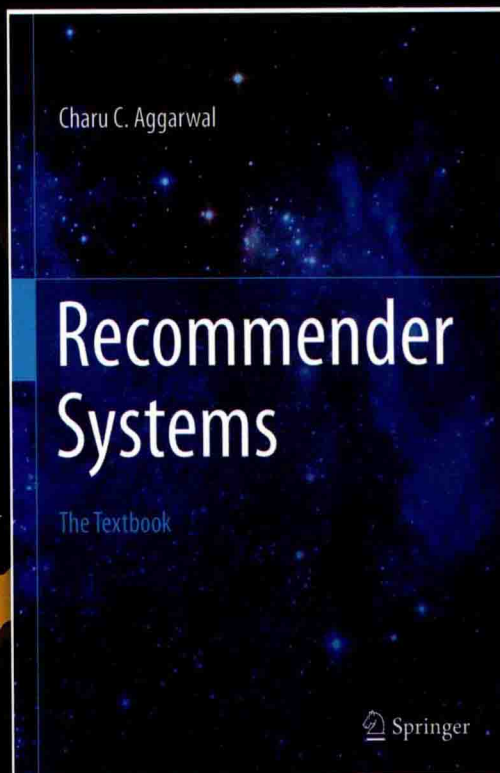


推荐系统

原理与实践

[美] 查鲁·C.阿加沃尔 (Charu C. Aggarwal) 著
黎玲利 尹丹 李默涵 王宏志 等译

Recommender Systems
The Textbook



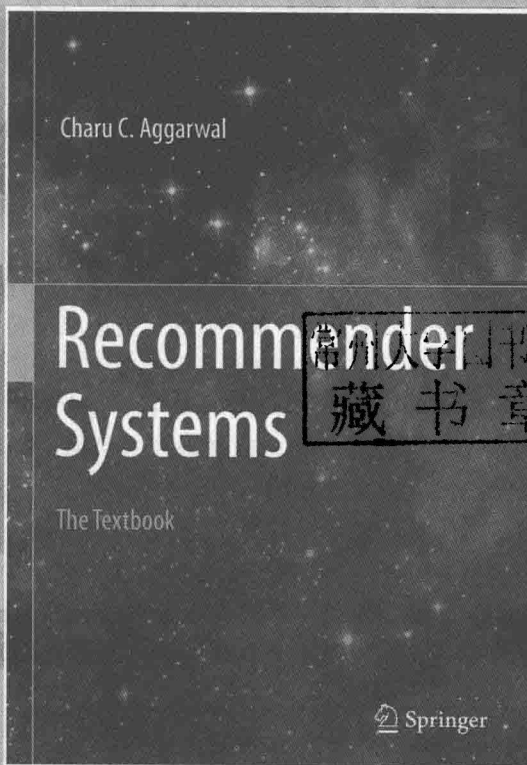
计 算 机 科 学 丛 书


推荐系统

原理与实践

[美] 查鲁·C.阿加沃尔 (Charu C. Aggarwal) 著
黎玲利 尹丹 李默涵 王宏志 等译

Recommender Systems
The Textbook



 机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

推荐系统：原理与实践 / (美) 查鲁·C. 阿加沃尔 (Charu C. Aggarwal) 著；黎玲利等译。
—北京：机械工业出版社，2018.5

(计算机科学丛书)

书名原文：Recommender Systems: The Textbook

ISBN 978-7-111-60032-9

I. 推… II. ①查… ②黎… III. 计算机网络—研究 IV. TP393

中国版本图书馆 CIP 数据核字 (2018) 第 110208 号

本书版权登记号：图字 01-2016-6250

Translation from the English language edition:

Recommender Systems: The Textbook

by Charu C. Aggarwal.

Copyright © Springer International Publishing Switzerland 2016.

Springer is part of Springer Nature.

All rights reserved.

本书中文简体字版由 Springer 授权机械工业出版社独家出版。未经出版者书面许可，不得以任何方式复制或抄袭本书内容。

本书介绍推荐系统的基本原理、方法和技术。不仅详细讨论了各类方法，还对同类技术进行了归纳总结，有助于读者对当前推荐系统研究领域有全面的了解。书中提供了大量的例子和习题来帮助读者深入理解和掌握相关内容。此外，本书还介绍了当前最新的研究方向，为读者进行推荐系统技术的研究提供参考。

本书既可以作为计算机及相关专业本科生和研究生的教材，也适合开发人员和研究人员阅读。

出版发行：机械工业出版社（北京市西城区百万庄大街 22 号 邮政编码：100037）

责任编辑：朱秀英

责任校对：殷虹

印刷：北京市兆成印刷有限责任公司

版次：2018 年 7 月第 1 版第 1 次印刷

开本：185mm×260mm 1/16

印张：24.25

书号：ISBN 978-7-111-60032-9

定价：129.00 元

凡购本书，如有缺页、倒页、脱页，由本社发行部调换

客服热线：(010) 88378991 88361066

投稿热线：(010) 88379604

购书热线：(010) 68326294 88379649 68995259

读者信箱：hzsj@hzbook.com

版权所有·侵权必究

封底无防伪标均为盗版

本书法律顾问：北京大成律师事务所 韩光 / 邹晓东

文艺复兴以来，源远流长的科学精神和逐步形成的学术规范，使西方国家在自然科学的各个领域取得了垄断性的优势；也正是这样的优势，使美国在信息技术发展的六十多年间名家辈出、独领风骚。在商业化的进程中，美国的产业界与教育界越来越紧密地结合，计算机学科中的许多泰山北斗同时身处科研和教学的最前线，由此而产生的经典科学著作，不仅擘划了研究的范畴，还揭示了学术的源变，既遵循学术规范，又自有学者个性，其价值并不会因年月的流逝而减退。

近年，在全球信息化大潮的推动下，我国的计算机产业发展迅猛，对专业人才的需求日益迫切。这对计算机教育界和出版界都既是机遇，也是挑战；而专业教材的建设在教育战略上显得举足轻重。在我国信息技术发展时间较短的现状下，美国等发达国家在其计算机科学发展的几十年间积淀和发展的经典教材仍有许多值得借鉴之处。因此，引进一批国外优秀计算机教材将对我国计算机教育事业的发展起到积极的推动作用，也是与世界接轨、建设真正的世界一流大学的必由之路。

机械工业出版社华章公司较早意识到“出版要为教育服务”。自1998年开始，我们就将工作重点放在了遴选、移译国外优秀教材上。经过多年的不懈努力，我们与 Pearson, McGraw-Hill, Elsevier, MIT, John Wiley & Sons, Cengage 等世界著名出版公司建立了良好的合作关系，从他们现有的数百种教材中甄选出 Andrew S. Tanenbaum, Bjarne Stroustrup, Brian W. Kernighan, Dennis Ritchie, Jim Gray, Alfred V. Aho, John E. Hopcroft, Jeffrey D. Ullman, Abraham Silberschatz, William Stallings, Donald E. Knuth, John L. Hennessy, Larry L. Peterson 等大师名家的一批经典作品，以“计算机科学丛书”为总称出版，供读者学习、研究及珍藏。大理石纹理的封面，也正体现了这套丛书的品位和格调。

“计算机科学丛书”的出版工作得到了国内外学者的鼎力相助，国内的专家不仅提供了中肯的选题指导，还不辞劳苦地担任了翻译和审校的工作；而原书的作者也相当关注其作品在中国的传播，有的还专门为其书的中译本作序。迄今，“计算机科学丛书”已经出版了近两百个品种，这些书籍在读者中树立了良好的口碑，并被许多高校采用为正式教材和参考书籍。其影印版“经典原版书库”作为姊妹篇也被越来越多实施双语教学的学校所采用。

权威的作者、经典的教材、一流的译者、严格的审校、精细的编辑，这些因素使我们的图书有了质量的保证。随着计算机科学与技术专业学科建设的不断完善和教材改革的逐渐深化，教育界对国外计算机教材的需求和应用都将步入一个新的阶段，我们的目标是尽善尽美，而反馈的意见正是我们达到这一终极目标的重要帮助。华章公司欢迎老师和读者对我们的工作提出建议或给予指正，我们的联系方式如下：

华章网站：www.hzbook.com

电子邮件：hzjsj@hzbook.com

联系电话：(010) 88379604

联系地址：北京市西城区百万庄南街1号

邮政编码：100037



华章科技图书出版中心

我们的时代已经由物品的缺乏时代进化到了丰富时代，随着物品的数量和种类越来越多，人的注意力成为稀缺资源，于是，推荐系统的重要性凸显了出来。推荐系统是一种预测用户对商品和信息的喜好或评分的模型，如何发现用户感兴趣的信息和商品是推荐系统要解决的问题，是用户从互联网上浩如烟海的信息中发现适合于自己信息和商品的重要渠道。

正因为推荐系统的重要性，它已经成为计算机科学中的一个热门领域，研究人员提出了大量模型和算法。推荐系统中需要考虑的因素很多，既要考虑效率，也要考虑有效性；既要考虑用户心理，也要考虑用户的行为；既要考虑商品和信息的外在属性，又要考虑商品和信息的相互关联。由于其综合性和复杂性，这个领域可以看成是数据库、自然语言处理、机器学习、信息检索、算法甚至心理学等领域的综合与交叉。

由于上述特点以及学科的快速发展，推荐系统的知识显得相当繁杂，从中梳理出一个明晰的知识体系对于学习者来说显得特别重要。本书就是能给推荐系统的学习者展示其完整体系的一本教材。

本书的作者 Charu Aggarwal 是数据库和数据挖掘领域知名专家，现就职于 IBM T. J. Watson Research Center，是 ACM Fellow、IEEE Fellow，其 H-index 达到 80。他在数据挖掘领域的多年的研究经历和深厚造诣保证了本书的广度、深度和厚度。

- **广度** 本书涵盖推荐系统的原理，并介绍推荐系统中的各类技术，大致可以分为协同过滤方法、基于内容的方法和基于知识的方法三类。除了推荐系统的相关主题以外，本书还探讨了和特定领域相关的技术，以及如何抵御攻击等高级话题。书中大量的示例和习题有助于读者对推荐系统技术的理解和掌握。
- **深度** 本书对推荐系统的内容介绍不仅仅流于技术层面，更深入阐述推荐技术背后的原理，特别是没有回避其中大量深入的数学方法，这保证了本书理论和技术的融合，使得读者知其然更知其所以然，做到理论和实际的衔接。
- **厚度** 推荐系统相关技术已经发展多年，并且现在仍然是研究热点。本书兼顾了推荐系统历史和发展，既介绍了一系列传统的推荐技术，也介绍了推荐系统最新的技术，体现了推荐系统发展的厚度。

由于这些特点，本书具有广泛的适用性，既适合作为初学者入门的教材，也适合作为进阶者深入学习的指导教材；既适合作为开发人员的参考资料，也适合作为研究人员的研究参考。

本书的翻译组织安排如下：黎玲利负责 1、2、6、11、12、13 章，前言，致谢及作者简介的翻译；尹丹负责 5、7、8、9、10 章的翻译；李默涵负责 3、4 章的翻译；王宏志组织了全书的翻译并进行统稿。参与翻译工作的人员还有哈尔滨工业大学的李东升、马靖昆、周昊天、王必聪同学，北京理工大学的郝俊卿、哈尔滨工程大学的张建川同学。另外，还要感谢黑龙江大学的邵鸿宇、冯博奇、徐绅宝等同学对译稿的审校。

限于译者水平，译文中存在许多不足，敬请读者批评指正。如有任何建议，请发送邮件至 lilingli@hlju.edu.cn。本书以及大数据相关资料将在“大数据与数据科学家”微信公众号和華章网站分享，欢迎读者关注。

译者

2018年6月

大自然呈现在我们面前的只是一头狮子的尾巴。但不要怀疑狮子的存在，尽管它因为身型巨大不能马上现出全身。

—Albert Einstein

随着 Web 成为商务和电子交易的重要媒介，推荐系统在 20 世纪 90 年代变得越来越重要。人们很早就认识到 Web 为个性化服务提供了空前的机会，这是其他渠道是不可能做到的。特别是 Web 为数据收集提供了便利，并且提供了一种非侵入式地推荐物品的用户界面。

自此以后，在公众眼中，推荐系统已经得到了显著的发展。这一事实的证据是，有许多会议和研讨会专门探讨该领域。会议 ACM Conference on Recommender Systems 特别值得一提，因为它为该领域定期贡献了许多前沿工作。推荐系统领域非常多样化，因为它能够使用各种类型的用户偏好数据和用户需求数据来做推荐。推荐系统中最著名的方法包括协同过滤方法、基于内容的方法和基于知识的方法。这三种方法构成了推荐系统研究领域的基本支柱。近年来，已经设计了一些专门的方法来针对不同的数据领域和上下文，例如时间、位置和社会信息。针对专门的场景提出了大量高级的方法，这些方法可以调整用于不同的应用领域，例如查询日志挖掘、新闻推荐和计算广告。本书的结构安排体现了这些重要的话题。本书的章节可以分为三类：

1) 算法和评估：这些章节讨论了推荐系统中的基本算法，包括协同过滤方法（第 2 和第 4 章）、基于内容的方法（第 4 章）和基于知识的方法（第 5 章）。这些方法的混合在第 6 章中讨论。第 7 章讨论了推荐系统评估。

2) 特定领域和上下文的推荐：推荐系统的上下文在提供有效推荐方面扮演了至关重要的角色。例如，一个用户想要用其位置作为附加的上下文（context）来查找饭店。推荐的上下文可以被看作是影响推荐目标的重要辅助信息。不同类型的域，例如时间数据、空间数据和社会数据，提供了不同类型的上下文。相关的方法在第 8~11 章中讨论。第 11 章也讨论了利用社会信息来增加推荐过程可信度的话题。最近的话题（如分解机和可信推荐系统）在这些章节中也有涉及。

3) 高级话题和应用：在第 12 章中，我们将从各个角度讨论推荐系统的健壮性，例如欺诈（shilling）系统、攻击模式及其防御。此外，近期的一些话题，例如排名学习、多臂赌博机、组推荐系统、多标准推荐系统和主动学习系统，将在第 13 章中讨论。该章的一个重要目标是向读者介绍当前研究的基本思想和原则。虽然不可能在一本书里对当前所有的研究技术细节进行讨论，但我们希望最后一章能够在高级话题方面为读者“破冰”。在这一章中，我们也研究了推荐技术的一些应用环境，例如新闻推荐、查询推荐和计算广告。本章还讨论了如何将前面章节中介绍的方法应用于各个不同的领域。

尽管本书是作为教科书来编写的，但仍有很多来自于工业界和学术界的读者。因此，

我们也从应用角度和文献角度来撰写此书。书中提供了大量的示例和习题，使得它可以被用作教科书。由于大部分推荐系统课程只涵盖基础话题，因此有关基础话题的章节和算法着重于课堂教学。另一方面，工业界人员也许发现讨论上下文敏感的推荐系统的章节很有用，因为在许多真实的应用中会有大量可用的上下文辅助信息。第 13 章的应用部分是特别为工业界人员编写的，不过教师也许会发现它在推荐课程中也是有用的。

最后，我们对所使用的符号进行简要的介绍。本书中一直使用一个 $m \times n$ 的评分矩阵，记为 \mathbf{R} ，其中 m 是用户的数量， n 是物品的数量。矩阵 \mathbf{R} 是不完整的，因为只有一部分项是已观测的。 \mathbf{R} 的第 (i, j) 项表示用户 i 对物品 j 的评分，当它是已观测项时，被记为 r_{ij} 。当项 (i, j) 是由推荐算法预测得到（而不是用户指定）时，被记为带“帽子”符号（即抑扬符号）的 \hat{r}_{ij} ，表示它是一个预测的值。向量用“上划线”来表示，例如 \bar{x} 或 \bar{y} 。

致谢

Recommender Systems: The Textbook

感谢在撰书期间妻子和女儿给予我的爱和支持，感谢父母给我持续的爱。

本书得到了很多人直接和间接的帮助，我很感激他们。在撰写本书时，我收到了许多同事的反馈，他们是 Xavier Amatriain、Kanishka Bhaduri、Robin Burke、Martin Ester、Bart Goethals、Huan Liu、Xia Ning、Saket Sathe、Jiliang Tang、Alexander Tuzhilin、Koen Versetrepren 和 Jieping Ye。感谢他们所提供的建设性反馈。这些年来，我从大量合作者那里受益良多。这些见解直接或间接地影响了本书。首先感谢多年来与我合作的 Philip S. Yu。其他重要的合作者还包括 Tarek F. Abdelzaher、Jing Gao、Quanquan Gu、Manish Gupta、Jiawei Han、Alexander Hinneburg、Thomas Huang、Nan Li、Huan Liu、Ruoming Jin、Daniel Keim、Arijit Khan、Latifur Khan、Mohammad M. Masud、Jian Pei、Magda Procopiuc、Guojun Qi、Chandan Reddy、Saket Sathe、Jaideep Srivastava、Karthik Subbian、Yizhou Sun、Jiliang Tang、Min-Hsuan Tsai、Haixun Wang、Jianyong Wang、Min Wang、Joel Wolf、Xifeng Yan、Mohammed Zaki、ChengXiang Zhai 和 Peixiang Zhao。我也要感谢导师 James B. Orlin 在早期对我的指导。

还要感谢我的经理 Nagui Halim 在我撰写此书时所提供的巨大支持。他的专业支持对我过去和现在的许多书都起着重要的作用。

最后，感谢 Lata Aggarwal 用微软 PowerPoint 软件帮我绘制了一些图片。



Charu C. Aggarwal 是位于纽约州约克城的 IBM T. J. Watson 研究中心的杰出研究人员 (DRSM)。他于 1993 年在印度坎普尔理工学院获得了学士学位, 1996 年在麻省理工学院获得了博士学位。他对数据挖掘领域有着广泛的研究。他在国际会议和期刊上发表了 300 余篇论文, 申请了 80 余项专利。他是 15 本书的作者或编辑, 包括一本数据挖掘教材和一本关于孤立点分析的综合性著作。由于他的专利的商业价值, 他曾三次被评为 IBM 的“发明大师”(Master Inventor)。由于提出了数据流上的生物恐怖威胁检测技术, 他获得了 2003 年 IBM 公司奖; 由于在隐私技术上的科学性贡献, 他获得了 2008 年 IBM 杰出创新奖; 由于在数据流和高维数据上的研究工作, 他分别于 2009 年和 2015 年两次获得了 IBM 杰出技术成就奖。他因为提出了基于冷凝的隐私保护数据挖掘技术而获得了 EDBT 2014 的时间检验奖。他还于 2015 年获得了 IEEE ICDM 研究贡献奖, 这是数据挖掘领域对具有突出贡献的研究的两个最高奖项之一。

他曾担任 IEEE 大数据会议 (2014) 的大会主席, ACM CIKM 会议 (2015)、IEEE ICDM 会议 (2015) 和 ACM KDD 会议 (2016) 的程序委员会主席。他从 2004 年到 2008 年担任了《IEEE Transactions on Knowledge and Data Engineering》的副主编。他是《ACM Transactions on Knowledge Discovery from Data》的副主编,《IEEE Transactions on Big Data》的副主编,《Data Mining and Knowledge Discovery Journal》的执行主编,《ACM SIGKDD Explorations》的主编,《Knowledge and Information Systems Journal》的副主编。他在 Springer 的刊物《Lecture Notes on Social Networks》的咨询委员会任职。他担任过 SIAM Activity Group on Data Mining 的副主席。由于在知识发现和数据挖掘算法方面的贡献, 他成为 SIAM、ACM 和 IEEE 的会士。

目 录

Recommender Systems: The Textbook

出版者的话	
译者序	
前言	
致谢	
作者简介	
第 1 章 推荐系统概述	1
1.1 引言	1
1.2 推荐系统的目标	2
1.2.1 推荐系统应用范围	5
1.3 推荐系统的基本模型	5
1.3.1 协同过滤模型	5
1.3.2 基于内容的推荐系统	10
1.3.3 基于知识的推荐系统	11
1.3.4 人口统计推荐系统	13
1.3.5 混合集成的推荐系统	14
1.3.6 对推荐系统的评价	14
1.4 推荐系统领域特有的挑战	14
1.4.1 基于上下文的推荐系统	14
1.4.2 时间敏感的推荐系统	15
1.4.3 基于位置的推荐系统	15
1.4.4 社交信息系统	15
1.5 高级论题和应用	16
1.5.1 推荐系统中的冷启动问题	17
1.5.2 抗攻击推荐系统	17
1.5.3 组推荐系统	17
1.5.4 多标准推荐系统	17
1.5.5 推荐系统中的主动学习	18
1.5.6 推荐系统中的隐私问题	18
1.5.7 应用领域	18
1.6 小结	18
1.7 相关工作	19
1.8 习题	20
第 2 章 基于近邻的协同过滤	21
2.1 引言	21
2.2 评分矩阵的关键性质	22
2.3 通过基于近邻的方法预测评分	24
2.3.1 基于用户的近邻模型	24
2.3.2 基于物品的近邻模型	29
2.3.3 高效的实现和计算复杂度	30
2.3.4 基于用户的方法和基于物品的方法的比较	31
2.3.5 基于近邻方法的优劣势	32
2.3.6 基于用户的方法和基于物品的方法的联合	33
2.4 聚类和基于近邻的方法	33
2.5 降维与近邻方法	34
2.5.1 处理偏差	35
2.6 近邻方法的回归模型视角	38
2.6.1 基于用户的最近邻回归	39
2.6.2 基于物品的最近邻回归	41
2.6.3 基于用户的方法和基于物品的方法的结合	42
2.6.4 具有相似度权重的联合插值	42
2.6.5 稀疏线性模型	43
2.7 基于近邻方法的图模型	45
2.7.1 用户-物品图	45
2.7.2 用户-用户图	47
2.7.3 物品-物品图	49
2.8 小结	50
2.9 相关工作	50
2.10 习题	51
第 3 章 基于模型的协同过滤	53
3.1 引言	53
3.2 决策和回归树	55
3.2.1 将决策树扩展到协同过滤	56
3.3 基于规则的协同过滤	57
3.3.1 将关联规则用于协同过滤	59
3.3.2 面向物品的模型与面向用户的模型	60

3.4 朴素贝叶斯协同过滤	61	关联性	112
3.4.1 处理过拟合	63	4.4.3 贝叶斯分类器	113
3.4.2 示例：使用贝叶斯方法处理 二元评分	63	4.4.4 基于规则的分类器	115
3.5 将任意分类模型当作黑盒来处理 ...	64	4.4.5 基于回归的模型	117
3.5.1 示例：使用神经网络作为黑盒 分类器	65	4.4.6 其他学习模型和比较概述 ...	118
3.6 潜在因子模型	67	4.4.7 基于内容的系统的解释	118
3.6.1 潜在因子模型的几何解释	68	4.5 基于内容的推荐与协同推荐	119
3.6.2 潜在因子模型的低秩解释	69	4.6 将基于内容的模型用于协同 过滤	120
3.6.3 基本矩阵分解原理	70	4.6.1 利用用户画像	120
3.6.4 无约束矩阵分解	71	4.7 小结	121
3.6.5 奇异值分解	84	4.8 相关工作	121
3.6.6 非负矩阵分解	88	4.9 习题	122
3.6.7 理解矩阵因子分解方法族	93	第5章 基于知识的推荐系统	123
3.7 集成因子分解和近邻模型	95	5.1 引言	123
3.7.1 基准估计：非个性化偏倚 中心模型	95	5.2 基于约束的推荐系统	127
3.7.2 模型的近邻部分	96	5.2.1 返回相关结果	128
3.7.3 模型的潜在因子部分	97	5.2.2 交互方法	129
3.7.4 集成近邻和潜在因子部分	97	5.2.3 排序匹配的物品	131
3.7.5 求解优化模型	98	5.2.4 处理不可接受的结果或 空集	131
3.7.6 关于精度的一些观察	98	5.2.5 添加约束	132
3.7.7 将潜在因子模型集成到任意 模型	99	5.3 基于案例的推荐系统	132
3.8 小结	99	5.3.1 相似性度量	134
3.9 相关工作	100	5.3.2 批评方法	138
3.10 习题	101	5.3.3 批评的解释	142
第4章 基于内容的推荐系统	103	5.4 基于知识的系统的持久个性化 ...	143
4.1 引言	103	5.5 小结	143
4.2 基于内容的系统的基本组件	104	5.6 相关工作	144
4.3 预处理和特征提取	105	5.7 习题	145
4.3.1 特征提取	105	第6章 基于集成的混合推荐系统 ...	146
4.3.2 特征表示和清洗	107	6.1 引言	146
4.3.3 收集用户的偏好	108	6.2 从分类角度看集成方法	149
4.3.4 监督特征选择和加权	108	6.3 加权型混合系统	151
4.4 学习用户画像和过滤	111	6.3.1 几种模型组合的方法	153
4.4.1 最近邻分类	111	6.3.2 对分类中的 bagging 算法的 调整	153
4.4.2 与基于案例的推荐系统的		6.3.3 随机性注入算法	155

6.4.1 为解决冷启动问题的切换机制	156	7.5.3 通过效用评估排名	181
6.4.2 桶模型	156	7.5.4 通过 ROC 曲线评估排名	184
6.5 级联型混合系统	156	7.5.5 哪种排名方式最好	186
6.5.1 推荐结果的逐步优化	157	7.6 评估指标的局限性	186
6.5.2 boosting 算法	157	7.6.1 避免评估游戏	188
6.6 特征放大型混合系统	159	7.7 小结	188
6.7 元级型混合系统	159	7.8 相关工作	188
6.8 特征组合型混合系统	160	7.9 习题	189
6.8.1 回归分析和矩阵分解	161	第 8 章 上下文敏感的推荐系统	191
6.8.2 元级特征	161	8.1 引言	191
6.9 交叉型混合系统	163	8.2 多维方法	192
6.10 小结	164	8.2.1 层级的重要性	194
6.11 相关工作	164	8.3 上下文预过滤: 一种基于降维的方法	196
6.12 习题	166	8.3.1 基于集成的改进	198
第 7 章 推荐系统评估	167	8.3.2 多级别的估计	199
7.1 引言	167	8.4 后过滤方法	199
7.2 评估范例	168	8.5 上下文建模	200
7.2.1 用户调查	168	8.5.1 基于近邻的方法	201
7.2.2 在线评估	169	8.5.2 潜在因子模型	202
7.2.3 使用历史数据集进行离线评估	170	8.5.3 基于内容的模型	207
7.3 评估设计的总体目标	170	8.6 小结	209
7.3.1 精确性	170	8.7 相关工作	209
7.3.2 覆盖率	171	8.8 习题	210
7.3.3 置信度和信任度	172	第 9 章 时间与位置敏感的推荐系统	211
7.3.4 新颖度	173	9.1 引言	211
7.3.5 惊喜度	173	9.2 时间协同过滤	212
7.3.6 多样性	174	9.2.1 基于新近的模式	213
7.3.7 健壮性和稳定性	174	9.2.2 处理周期性上下文	215
7.3.8 可扩展性	174	9.2.3 将评分建模为时间的函数	216
7.4 离线推荐评估的设计要点	175	9.3 离散时间模型	220
7.4.1 Netflix Prize 数据集的案例研究	175	9.3.1 马尔可夫模型	220
7.4.2 为训练和测试分解评分	177	9.3.2 序列模式挖掘	223
7.4.3 与分类设计的比较	178	9.4 位置感知推荐系统	224
7.5 离线评估的精确性指标	178	9.4.1 偏好位置	225
7.5.1 度量预测评分的精确性	178	9.4.2 旅行位置	227
7.5.2 通过相关性评估排名	180	9.4.3 结合偏好位置与旅行位置	227
		9.5 小结	227

9.6 相关工作	227	中心的方法	259
9.7 习题	229	11.3.1 收集数据来建立信任 网络	260
第 10 章 网络中的结构化推荐	230	11.3.2 信任的传播和聚合	261
10.1 引言	230	11.3.3 没有信任传播的简单推荐	262
10.2 排序算法	231	11.3.4 TidalTrust 算法	262
10.2.1 PageRank	231	11.3.5 MoleTrust 算法	264
10.2.2 个性化 PageRank	234	11.3.6 信任游走算法	265
10.2.3 基于近邻的方法应用	235	11.3.7 链接预测法	266
10.2.4 SimRank	239	11.3.8 矩阵分解法	268
10.2.5 搜索与推荐的关系	240	11.3.9 社交推荐系统的优点	271
10.3 使用集合分类的推荐	240	11.4 社交推荐系统中的用户交互	272
10.3.1 迭代分类算法	241	11.4.1 大众分类法的代表	273
10.3.2 使用随机游走的标签传播	242	11.4.2 社会性标签系统中的协同 过滤	275
10.3.3 社交网络中协同过滤的 适用性	243	11.4.3 选择有价值的标签	276
10.4 推荐好友: 链接预测	243	11.4.4 无评分矩阵的社会性标签 推荐	276
10.4.1 基于近邻的方法	243	11.4.5 使用评分矩阵的社会性标签 推荐	281
10.4.2 Katz 度量	244	11.5 小结	285
10.4.3 基于随机游走的度量	245	11.6 相关工作	285
10.4.4 作为分类问题的链接 预测	245	11.7 习题	287
10.4.5 链接预测的矩阵分解	246	第 12 章 抵抗攻击的推荐系统	288
10.4.6 链接预测和协同过滤的 关联	249	12.1 引言	288
10.5 社会影响力分析和病毒式 营销	251	12.2 对攻击模型中的权衡的理解	289
10.5.1 线性阈值模型	253	12.2.1 量化攻击的影响	292
10.5.2 独立级联模型	253	12.3 攻击类型	293
10.5.3 影响力函数评估	253	12.3.1 随机攻击	294
10.5.4 社交流中的目标影响力分析 模型	254	12.3.2 均值攻击	294
10.6 小结	255	12.3.3 bandwagon 攻击	295
10.7 相关工作	255	12.3.4 流行攻击	295
10.8 习题	256	12.3.5 爱/憎攻击	296
第 11 章 社交和以信任为中心的推荐 系统	257	12.3.6 反向 bandwagon 攻击	296
11.1 引言	257	12.3.7 探测攻击	296
11.2 社交上下文的多维模型	258	12.3.8 分段攻击	296
11.3 以网络为中心的方法和以信任为 中心的方法	259	12.3.9 基本推荐算法的效果	297
		12.4 探测推荐系统中的攻击	298
		12.4.1 单体攻击画像的探测	299

12.4.2 群体攻击画像的探测	300	13.4.1 协同和基于内容的系统	316
12.5 健壮推荐设计策略	301	13.4.2 基于知识的系统	317
12.5.1 用 CAPTCHA 防止自动 攻击	302	13.5 多标准推荐系统	317
12.5.2 使用社会信任	302	13.5.1 基于近邻的方法	318
12.5.3 设计健壮的推荐算法	302	13.5.2 基于集成的方法	319
12.6 小结	305	13.5.3 无整体评分的多标准系统	320
12.7 相关工作	305	13.6 推荐系统中的主动学习	320
12.8 习题	306	13.6.1 基于异质性的模型	321
第 13 章 推荐系统高级主题	307	13.6.2 基于性能的模型	322
13.1 引言	307	13.7 推荐系统中的隐私	322
13.2 排名学习	308	13.7.1 基于冷凝的隐私	323
13.2.1 成对排名学习	309	13.7.2 高维数据的挑战	323
13.2.2 列表排名学习	310	13.8 一些有趣的应用领域	324
13.2.3 与其他领域中排名学习方法 的比较	311	13.8.1 门户内容个性化	324
13.3 多臂赌博机算法	311	13.8.2 计算广告与推荐系统	326
13.3.1 朴素算法	313	13.8.3 互惠推荐系统	329
13.3.2 ϵ 贪心算法	313	13.9 小结	331
13.3.3 上限方法	314	13.10 相关工作	332
13.4 组推荐系统	315	参考文献	334
		索引	368

推荐系统概述

很多人获得过建议，却只有智者从中获益。

——Harper Lee

1.1 引言

作为电子和商务交易的媒介，Web 如今扮演着越来越重要的角色并推动了推荐系统技术的发展。其中一个重要的作用是能够让用户轻松地提供他喜欢或不喜欢的反馈。例如，Netflix（一家在线影片租赁提供商）的用户只需要简单地动动鼠标就能提供反馈。评分是一种提供反馈的典型方法，在某个特定的评分系统（例如五星评分系统）中，用户可以选择不同大小的数值来说明自己对不同物品的满意程度。

其他形式的反馈不像评分一样清晰明了，但却更容易采集。例如，可将用户在网上购买或是浏览一件物品的行为视为对该物品的认可。获取这类反馈形式的数据十分容易，这种方法被 Amazon.com 等网上商家广泛采用。推荐系统的基本思想是利用这些不同来源的数据来推断顾客的喜好。推荐系统面向的对象称为用户（user），推荐的产品称为物品（item）。由于用户曾经的兴趣喜好通常预示着未来的选择，因此推荐分析也通常是基于先前用户与物品之间的关系。但仍有一个特例——基于知识的推荐系统是根据用户指定需求而非用户的历史记录进行推荐。

那么，推荐算法背后隐藏着什么基本原则呢？以用户为中心的活动和以物品为中心的活动之间存在着显著的依赖关系。例如，对一个喜欢看历史纪录片的人来说，与动作片相比他更会对其他历史纪录片或者教育片感兴趣。在很多情况下，不同类别的物品可能显示出明显的相关性，可以利用这一点做出更精确的推荐。此外，这种关联可能表现在个别物品上而不是一类物品上。这些关联可以用数据驱动方式从评分矩阵中学习（learn）得到，产生的模型可用来预测目标用户的行为。单个用户评过分数的物品数量越多，对其做出准确预测就越容易。对目标用户行为进行预测的学习模型有很多种。例如，大量用户的购买信息或评分行为可以用于对用户的聚类，使得对相似产品感兴趣的被归为一组。同类群体的爱好与行为可以用来为组内个体做推荐。

上面描述的方法是基于一类非常简单的推荐算法——近邻模型（neighborhood model）。这类算法还属于更大的一类算法模型——协同过滤（collaborative filtering）。“协同过滤”是指协同处理大量用户的评分来预测遗失的评分。实际上，推荐系统可以变得更复杂，数据更丰富并包含大量的附加数据类型。例如，在基于内容的推荐系统中，用户的评分和物品的描述信息被用来做预测。其基本思想是根据其用户以往评价过或访问过的物品属性（attribute）对用户的兴趣建模。另一种推荐系统是基于知识的系统（knowledge-based system），用户先阐明他们的兴趣，系统结合用户的兴趣和相关领域知识来做推荐。在更高级的模型中，上下文数据，例如时间信息、外部知识、位置信息、社交信息或是网络信息等都可能被用于预测。

本书将会讲述所有基本类型的推荐系统，包括协同系统、基于内容的系统和基于知识

的系统。我们还将探讨在不同领域推荐系统的基础模型与高级模型，学习推荐系统健壮性的各个方面，例如攻击模型、可信赖模型。此外，还会介绍推荐系统的多种评价模型和混合模型。本章是对推荐系统领域各类工作的概述并将各种话题与本书各个章节关联起来。

本章内容安排如下：1.2节探讨推荐系统的主要目标；1.3节介绍推荐系统中用到的基础模型与评价方法；1.4节讨论推荐系统在不同领域的应用；1.5节讲述推荐系统的高级模型；1.6节是本章的小结。

1.2 推荐系统的目标

在讨论推荐系统的目标之前，我们先介绍推荐问题的几种不同表述方式。下面是两种主要的模型：

1) 预测模型：第一种方法是对用户-物品组合的评分值进行预测。该方法假设描述用户对物品喜好的训练数据是可用的。对于 m 个用户和 n 件物品，这个训练集相当于一个 $m \times n$ 的不完全矩阵，矩阵中的已知值（或观测值）被用来训练。矩阵中的缺失值（或未观测值）则通过这个训练模型进行预测。因为是根据不完整的数值矩阵用学习算法预测出剩余的未知值，所以这个问题又被称作矩阵补全问题。

2) 排名模型：实际上，对用户做推荐并不需要预测出用户对具体物品的评分。商家可能希望向特定的用户推荐前 k 种（top- k ）物品或者是为某个指定物品确定前 k 个（top- k ）感兴趣的物品。虽然这两种算法极其类似，但是对 top- k 物品的计算比确定 top- k 用户要应用普遍，因此本书中我们只讨论对 top- k 物品的计算。这个问题也被叫作 top- k 推荐问题，它是推荐问题的排名模型。

在第二种情况下，对评分的准确值的预测并不重要。由于排名模型可以由第一种预测模型得出结果后再排序得到，所以第一种模型的使用更加普遍。但是在很多情况下，直接设计算法解决排名问题更加自然，也更加简单。这类方法会在第13章中讲到。

推荐系统毕竟是商家用来提高利润的，所以其主要目的是增加产品销量。通过把仔细筛选后的物品推荐给用户，推荐系统能使相关物品得到用户的关注，从而达到增加销量、提高利润的目的。尽管主要目的是盈利，但要实现其功能，方法并不是所想的那么直观。为了实现商业性盈利，一般推荐系统操作和技术上的目标如下：

1) 相关性：推荐系统最重要的操作目标是推荐与用户相关的物品。用户更可能消费那些他们觉得有趣的物品。尽管相关性是推荐系统的主要操作目标，但并不充分。因此，我们下面会讨论一些不如相关性重要但仍具有很大影响力的其他操作目标。

2) 新颖性：如果所推荐的物品是用户从没见过的，那么推荐系统确实很有用。例如，用户喜欢类型的流行电影很少会让用户眼前一亮。反复推荐受欢迎的物品也可能导致销售的多样性降低^[203]。

3) 意外性：意外性是指所推荐的物品出乎意料^[229]。幸运的发现相比于明显的建议要温和得多。意外性不同于新颖性的地方在于其能真正让用户感到惊喜，而不是简单地推荐一些之前没见过的东西。通常情况下，用户可能只是消费一类特定的物品，然而并不排除同时存在着使他们惊喜的物品。和新颖性不同，意外性注重于发现这类推荐物品。

例如，如果隔壁新开了一家印度菜馆，推荐给一个平常就吃印度菜的顾客，他大概只会觉得新颖而不一定惊喜。另一方面，同样是这名顾客，如果向他推荐埃塞俄比亚菜，尝试之前，这名顾客并不知道是否喜欢这种食物，这种推荐就是意料之外的。意外性除了有