



计 算 机 科 学 丛 书

原书第2版



推荐系统

技术、评估及高效算法

弗朗西斯科·里奇 (Francesco Ricci)

[美] 利奥·罗卡奇 (Lior Rokach) 著

布拉哈·夏皮拉 (Bracha Shapira)

李艳民 吴宾 潘微科 刘淇 蒋凡 等译

Recommender Systems Handbook
Second Edition

Francesco Ricci · Lior Rokach
Bracha Shapira *Editors*

Recommender Systems Handbook

Second Edition



机械工业出版社
China Machine Press

计 算 机 科 学 丛 书

原书第2版

推荐系统

技术、评估及高效算法

弗朗西斯科·里奇 (Francesco Ricci)

[美] 利奥·罗卡奇 (Lior Rokach) 著

布拉哈·夏皮拉 (Bracha Shapira)

李艳民 吴宾 潘微科 刘淇 蒋凡 等译

Recommender Systems Handbook
Second Edition

Francesco Ricci · Lior Rokach
Bracha Shapira *Editors*

Recommender
Systems
Handbook

Second Edition

 Springer



机械工业出版社
China Machine Press

图书在版编目 (CIP) 数据

推荐系统：技术、评估及高效算法 (原书第 2 版) / (美) 弗朗西斯科·里奇 (Francesco Ricci) 等著；李艳民等译. —北京：机械工业出版社，2018.6
(计算机科学丛书)

书名原文：Recommender Systems Handbook, Second Edition

ISBN 978-7-111-60075-6

I. 推… II. ①弗… ②李… III. 计算机网络 IV. TP393

中国版本图书馆 CIP 数据核字 (2018) 第 115533 号

本书版权登记号：图字 01-2016-4522

Translation from the English language edition: *Recommender Systems Handbook* edited by Francesco Ricci, Lior Rokach and Bracha Shapira.

Copyright © Springer Science+Business Media LLC 2011, 2015.

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。

本书汇聚不同领域专家学者的理论成果和实践经验，全面介绍推荐系统的主要概念、理论、趋势、挑战和应用，详细阐释如何支持用户决策、计划和购买过程。书中既详细讲解了经典方法，又介绍了一些新的研究成果，内容涵盖人工智能、人机交互、信息技术、数据挖掘、统计学、自适应用户界面、决策支持系统、市场和客户行为等。无论是从事技术开发的读者，还是从事产品营销的读者，都能从中受益。

全书分五部分，共 28 章。第 1 章是概述，系统介绍推荐系统的概念、功能、应用领域以及当前应用过程中遇到的问题与挑战。第一部分 (第 2~7 章) 展示如今构建推荐系统最流行和最基础的技术，如协同过滤、基于语义的方法、数据挖掘方法和基于情境感知的方法。第二部分 (第 8~10 章) 主要关注离线和真实用户环境下用于评估推荐质量的技术及方法。第三部分 (第 11~17 章) 包括一些推荐技术多样性的应用，首先简述与工业实现和推荐系统开发相关的一般性问题，随后详细介绍推荐系统在各领域中的应用：音乐、学习、移动、社交网络及它们之间的交互。第四部分 (第 18~21 章) 包含探讨一系列问题的文章，这些问题包括推荐的展示、浏览、解释和视觉化以及人工决策与推荐系统相关的重要问题。第五部分 (第 22~28 章) 收集了一些关于高级话题的文章，例如，利用主动学习技术来引导新知识的学习，构建能够抵挡恶意用户攻击的健壮推荐系统的合适技术，以及结合多种用户反馈和偏好来生成更加可靠的推荐系统。

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：缪 杰

责任校对：殷 虹

印 刷：中国电影出版社印刷厂

版 次：2018 年 7 月第 1 版第 1 次印刷

开 本：185mm×260mm 1/16

印 张：41

书 号：ISBN 978-7-111-60075-6

定 价：139.00 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：hzjsj@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



推荐序一

Recommender Systems Handbook, Second Edition

在大数据和人工智能时代，基于海量用户行为数据、机器学习技术和智能推荐算法的个性化服务已经实实在在地给老百姓的日常生活带来了便利。也许你曾有过或听过类似下面的感慨：

“京东商城居然把我的购物喜好猜得那么准！”

“网易云音乐今天的推荐歌单实在是太走心了，必须给五星好评！”

“百度主页上的新闻总能引起我的兴趣。”

.....

这些贴心得令人惊叹的个性化服务无一不是应用人工智能、机器学习和智能推荐等系统和技术的结果。显然，高质量的个性化服务是非常重要的，不仅能使用户更加满意，也能为商家带来更多利润。如今，国内互联网领域的领先企业（如京东、淘宝、网易、百度、豆瓣和今日头条等），都在自家的应用场景上使用了前沿的推荐系统和技术，并将自身的实践经验反哺学术界，进一步推动该领域的科学研究。随着用户行为数据的爆发式增长和人工智能、机器学习等技术和算法的不断创新，智能推荐系统和技术可望在自动驾驶、机器人等更多领域大展拳脚。在这样的背景下，社会对智能推荐系统和技术从业者的需求也急剧增加，而本译著的出版正好顺应了这个需求。

不论是从科学研究还是从工业实践的角度出发，我都很欣慰能够看到《Recommender Systems Handbook》这样包含大量人工智能和机器学习等技术与算法的大部头著作被翻译成中文。一方面，它在推荐系统领域的重要地位在全世界是得到公认的；另一方面，原著的语言多少给国内有志于从事该领域的人员造成了一定的困难，译著的出版有利于降低这些非技术性因素造成的门槛。

这本书在内容上兼顾了广度和深度，包含了该领域多年的理论成果和实践经验，特别是对应用人工智能和机器学习等技术与算法的总结，较为全面地介绍了智能推荐系统和技术的核心概念、原理、前沿技术、未来趋势和应用等。我相信广大智能推荐系统和技术从业者、科研人员、高年级本科生和研究生都能从中得到启发并获益。

杨强

香港科技大学计算机科学及工程学系讲座教授

作为一个推荐系统相关领域的科研工作者，本人一直在关注推荐系统的研究进展，近来欣然获悉《推荐系统：技术、评估及高效算法（原书第2版）》译稿已经完成。第2版相较于第1版有较大范围的修改与梳理，新增了12章内容，同时对剩余章节也进行了内容更新。感谢李艳民、吴宾、潘微科、刘淇和蒋凡老师等译者细致而浩繁的翻译工作，大大降低了从业者全面了解推荐系统最新研究进展的门槛。同时，也从另一个角度说明当下中国研究者和从业者在推荐系统领域的前瞻性关注，博采众长以自强。

自1995年Marko Balabanovic等人在美国人工智能协会上推出个性化推荐系统LIRA以来，推荐系统的发展日新月异，由早期单一、简单的推荐系统，发展成为融合了大数据、云计算和深度学习等多领域知识的精准、高效的推荐服务。在与工业界的广泛接触和深入合作中，我发现不仅是学术界对推荐系统保持着关注与投入，工业界在发展推荐系统技术、构建自己的推荐系统方面更是不惜余力。如阿里的商品推荐、腾讯的好友推荐、百度的搜索关联推荐、科大讯飞的个性化学习推荐等，为千千万万的互联网用户提供了贴心和省心的信息过滤服务。尤其是看到我培养的学生将实验室多年来在推荐系统方面的积累与前瞻性思考应用于各大公司的推荐系统，做到了学以致用，并得到业界的广泛认可，很是欣慰。

近年来，随着移动互联网的大范围普及和应用、用户生活理念的转变和升级，在丰富推荐系统使用场景、提升推荐系统重要性的同时，也为推荐系统研究者和从业者带来了新的挑战，如数据规模的增大、更新频率的加快以及用户对隐私关注度的提高等。如何把握这些新机遇，应对这些新挑战，需要每一个推荐系统相关人员进行深入思考与研究。就我们实验室而言，虽然在推荐系统领域深耕多年，与业界也开展了广泛的合作与成果转化，但信息推荐技术的发展日新月异，断不敢停止学习。而《推荐系统：技术、评估及高效算法》这本书能够为从业人员提供一个非常好的研究基础和思考视角。本书在内容构成上不仅做到了脚踏实地，而且做到了仰望星空。既介绍了推荐系统的基本概念、理论、方法和案例，又展示了推荐系统的趋势与挑战。可以帮助从业者很好地夯实推荐系统的技术基础，同时也会拓宽从业者的思考维度。

相信这本书的出版将有助于推荐系统研究者以及从业者为用户提供更知心的推荐服务，真正做到“心有灵犀一点通”。

陈恩红

2018年2月于中国科学技术大学

当今我们身处一个数据爆炸的时代。2017 年全球数据总存储量估计是 16 ~ 20ZB，再过 8 年，这个量还会增长 10 倍，接近甚至超过 200ZB。拥有海量数据，并不等于我们就能够掌握和利用它们。实际上，大数据带来的最直接挑战就是如何让普通消费者在数据汪洋中找到自己需要的内容。

目前来看，搜索引擎和推荐系统是解决上述“信息过载”问题最好的两种手段。推荐系统在两个方面和搜索引擎有重大差异。首先，推荐系统给出的推荐结果都是个性化的，不同的消费者看到的结果一般而言差异很大。其次，用户在搜索的时候对于自己的需求大致是清楚的，但是推荐系统可以通过分析消费者以往点击或购买的记录，以及其他用户点击或购买的记录，给出一些让消费者意想不到甚至以前从未了解的推荐，而这些推荐往往却是消费者喜欢的。

我自己和推荐系统很有缘分。在瑞士读博士的时候，导师给出的第一个任务就是去参加推荐系统领域迄今为止最具影响力的全球赛事——Netflix 推荐算法大赛。后来做图挖掘方面的研究，最早也是从推荐系统开始的。我后来作为联合创始人第一次创业的企业——北京百分点科技——也是从做推荐系统开始的。

不管你是对推荐系统感兴趣的学生、业界技术同行、研究人员，还是企业相关管理人员，了解推荐系统都最好是从阅读系统性的综述和评论开始。我针对这个问题写过两篇综述：一是 2009 年在《自然科学进展》上写过一篇短综述《个性化推荐系统的研究进展》，主要针对中文读者；二是 2012 年在《Physics Reports》写过一篇长综述《Recommender Systems》，主要针对物理领域对这个问题感兴趣的读者。在计算机领域，我认为最好的入门读物是 Adomavicius 和 Tuzhilin 2005 年在 IEEE TKDE 发表的题为《Towards the next generation recommender systems》的综述。

在当前，说起对整个推荐系统研究领域影响最大、起到承上启下作用的出版物，毫无疑问要算 2011 年 Springer 出版社出版的《Recommender Systems Handbook》。该书实际上是当时全球很多著名研究团队在其擅长方向上综述论文的汇编，里面很多章节的引用都超过 1000 次，甚至数千次。我记得这本书很多章节的预印本 2010 年就可以下载了，我当时就打印出来装订出了缺几个章节的 Handbook，其中基于文本的推荐、社会化推荐和推荐系统的评价等几个章节对我后续的研究工作提供了重要的借鉴和参考，我的很多同事和学生也受益匪浅。最近得知本书第 2 版中译本即将出版，我非常高兴，很多以前苦读的回忆又历历在目。

我觉得今天中国的学生和学者是幸运的，可以用自己熟悉亲切的语言，花很少的钱，读到具有超高质量的学术专著。特别感谢出版社愿意引进这本专业性很强的著作，也要特别感谢翻译团队愿意翻译这本书，它在知识传播方面的价值非常大。希望未来出版社还能在第一时间引进更多如本书一般世界公认的重要著作。

周涛教授

电子科技大学大数据研究中心主任

推荐系统最初只是计算机学术领域的一个研究方向，在人工智能时代则越来越成为影响每个人日常生活的重要因素。几乎所有涉及需要解决信息过载和个性化问题的商业应用都会出现推荐系统的身影，凡是有志于进入数据智能领域的工程师和产品经理，也都将学习推荐系统算法和思想作为理解这一领域的必经之路。然而在学习推荐系统方面，最全面最权威的著作则是这本《Recommender Systems Handbook》。

《Recommender Systems Handbook》自 2010 年第 1 版面世以来就成为推荐系统爱好者最为推崇的一本大师级著作。这不光是因为其内容覆盖了推荐系统知识体系的方方面面，而且全书 25 章都是由每个子领域内最富有权权威的多名专家合力而作，体现了在这些方面的最高科研成果。这一版的中文版也由我们在机械工业出版社的支持下在 2015 年翻译成书，首次将这一经典著作引入中文科技社区，让更多希望探究推荐系统领域原汁原味算法和思想的爱好者们有了可以细细研读的资料。

然而科研成果的突飞猛进和应用产品的高速发展超出了所有人的想象，推荐系统越来越成为人工智能领域的一门显学。尤其是在国内 2012 年以来随着移动互联网经济的高速发展，新闻资讯、知识图谱、在线视频音乐、网上购物、O2O、共享经济一波接一波的技术革新和商业创新，极大丰富了可以推荐的物品，极大降低了计算推荐的成本，极大拓宽了适合推荐的场景，让推荐系统技术进一步在人工智能时代变得更加重要，也让人们更加迫切需要深入地理解推荐系统算法和思想，改变原先因为时代局限而对推荐系统技术形成的还不够成熟的见解。

有鉴于此，《Recommender Systems Handbook》在 2015 年及时更新出版了第 2 版。新版在旧版 25 章的基础上扩展到了 28 章，其中有 12 章是新增内容，反映了这 5 年来推荐系统领域研究成果的最新变化；其余 16 章也都对内容进行了大幅度的重新梳理，并增加了最新的案例说明和领域进展。当然这也给我们译者带来了更大挑战，几乎是又要重新翻译一本新的书籍。这既是一座高耸的山峰横亘在面前，催促我们去攀登、去学习专家学者们最新的研究成果；这也是一枚闪亮的奖章佩戴在胸前，奖励我们所专注的是时变时新的领域、探索未有穷期。

在这里还是要感谢所有参与本书翻译、校订和审核的同学：李艳民、吴宾等。

感谢机械工业出版社几年如一日的支持和帮助，让我们能为技术社区奉上经典大作。

我又想起了曾经在第 1 版译者序中写下的话：很幸运能和这一批不计辛劳、只求学问的小伙伴们一起完成了这个心愿。翻过高山，收入眼帘的就是一马平川的美景；架起云梯，就能帮助更多的小伙伴攻城拔寨无往不利。

《Recommender Systems Handbook》第 2 版中文版终于可以面世了，我们用专注和坚韧将它唤醒，是预感它的魔力能够召唤来更多有志于此的攻城狮，让技术的力量改变世界。

蒋凡

百度外卖技术委员会主席、首席架构师

拉扎斯集团高级科学家

一个偶然的机会，我收到了来自国内同行的邮件，问我是否有兴趣参与翻译一本关于推荐系统的英文书籍。在看到这封邮件的时候，我突然有了种如果我接受这份邀请就是在一定程度上为祖国的科研事业做贡献的自豪感。既然如此，那还有什么好犹豫的呢？事实上，翻译这本

书也让我对推荐系统这一领域，尤其是其中的隐私保护问题有了更深层次的认识，这是对我自身的一种升华。很高兴能够和大家一起完成这一艰巨而又伟大的工作，也很感谢在此过程中为我们所有翻译人员提供帮助与支持的李艳民和蒋凡老师。这本书的最终出版离不开任何一位为之付出努力的参与者，希望本书能对推荐系统的爱好者和研究人员有所启发。

王喜玮

美国东北伊利诺伊大学计算机科学系助理教授

经过近十几年的发展，推荐系统已经成了一个相对独立并且成熟的领域，而《Recommender Systems Handbook》一书对此领域的研究工作与应用场景进行了非常系统与全面的总结。鉴于此，本着能够使更多中文社区的同伴们从此书中受益的想法，此书的翻译工作应运而生。

很荣幸有机会与多位专家学者共同参与本书第2版的翻译工作。多位参与者在用词习惯上难免存在差异，也感谢李艳民、吴宾等人对全书用词统一方面的审核与校订。

姚远

南京大学计算机科学与技术系助理研究员

推荐系统是为用户推荐所需物品的软件工具和技术。提供的推荐旨在通过各种决策过程来支持用户，例如，买什么物品、听什么歌曲或读什么新闻。推荐系统的价值在于帮助用户解决信息过载和做出更好的选择，也是现在互联网领域最强大和最流行的信息发现工具之一。因此，人们提出了各式各样的推荐技术，并在过去的 10 年中将其中很多方法成功地运用在商业环境。

推荐系统的发展需要多学科的支持，涉及来自各个领域的专家知识，如人工智能、人机交互、数据挖掘、统计学、决策支持系统、市场营销和消费者行为学等。

本书第 1 版在四年前出版，并受到了推荐系统社区的一致好评。伴随着推荐系统研究的日新月异，这些好评激励我们来更新本书。本书第 2 版重新整理了第 1 版中各章节的内容并在相应章节融入了该领域的新进展。本书基于第 1 版做了较大修订；大约有一半的章节是新增的，并且保留的章节相比第 1 版也做了相应更新。

尽管第 2 版发生了较大修订，但本书的目标始终不渝。本书呈现了基础知识和更高级的话题两方面，通过展示推荐系统的主要概念、理论、方法论、趋势、挑战和应用等连贯而又统一的知识体系，帮助读者从差异中梳理出头绪。这是目前唯一一本全面阐述推荐系统的书，完全涵盖推荐系统主要技术的多个方面。本书中的丰富信息和实践内容为研究人员、学生和行业中的实践者提供了一个有关推荐系统的全面且简洁方便的参考源。

本书不仅详细地介绍了推荐系统研究的经典方法，同时也介绍了最近发表的新方法及其扩展。本书由五部分组成：推荐系统的技术、评估、应用、人机交互及高级话题。第一部分展示了如今构建推荐系统的最流行和最基础的技术，如协同过滤、基于语义的方法、数据挖掘方法和基于情境感知的方法。第二部分主要关注离线和真实用户环境下用于评估推荐质量的技术及方法。第三部分包括一些推荐技术多样性的应用，首先简述与工业实现和推荐系统开发相关的一般性问题，随后详细介绍推荐系统在各领域中的应用：音乐、学习、移动、社交网络及它们之间的交互。第四部分包含探讨一系列问题的文章，这些问题包括推荐的展示、浏览、解释和视觉化以及人工决策与推荐系统相关的重要问题。第五部分收集了一些关于高级话题的文章，例如，利用主动学习技术来引导新知识的学习，构建能够抵挡恶意用户攻击的健壮推荐系统的合适技术，以及结合多种用户反馈和偏好来生成更加可靠的推荐系统。

我们要感谢所有为本书做出贡献的作者。感谢所有审阅人员提出的慷慨意见及建议。特别感谢 Susan Lagerstrom-Fife 和 Springer 的成员，感谢他们在写这本书过程中的合作。最后我们希望这本书有助于这一学科的发展，为新手提供一个卓有成效的学习方案，能够激起更多专业人士有兴趣参与本书所讨论的主题，使这个具有挑战性的领域能够硕果累累，长足发展。

Francesco Ricci

Lior Rokach

Bracha Shapira

译者简介

Recommender Systems Handbook, Second Edition

李艳民 电商推荐和搜索行业资深从业人员，在推荐系统和电商搜索方向有着丰富经验，曾就职于京东和百度，负责搜索和推荐相关产品业务，《推荐系统：技术、评估及高效算法》第1版译者。

吴 宾 郑州大学产业技术研究院在读博士。研究方向为机器学习和模式识别，具体包括社交网络、推荐系统、深度学习等。参与国家自然科学基金项目三个，曾参与软件学报、计算机研究与发展、TKDE、TSC、WWWJ等多个国内外著名学术期刊的审稿工作。著名推荐系统开源库 LibRec 的前核心成员，《推荐系统：技术、评估及高效算法》第1版译者。

刘 淇 中国科学技术大学计算机学院副教授，中国计算机学会(CCF)大数据专家委员会委员、中国人工智能学会机器学习专委会委员。主要研究数据挖掘与知识发现、机器学习方法及其应用，着重于针对用户行为数据(如消费数据、社交数据、教育数据等)的建模和应用研究。在重要国际学术会议和期刊共发表论文60余篇，2011年获得数据挖掘领域顶级国际会议之一IEEE ICDM的最佳研究论文奖，还获得过中科院院长特别奖、KSEM 2013最佳论文奖以及SDM 2015最佳论文候选奖、中科院优博等重要学术奖励，入选中科院青年创新促进会。主持了多项国家、省部级以及与知名公司(如微软、腾讯、科大讯飞)的合作研究项目。担任了CCF大数据学术会议(BigData)2015~2017的宣传主席，还是包括IJCAI、KDD、WWW、AAAI、ICDM、CIKM等国际会议的程序委员会委员以及国际期刊TKDE、TKDD、TC、TSMC-C、TIST等的审稿人、FCS青年AE。

潘微科 深圳大学计算机与软件学院副教授，深圳市海外高层次人才，“深圳大学腾讯益友奖”优秀班主任。2005年毕业于浙江大学，获学士学位；2012年毕业于香港科技大学，获博士学位。研究方向为数据挖掘和人工智能，具体包括迁移学习、智能推荐和机器学习等。已在相关领域发表30余篇科研论文，主持国家自然科学基金等多个科研项目。(曾)担任国际著名学术杂志《IEEE Intelligent Systems》的客座编委、国际著名学术期刊《ACM Transactions on Intelligent Systems and Technology》的信息官和多个国际顶级学术会议的程序委员会委员。曾在2015年龙星计划开放课程中担任讲师。

蒋 凡 现任百度外卖技术委员会主席、首席架构师，拉扎斯集团高级科学家。主要研究方向为物流调度、个性化推荐、智能营销和画像建模。2006年硕士毕业于中国科学技术大学，后加入百度网页搜索团队，设计开发相关性核心算法和跨语言搜索基础架构。2012年加入百度知心团队，设计开发新一代的知识图谱推荐引擎，改变百度搜索右侧的展现形态。2014年加入百度外卖团队，负责物流调度、个性化推荐、智能营销和画像建模等多个技术方向。著有《智能增长》(人民邮电出版社)，译有《推荐系统》(人民邮电出版社)和《推荐系统：技术、评估及高效算法》(机械工业出版社)。荣获2017年吴文俊人工智能科学技术创新一等奖。

参与本书翻译的所有译者名单如下(排名不分先后)：李艳民，吴宾，刘淇，潘微科，蒋凡，刘梦思，郭贵冰，姚远，蔡婉铃，胡聪，高全力，曹瑞，李鑫，孙明明，刘金木，郑勇，朱郁筱，刘俊涛，陈义，温颖，史艳翠，徐斌，吴金龙，阴红志(Hongzhi Yin)，王雪丽，赖博先，谢妍，Li. Chen，吴雯(Carrie Wu)，王喜玮，曾子杰，梁华东，黄山山，何爱龙，李喆，张朋飞，蔡凯伟，朱亚东，姜铮铮，李斌，李聪。

出版者的话
 推荐序一
 推荐序二
 推荐序三
 译者序
 前言
 译者简介

第 1 章 推荐系统：简介和挑战 1

1.1 简介 1

1.2 推荐系统的功能 3

1.3 数据和知识来源 5

1.4 推荐技术 7

1.5 推荐系统评估 10

1.6 推荐系统应用 11

1.7 推荐系统与人机交互 13

1.8 高级话题 14

1.9 挑战 16

1.9.1 偏好获取与分析 16

1.9.2 交互 17

1.9.3 新的推荐任务 18

参考文献 19

第一部分 推荐系统技术

第 2 章 基于邻域的推荐方法综述 ... 24

2.1 简介 24

2.1.1 基于邻域方法的优势 25

2.1.2 目标和概要 26

2.2 问题定义和符号 26

2.3 基于邻域的推荐 27

2.3.1 基于用户的评分预测 28

2.3.2 基于用户的分类预测方法 ... 28

2.3.3 回归与分类 29

2.3.4 基于物品的推荐 29

2.3.5 基于用户和基于物品的推荐方法的比较 30

2.4 基于邻域方法的要素 31

2.4.1 评分标准化 31

2.4.2 相似度权重的计算 33

2.4.3 邻域的选择 37

2.5 高级进阶技术 37

2.5.1 基于图的方法 38

2.5.2 基于学习的方法 40

2.6 总结 44

参考文献 44

第 3 章 协同过滤方法进阶 48

3.1 简介 48

3.2 预备知识 49

3.2.1 基准预测 49

3.2.2 Netflix 数据 50

3.2.3 隐式反馈 51

3.3 矩阵分解模型 51

3.3.1 SVD 52

3.3.2 SVD++ 53

3.3.3 时间敏感的因子模型 54

3.3.4 比较 57

3.3.5 小结 58

3.4 基于邻域的模型 59

3.4.1 相似度度量 59

3.4.2 基于相似度的插值 60

3.4.3 联合派生插值权重 61

3.4.4 小结 63

3.5 增强的基于邻域的模型 63

3.5.1 全局化的邻域模型 64

3.5.2 因式分解的邻域模型 67

3.5.3 基于邻域模型的动态时序 ... 71

3.5.4 小结 72

3.6 基于邻域的模型和因子分解模型
 的比较 73

参考文献 75

第 4 章 基于内容的语义感知推荐
 系统 77

4.1 简介 77

4.2 基于内容的推荐系统概述	77
4.2.1 基于关键词的向量空间模型 ...	79
4.2.2 用户特征学习的方法	80
4.2.3 基于内容过滤的优缺点	81
4.3 自上而下的语义方法	82
4.3.1 基于本体资源的方法	83
4.3.2 基于非结构化或半结构化 百科知识的方法	84
4.3.3 基于关联开放数据的方法 ...	86
4.4 自下而上的语义方法	90
4.4.1 基于判别式模型的方法	90
4.5 方法比较与小结	94
4.6 总结与未来挑战	95
致谢	96
参考文献	96
第5章 基于约束的推荐系统	103
5.1 简介	103
5.2 推荐知识库的开发	105
5.3 推荐过程中的用户导向作用	108
5.4 计算推荐结果	113
5.5 实际应用的经验	114
5.6 未来的研究方法	116
5.7 总结	118
参考文献	118
第6章 情境感知推荐系统	123
6.1 简介和动机	123
6.2 推荐系统中的情境	124
6.2.1 什么是情境	124
6.2.2 推荐系统中模型化情境 信息的表征性方法	125
6.2.3 推荐系统中主要的情境 信息建模方法	127
6.2.4 获取情境信息	130
6.3 结合具有代表性情境的推荐 系统范式	131
6.3.1 情境预过滤	133
6.3.2 情境后过滤	136
6.3.3 情境建模	137
6.4 讨论和总结	138
致谢	140

参考文献	140
------------	-----

第7章 推荐系统中的数据挖掘 方法	145
7.1 简介	145
7.2 数据预处理	146
7.2.1 相似度度量方法	146
7.2.2 抽样	147
7.2.3 降维	148
7.2.4 去噪	150
7.3 监督学习	150
7.3.1 分类	150
7.3.2 分类器的集成	157
7.3.3 评估分类器	157
7.4 无监督学习	159
7.4.1 聚类分析	159
7.4.2 关联规则挖掘	161
7.5 总结	162
参考文献	163

第二部分 推荐系统评估

第8章 推荐系统的评估	170
8.1 简介	170
8.2 实验设置	171
8.2.1 离线实验	172
8.2.2 用户调查	173
8.2.3 在线评估	175
8.2.4 得出可靠结论	176
8.3 推荐系统属性	178
8.3.1 用户偏好	179
8.3.2 预测精度	179
8.3.3 覆盖率	186
8.3.4 置信度	187
8.3.5 信任度	188
8.3.6 新颖性	188
8.3.7 惊喜度	189
8.3.8 多样性	190
8.3.9 效用	191
8.3.10 风险	191
8.3.11 健壮性	192
8.3.12 隐私	192

8.3.13 适应性	193
8.3.14 可扩展性	193
8.4 结论	193
参考文献	194

第9章 使用用户实验评估推荐

系统	198
9.1 简介	198
9.2 理论基础与现有工作	199
9.2.1 理论基础: Knijnenburg 等人 提出的评估框架	199
9.2.2 现有以用户为中心的研究 概览以及有前景的方向	201
9.3 实践指南	203
9.3.1 研究模型	203
9.3.2 参与者	206
9.3.3 实验操控	207
9.3.4 测量	209
9.3.5 统计评估	214
9.4 结论	219
参考文献	221

第10章 对推荐结果的解释:

设计和评估	228
10.1 简介	228
10.2 推荐设计的呈现和交互	229
10.2.1 推荐呈现	229
10.2.2 偏好提取	230
10.3 解释风格	231
10.3.1 基于协作的解释风格	232
10.3.2 基于内容的解释风格	233
10.3.3 基于案例推理(CBR)的 解释风格	234
10.3.4 基于知识和基于效用的 解释风格	234
10.3.5 基于人口统计的解释 方式	235
10.4 目标和度量	235
10.4.1 系统如何工作: 透明性 ...	237
10.4.2 允许用户告诉系统它是 错误的: 被理解	237
10.4.3 增加用户对系统的信任: 信任度	238

10.4.4 说服用户尝试或购买: 说服力	239
10.4.5 帮助用户充分地决策: 有效性	239
10.4.6 帮助用户快速制定决策: 效率	240
10.4.7 系统满意度	241
10.5 未来的方向	242
10.5.1 推荐的社会性	242
10.5.2 解释、偶然性和过滤 泡泡	242
10.5.3 应该何时展现推荐解释 ...	242
10.5.4 推荐解释: 有益还是 有害	242
参考文献	243

第三部分 推荐系统应用

第11章 工业界的推荐系统:

Netflix 案例分析	248
11.1 简介	248
11.2 推荐系统在工业界中的应用 ...	248
11.3 Netflix 大奖赛	250
11.4 评分预测之外的推荐工作	251
11.4.1 推荐无处不在	251
11.4.2 排序	253
11.4.3 页面优化	254
11.5 数据和模型	255
11.5.1 数据	255
11.5.2 模型	256
11.6 消费者数据科学	256
11.7 架构	258
11.7.1 事件和数据分布系统	259
11.7.2 线下、近似在线、线上 计算	260
11.7.3 推荐结果	261
11.8 可扩展的研究方向	261
11.8.1 隐式反馈	262
11.8.2 个性化排序学习	262
11.8.3 全页优化	262
11.8.4 情景推荐	263
11.8.5 评测及标准	263

11.8.6 类别不平衡及其效应 263

11.8.7 社交推荐 264

11.9 结论 264

参考文献 265

第 12 章 辅助学习的推荐系统

 综述 271

12.1 简介 271

12.2 技术增强学习 272

12.3 技术增强学习推荐系统的
 分类框架 273

12.4 方法综述 274

 12.4.1 方法和技术增强学习推荐
 系统综述 274

 12.4.2 框架分析 278

12.5 结论 284

致谢 287

参考文献 287

第 13 章 音乐推荐系统..... 294

13.1 简介 294

13.2 基于内容的音乐推荐 295

 13.2.1 元数据信息 295

 13.2.2 音频内容 296

13.3 基于上下文的音乐推荐 298

 13.3.1 环境相关的上下文 298

 13.3.2 用户相关的上下文 300

 13.3.3 在音乐推荐系统中结合
 上下文信息 300

13.4 混合音乐推荐 301

 13.4.1 结合内容与上下文
 描述符 302

 13.4.2 结合协同过滤与内容
 描述符 302

 13.4.3 结合协同过滤与上下文
 描述符 304

13.5 自动生成播放列表 305

 13.5.1 并行和序列消费 305

 13.5.2 播放列表评估 306

 13.5.3 播放列表生成算法 307

13.6 数据集和评估 308

 13.6.1 评估方法 309

 13.6.2 Yahoo! Music 数据集和
 KDD Cup 2011 310

 13.6.3 百万级别歌曲数据集(MSD)
 和 MSD Challenge 2012 ... 310

 13.6.4 Last.fm 数据集: 360K/1K
 用户 311

 13.6.5 MusicMicro 和百万级别
 音乐数据集(MMTD) 311

 13.6.6 AotM-2011 312

13.7 总结与挑战 312

参考文献 312

第 14 章 剖析基于位置的移动

 推荐系统 320

14.1 简介 320

14.2 移动推荐系统的数据 321

 14.2.1 发现兴趣点和位置偏好 ... 322

 14.2.2 根据智能手机传感器的
 行为推导 323

14.3 移动应用中计算推荐的方法 ... 323

 14.3.1 推荐形式化概述 324

 14.3.2 推荐场所的算法 325

14.4 移动推荐的评估 326

14.5 结论和未来方向 327

参考文献 328

第 15 章 社会化推荐系统 332

15.1 简介 332

15.2 内容推荐 332

 15.2.1 关键领域 333

 15.2.2 群组推荐 334

 15.2.3 案例研究: 企业社交媒体
 推荐 335

 15.2.4 小结 337

15.3 人物推荐 338

 15.3.1 推荐连接的人 339

 15.3.2 推荐陌生人 341

 15.3.3 推荐被关注者 342

 15.3.4 相关研究领域 343

 15.3.5 小结 343

15.4 讨论 343

15.5 新兴领域和开放性挑战 345

15.5.1 新兴领域	345
15.5.2 开放性挑战	346
参考文献	346
第16章 人与人之间的相互推荐 ...	351
16.1 简介	351
16.2 互惠推荐与传统推荐	351
16.3 关于人与人推荐的已有工作 ...	352
16.3.1 社交网络	352
16.3.2 师徒匹配	353
16.3.3 工作推荐	353
16.3.4 在线婚恋	354
16.4 在线婚恋系统案例分析	355
16.4.1 一种基于内容—协同双向 在线婚恋推荐系统	356
16.4.2 显式与隐式的用户偏好 ...	359
16.5 总结与未来工作	362
参考文献	363
第17章 社交网络搜索中的协作、 信用机制和推荐系统	365
17.1 简介	365
17.2 网络搜索的历史简介	366
17.3 网络搜索的未来	368
17.3.1 个性化网络搜索	368
17.3.2 协同信息检索	371
17.3.3 关于信誉与推荐	372
17.3.4 向社交搜索前进	373
17.4 案例研究1: HeyStaks——社交 搜索用例	373
17.4.1 HeyStaks 系统	374
17.4.2 HeyStaks 推荐引擎	374
17.4.3 评估	376
17.5 案例研究2: 社会搜索的信誉 模型	377
17.5.1 从活动到信誉	378
17.5.2 信誉作为协同	378
17.5.3 实例	379
17.5.4 基于图的信誉模型	379
17.5.5 从用户信誉到结果推荐 ...	380
17.5.6 评估	381
17.6 搜索未来	383

17.6.1 从搜索到发现	383
17.6.2 在传感器丰富的移动世界 中搜索	383
致谢	384
参考文献	384

第四部分 人机交互

第18章 人类决策过程与推荐

系统	392
18.1 简介	392
18.2 选择模式和推荐	393
18.2.1 基于属性的选择	394
18.2.2 基于结果的选择	395
18.2.3 基于经验的选择	395
18.2.4 基于社会的选择	396
18.2.5 基于策略的选择	396
18.2.6 基于反复试错的选择	397
18.2.7 混合模式的选择	397
18.2.8 什么是一个好的选择	398
18.3 支持选择策略和推荐	398
18.3.1 代表选择者的评估	398
18.3.2 选择过程的建议	399
18.3.3 获取信息与经验	399
18.3.4 代表选择的情景	400
18.3.5 合并和计算	400
18.3.6 设计域	400
18.3.7 支持策略的结束语	400
18.4 论证和解释	400
18.4.1 论证	400
18.4.2 推荐的可解释性	401
18.5 偏好和评级	402
18.5.1 什么是“偏好”	402
18.5.2 评分反映了什么	403
18.6 消除选择过载	405
18.7 支持性的反复试错法	405
18.7.1 有稳定评判标准的反复试 错法	406
18.7.2 有变化评判标准的反复试 错法	407
18.8 处理在选择过程中潜在噪声 造成的影响	407
18.8.1 上下文效应	407

18.8.2	顺序效应	408
18.8.3	框架效应	408
18.8.4	启动效应	409
18.8.5	默认值	409
18.9	总结	410
	致谢	410
	参考文献	411

第 19 章 推荐系统中的隐私问题 ... 415

19.1	简介	415
19.2	推荐系统中的隐私风险	416
19.2.1	由推荐系统导致的风险	417
19.2.2	由普通用户导致的风险	419
19.2.3	由外部人员导致的风险	420
19.2.4	小结	420
19.3	隐私保护方法	421
19.3.1	面向架构与系统设计的方案	422
19.3.2	面向算法的方案	425
19.3.3	基于政策的方案	429
19.4	人的因素和对于隐私的认知	430
19.4.1	透明化与管控的局限性	431
19.4.2	隐私助推	432
19.4.3	隐私自适应	433
19.5	总结与思考	434
	参考文献	435

第 20 章 影响推荐系统可信度评估的来源因素 ... 442

20.1	简介	442
20.2	在线来源的可信度评估	442
20.3	推荐系统作为社交参与者	443
20.4	人际交互中的来源因素	443
20.4.1	来源可信度	443
20.4.2	来源线索	444
20.5	人-技术交互中的来源因素	445
20.6	用户与推荐系统交互的来源因素	446
20.6.1	推荐系统类型	447
20.6.2	输入特征	447
20.6.3	过程特征	447

20.6.4	输出特征	448
20.6.5	嵌入式智能代理特征	448
20.6.6	新兴社交技术的影响	449
20.7	讨论	449
20.8	影响	450
20.9	未来研究方向	451
	参考文献	451

第 21 章 用户性格和推荐系统 ... 459

21.1	简介	459
21.2	什么是性格特征	460
21.2.1	五因素性格模型	460
21.2.2	其他性格模型	461
21.2.3	用户性格与用户偏好如何相关	462
21.3	性格采集	462
21.3.1	显式性格采集方法	463
21.3.2	隐式性格采集方法	464
21.3.3	推荐系统的线下实验数据集	466
21.4	如何在推荐系统中使用用户性格	466
21.4.1	解决新用户的问题	467
21.4.2	多样性/意外收获	468
21.4.3	跨域推荐	469
21.4.4	群组推荐	469
21.5	难题和挑战	469
21.5.1	非侵入式方法获取性格信息	469
21.5.2	大规模数据集	470
21.5.3	跨域应用	470
21.5.4	多样性	470
21.5.5	隐私问题	470
21.6	总结	471
	致谢	471
	参考文献	471

第五部分 高级话题

第 22 章 组推荐系统：聚合、满意度和组属性 ... 476

22.1	简介	476
------	----	-----