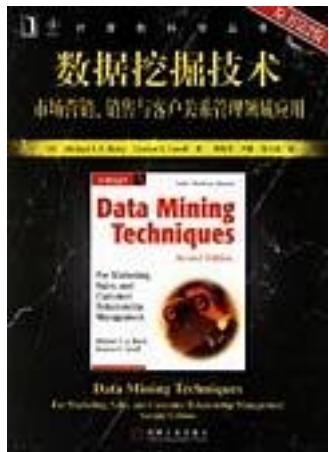


# 数据挖掘技术



[数据挖掘技术 下载链接1](#)

著者:Gordon S.Linoff

出版者:清华大学出版社

出版时间:2013-3

装帧:平装

isbn:9787302310143

《数据挖掘技术:应用于市场营销、销售与客户关系管理(第3版)》内容简介：谁将是忠实的客户？谁将不是呢？哪些消息对哪些客户细分最有效？如何最大化客户的价值？如何将客户的价值最大化？《数据挖掘技术:应用于市场营销、销售与客户关系管理(第3版)》提供了强大的工具，可以从上述和其他重要商业问题所在的公司数据库中提取它们的答案。自《数据挖掘技术:应用于市场营销、销售与客户关系管理(第3版)》第1版问世以来，数据挖掘已经日益成为现代商业不可缺少的工具。在这个最新版本中，作者对每个章节都进行了大量的更新和修订，并且添加了几个新的章节。《数据挖掘技术:应用于市场营销、销售与客户关系管理(第3版)》保留了早期版本的重点，指导市场分析师、业务经理和数据挖掘专家利用数据挖掘方法和技术来解决重要的商业问题。在不牺牲准确度的前提下，为了简单起见，即使是复杂的主题，作者也进行了简洁明了的介绍，并尽量减少对技术术语或数学公式的使用。每个技术主题都通过案例研究和源自作者经验的真实案例进行说明，每章都包含了针对从业者的宝贵提示。书中介绍的新技术和更为深入的技术包括：线性和逻辑回归模型、增量响应（提升）建模、朴素贝叶斯模型、表查询模型、相似度模型、径向基函数网络、期望值最大化（EM）聚类和群体智慧。新的章节专门讨论了数据准备、派生变量、主成分分析和其他变量减少技术，以及文本挖掘。

在建立了全面的数据挖掘应用业务环境，并介绍了所有数据挖掘项目通用的数据挖掘方法论的各个方面之后，《数据挖掘技术:应用于市场营销、销售与客户关系管理(第3版)》详细介绍了每个重要的数据挖掘技术。

## 作者介绍:

Gordon S.Linoff和Michael J.A.Berry在数据挖掘领域的知名度众所周知。他们是Data Miners公司——一家从事数据挖掘的咨询公司——的创始人，而且他们已经共同撰写了一些在该领域有影响力和得到广泛阅读的书籍。他们共同撰写的第一本书是Data Mining

Techniques的第一个版本，于1997年出版。自那时起，他们就一直积极地挖掘各种行业里的数据。持续的实践分析工作使得两位作者能够紧跟数据挖掘、预测以及预测分析领域的快速发展。Gordon和Michael严格地独立于供应商。通过其咨询工作，作者接触了所有主要软件供应商（以及一些小的供应商）的数据分析软件。他们相信好的结果不在于是采用专用的还是开源的软件，命令行的还是点击的软件，而是在于创新思维和健全的方法。

Gordon和Michael专注于数据挖掘在营销和客户关系管理方面的应用——例如，为交叉销售和向上销售改进推荐，预测未来的用户级别，建模客户生存期价值，根据用户行为对客户进行划分，为访问网站的客户选择最佳登录页面，确定适合列入营销活动的候选者，以及预测哪些客户处于停止使用软件包、服务或药物治疗的风险中。Gordon和Michael致力于分享他们的知识、技能以及对这个主题的热情。当他们自己不挖掘数据时，他们非常喜欢通过课程、讲座、文章、现场课堂，当然还有你要读的这本书来教其他人。经常可以发现他们在会议上发言和在课堂上授课。作者还在blog.data-miners.com维护了一个数据挖掘的博客。

Gordon生活在曼哈顿。在本书之前，他最近的一本书是Data Analysis Using SQL and Excel，已经由Wiley于2008年出版。

Michael生活在马萨诸塞州剑桥市。他除了在Data Miners从事咨询工作之外，还在波士顿大学卡罗尔管理学院讲授市场营销分析（Marketing Analytics）课程。

## 目录: 第1章 什么是数据挖掘以及为什么要进行数据挖掘

### 1.1 什么是数据挖掘

#### 1.1.1 数据挖掘是一项业务流程

#### 1.1.2 大量的数据

#### 1.1.3 有意义的模式和规则

#### 1.1.4 数据挖掘和客户关系管理

### 1.2 为什么是现在

#### 1.2.1 数据正在产生

#### 1.2.2 数据正存在于数据仓库中

#### 1.2.3 计算能力能够承受

#### 1.2.4 对客户关系管理的兴趣非常强烈

#### 1.2.5 商业的数据挖掘软件产品变得可用

### 1.3 数据挖掘人员的技能

### 1.4 数据挖掘的良性循环

### 1.5 业务数据挖掘的案例研究

#### 1.5.1 识别美国银行的业务挑战

#### 1.5.2 应用数据挖掘

#### 1.5.3 对结果采取行动

#### 1.5.4 度量数据挖掘的影响

### 1.6 良性循环的步骤

1.6.1 识别业务机会

1.6.2 将数据转换为信息

1.6.3 根据信息采取行动

1.6.4 度量结果

1.7 良性循环上下文中的数据挖掘

1.8 经验教训

第2章 数据挖掘在营销和客户关系管理中的应用

2.1 两个客户生存周期

2.1.1 客户个人生存周期

2.1.2 客户关系生存周期

2.1.3 基于订阅的关系和基于事件的关系

2.2 围绕客户生存周期组织业务流程

2.2.1 客户获取

2.2.2 客户激活

2.2.3 客户关系管理

2.2.4 赢回

2.3 数据挖掘应用于客户获取

2.3.1 识别好的潜在客户

2.3.2 选择通信渠道

2.3.3 挑选适当的信息

2.4 数据挖掘示例：选择合适的地方做广告

2.4.1 谁符合剖析

2.4.2 度量读者群的适应度

2.5 数据挖掘改进直接营销活动

2.5.1 响应建模

2.5.2 优化固定预算的响应

2.5.3 优化活动收益率

2.5.4 抵达最受信息影响的人

2.6 通过当前客户了解潜在客户

2.6.1 在客户成为“客户”以前开始跟踪他们

2.6.2 收集新的客户信息

2.6.3 获取时间变量可以预测将来的结果

2.7 数据挖掘应用于客户关系管理

2.7.1 匹配客户的活动

2.7.2 减少信用风险

2.7.3 确定客户价值

2.7.4 交叉销售、追加销售和推荐

2.8 保留

2.8.1 识别流失

2.8.2 为什么流失是问题

2.8.3 不同类型的流失

2.8.4 不同种类的流失模型

2.9 超越客户生存周期

2.10 经验教训

第3章 数据挖掘过程

3.1 会出什么问题

3.1.1 学习的东西不真实

3.1.2 学习的东西真实但是无用

3.2 数据挖掘类型

3.2.1 假设检验

3.2.2 有指导数据挖掘

3.2.3 无指导数据挖掘

3.3 目标、任务和技术

3.3.1 数据挖掘业务目标

3.3.2 数据挖掘任务

3.3.3 数据挖掘技术

3.4 制定数据挖掘问题：从目标到任务再到技术

3.4.1 选择广告的最佳位置

3.4.2 确定向客户提供的最佳产品

3.4.3 发现分支或商店的最佳位置

3.4.4 根据未来利润划分客户

3.4.5 减少暴露于违约的风险

3.4.6 提高客户保留

3.4.7 检测欺诈性索赔

3.5 不同技术对应的任务

3.5.1 有一个或多个目标

3.5.2 目标数据是什么

3.5.3 输入数据是什么

3.5.4 易于使用的重要性

3.5.5 模型可解释性的最重要性

3.6 经验教训

第4章 统计学入门：关于数据，你该了解些什么

4.1 奥卡姆 (Occam) 剃刀

4.1.1 怀疑论和辛普森悖论

4.1.2 零假设 (Null Hypothesis)

4.1.3 p-值

4.2 观察和度量数据

4.2.1 类别值

4.2.2 数值变量

4.2.3 更多的统计思想

4.3 度量响应

4.3.1 比例标准误差

4.3.2 使用置信区间比较结果

4.3.3 利用比例差异比较结果

4.3.4 样本大小

4.3.5 置信区间的真正含义是什么

4.3.6 实验中检验和对照的大小

4.4 多重比较

4.4.1 多重比较的置信水平

4.4.2 Bonferroni修正

4.5 卡方检验

4.5.1 期望值

4.5.2 卡方值

4.5.3 卡方值与比例差异的比较

4.6 示例：区域和开局卡方

4.7 案例研究：利用A/B检验比较两种推荐系统

4.7.1 第一个指标：参与会话

4.7.2 第二个指标：每个会话的日收益

4.7.3 第三个指标：每天谁取胜

4.7.4 第四个指标：每个会话的平均收益

.....

第5章 描述和预测：剖析与预测建模

第6章 使用经典统计技术的数据挖掘

第7章 决策树

第8章 人工神经网络

第9章 最近邻方法：基于记忆的推理和协同过滤

第10章 了解何时应担忧：使用生存分析了解客户

第11章 遗传算法与群体智能

- 第13章 发现相似的岛屿：自动群集检测
- 第14章 其他的群集检测方法
- 第15章 购物篮分析和关联规则
- 第16章 链接分析
- 第17章 数据仓库、OLAP、分析沙箱和数据挖掘
- 第18章 构建客户签名
- 第19章 派生变量：使数据的含义更丰富
- 第20章 减少变量数量的技术
- 第21章 仔细聆听客户所述：文本挖掘
- · · · · (收起)

[数据挖掘技术 下载链接1](#)

## 标签

数据挖掘

数据分析

营销

CRM

商业智能

市场

大数据

客户

## 评论

很想找给这本书4星以上的人聊一聊，问问觉得写得好是因为自己底子厚还是真的觉得翻译得不错，在我看这译者数据挖掘懂不懂不好说，其他知识几乎一塌糊涂。内容还可以，不太注重数学过程，就讲讲原理，现在工具多，也无所谓。对神经网络部分失望，

其他不表，我看了5天，2天浪费在翻译上，一言难尽。

这本介绍的没有重点，如果偏实践业务，应该对多些案例，如果偏技术，那就应该更深入，并且翻译的实在是太差了，各种拗口不知所云。

尼玛，花了很长时间才啃完的一本书

广告、销售、运营感觉都是广义的客户关系管理的一部分，在技术上手段也有很多相似之处

翻译问题？根本读不进去，只能看图。

系统详细，值得入手

一些概念和方法论值得学习应用

内容深度不够

看到“集中度”翻译成“浓度”，我就放弃了。。。

很适合刚学了一大堆数学不知道怎么应用的初学者人群，但是由于出版时间限制，一些问题还需要多体谅。

趁着有整块时间终于读完了。。基本覆盖了data

mining能用到的各种方法，也只是粗浅的介绍，具体推导过程肯定不会一个一个详述，技术本身不重要，怎么与业务结合才是需要重点培养sense的。最后，翻译嘛，有的地方还是各种拗口。

门外汉入门数据挖掘，必读书

英文原著是值得给5星经典好书，中文版的问题还是出在翻译上，没能做到准确、流畅，让原本简单的模型描述，变得晦涩。这点上，机械工业出版社的《商务和经济统计》给出了很好的示范，期待数据科学类专业书籍的中文翻译能更进一步

广泛而不深入，随便翻翻吧

大部头，CRM相关业务比较多，数据挖掘的算法和实现没有涉及。适合业务分析师看的书。

翻译成这样也是佩服

一刷成功 这是我离数据挖掘最近的一次，也是我离数据挖掘最远的一次

好书一本

这本书为什么打分这么高，可能每个人偏重不一样吧。以后有时间再看一遍。

这本书翻译的实在太烂了，看了5章，弃坑

---

[数据挖掘技术 下载链接1](#)

## 书评

书本身非常好，但翻译很差。outbound在营销中明明是“外呼”，也就是外呼营销，“打电话营销的意思，书里居然翻译成“出站”，相当无语。。客户关系生命周期里，“潜在客户”、“新客户”、“已建立的客户”、“前客户”这几个名词翻译的也不敢恭维，应该叫“潜在客户”、“新客…

---

内容4分，翻译-1分

翻译实在太差，不如直接看英文版，清华大学出版社的烂名声果然不是盖的  
“古代水手学会了如何避免为保护西西里和意大利大陆之间狭窄海峡的锡拉岩礁岩石和卡律布迪斯漩涡”这一看就知道不是人翻译出来的。

---

[数据挖掘技术 下载链接1](#)