

[门户](#) [数据·中国](#) [专才计划](#) [特训营](#) [培训](#) [课程](#) [专业](#) [企业服务](#) [论坛](#) [奖学金](#) [大数据](#) [商业智能](#) [量化投资](#) [创业](#) [专家团](#) [关于我们](#)[知识图谱实战（第一期）](#)[深度学习框架Tensorflow学...](#)[R语言数据分析、展现与实例...](#)[Go语言实战编程（第三期）](#)[深入浅出Spring（第五期）](#)[大数据算法导论（第14期）](#)[Python数据分析案例实战（第...](#)[Python金融投资分析实践（第...](#)[查看内容](#)

《您可能感兴趣的课程》

受欢迎的机器学习应用标准数据集

炼数成金_小数 | 查看: 23469 | 评论: 0 | 原作者: Jason Brownlee | 来自: 新智元

是用许多不同的数据集来练习。因为对不同的问题，需要有不同的数据准备和建模方法。本文介绍了10个最受欢迎的标准机器学习数据集，可以用每个数据集均按照一定的格式介绍，以使读者 ...

[商业智能](#) [ETL](#) [时间序列](#)

对于机器学习的关键是用许多不同的数据集来练习。因为对不同的问题，需要有不同的数据准备和建模方法。本文介绍了10个更受欢迎的标准机器学习数据集，可以用作练习的资源。

每个数据集均按照一定的格式介绍，以使读者相对容易比较，为他们的特定练习任务选择数据集或建模方法。

格式：

热门频道

大数

量化投资

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

炼数成金个性化教育产品“专才计划”
上线，导师制私人订制尊贵服务

即将开课

- Python金融投资分析实践（第六期）
- 大数据算法导论（第14期）
- Go语言实战编程（第三期）

名称：如何引用数据集

问题类型：是回归问题还是分类问题

输入和输出：输入和输出特征的数量和名称

性能：使用零规则算法（Zero Rule Algorithm）的基准性能比较，以及已知的较佳性能

示例：原始数据前5行的快照

链接：下载数据集及了解更多的链接

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

数据集的列表。每个数据集都不大，能够适应内存或使用电
脑由表格数据组成，并且没有（显式的）缺失值。

鲍鱼数据集

电离层数据集

小麦种子数据集

波士顿房价数据集

瑞典汽车保险数据集

- 深度学习框架Tensorflow学习与应用（第五
- 知识图谱实战（第一期）
- 【免费公开课】Java经验谈（第三期）
- R语言数据分析、展现与实例（第30期）
- 深入浅出Spring（第五期）
- Python数据分析案例实战（第六期）
- DevSecOps安全交付应用实战（第三期）
- 机器读心术之神经网络与深度学习（第九
- Spark企业级大数据项目实战（第一期）

最新活动

充值700送500，仅20个名额

• 【免费公开

• Python数据

人工智能前沿人才培养计划

• Node.js项目

郭一军“新DBA”精英计划

• 金融的人工

• 黄美灵的S

葛一鸣JAVA实战精英培养计划

• 人工智能前

精品小班特训营早报优惠多

• 开源计算机

联系客服

• 深入浅出设

4008-010-006

• 【免费公开

• 深入浅出O

QQ交谈

• 深度学习框

QQ交谈

• Excel数据分



客服微信

- Python机器学习Kaggle案例实战（第四期）
- 大数据必知的java基础（第八期）
- 机器读心术之文本挖掘与自然语言处理（第
- django打造电商项目（第一期）
- 【免费公开课】Julia快速数据分析（第三期）
- 让服务飞起来：实时计算及其应用（第二

瑞典汽车保险数据集（ Swedish Auto Insurance Dataset ）包含了对所有索赔要求的总赔付预测，以千瑞典克朗计，给定的条件是索赔要求总数。

这是一个回归问题。它由 63 个观察值组成，包括1个输入变量和1个输出变量。变量名分别是：

索赔要求数量

对所有索赔的总赔付 以千瑞典克朗计

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

均方根误差（RMSE）约为 72.251 千克朗。

新智元

图：

- ELKStack及Solr企业级搜索引擎实战（第七
- Python机器学习（第五期）
- 深入JVM内核—原理、诊断与优化（第14
- 大数据的统计学基础（第21期）
- 从0到1搭建企业级大数据平台（第一期）
- Oracle 12c特性解读-容器数据库和灾备（第四
- Hadoop数据分析平台（第46期）
- MySQL DBA从小白到大神实战（第七期）
- MATLAB计算机视觉与深度学习实战（第二

最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

联系客服

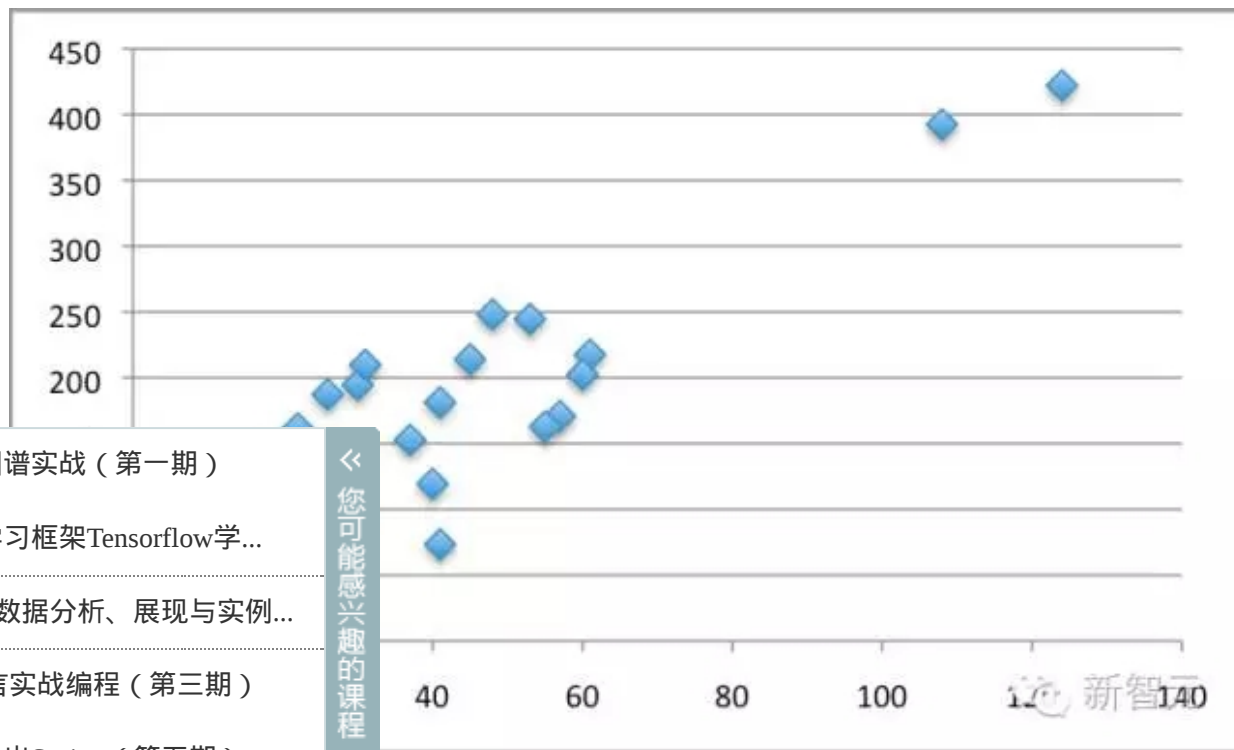
4008-010-006

QQ交谈

QQ交谈



客服微信



知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

AbI

葡萄酒质量数据集（Wine Quality Dataset）涉及根据每种葡萄酒的化学度量值来预测白葡萄酒的质量。

它是一个多类分类问题，但也可以定义为回归问题。每个类的观察值数量不均等。一共有 4898 个观察值，11 个输入变量和一个输出变量。变量名如下：

非挥发性酸度



热门文章

- 2018年值得
- 开源DAWI
- 理解XGB
- 谷歌用ML
- 2018 AI趋势
- Jeff Dean出
- 基于Docker
- Andrej Kar
- 为什么XG

最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

挥发性酸度

柠檬酸

残留糖

氯化物

游离二氧化硫

总二氧化硫

浓度

pH值

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

均方根误差（RMSE）为 0.148 的质量分数。

45,45,170,1.001,3,0.45,8.8,6
 49,14,132,0.994,3.3,0.49,9.5,6
 5,30,97,0.9951,3.26,0.44,10.1,6
 058,47,186,0.9956,3.19,0.4,9.9,6
 058,47,186,0.9956,3.19,0.4,9.9,6

比马印第安人糖尿病数据集

比马印第安人糖尿病数据集（Pima Indians Diabetes Dataset）涉及根据医疗记录预测比马印第安人5年内糖尿病的发病情况。

它是一个二元分类问题。每个类的观察值数量不均等。一共有 768 个观察值，8个输入变量和1个输出变量。缺失值通常用零值编码。变量名如下：

深入浅出设计模式

码农通往架构师的阶梯，
致力传授最精髓设计思想

最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

怀孕次数

口服葡萄糖耐受试验中，2小时的血浆葡萄糖浓度。

舒张压（mm Hg）

三头肌皮肤褶层厚度（mm）

2小时血清胰岛素含量（ $\mu\text{U/ml}$ ）

体重指数（体重，kg / （身高，m）²）

糖尿病家族史

年龄（岁）

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

是约 65% 的分类准确率，较佳结果达到约 77% 的分类准确

627,50,1
51,31,0
72,32,1
167,21,0
2 288 33 1

新智元

q8

声纳数据集

声纳数据集（Sonar Dataset）涉及预测根据给定声纳从不同角度返回的强度预测目标物体是岩石还是矿井。

它是一个二元分类问题。每个类的观察值数量不均等。一共有208个观察值，60个输入变量和1个输出变量。变量名如下：

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

从不同角度返回的声纳

... ..

类（M为矿井，R为岩石）

预测最普遍类的基准性能是约 53% 的分类准确率，较佳结果达到约 88% 的分类准确率。

该数据集前五行的示例如下：

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

0.0207,0.0954,0.0986,0.1539,0.1601,0.3109,0.2111,0.1609,0.1582
0.0689,0.1183,0.2583,0.2156,0.3481,0.3337,0.2872,0.4918,0.6552
0.1083,0.0974,0.2280,0.2431,0.3771,0.5598,0.6114,0.6332,0.7060
0.0205,0.0205,0.0368,0.1098,0.1276,0.0598,0.1264,0.0881,0.1992
0.0394,0.0590,0.0649,0.1209,0.2467,0.3564,0.4459,0.4152,0.3952

P7

dataset) 涉及根据给定钞票的数个度量的照片预测是真钞还是

每个类的观测值数量不均等。一共有 1372 个观察值，4个输入

变量和1个输出变量。变量名如下：

小波变换图像（连续）

小波偏斜变换图像（连续）

小波峰度变换图像（连续）

图像熵（连续）。

类（0 为真钞，1 为假钞）

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

预测最普遍类的基准性能是约 50% 的分类准确率。

该数据集前5行的示例如下：

```
1 3.6216,8.6661,-2.8073,-0.44699,0
2 4.5459,8.1674,-2.4586,-1.4621,0
3 3.866,-2.6383,1.9242,0.10645,0
4 3.4566,9.5228,-4.0112,-3.5944,0
5 0.32924,-4.4552,4.5718,-0.9888,0
6 4.3684,9.6718,-3.9606,-3.1625,0
```

新智元

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

《Iqo

您可能感兴趣的课程

vers Dataset) 涉及根据鸢尾花的测量数据预测花卉品种。

每个类的观察值数量是均等的。一共有 150 个观察值，4个输出变量名如下：

花瓣长度（cm）

花瓣宽度（cm）

类（Iris Setosa，Iris Versicolour，Iris Virginica）

预测最普遍类的基准性能是约 26% 的分类准确率。

该数据集前5行的示例如下：

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信


```

1 5.1,3.5,1.4,0.2,Iris-setosa
2 4.9,3.0,1.4,0.2,Iris-setosa
3 4.7,3.2,1.3,0.2,Iris-setosa
4 4.6,3.1,1.5,0.2,Iris-setosa
5 5.0,3.6,1.4,0.2,Iris-setosa

```



下载地址：<http://t.cn/Rf8GeUq>

鲍鱼数据集

鲍鱼数据集（Abalone Dataset）涉及根据鲍鱼个体的测量数据来预测鲍鱼的年龄（环的

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

class classification）问题，但也可以作为回归问题。每个类的数据集有 4177 个观察值，8个输入变量和1个输出变量。变量名

剥壳重量

内脏重量

壳重

环的数量

预测最普遍类的基准性能是约 16% 的分类准确率，预测平均值的基准性能的均方根误差（RMSE）是约 3.2 个环。

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

该数据集前5行的示例如下：

```
1 1,0,0.99539,-0.05889,0.85243,0.02306,0.83398,-0.37708,1,0.03760,0.85243,-0.17755,0.
2 1,0,1,-0.18829,0.93035,-0.36156,-0.10868,-0.93597,1,-0.04549,0.50874,-0.67743,0.344
3 1,0,1,-0.03365,1,0.00485,1,-0.12062,0.88965,0.01198,0.73082,0.05346,0.85443,0.00827
4 1,0,1,-0.45161,1,1,0.71216,-1,0,0,0,0,0,-1,0.14516,0.54094,-0.39336,-1,-0.54467,-
5 1,0,1,-0.02401,0.94140,0.06531,0.92106,-0.23255,0.77152,-0.16399,0.52798,-0.20275,0
```

下载地址：<http://t.cn/Rf8GDdu>

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

Dataset) 需要根据给定的电离层中的自由电子的雷达回波预

每个类的观察值数量不均等，一共有 351 个观察值，34 个输
量名如下：

预测最普遍类的基准性能是约 64% 的分类准确率，较佳结果达到约 94% 的分类准确率。

该数据集前5行的示例如下：

```
1 1,0,0.99539,-0.05889,0.85243,0.02306,0.83398,-0.37708,1,0.03760,0.85243,-0.17755,0.
2 1,0,1,-0.18829,0.93035,-0.36156,-0.10868,-0.93597,1,-0.04549,0.50874,-0.67743,0.344
3 1,0,1,-0.03365,1,0.00485,1,-0.12062,0.88965,0.01198,0.73082,0.05346,0.85443,0.00827
4 1,0,1,-0.45161,1,1,0.71216,-1,0,0,0,0,0,-1,0.14516,0.54094,-0.39336,-1,-0.54467,-
5 1,0,1,-0.02401,0.94140,0.06531,0.92106,-0.23255,0.77152,-0.16399,0.52798,-0.20275,0
```

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

下载地址：<http://t.cn/Rf8GFY4>

小麦种子数据集

小麦种子数据集（Wheat Seeds Dataset）涉及对不同品种的小麦种子进行预测，给定的是种子的计量数据。

它是一个二元分类问题。每个类的观察值是均等的，一共 210 个观察值，7个输入变量和1个输出变量。变量名如下：

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

预测最普遍类的基准性能是约 28% 的分类准确率。

数据集前5行的示例如下：

```
1 15.26,14.84,0.871,5.763,3.312,2.221,5.22,1
2 14.88,14.57,0.8811,5.554,3.333,1.018,4.956,1
3 14.29,14.09,0.905,5.291,3.337,2.699,4.825,1
4 13.84,13.94,0.8955,5.324,3.379,2.259,4.805,1
5 16.14,14.99,0.9034,5.658,3.562,1.355,5.175,1
```

下载地址：<http://t.cn/RfHHbzw>

最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

波士顿房价数据集

波士顿房价数据集（ Boston House Price Dataset ）包含对房价的预测，以千美元计，给定的条件是房屋及其相邻房屋的详细信息。

该数据集是一个回归问题。每个类的观察值数量是均等的，共有 506 个观察，13 个输入变量和1个输出变量。变量名如下：

| | |
|----------------------|-------------|
| 知识图谱实战（第一期） | 《您可能感兴趣的课程》 |
| 深度学习框架Tensorflow学... | |
| R语言数据分析、展现与实例... | |
| Go语言实战编程（第三期） | |
| 深入浅出Spring（第五期） | |
| 大数据算法导论（第14期） | |
| Python数据分析案例实战（第... | |
| Python金融投资分析实践（第... | |

CRATIO：城镇中自有房屋的比例。
INDUS：工业用地的比例。
NOX：如果边界是河流，则为1；否则为0）。
PTRATIO：城镇自有房屋比例。
LSTAT：区域的加权距离。
DIS：指数。
TAX：财产税率。

PTRATIO：城镇师生比例。
B：1000（Bk-0.63）^ 2，其中 Bk 指代城镇中黑人的比例。
LSTAT：人口中地位低下者的比例。
MEDV：自住房的平均房价，以千美元计。

预测平均值的基准性能的均方根误差（RMSE）是约 9.21 千美元。

>> 最新活动

- 充值700送500，仅20个名额
- 人工智能前沿人才培养计划
- 郭一军“新DBA”精英计划
- 葛一鸣JAVA实战精英培养计划
- 精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

数据集前5行的示例如下：

```
1 0.00632 18.00 2.310 0 0.5380 6.5750 65.20 4.0900 1 296.0 15.30 396.90 4.98 24.00
2 0.02731 0.00 7.070 0 0.4690 6.4210 78.90 4.9671 2 242.0 17.80 396.90 9.14 21.60
3 0.02729 0.00 7.070 0 0.4690 7.1850 61.10 4.9671 2 242.0 17.80 392.0 4.03 54.70
4 0.03237 0.00 2.180 0 0.4580 6.9980 45.80 6.0622 3 222.0 18.70 394.6 2.94 33.40
5 0.06905 0.00 2.180 0 0.4580 7.1470 54.20 6.0622 3 222.0 18.70 396.90 5.33 36.20
```

下载地址：<http://t.cn/RfHTAgY>

时间序列数据集

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

数据集上应用。这些属于需要预测数值或分类的问题，但数据
7个标准时间序列数据集，可用于使用机器学习进行时间序

数据集称为单变量数据集（univariate datasets），其优点是：

的比较；

以下是4个单变量时间序列数据集，均可从datamarket上下载。

洗发水销售数据集（Shampoo Sales Dataset）

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

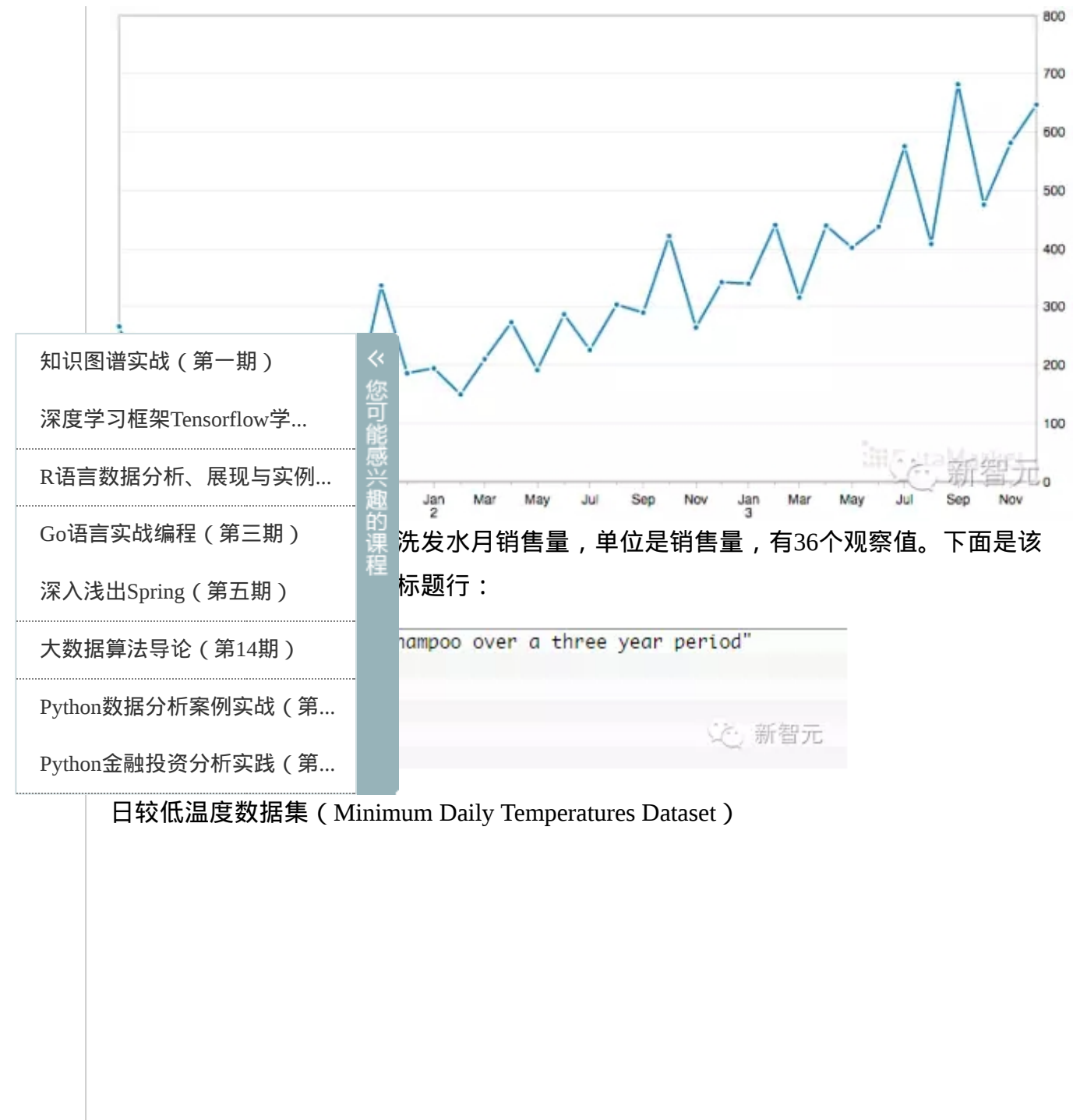
4008-010-006

QQ交谈



QQ交谈

客服微信



>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

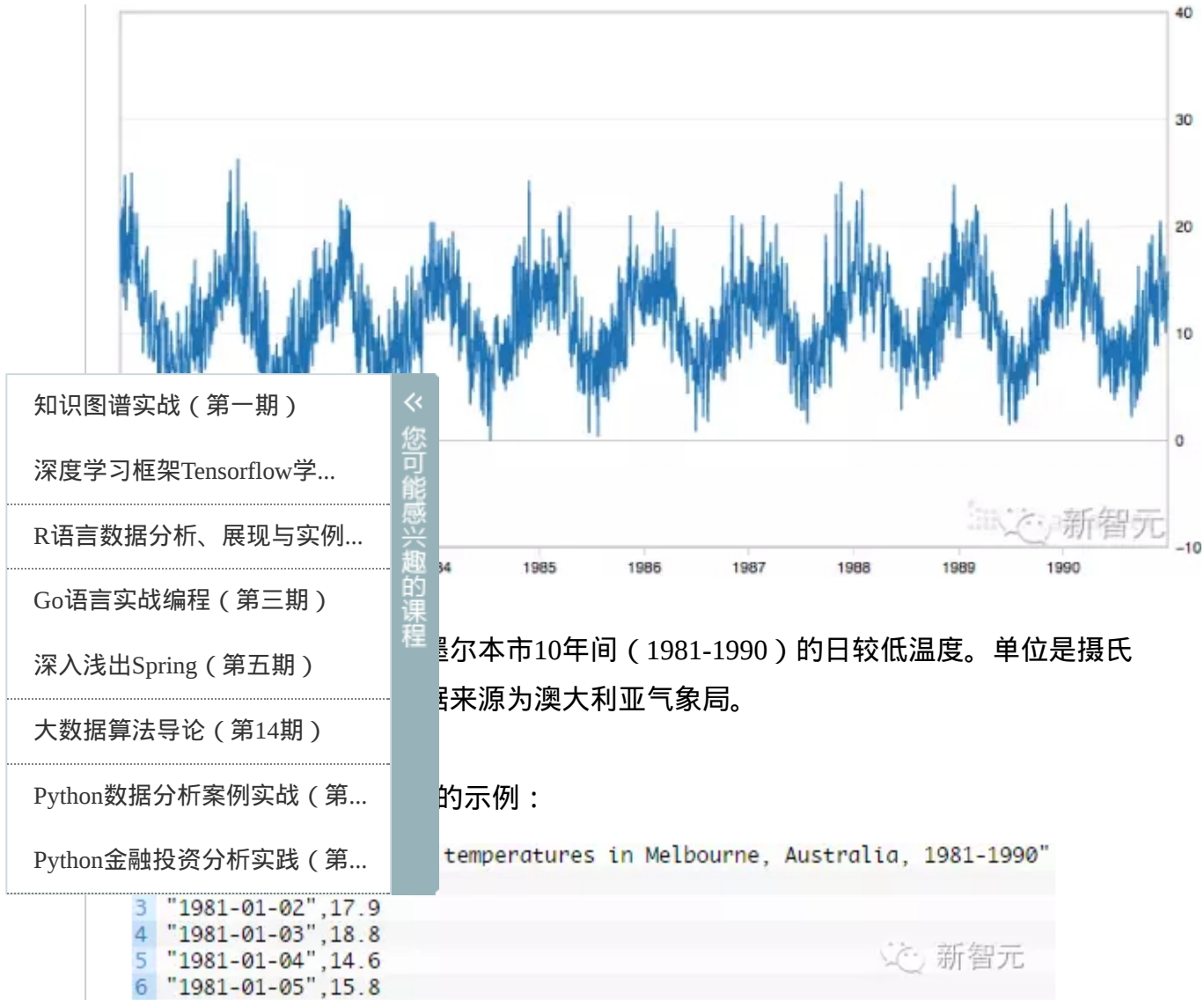
4008-010-006

QQ交谈

QQ交谈



客服微信



每月太阳黑子数数据集（Monthly Sunspot Dataset）

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

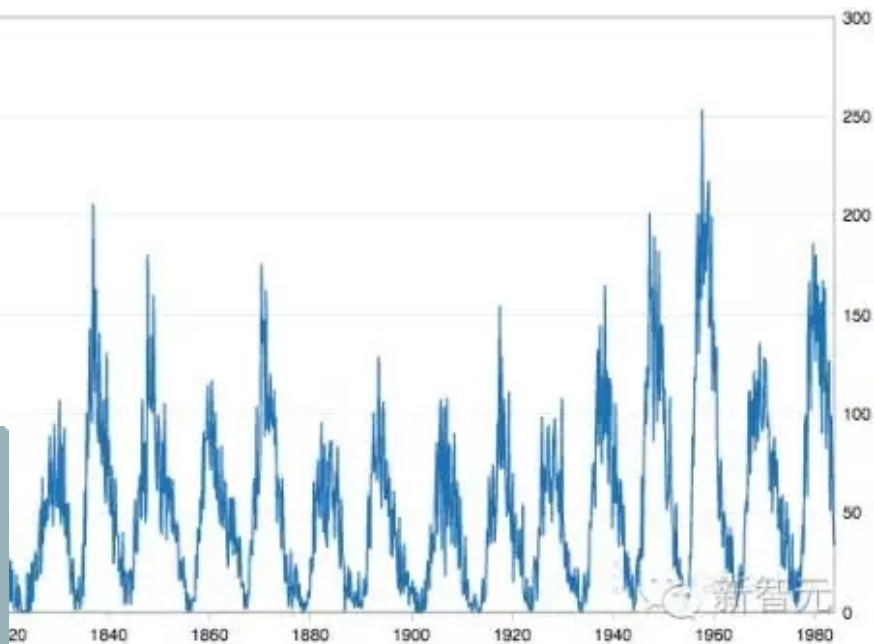
您可能感兴趣的课程

1749-1983) 观测到的每月太阳黑子数量。单位是太阳黑子数
数据集的来源为 Andrews & Herzberg (1985)。

only sunspot numbers 1749-1983"

```
3 "1749-02",62.6  
4 "1749-03",70.0  
5 "1749-04",55.7  
6 "1749-05",85.0
```

每日女婴出生人数数据集



最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信



多变量时间序列数据集

多变量数据集（Multivariate datasets）通常更具挑战性，多变量时间序列数据的主要来源是 UCI 机器学习库（<http://archive.ics.uci.edu/ml/>），下文推荐的3个数据集均可下载。

>> 最新活动

- 充值700送500，仅20个名额
- 人工智能前沿人才培养计划
- 郭一军“新DBA”精英计划
- 葛一鸣JAVA实战精英培养计划
- 精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈

客服微信

EEG 人眼状态数据集

该数据集描述个体的 EEG 数据，以及他们的眼睛是睁着还是闭着。这个问题是为了根据跟定的 EEG 数据预测眼睛的状态。

这是一个分类预测模型问题，共有14980个观察值和15个输入变量。分类值“1”表示眼睛闭着，“0”表示眼睛睁开。数据按时间排序，记录观察结果的时间是117秒。

下面是数据集前5行的示例：

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

您可能感兴趣的课程

```
23,4148.21,4350.26,4586.15,4096.92,4641.03,4222.05,4238.46,421
5,4148.72,4342.05,4586.67,4097.44,4638.97,4210.77,4226.67,420
8,4156.41,4336.92,4583.59,4096.92,4630.26,4207.59,4222.05,420
1,4155.9,4343.59,4582.56,4097.44,4630.77,4217.4,4235.38,4210
1,4151.28,4347.69,4586.67,4095.9,4627.69,4210.77,4244.1,4212.
```

cy Detection Dataset)

特征的数据，目的是预测房间是否在使用中。数据集包含几个小时的观察，属于分类预测问题。数据集包括7个特征值，例如等。

包括标题行：

```
"Humidity","Light","CO2","HumidityRatio","Occupancy"
```

```
2 "1","2015-02-04 17:51:00",23.18,27.272,426,721.25,0.00479298817650529,1
3 "2","2015-02-04 17:51:59",23.15,27.2675,429.5,714,0.00478344094931065,1
4 "3","2015-02-04 17:53:00",23.15,27.245,426,713.5,0.00477946352442199,1
5 "4","2015-02-04 17:54:00",23.15,27.2,426,708.25,0.00477150882603175,1
6 "5","2015-02-04 17:55:00",23.1,27.2,426,704.5,0.00475699293331518,1
7 "6","2015-02-04 17:55:59",23.1,27.2,419,701,0.00475699293331518,1
```

臭氧水平检测数据集

这个数据集描述了6年期间的地面臭氧浓度数据，目的是预测是否“臭氧日”。数据集包含2,536个观察值，73个特征。这是分类预测问题，类别值为“1”表示这天是臭氧日，

最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

为“0”表示正常日。

下面是前5行的示例：

```
1 1/1/1998,0.8,1.8,2.4,2.1,2.2,1.5,1.7,1.9,2.3,3.7,5.5,5.1,5.4,5.4,4.7,4.3,3.5,3.5,
2 1/2/1998,2.8,3.2,3.3,2.7,3.3,3.2,2.9,2.8,3.1,3.4,4.2,4.5,4.5,4.3,5.5,5.1,3.8,3,2.6,
3 1/3/1998,2.9,2.8,2.6,2.1,2.2,2.5,2.5,2.7,2.2,2.5,3.1,4.4,4.6,5.6,5.4,5.2,4.4,3.5,
4 1/4/1998,4.7,3.8,3.7,3.8,2.9,3.1,2.8,2.5,2.4,3.1,3.3,3.1,2.3,2.1,2.3,3.8,2.8,2.4,1.
5 1/5/1998,2.6,2.1,1.6,1.4,0.9,1.5,1.2,1.4,1.3,1.4,2.2,2.3,3.3,3.1,3.1,2.3,3,2.4,2.8,2.
6 1/6/1998,3.1,3.5,3.3,2.5,1.6,1.7,1.6,1.6,2.3,1.8,2.5,3.9,3.4,2.7,3.4,2.5,2.2,4.4,4.
```

总结

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

的标准数据集，你可以用它们来进行机器学习的应用练习。

Weka，scikit-learn 或 R）

高多少。

产生价值的办法，实际应用案例分享与讨论，分析工具，ETL

工具，数据仓库，数据挖掘工具，报表系统等全方位知识

QQ群：81035754



鲜花



握手



雷人



路过



鸡蛋

邀请

分享

收藏

分享到新浪微博

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

上一篇：图解数据科学从业者薪资状况

下一篇：值得膜拜的三个数据分析案例

最新评论

评论

知识图谱实战（第一期）

深度学习框架Tensorflow学...

R语言数据分析、展现与实例...

Go语言实战编程（第三期）

深入浅出Spring（第五期）

大数据算法导论（第14期）

Python数据分析案例实战（第...

Python金融投资分析实践（第...

<< 您可能感兴趣的课程

微博

还可以输入140字

还没有人评论过，赶快抢沙发吧！

>> 最新活动

充值700送500，仅20个名额

人工智能前沿人才培养计划

郭一军“新DBA”精英计划

葛一鸣JAVA实战精英培养计划

精品小班特训营早报优惠多

>> 联系客服

4008-010-006

QQ交谈

QQ交谈



客服微信

关于我们

关于我们

新手指南

教育模式

企业合作

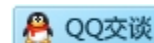
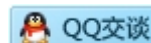
企业会员

联系我们

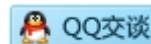
电话：4008-010-006

[讲师招募](#)[选课流程](#)[内训合作](#)[邮箱：kefu@dataguru.cn](mailto:kefu@dataguru.cn)[校园大使](#)[学费返还](#)[媒体合作](#)

客服QQ:

[版权声明](#)[奖学金激励](#)[代理合作](#)

售后QQ:



订阅号



小程序

[关于我们](#) [手机版](#) [友情链接](#) [站点统计](#) [文本模式](#) [小游戏](#)

版权所有 广州市皓岚信息技术有限公司 合作伙伴 中山大学海量数据与云计算研究中心



粤公网安备 44011302000975号 粤ICP备08028958号

Copyright 2011-2018 dataguru.cn All Right Reserved.

[知识图谱实战（第一期）](#)[深度学习框架Tensorflow学...](#)[R语言数据分析、展现与实例...](#)[Go语言实战编程（第三期）](#)[深入浅出Spring（第五期）](#)[大数据算法导论（第14期）](#)[Python数据分析案例实战（第...](#)[Python金融投资分析实践（第...](#)

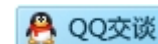
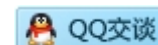
《您可能感兴趣的课程》

» 最新活动

[充值700送500，仅20个名额](#)[人工智能前沿人才培养计划](#)[郭一军“新DBA”精英计划](#)[葛一鸣JAVA实战精英培养计划](#)[精品小班特训营早报优惠多](#)

» 联系客服

4008-010-006



客服微信