

《社交网站的数据挖掘与分析》

图书基本信息

书名：《社交网站的数据挖掘与分析》

出版时间：2012-2

作者：Matthew A. Russell

页数：301

译者：师蓉

版权说明：本站所提供下载的PDF图书仅提供预览和简介以及在线试读，请支持正版图书。

更多资源请访问：www.tushu000.com

《社交网站的数据挖掘与分析》

内容概要

《社交网站的数据挖掘与分析》

作者简介

马修·罗塞尔（Matthew A. Russell），Digital Reasoning Systems公司的技术副总裁和Zaffra公司的负责人，是热爱数据挖掘、开源和Web应用技术的计算机科学家。他也是《Dojo: The Definitive Guide》（O'Reilly出版社）的作者。在LinkedIn上联系他或在Twitter上关注@ptwobrussell，可随时关注他的最新动态。

《社交网站的数据挖掘与分析》

书籍目录

前言	1
第1章绪论：Twitter 数据的处理	9
Python 开发工具的安装	9
Twitter 数据的收集和处理	11
小结	24
第2章微格式：语义标记和常识碰撞	26
XFN 和朋友	27
使用XFN 来探讨社交关系	29
地理坐标：兴趣爱好的共同主线	37
（以健康的名义）对菜谱进行交叉分析	41
对餐厅评论的搜集	43
小结	45
第3章邮箱：虽然老套却很好用	47
mbox：Unix 的入门级邮箱	48
mbox+CouchDB= 随意的Email 分析	54
将对话线程化到一起	70
使用SIMILE Timeline 将邮件“事件”可视化	79
分析你自己的邮件数据	82
小结	84
第4章Twitter：朋友、关注者和Setwise 操作	85
REST 风格的和OAuth-Cladded API	86
干练而中肯的数据采集器	90
友谊图的构建	108
小结	116
第5章Twitter：tweet，所有的tweet，只有tweet	118
笔PK 剑：和tweet PK 机枪（!?!）	118
对tweet 的分析（每次一个实体）	121
并置潜在的社交网站（或#JustinBieber VS #TeaParty）	144
对大量tweet 的可视化	155
小结	163
第6章LinkedIn：为了乐趣（和利润？）将职业网络聚类	164
聚类的动机	165
按职位将联系人聚类	167
获取补充个人信息	183
从地理上聚类网络	188
小结	192
第7章Google Buzz：TF-IDF、余弦相似性和搭配	194
Buzz=Twitter+ 博客（??）	195
使用NLTK 处理数据	198
文本挖掘的基本原则	201
查找相似文档	208
在二元语法中发Buzz	215
利用Gmail	221
在中断之前试着创建一个搜索引擎.....	225
小结	226
第8章博客及其他：自然语言处理（等）	228
NLP：帕累托式介绍	228

《社交网站的数据挖掘与分析》

使用NLTK的典型NLP管线231
使用NLTK检测博客中的句子234
对文件的总结237
以实体为中心的分析：对数据的深入了解245
小结256
第9章Facebook：一体化的奇迹257
利用社交网络数据258
对Facebook数据的可视化274
小结294
第10章语义网：简短的讨论296
发展中的变革296
人不可能只靠事实生活297
期望301

《社交网站的数据挖掘与分析》

精彩短评

- 1、好书,介绍了很多python工具包和社交网站的api,就是twitter/facebook什么的还得翻墙,太纠结了...
- 2、3本书共比**便宜了14元!
- 3、很烂的一本书
- 4、可视化工具和Python编程实践
- 5、没什么深入的技术,多在介绍 api
- 6、使用python处理web上的数据。例子加说明。
- 7、第一次买,感觉不错,都是经典的书,内容么话说,质量也还好
- 8、感觉略坑,原书好象是2010年写的,翻到国内都2012年了, twitter的API都不知道改了几次了吧【笑而不语
- 9、虽然翻译差强人意但是这本书的概念很好。
- 10、拿着随手翻翻吧,能知道一些数据挖掘方面的工具。除此好像没有什么东西
- 11、寒假让我哥帮我打印了一份英文原版,装逼翻了一页就扔书架上,原来该相遇的还是会相遇,连书都是这样。
- 12、单趟地铁大致翻完了 也不怎么有意思 但是地铁也坐了太久吧擦
- 13、networkx实现图论, NLTK做语义分析, CouchDB,redis数据存儲, lucene搜索.....好多轮子,不要重复发明轮子
- 14、Facebook、Twitter和LinkedIn产生了大量宝贵的社交数据,但是怎样才能找出谁通过社交媒介正在进行联系?他们在讨论些什么?或者他们在哪儿?本书简洁而且具有可操作性,它将揭示如何回答这些问题甚至更多的问题。你将学到如何组合社交网络数据、分析技术,如何通过可视化帮助你找到你一直在社交世界中寻找的内容,以及你闻所未闻的有用信息。本书每章都介绍了在社交网络的不同领域挖掘数据的技术,这些领域包括博客和电子邮件。你所需要具备的就是一定的编程经验和学习基本的Python工具的意愿。通过本书,你将 . 获得对社交网络世界的直观认识 . 使用GitHub上灵活的脚本来获取从诸如Twitter、Facebook和LinkedIn等社交网络API中的数据 . 学习如何应用便捷的Python工具来交叉分析你所收集的数据 . 通过XFN探讨基于微格式的社交联系 . 应用诸如TF-IDF、余弦相似性、搭配分析、文档摘要、派系检测之类的先进挖掘技术 . 通过基于HTML 5和JavaScript工具包的网络技术建立交互式可视化
- 15、翻译质量不高
- 16、虽然只了解了点皮毛,但如果深层研究下去,应该会挺有意思的。
- 17、这是目前最热门的方向,好好学习一下,希望能有大的帮助
- 18、数学无力.....
- 19、如题,书里代码不少,偏重应用,对社交网络数据挖掘感兴趣的童鞋可以看看~
- 20、有大段的代码,不适合我这种随便看看的心态。。。而且很多其实已经过时了,twitter api不一样了, buzz也关掉了。 anyway,大致了解了一下概念,真的要写程序分析的时候再查。
- 21、有技术有内容,有讲解有代码
- 22、例子都是twitter,没有微博,要是再本土化一些就更好了
- 23、水书,这个系列越来越让人失望
- 24、这种书豆瓣阅读居然还要付费,翻译较差。Python不太好用。
- 25、不错的书,很好很强大,只是还没有看
- 26、Python简单易懂
- 27、实用
- 28、书很好,发货也挺及时~
- 29、SNS大数据分析必读,但是引用了Twitter和Facebook的数据,要通过网络访问需要一点额外的功夫,这个你懂的。LinkedIn的数据时而可用时而不可用。
- 30、送给小朋友的,有趣且增加知识量,不错
- 31、这个是从从事微博等社交网络挖掘者的利器。不足之处是API运行不太稳定
- 32、python、、、
- 33、不错,值得读读,就算API访问有难度,工具包的介绍,以及代码的编写还是值得学习的。

《社交网站的数据挖掘与分析》

- 34、看见是Python~结果发现难度大了点，所以给同学看了~@KPSN~
- 35、很有深度不太好理解
- 36、技术点不多
- 37、很多案例，主要讲facebook的，很实用。
- 38、比较专业的书 挖掘很好 可惜分析没有
- 39、书还是挺技术的,我承认我买错了
- 40、已弃疗，没VPN看这本书，也就黄粱一梦。
- 41、暂无~~
- 42、书不错很新很新鲜里面的各种也不错
- 43、社交网站之间的竞争，归根结底是数据之争。针对社交网站的数据挖掘与分析，无疑能赢得更多的竞争筹码。
- 44、社会计算必备图书
- 45、。。。|毕业设计|cmm老师
- 46、就喜欢实践性的，nltk那本书也如此
- 47、还灭有看
- 48、刚刚到手，正在研读。
- 49、面向程序员的书，没啥意义
- 50、第二版更好，及时更新解决了twitter API的问题。
<https://github.com/ptwobrussell/Mining-the-Social-Web-2nd-Edition>
- 51、仅仅粗读部分章节，留待以后细读
- 52、读晚了.....twitter API 更新到很不友好的地步了.....
- 53、先前下载一本英文版的，还是看中文的舒服，其中的代码是用python写的，主要是对faceook做数据挖掘与分析。
- 54、主要是分析的国外三大网站的数据！
- 55、Mining the Social Web 中文版。侧重介绍如何获取数据。
- 56、结合数据挖掘的一些算法，运用很多开源工具和开放api进行实践。总体没有集体智慧编程有深度
- 57、神翻
- 58、当时做微博数据挖掘时读的，很喜欢Buzz那一章
- 59、太浅了。根本不能叫数据挖掘。
- 60、因为它是用python的，而且又是得奖的
- 61、买过来看一下，需要用到社交网络挖掘方面的知识
- 62、甜点而非主菜！
- 63、买了，还没看，看后评价
- 64、用来入门不错，实习的时候做大数据，其中实验讲的也不错。
- 65、这本书的内容真的很不错，但是感觉中文的翻译并不是很好。
- 66、现在火爆的社交网站，其中的数据的价值更是至高无上的
- 67、虽然不深 但是可以带着入门
- 68、但在中国测试，要翻墙去twitter
- 69、感觉有点过时，没怎么看。现在在黄tao手上
- 70、一本理论与实际紧密结合的好书，打开眼界。
- 71、《社交网站的数据挖掘与分析》这本书很不错嘛。
- 72、数据获取部分非常实用，分析就太小儿科了。边实验边看书非常有效率。由于SNS的API经常改，这书已过期需要看第二版，第二版改用了ipython notebook
- 73、对分析算法讲的不多，国外主要的社交网络的数据挖掘都有覆盖，希望将来有机会自己能做一些类似的针对豆瓣，微博，人人之类的数据挖掘
- 74、帮助他人买的。
- 75、没有给发票，书还没看，应该可以吧。。
- 76、介绍外国网站的api，而且还是一些已经过时的api。随手翻翻，了解数据挖掘的概念即可，精读意义不是很大。

《社交网站的数据挖掘与分析》

- 77、还没读完，读完再补
- 78、入门好书
- 79、这本书介绍的api已经跟不上节奏了...第一章弃
- 80、从中也可体味社交网站的细微差别。
- 81、有点难度，需要有py语言基础
- 82、书的内容很好，需要消耗很久哦
- 83、这本书真是超级好，介绍了大量非常有用的工具，而且有很多我从前没听过的新概念、新方法，读完之后可谓是受益良多。书中内容要借助python来实现，所以要有python基础，但是这点也不是关键，因为python还是比较好学的，现学都不难。虽然讲的是外国的社交网站，而且国内还不能够直接看到，但是移花接木，生活中还是可以用到的。
- 84、名不副实
- 85、喜欢大数据的仁兄值得一看~~~
- 86、我的毕设主题
- 87、
- 88、各大#API使用指南#
- 89、没什么太特别的地方
- 90、好书，喜欢
- 91、拿到书了，挺满意的。期待阅读有收获~
- 92、代码无注释，API失效，上Twitter还得找代理。国内玩家表示自己等级太低，杠不住
- 93、感觉翻译的不是很好，再加上个人对python不是很了解，对社交网络的挖掘也是刚刚接触，看着有些吃力。。。
- 94、书不错，使用python的一些工具来数据挖掘
- 95、略out of date，总的data exploration techniques还是挺有用的
- 96、货刚到，现在开始读
- 97、很好，再说也还是这，呵呵
- 98、能翻译的这么烂也是本事。内容上有意思的是方向上的东西，代码基本略过了。
- 99、社交分析非常棒的实践书籍，还介绍了非常多的开源工具比如NLTK、CouchDB、redis、一些可视化工具等。
- 100、自己查找(学到了蛮多工具包的)
- 101、感觉是好书，但还没看，使用的语言，之前又不懂。以后有时间再慢慢看了

《社交网站的数据挖掘与分析》

版权说明

本站所提供下载的PDF图书仅提供预览和简介，请支持正版图书。

更多资源请访问:www.tushu000.com