

《集体智慧编程》

图书基本信息

书名：《集体智慧编程》

13位ISBN编号：9787121254433

出版时间：2015-3

作者：Toby Segaran

页数：356

译者：莫映,王开福

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

《集体智慧编程》

内容概要

《集体智慧编程》以机器学习与计算统计为主题背景，专门讲述如何挖掘和分析Web上的数据和资源，如何分析用户体验、市场营销、个人品味等诸多信息，并得出有用的结论，通过复杂的算法从Web网站获取、收集并分析用户的数据和反馈信息，以便创造新的用户价值和商业价值。全书内容翔实，包括协作过滤技术（实现关联产品推荐功能）、集群数据分析（在大规模数据集中发掘相似的数据子集）、搜索引擎核心技术（爬虫、索引、查询引擎、PageRank算法等）、搜索海量信息并进行分析统计得出结论的优化算法、贝叶斯过滤技术（垃圾邮件过滤、文本过滤）、用决策树技术实现预测和决策建模功能、社交网络的信息匹配技术、机器学习和人工智能应用等。

《集体智慧编程》是Web开发者、架构师、应用工程师等的绝佳选择。

《集体智慧编程》

书籍目录

| | |
|-------------------------------|------|
| 前言..... | viii |
| 第1章 集体智慧导言..... | 1 |
| 什么是集体智慧..... | 2 |
| 什么是机器学习..... | 3 |
| 机器学习的局限..... | 4 |
| 真实生活中的例子..... | 5 |
| 学习型算法的其他用途..... | 5 |
| 第2章 提供推荐..... | 7 |
| 协作型过滤..... | 7 |
| 搜集偏好..... | 8 |
| 寻找相近的用户..... | 9 |
| 推荐物品..... | 15 |
| 匹配商品..... | 17 |
| 构建一个基于del.icio.us的链接推荐系统..... | 19 |
| 基于物品的过滤..... | 22 |
| 使用MovieLens数据集..... | 25 |
| 基于用户进行过滤还是基于物品进行过滤..... | 27 |
| 练习..... | 28 |
| 第3章 发现群组..... | 29 |
| 监督学习和无监督学习..... | 29 |
| 单词向量..... | 30 |
| 分级聚类..... | 33 |
| 绘制树状图..... | 38 |
| 列聚类..... | 40 |
| K-均值聚类..... | 42 |
| 针对偏好的聚类..... | 44 |
| 以二维形式展现数据..... | 49 |
| 有关聚类的其他事宜..... | 53 |
| 练习..... | 53 |
| 第4章 搜索与排名..... | 54 |
| 搜索引擎的组成..... | 54 |
| 一个简单的爬虫程序..... | 56 |
| 建立索引..... | 58 |
| 查询..... | 63 |
| 基于内容的排名..... | 64 |
| 利用外部回指链接..... | 69 |
| 从点击行为中学习..... | 74 |
| 练习..... | 84 |
| 第5章 优化..... | 86 |
| 组团旅游..... | 87 |
| 描述题解..... | 88 |
| 成本函数..... | 89 |
| 随机搜索..... | 91 |
| 爬山法..... | 92 |
| 模拟退火算法..... | 95 |
| 遗传算法..... | 97 |
| 真实的航班搜索..... | 101 |

| | |
|------------------------------|------------|
| 涉及偏好的优化..... | 106 |
| 网络可视化..... | 110 |
| 其他可能的应用场合..... | 115 |
| 练习..... | 116 |
| 第6章 文档过滤..... | 117 |
| 过滤垃圾信息..... | 117 |
| 文档和单词..... | 118 |
| 对分类器进行训练..... | 119 |
| 计算概率..... | 121 |
| 朴素分类器..... | 123 |
| 费舍尔方法..... | 127 |
| 将经过训练的分类器持久化..... | 132 |
| 过滤博客订阅源..... | 134 |
| 对特征检测的改进..... | 136 |
| 使用Akismet..... | 138 |
| 替代方法..... | 139 |
| 练习..... | 140 |
| 第7章 决策树建模..... | 142 |
| 预测注册用户..... | 142 |
| 引入决策树..... | 144 |
| 对树进行训练..... | 145 |
| 选择最合适的拆分方案..... | 147 |
| 以递归方式构造树..... | 149 |
| 决策树的显示..... | 151 |
| 对新的观测数据进行分类..... | 153 |
| 决策树的剪枝..... | 154 |
| 处理缺失数据..... | 156 |
| 处理数值型结果..... | 158 |
| 对住房价格进行建模..... | 158 |
| 对“热度”评价进行建模..... | 161 |
| 什么时候使用决策树..... | 164 |
| 练习..... | 165 |
| 第8章 构建价格模型..... | 167 |
| 构造一个样本数据集..... | 167 |
| k-最近邻算法..... | 169 |
| 为近邻分配权重..... | 172 |
| 交叉验证..... | 176 |
| 不同类型的变量..... | 178 |
| 对缩放结果进行优化..... | 181 |
| 不对称分布..... | 183 |
| 使用真实数据——eBay API..... | 189 |
| 何时使用k-最近邻算法..... | 195 |
| 练习..... | 196 |
| 第9章 高阶分类：核方法与SVM..... | 197 |
| 婚介数据集..... | 197 |
| 数据中的难点..... | 199 |
| 基本的线性分类..... | 202 |
| 分类特征..... | 205 |
| 对数据进行缩放处理..... | 209 |

《集体智慧编程》

| | |
|--------------------|-----|
| 理解核方法..... | 211 |
| 支持向量机..... | 215 |
| 使用LIBSVM..... | 217 |
| 基于Facebook的匹配..... | 219 |
| 练习..... | 225 |
| 第10章 寻找独立特征..... | 226 |
| 搜集一组新闻..... | 227 |
| 先前的方法..... | 231 |
| 非负矩阵因式分解..... | 232 |
| 结果呈现..... | 240 |
| 利用股票市场的数据..... | 243 |
| 练习..... | 248 |
| 第11章 智能进化..... | 250 |
| 什么是遗传编程..... | 250 |
| 将程序以树形方式表示..... | 253 |
| 构造初始种群..... | 257 |
| 测试题解..... | 259 |
| 对程序进行变异..... | 260 |
| 交叉..... | 263 |
| 构筑环境..... | 265 |
| 一个简单的游戏..... | 268 |
| 更多可能性..... | 273 |
| 练习..... | 276 |
| 第12章 算法总结..... | 277 |
| 贝叶斯分类器..... | 277 |
| 决策树分类器..... | 281 |
| 神经网络..... | 285 |
| 支持向量机..... | 289 |
| k-最近邻..... | 293 |
| 聚类..... | 296 |
| 多维缩放..... | 300 |
| 非负矩阵因式分解..... | 302 |
| 优化..... | 304 |
| 附录A：第三方函数库..... | 309 |
| 附录B：数学公式..... | 316 |
| 索引..... | 323 |

《集体智慧编程》

精彩短评

- 1、在这里，代码比数学还要难懂，不讲数学就是耍流氓。再浅显也只能用做入门书。
- 2、github 完成作业：<https://github.com/wanggang3333/CollectiveIntelligence> 虽然是2015年版本，但是已略显陈旧。
- 3、粗读了一遍，通过实例讲解理论，适合刚接触机器学习、数据挖掘方向的读者，很好的入门书籍，浅显易懂，很多细节看一遍看不懂，还需要再多看几遍
- 4、读过电子版的先前印批9787121075391。2015.11.5jd
- 5、看完之后才明白别人说的代码比较老是什么意思，讲的还是不错的，浅显易懂，就是要选择性阅读，不要纠结于原始代码能否运行，能力理解方法和代码实现就行，个别跟着敲一敲，入门兴趣书
- 6、包含了很多python语言代码实现，互联中信息的处理和应用实例，总结了各种处理的方法，最后还有遗传算法类似机器编程的实现，理论结合实际，以后如果有需要，可以再对其中的方法和实例进入更深入的研究和应用
- 7、数据挖掘算法入门好书，浅显易懂，应该早几年看它的代码的例子有点弱，有些打印的程序片段感觉是凑字数的
- 8、字典：setdefault。None关键字。''.join
- 9、和“机器学习实战”有的一拼，就是随书的很多资源都404了好多demo所以很多demo没有run起来~

《集体智慧编程》

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