。下图是在当前参数取值下的beta曲线，可以看到其收敛良好：



(3)最后建立Sigmoid预测函数，对测试集数据进预测，得到混淆矩阵如下：

```
1 [[ 4.  1.]
2  [ 1.  3.]
```

可以看出其总体预测精度（7/9 ≈ 0.78）与调用sklearn库得出的结果相当。

(4)采用随机梯度下降法来优化：上面采用的是全局定步长梯度下降法（称之为批量梯度下降），这种方法在可能会面临收敛过慢和收敛曲线波动情况的同时迭代需要全局计算，计算量随数据量增大而急剧增大。所以尝试采用随机梯度下降来改善参数迭代寻优过程。

随机梯度下降法的核心思想是增量学习：一次只用一个新样本来更新回归系数，从而形成在线流式处理。

同时为了加快收敛，采用变步长的策略，h随着迭代次数逐渐减小。

给出变步长随机梯度下降法的代码如下：

```
1 def gradDscnt_2(X, y): #implementation of stochastic gradDscnt algorithms
2     ...
3     @param X: X is the variable matrix
4     @param y: y is the label array
```

```
5 @return: the best parameter estimate of 3.27
6 '''
7 import matplotlib.pyplot as plt
8
9 m, n = np.shape(X)
10 h = 0.5 # step length of iterator and initial
11 beta = np.zeros(n) # parameter and initial
12 delta_beta = np.ones(n) * h
13 llh = 0
14 llh_temp = 0
15
16 for i in range(m):
17     beta_temp = beta.copy() # for partial derivative
18
19     for j in range(n):
20         h = 0.5 * 1 / (1 + i + j) # change step length of iterator
21         beta[j] += delta_beta[j]
22         llh_tmp = likelihood_sub(X[i], y[i], beta)
23         delta_beta[j] = -h * (llh_tmp - llh) / delta_beta[j]
24         beta[j] = beta_temp[j]
25
26     beta += delta_beta
27     llh = likelihood_sub(X[i], y[i], beta)
28
29 return beta
```

得出混淆矩阵：

```
1 [[ 3.  2.]
2  [ 0.  4.]
```

从结果看到的是：由于这里的西瓜数据集并不大，所以随机梯度下降法采用一次遍历所得的结果不太好，参数也没有完成收敛。这里只是给出随机梯度下降法实现样例，这种方法在大数据集下相比批量梯度法应会有明显的优势。

参考链接：

由于这是本书第一个编程，索引资料较多，择其重要的一些列出如下：

- [Introduction to Machine Learning with Python and Scikit-Learn](#)
- [scikit-learn官方主页](#)
- [matplotlib官方主页](#)
- [随机梯度下降（Stochastic gradient descent）和批量梯度下降（Batch gradient descent）的公式对比、实现对比](#)
- [机器学习算法与Python实践之（七）逻辑回归（Logistic Regression）](#)
- [plot decision boundary matplotlib Ask（matplotlib决策区域和边界绘制](#)
- [Python数据可视化——散点图](#)



几乎没人知道！微信新出的这个赚钱功能简直赚疯了！

富洋松旺 · 熾燚

## Python爬虫全栈教学，零基础教你成编程大神


零基础学爬虫，你要掌握学习那些技能？

想对作者说点什么？

我来说两句

**qq\_40530372**： 利用随机梯度下降法求出的混淆矩阵跟您给出的不一样，用你的GitHub上的源码 （21小时前 #6楼）

**xiaocuishuoshu**： plt.scatter(X[y == 0,0], X[y == 0,1], marker = 'o', color = 'k', s=100, label = 'bad') plt.scatter(X[y == 1,0], X[y == 1,1], marker = 'o', color = 'g', s=100, label = 'good')主，命令中[y == 0,0], X[y == 0,1]是什么意思？在怎么有两个取值？ （3周前 #5楼）


- 

m0\_37709336:

您好，有个小疑问：GradDecent函数（批量梯度下降函数）第18行:beta\_tmp = beta和第25行:beta[j] = beta\_tmp[j]。我理解的意义为：只变化一个维度x，其他维度x不变求得在这个维度的delta\_beta。但是有一点很疑惑，beta\_tmp = beta,使得这两个变量指向同一个底层矩阵对象，即使得：beta[j] = beta\_tmp[j]失效了（我的beta\_tmp = beta,使得beta\_tmp赋值为一个指针，其值等于beta，即指向同一个对象。使得，beta[j] = beta\_tmp[j]失效，因为beta\_tmp[j]被21行的beta[j] += delta\_beta[j]改第18行改成：beta\_tmp = np.array(beta)可能才是您的原本意义，但是这样求得的beta和您给出的beta收敛曲线中得到的值不一样。对这点很疑惑。

（6个月前

#4楼


查看回复(1)
- 
- 

项慈航:

公式中是beta转秩乘x，你实现为何是相反的？暂时数学比较渣，困惑多时，麻烦指点下

（1年前


#3楼

查看回复(1)
- 
- 

飞雨新风:

图上面我没找到（0.243,0.267）这个点

（1年前

#2楼
- 
- 

飞雨新风:

数据点有(0.243，0.267)但是到了你这里有没有了，你用错数据了~~~

（1年前

#1楼

查看回复(1)

- 机器学习(周志华) 参考答案 第三章 线性模型 3.3

机器学习(周志华) 参考答案 第三章 线性模型 3.3机器学习(周志华西瓜书) 参考答案 ...

1.1万

来自: 我的博客
- 西瓜书 习题3.3 编程实现对数几率回归，梯度下降法

最近入坑上道了，跟着周志华老师的《机器学习》，先搞个课后题练练手。我电脑...

1301

来自: 世靖的码场
- 周志华机器学习，3.3编程实现对率回归，并给出西瓜数据集3.0a上的结果

3.3编程实现对率回归，并给出西瓜数据集3.0a上的结果数据集： 1 0.697 0.460 1 2 ...

4278

来自: zjy\_lilas的博客

- 30W年薪的人工智能工程师只是“白菜价”？

机器学习|深度学习|图像处理|自然语言处理|无人驾驶，这些技术都会吗？看看真正的人工智能师都会那些关键...

区块链开发八周学会，小白程序员都能学？

区块链DApp开发学习大纲免费领

- Linux Socket编程实战第1季第1部分

适合人群：1、有C语言基础； 2、对网络通讯感兴趣的人员； 3、从事网络通讯的技术人员； 4、在校学生； ,...

学院

- 用对数几率回归实现周志华《机器学习》习题3.3西瓜分类，python编程

数据集如下，要求根据西瓜的两个属性x1(密度)，x2(含糖率)实现对西瓜好瓜（1） ...

2716

来自: 大树挖掘工的博客
- 《机器学习》周志华 课后习题3.3： 编程实现对率回归,并给出西瓜数据集 ...

数据如下： python 代码如下：#!/usr/bin/env python3 #-\*- coding: utf-8 -\*- """ Cre...

1309

来自: llwleon的博客

- 《机器学习（周志华）》习题3.3答案

编程实现对率回归，并给出西瓜数据集3.0@上的结果。对率回归即逻辑回归，可以...

2891

来自: 勿忘初衷

- 机器学习（周志华）习题3.3

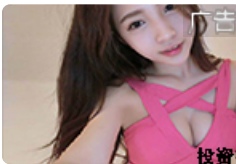
本人菜鸟一枚，由于需要完成作业，所以尝试使用机器学习工具库去解决该题（周...

336

来自: CHNguoshiwush...



几乎没人知道！微信新出的这个赚钱功能简直赚疯了！  
富洋松旺 · 熾燚



劲爆！石家庄25岁美女用微信做这个，1个月存款吓呆父母！  
煜隆投资 · 熾燚

- 你走过最长的路,就是机器学习过程中的弯路 - AI科技大...\_CSDN博客

营长的一位转型AI的朋友,最近对营长抱怨,“走过的最远的路,就是机器学习过程中的弯路”,然后开始各种blabl...

- 机器学习系列(一)——机器学习简介 - CSDN博客

前前后后接触机器学习也有一年时间,但一直没有系统整理总结过。从本篇博客开始,将记录下我的学习与参考资料,系列按照李宏毅的机器学习课程,吴恩达的机器学习课程...

- 你走过最长的路,就是机器学习过程中的弯路 - AI科技大...\_CSDN博客

营长的一位转型AI的朋友,最近对营长抱怨,“走过的最远的路,就是机器学习过程中的弯路”,然后开始各种blabl...

- 机器学习系列(一)——机器学习简介 - CSDN博客

前前后后接触机器学习也有一年时间,但一直没有系统整理总结过。从本篇博客开始,将记录下我的学习与...

- 机器学习(周志华) 参考答案 第三章 线性模型

机器学习(周志华) 参考答案 第三章 线性模型 机器学习(周志华西瓜书) 参考答案 总目...

2万

来自: 我的博客



3.3编程实现对率回归

338

&quot;&quot;&quot; Author: Victoria Created on: 2017.9.14 11:00 &quot;&quot;&quot; 来自：周博的博客



Snoopy\_Yuan

关注

40篇文章



carson0408

关注

166篇文章



steve\_99

关注

75篇文章

《机器学习有意思! 01》 - 世界上最简单的机器学习入门

本文首发于https://jizhi.im/blog/post/ml\_is\_fun\_01 你是否也曾听人们谈起机器学习但是只有一个朦胧的概念?你是否厌倦了在同事的高谈阔论中颓然欲睡?此诚求...

图解十大经典机器学习算法入门 - jrunw的博客 - CSDN博客

传统的机器学习算法包括决策树、聚类、贝叶斯分类、支持向量机、EM、Adaboost等等。这篇文章将对常用算法做常识性的介绍,没有代码,也没有复杂的理论推导,就是图解...

《机器学习有意思! 01》 - 世界上最简单的机器学习入门

本文首发于https://jizhi.im/blog/post/ml\_is\_fun\_01 你是否也曾听人们谈起机器学习但是只有一个朦胧的概念?你...

图解十大经典机器学习算法入门 - jrunw的博客 - CSDN博客

传统的机器学习算法包括决策树、聚类、贝叶斯分类、支持向量机、EM、Adaboost等等。这篇文章将对常用...

对率回归的实验

141

对数几率回归在python中的实现 在做分类任务时，需要找一个单调可微函数将分类... 来自：qq\_28915885的...

机器学习(周志华)-python编程练习-习题3-3

19

习题3.3 编程实现对率回归，并给出西瓜数据集 3.0α 上的结果. 数据集3.0α sn den... 来自：bebusy的专栏

机器学习-周志华-课后习题答案-线性模型

1325

3.1试分析在什么情况下，在以下式子中不比考虑偏置项b。答：在线性回归中，所... 来自：天台的猫爷爷的...



石家庄90后小伙在家无聊玩微信，存款惊呆父母

东平商贸 · 熯熯



劲爆，靠死工资怎么买房买车，聪明人是这样赚钱！

投资 · 熯熯

matlab自定义函数的几种方法

2.6万

1、函数文件+调用命令文件：需单独定义一个自定义函数的M文件; 2、函数文件+... 来自：yuxiaoxi21的博客

Matlab中如何定义函数

1.3万

1.在m文件中，function y=f(x) %函数的声明 y=x^2就是建立了一个y=x2y=x^2的函... 来自：Katherine\_S的专...

调用自己编写的matlab函数

4.2万

在matlab中调用自定义的函数 来自：without\_scruple...

对数几率回归（Logistic Regression）总结

4966

逻辑回归logistic regression，虽然名字是回归，但是实际上它是处理分类问题的算... 来自：code\_caq的博客

周志华《机器学习》习题3.3

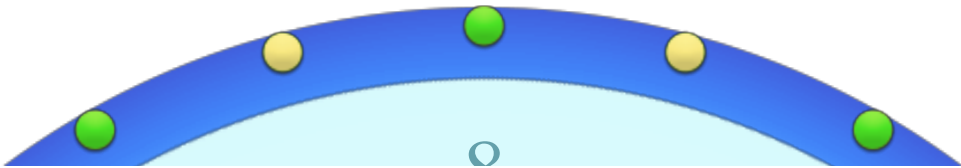
748

编程实现对数回归，并给出西瓜数据集3.0α\alpha上的结果。既然是编程实现，就... 来自：WUTab的博客

哈佛大学录取分

百度广告

广告



3分钟了解入门「机器学习」该学习什么?(上) - GitChat...\_CSDN博客



• 一张图看懂AI、机器学习 和深度学习的区别 - dukai392的...\_CSDN博客

要搞清它们的关系,最直观的表述方式就是同心圆,最先出现的是理念,然后是机器学习,当机器学习繁荣之后就出现了深度学习,今天的AI大爆发是由深度学习驱动的。 从衰败...

• 3分钟了解入门「机器学习」该学习什么?(上) - GitChat...\_CSDN博客

本文来自作者 粽子 在 GitChat 上分享「零基础的新手,如何入门机器学习?」,「阅读原文」查看交流实录 「文...

• 一张图看懂AI、机器学习 和深度学习的区别 - dukai392的...\_CSDN博客

要搞清它们的关系,最直观的表述方式就是同心圆,最先出现的是理念,然后是机器学习,当机器学习繁荣之后就出...

• 机器学习 实战(用Scikit-learn和TensorFlow进行机器学习)(五) 👁 1822

上几节讲述了真实数据集在回归问题以及分类问题上的总流程，但是对于模型的选... 来自: [fjl\\_CSDN的博客](#)

• 逻辑回归（Logistic regression）详解-并用scikit-learn训练逻辑回归拟合lr... 👁 4.9万

引言这篇文章主要介绍逻辑回归背后的一些概率概念，给你一些直观感觉关于它的... 来自: [Xurtle](#)

• Logistic Regression（逻辑回归） 👁 806

Logistic回归思想，Python实现，应用。Logistic 回归是与线性回归相对应的一种分... 来自: [静敬澹一](#)

• 机器学习之线性回归及代码示例 👁 7252

一、线性回归线性回归一般用来做连续值的预测，预测的结果为一个连续值。因训... 来自: [cxmscb的博客](#)

• 周志华《机器学习》课后习题解答系列（五）：Ch4 - 决策树 👁 4605

本章讲述决策树的相关内容，包括决策树的生成、剪枝，连续值、缺失值的处理， ... 来自: [Snoopy\\_Yuan技...](#)



十一月靠死工资怎么买房买车，聪明人是这样赚钱！

投资 · 煊燊



去银行转账，发现老婆的余额，瞬间吓坏了！！！！

中乐咨询 · 煊燊

• 机器学习（周志华）第四章习题解答 👁 1650

转自：http://blog.csdn.NET/wzmsltw/article/details/51059394 本文是对周志华的《... 来自: [carson0408的博...](#)

• 《机器学习》周志华习题4.3答案 👁 3615

《机器学习》周志华著，第四章课后习题4.3答案 来自: [Just Do IT](#)

• 机器学习(周志华) 参考答案 第五章 神经网络 👁 1.1万

机器学习(周志华) 参考答案 第五章 神经网络机器学习(周志华西瓜书) 参考答案 总目... 来自: [我的博客](#)

• 机器学习(周志华) 参考答案 第六章 支持向量机 6.9 👁 2750

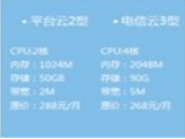
机器学习(周志华) 参考答案 第六章 支持向量机 6.9机器学习(周志华西瓜书) 参考答... 来自: [我的博客](#)

• 

下载

 编程实现对率回归，并给出西瓜数据集3.0a上的结果。 10-14

机器学习的第一次作业。周志华老师教材的p3.3。python实现对率回归，已给数据集



免费永久云主机

百度广告

广告

• 机器学习的方法 - 修炼之路 - CSDN博客

机器学习(machine learning)是一门多领域交叉学科,涉及了概率论、统计学、算法复杂度等多门学科。专门研究计算机怎样模拟或实现人的学习行为,它能够发现和挖掘数据所...

• 3分钟了解入门「机器学习」该学习什么?(下) - GitChat...\_CSDN博客

本文来自作者 刘明 在 GitChat 上分享「机器学习/深度学习书单推荐及学习方法」,「阅读原文」查看交流实录 「文末高能」 编辑 | 坂本 写在前面 本人是...

机器学习的方法 - 修炼之路 - CSDN博客

机器学习(machine learning)是一门多领域交叉学科,涉及了概率论、统计学、算法复杂度等多门学科。专门研究...

3分钟了解入门「机器学习」该学习什么?(下) - GitChat...\_CSDN博客

本文来自作者 刘明 在 GitChat 上分享「机器学习/深度学习书单推荐及学习方法」,「阅读原文」 查看交流实录 ...

机器学习作业1 - 对率回归（逻辑回归）

1919

使用10折交叉验证法和留一法评测对率回归分类器标题有点长哈.....这是第一次作...

来自： Dapan同学

逻辑回归/对数几率回归--西瓜书、统计学习总结

251

广义模型可以解决分类任务。只需找到一个单调可微函数将分类任务的真实标记y与...

来自： zhangdamengcs...

对数几率回归Logistic Regression（Matlab）

6769

这里的数据均来源于吴恩达老师机器学习的课程。

上一篇内容是线性回归，利...

来自： XLM11的专栏

机器学习—西瓜书-chapter3—对率回归

215

编程实现对率回归，并给出西瓜数据集3.0α上的结果 对率函数是任意阶可导的凸函...

来自： 糖糖糖豆的博客

《机器学习》(周志华)课后习题参考答案

3325

目录： 周志华《机器学习》课后习题解答系列（二）： Ch1 - 绪论周志华《机器学...

来自： kchai31的博客



几乎没人知道！微信新出的这个赚钱功能简直赚疯了！

爱尚庄园 · 熯熯



几乎没人知道！手机新出的这个赚钱功能简直赚疯了！

上海永安 · 熯熯

matlab如何在文中定义函数

9514

matlab如何在文中定义函数 通常我们多是将函数单独在编写在单个的m文件中，再...

来自： lusongno1的博客

第三章 线性模型--机器学习（周志华）参考答案

834

原文的链接 机器学习(周志华) 参考答案 第三章 线性模型 机器学习(周志华西瓜书) 参...

来自： qq\_35218763的...

周志华 《机器学习》之 第三章（线性模型）概念总结

2699

阅读之后，根据周志华老师对本章节的安排，首先从线性模型的基本形式入手，逐...

来自： 不系之舟的专栏...

使用sklearn做各种回归

9274

使用sklearn做各种回归 基本回归：线性、决策树、SVM、KNN 集成方法：随机森...

来自： Yeoman92的博客

西瓜书《机器学习》课后答案——Chapter3\_3.4

1617

选择两个UCI数据集，比较10折交叉验证法和留一法所估计出的对率回归的错误率...

来自： CodeTutor



十一月靠死工资怎么买房买车，聪明人是这样赚钱！

投资 · 熯熯



2018年微赚钱策略大公开，第一批网民已经赚翻了！

嘉超电子 · 熯熯

机器学习入门好文,强烈推荐 - Ritter Liu的专栏 - CSDN博客

转自 飞鸟各投林 史上最强---机器学习经典总结---入门必读---心血总结---回味无穷 让我们从机器学习...

机器学习基本概念梳理 - ForLearning - CSDN博客

1. 什么是机器学习?权威定义: Arthur samuel: 在不直接针对问题进行编程的情况下,赋予计算机学习能力的一个研究领域。 Tom Mitchell: 对于某类任务T和性能度量P,...

机器学习入门好文,强烈推荐 - Ritter Liu的专栏 - CSDN博客

转自 飞鸟各投林 史上最强---机器学习经典总结---入门必读---心血总结---回味无穷 让我们从机器学习...



机器学习基本概念梳理 - ForLearning - CSDN博客

1. 什么是机器学习?权威定义: Arthur samuel: 在不直接针对问题进行编程的情况下,赋予计算机学习能力的一个...

周志华《机器学习》课后习题解答系列（五）：Ch4.3 - 编程实现ID3算法

这里采用了自己编程的方式实现ID3算法，并基于西瓜数据集生成了决策树，评估了... 来自：Snoopy\_Yuan技...

西瓜书《机器学习》课后答案——Chapter3\_3.5

编程实现线性判别分析，并给出西瓜数据集3.0alpha上的结果。"" Author: Victoria ... 来自：CodeTutor

机器学习(周志华) 参考答案 第九章 聚类 9.10

机器学习(周志华) 参考答案 第九章 聚类 9.10机器学习(周志华西瓜书) 参考答案 总目... 来自：我的博客

《机器学习》--周志华版（西瓜书）--课后参考答案

开始学习周志华版的《机器学习》，将别人写的课后习题的参考答案保存下来供参... 来自：这里记录着我一...

logistic回归概率详解

logistic回归概率详解 上一篇我们介绍了线性代数的基本知识，并以PCA作为案例进... 来自：qq\_40213457的...

免费永久云主机

百度广告

网店店铺转让

百度广告

广告

你应该知道的7种回归方法

本文是我从国外网站翻译而来的文章，如有错误之处，敬请指出！ 原文标题：7 Ty... 来自：Steve lock的博客...

机器学习(周志华西瓜书) 参考答案 总目录

机器学习(周志华西瓜书)参考答案总目录 从刚开始学习机器学习到现在也有几个月... 来自：我的博客

周志华《机器学习》读书笔记（一）

本书前几章讲的都是基本术语，最硬核的数学部分很少，所以比较简单。 机器学习... 来自：garrulousabyss...

周志华机器学习笔记（一）

新人一枚，既是机器学习的初学者，也是首次发博客。谨以此记录我的学习体会， ... 来自：baidu\_3840152...

机器学习(周志华)习题解答1.1-1.3: 理解假设和版本空间

本文介绍版本空间，假设空间的概念并举例求解。另外介绍和简单证明“没有免费的... 来自：走过的都是未来



免费永久云主机

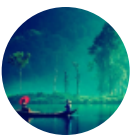
百度广告

机器学习(周志华) 参考答案 第一章 绪论

机器学习(周志华) 参考答案 第一章 假设空间指的是问题所有假设组成的空间，我们... 来自：我的博客

逻辑回归模型推导及梯度下降

这里的逻辑回归模型，除了重要要放在回归上，还要看到逻辑，所谓的逻辑其实就... 来自：哆啦咪~fo



Snoopy\_Yuan

关注

原创 40 粉丝 415 喜欢 70 评论 130

等级： 博客 4 访问：12万+



积分： 1334      排名： 4万+

勋章：



如何自学编程



更多信息

**Email:** pn\_yuan@163.com

**GitHub:** <https://github.com/PnYuan>

最新文章

明星经纪人网站

Kaggle滑水 - CTR预估 (FM\_FEM)



Kaggle滑水 - CTR预估 (GBDT-LR)

Kaggle滑水 - CTR预估 (LR)

Kaggle滑水 - 泰坦尼克之灾 (决策树)

深度学习基础 - 对象检测 (CNN+滑窗+YOLO)

归档

2018年6月	3篇
2018年4月	2篇
2018年3月	3篇
2017年10月	2篇
2017年7月	6篇

[展开](#)

个人分类

机器学习	33篇
深度学习	8篇
概率图	1篇
数据挖掘	9篇
Hadoop	1篇

[展开](#)

最新评论

周志华《机器学习》课后习题解答系列...  
qq\_40530372: 利用随机梯度下降法求出的混淆矩阵跟您给出的不一样，用你的GitHub上的源码

周志华《机器学习》课后习题解答系列...  
u012016803: CART\_watermelon.py源码报错  
<class> Traceback (m...

周志华《机器学习》课后习题解答系列...  
huazeci: 真是有心人

周志华《机器学习》课后习题解答系列...  
catherined: [reply]liucheng\_34[/reply] 感谢提示！

周志华《机器学习》课后习题解答系列...

mmm\_jsw：作者出现了一处笔误，应该是：  $n1 = (2+1)(3+1)(3+1)+1 = 49$ 种

开发一个app多少钱

广告



雨水收集系统



一点点加盟



创意产品设计



工业键盘



联系我们



扫码联系客服



官方公众号

- QQ客服
- kefu@csdn.net
- 客服论坛
- 400-660-0108
- 工作时间 8:00-22:00

[关于我们](#) | [招聘](#) | [广告服务](#) | [网站地图](#)  
 百度提供站内搜索 京ICP证09002463号  
©2018 CSDN版权所有

网络110报警服务    经营性网站备案信息  
北京互联网违法和不良信息举报中心  
中国互联网举报中心

联系我们

- QQ客服
- kefu@csdn.net
- 客服论坛
- 400-660-0108
- 工作时间 8:00-22:00

[关于我们](#) | [招聘](#) | [广告服务](#) | [网站地图](#)  
 百度提供站内搜索 京ICP证09002463号  
©2018 CSDN版权所有

网络110报警服务    经营性网站备案信息  
北京互联网违法和不良信息举报中心  
中国互联网举报中心