

仅供非商业用途或交流学习使用

中国科学技术大学教授 陈恩红
思科大中华区副总裁 方剑斌
《人都是产品经理》作者 苏杰

联合
力荐

Broadview®
www.broadview.com.cn



数据产品经理

从零基础到令人惊艳



中国工信出版集团

电子工业出版社
PUBLISHING HOUSE OF ELECTRONICS INDUSTRY
http://www.phei.com.cn



作者简介



李鑫

于中国科学技术大学获计算机科学博士学位，悉尼科技大学访问学者，大数据分析与应用安徽省重点实验室研究员，中国互联网协会青年专家。现任科大讯飞研究院研究主管，先后负责大数据与人工智能技术在教育、脑科学等领域落地的业务，参与多个国家863、国家自然科学基金以及国家重点研发项目，并在国际知名学术会议与期刊发表论文近30篇。《推荐系统》一书译者之一。





数据产品经理

从零经验到令人惊艳



电子工业出版社

Publishing House of Electronics Industry

北京·BEIJING



内 容 简 介

当产品经理遇上大数据时代，数据产品经理应运而生。新时代的新岗位自然也有新要求。数据思维、数据预处理、数据统计、数据挖掘、数据可视化等是产品经理的必备技能。懂产品、懂运营、懂市场、懂表达、懂管理则是数据分析师的技能外延。本书正是为有志于从事数据产品岗位的人士提供掌握上述技能的必修课。

让我们通过本书，在大数据的浪潮中乘科技与人文的扁舟，驶过数据产品经理的港湾，驶向数据科学家的彼岸。

未经许可，不得以任何方式复制或抄袭本书之部分或全部内容。
版权所有，侵权必究。

图书在版编目（CIP）数据

数据产品经理必修课：从零经验到令人惊艳 / 李鑫著.
—北京：电子工业出版社，2018.4
ISBN 978-7-121-33695-9

I. ①数… II. ①李… III. ①数据处理—产品设计 IV. ①TP274

中国版本图书馆 CIP 数据核字（2018）第 029472 号

策划编辑：孙奇俏

责任编辑：张春雨

印 刷：三河市双峰印刷装订有限公司

装 订：三河市双峰印刷装订有限公司

出版发行：电子工业出版社

北京市海淀区万寿路 173 信箱

邮编：100036

开 本：720×1000 1/16 印张：20.25

字数：350 千字

版 次：2018 年 4 月第 1 版

印 次：2018 年 4 月第 1 次印刷

定 价：69.00 元

凡所购买电子工业出版社图书有缺损问题，请向购买书店调换。若书店售缺，请与本社发行部联系，联系及邮购电话：（010）88254888，88258888。

质量投诉请发邮件至 zltz@phei.com.cn，盗版侵权举报请发邮件至 dbqq@phei.com.cn。

本书咨询联系方式：010-51260888-819，faq@phei.com.cn。



推荐序 1

我在中科大长期为本科生和研究生讲授人工智能相关课程，每次讲到人工智能的历史，就会提到人工智能之父——艾伦·图灵（Alan Turing），他也是英国著名的数学家和逻辑学家。二战时期，图灵曾协助军方破解了德国的著名密码系统“英格玛”（Enigma），从而扭转了战局，帮助盟军取得了二战的胜利。军事行动中破解密码的行为，其实就是在获取不对称的信息，从而占得先机，拔得头筹。

在和平年代，尽管没有烽火和狼烟，但商业环境中围绕信息的竞争却一点也不逊色于军事行动。掌握行业动态信息的基金公司可以把握股市的走向，从量化交易中获得利润；了解客户喜好的电子商务公司能够绘制用户画像，用精准推荐扩大流量；拥抱数据智能的政府机构可以打通数据孤岛，在海量数据指导下智慧治理。习近平总书记高屋建瓴总结性地指出“信息掌握的多寡成为国家软实力和竞争力的重要标志”“谁掌握了数据，谁就掌握了主动权”。

无论是量化交易程序，还是个性化精准推荐系统与智慧治理，都属于数据产品的范畴。数据除了在金融、电商、政府治理等领域有用武之地，其与实体经济也已深度融合，在各行各业都形成了增长点和新动能。然而利用好数据，做出适用于行业与特定领域的数据产品并非易事，这需要企业、机构员工都需具备一定的数据思维。

在我看来，正是由于海量的数据极大地扩展了人们的视野，传统的搜索才让位于个性化的精准推荐；正是由于数据快速的变化，人们才会目不暇接，甚至尚未从上一条数据所提供的信息中缓过神来，又要快速地投入到对下一条数据的处理。数据就这样从体量和速度上使得人们的认知有了盈余，从而我们可以在一个信息爆炸的时代站在历史的高度和宏观的尺度上，更为深入地理解、思考、判断。

我们为整个互联网贡献了语音、图像、文本、视频以及地理位置数据，作为回馈，也享受了互联网基于这些数据而为我们提供的全方位、多角度、便捷的服务。无论是从移动互联网时代走过来的数字移民，还是生长在万物互联时代的数字原住民，都具有一幅全息、多维的数字用户画像。参与得越深刻，画像越清晰。毫不夸张地说，现实世界的物理空间与网络世界的虚拟空间的界限越来越模糊，



数字化的自己已经逼近了真实的自己，数字化的事实也很有可能就是事实。多维数据的真实性越来越受到重视且不容置疑。

数据思维就像是现代化社会的一种方法论，它既是文化，也是工具；它既有阳春白雪的理论体系，也有下里巴人的实践指南。这一虚一实之间，顶天与立地之间，都润物细无声地透露着数据的价值。

本书作者曾在我的科大实验室中度过了7年的学生时光，学生时期他就善于深入浅出深奥知识，并分享给身边的人。这本书也延续了这个风格，原本需要高等数学基础的数据挖掘技术竟可以通过直白、简单的语言和比喻来说清楚，束之高阁的专业知识就这样走向大众，为更多的人服务。

书中详细介绍了大数据的来龙去脉，以及数据技能的方方面面。这既可以为企业中从事数据相关工作人员提供思维地图，也可以为其他机构中想要跨界了解数据世界的人提供一扇窗。另外，书中充满了关于数据历史、思维、技术周边的奇闻轶事，读起来想必不会枯燥，由此也可见作者广泛的阅读经历。

“纸上得来终觉浅，绝知此事要躬行”，要想发挥数据的价值，除了从思想上拥抱数据，更要掌握数据变换的规律，知晓数据挖掘的算法，运用数据产品的技巧，锻造数据人的自身修养，用数据指导实际工作。

希望《数据产品经理必修课：从零经验到令人惊艳》这本书能够带给你所需要的思维与技巧，也希望更多产业界的朋友能够做好合格的数据人，开发出优秀的数据产品，拥有惊艳的数据业绩！

陈恩红

中国科学技术大学教授

安徽省计算机学会理事长

安徽省大数据产业联盟理事长



推荐序 2

60年前，苏联第一颗人造地球卫星被送入太空。作为冷战的响应，美国国防部组建了高级研究计划局（ARPA），开始将科技用于军事。11年后，ARPA组网连接了美国西部的若干个高校，成为了互联网的雏形。1983年，美国国家自然科学基金会（NSF）投资建设了连接更多节点的广域网 NSFNet。1984年12月，思科系统公司正式成立。

正如公司的标志性图标“旧金山大桥”所展示的那样，思科成立之初就致力于连接世界。从有线交换机到移动路由器，连接的自由增加了；从协作终端到在线视频会议，连接的内容丰富了；从互联网到物联网，连接的品类增多了。一切围绕连接，一切为了连接，一切走向连接，这似乎是技术发展亘古不变的主题。

在云计算时代，更多的物理资源被连接，而后数据孤岛的连接与打通又召唤来了蓬勃发展的大数据时代。数据的沉淀与喂养使得更多的智能算法被跨界连接在一起，成就了现在的人工智能时代。当下，每一个身处信息与通信行业的人都需要、也应该走出自己的小圈子，与周围更多的人和事物连接。

作为行业的典型，数据科学家是当今业界非常热门的岗位，但真正符合其职业要求的却少之又少。原因就在于数据科学连接了计算与管理，连接了技术与商业。跨学科、跨领域、跨门类的知识连接让很多人望而生畏。

“世上无难事，只要肯登攀”，数据科学家也许遥不可及，但数据产品经理却更容易上手。正如作者在书中所述：“数据产品经理既可以是拥有产品思维的数据分析师，也可以是掌握数据分析技能的产品经理。”只要能够做好数据技能与产品思维的连接，你就是一个合格的数据产品经理。这本书提供的就是这种连接。

我与作者相识8年，尽管当时“大众创业，万众创新”的号召还没有被提出，但还是学生的他就已经开始在宿舍“捣鼓”一个叫作“心愿FM”的App创业项目。这款App提供类似校园社交的功能，其本质就是建立校园中异性之间的连接。时光荏苒，尽管这个项目已经不复存在，但作者却将“连接”的精神保留了下来，成为了一名跨界连接的专家。



IV 数据产品经理必修课：从零经验到令人惊艳

阅读这本《数据产品经理必修课》，透过关于技术的字里行间，我看到了人文的斑驳倩影，不得不说，作者在写作上也将技术与人文进行了一次连接。

面向未来，我笃定地坚信，只有不断学习、自我更新的人才有可能胜出，成为企业需要的人才。学习既包括深度的探究，也包括广度的见识。无论是在大公司还是创业公司，是国企还是外企，了解一门新的知识总是有百益而无一害的，更何况是紧跟时代节奏的数据知识。

希望这本书能够真正帮助到有志于从事数据科学与数据产品工作的人，也希望它能成为读者和作者连接与沟通交流的纽带。

方剑斌

思科大中华区副总裁



前言

Preface

数据产品是什么

倘若我提出这样一个问题：数学中的 1、2、3 分别代表什么？你心中会有什么样的思考，又会有什么样的答案呢？早在 2500 年前的古希腊，毕达哥拉斯学派就已经给出了答案：点、直线与平面是对这三个数字的几何描述；源头、两性与稳定则是对它们深层次内涵的诠释。由此看来，数学有着神秘的意义。难怪古希腊数学家普洛克拉斯会说：“哪里有数学，哪里就有美。”

数据一词事实上是按照宾语前置的方式来构词的，所以我们可以理解为“据数”或“以数为据”，意思是把数据当成考究的凭证。正如数学是人类早期复杂贸易催生的结果，数据先天也带有商业的属性。从结绳记账到珠算发明，从证券股票到数字广告，我们甚至可以模仿先贤的口吻说道：“哪里有数据，哪里就有商业。”

历史的车轮已滚过千年，但数据的概念并未行将迟暮，垂垂老矣，反而老当益壮，焕发生机，这都要归功于“大数据”概念的产生。对于大数据这一概念，行业中有人将其归功于某家公司，有人将其产生与某位学者联系起来，但他们更多地是这个概念的精神作者，大数据真正的作者应该是接受并使用它的人，从这点来说，消费者才是其真正的衣食父母。

然而数据毕竟是一个虚构的概念，当我们谈论数据的时候，我们并没有办法在物理世界中找到一个实物来说明其客观存在。也许你会拿出刚打印出的报表，并反驳我说：“难道这不是数据吗？”可是你指的究竟是白色的纸还是黑色的油墨呢？由此可见，数据之名，需要借以载体之实才可以发挥价值，我们身边的产品就是这样一类载体。这么看来，打印出来的报表，包括邮寄上门的水电费单据，

都是数据产品。

数据产品的内涵应该不止这么肤浅，要不然岂不是“人人都懂数据产品”了？数据产品最为重要和关键的价值是驱使行动。水电费单据驱使我们缴费，推荐系统驱使我们阅读，财务指标驱使公司制定战略，就连菜市场用于标注商品价格的黑板也能驱使人们采购。如此看来，一个不能够驱使行动的数据产品其价值可能要大打折扣了。

让我们再次回顾一下数据产品这个概念，人的行动产生贸易，贸易产生数据，数据通过产品展现，数据产品驱动人的行动，人的行动又产生数据……周而复始，形成闭环。这才是一个完整的数据产品。

为什么要写此书

有关产品经理的著作有很多，仅在 2017 年，我拜读过的就不下 15 本。从琳达·哥乔斯到苏杰，从乔克·布苏蒂尔、卢克·米勒到陈峻锐、闫荣、后显慧、刘飞，我的产品思维便是从他们的荟萃中汲取的营养，因此你也许会从本书中读到他们的部分观点。关于大数据技术的书更是数不胜数，倘若将概念扩展到数据挖掘与机器学习领域，光是近几年出版的书籍便已不胜枚举。

如此说来，市场上似乎并不缺少有关数据产品的书，读者只需要兼读两者即可。那么为什么我又要写这本书呢？

著书并非我的本意，我的初心只是分享，这些内容最初被我写在我的个人头条号上，因为我信仰“分享是最好的学习方式”。在分享的过程中当然会受到质疑，但回答质疑乃至承认错误也是学习与进步的一部分。著书不过是分享的一种渠道罢了。

依我有限的阅读量和浅见来看，大多数技术类书籍充满了大量面向“圈内人”的专业术语，每一个术语都像是横梗在读者通往知识彼岸道路上的一座大山。并不是所有的读者都需要，或者愿意，甚至有能力“逢山开路”。“知识若庞杂到无法在民众中普及，则极易沦为经院哲学，甚至演化为民众对权威的盲目迷信”，威尔·杜兰特如是说。因此美国历史学家詹姆斯·哈维·罗宾逊号召“拆除壁垒，还知识于民众”，我仅仅是众多拥趸者之一罢了。

恰逢此时，我拜读了英国作家赫伯特·乔治·威尔斯（Herbert George Wells）的

《世界史纲》，这本装订成上下册的历史界的“红宝书”可以算是开了“概论”的先河，行文有趣，笔法生动，让我一个在历史方面十分愚钝的人也感起兴趣来。我也拜读了吴军老师与涂子沛老师几乎所有的著作，书中所介绍的科技背后的历史桥段让我反复咂摸。《吴晓波频道》和《罗辑思维》对我也颇有启迪。或许，将理性的数据用感性的故事进行呈现，是一种更易于让读者接纳的方式。我不禁这样想。

数据冷酷得像一个法官，故事却像富有温情的妇人，理性与感性的矛盾不言而喻。不仅如此，科技与人文的较量，非虚构与文学的角力，也都正在进行。尽管美国作家菲茨杰拉德告诉过我们：“测验一个人的智力是否属于上乘，只看脑子里能否同时容纳两种相反的思想，而无碍于其处世行事。”但知易行难。

为了获得这样的智慧，我们需要找到途径和方法。好在查理·芒格、赫伯特·西蒙（司马贺）以及小泉英明各自都著书立说，为我们提供了工具与方向。这是一种被称为多学科交叉的思考方式，我也希望用这种方式来写作与分享，让这本书在丰富内容之余兼具有趣的灵魂。

本书内容

全书分四大部分，共计 15 章，每部分及每章的具体内容如下。

第一部分 产品经理的前世今生

第 1 章 产品经理的前世

产品经理一词究竟是何意义？该岗位从何而来？广义与狭义上有何区别？当前互联网行业的产品经理究竟做些什么？本章将为你一一揭晓。

第 2 章 产品经理的今生

提出管理动机、广义竞品分析与交互设计这样的对产品经理的更高要求。另外，本章也将梳理产品经理在产品、团队、公司层面必须迈过的雷池。

第 3 章 产品经理的入行

“宽进严出”是产品经理岗位的特色，低门槛使得初入此行甚是轻松，高要求则使得出类拔萃愈加困难。本章将介绍从事产品经理的人员该具备什么样的视野，又该在工作中如何学习。

第二部分 古往今来的数据思维

第4章 历史中的数据思维

数据埋点、数据可视化、数据产品落地、数据驱动决策、利用数据降本增效、统计分析以及打通数据孤岛这些老生常谈的话题似乎可以在过往的岁月中找到关联。本章向历史溯源，给出老概念的新故事，诉说新时代的旧往事。

第5章 行业拥抱数据思维

从蓝色星球到 960 万平方公里的泱泱大国，从与政府密切联系的科教文卫体到与民生息息相关的衣食住行，本章将介绍大数据渗透到的每一处角落。

第6章 当产品经理遇见数据思维

当产品经理遇见数据思维的时候，不仅有“眼前的苟且”——数据产品经理，还有“诗和远方”——数据科学家。从现实到理想的距离，本章将为你搭梯。

第三部分 数据产品经理的技能进阶

第7章 面向产品经理的数据预处理

“磨刀不误砍柴工”，做好数据预处理可以为数据分析与挖掘过程节省许多时间。本章从数据清洗、数据集成、数据变换以及数据规约四个角度全面阐述数据预处理的相关知识。

第8章 面向产品经理的统计分析

本章从非时序数据与时序数据、分类数据与连续数据的角度，介绍数据统计与数据分析的概念及技巧。

第9章 面向产品经理的数据挖掘

本章旨在讲清楚数据挖掘的方方面面，内容包括回归、分类、聚类、关联分析和时间序列分析等数据挖掘算法，以及在此基础上的集成学习、文本挖掘、社交挖掘、排序算法、推荐系统以及用户画像。最后以这些算法中蕴含的哲学内涵作为结束。

第10章 面向产品经理的数据可视化

数据可视化是“技术与美”的最好结合。本章分别介绍“高大上”与“接地

气”的两类数据可视化，并从数据展现和逻辑修饰两个层面介绍识别数据“说谎”的技巧，最后给出数据可视化的终极形式——数据报告的制作方法。

第 11 章 向数据科学家再迈一步

本章介绍与数据产品岗位相关的另外三个岗位，分别是运营、研发与市场。在通往数据科学家的铸鼎之路上，这三足必不可少。

第四部分 数据产品经理的自我修养

第 12 章 学习力：借方法论加速

本章系统介绍工作中遇到的各种方法论，并概括总结诸多方法论的“模板”，最后提出笔者自己的学习方法论，帮助读者建立工作中的“理论自信”。

第 13 章 表达力：用逻辑学帮衬

本章从“为国考正名”谈到“名著中的名言警句”，详细介绍产品经理日常工作与案头写作的心得，以及汇报与分享中极具感染力的“故事思维”。

第 14 章 领导力：以经济学诠释

本章在经济学原理中找到管理学中领导力的跨学科基础——无论是团队制定目标时的“举旗定向”，还是实际工作时的“谋篇布局”；无论是团队内部配合的“取长补短”，还是团队之间协作的“互利共赢”。

第 15 章 软实力：靠心理学打造

本章通过心理学中的若干现象与小实验，介绍帮助数据产品经理灵活驾驭本职工作的若干心得与技巧。

联系作者

无论是随意翻阅至此，还是计划花一番功夫细细研读，从你捧起此书的那一刻起，我相信你我之间便产生了神奇的联系与耦合。书，就是我们之间的桥梁与纽带。我是本书的唯一作者，这意味着所有书中可能存在的疏漏与谬误也尽归于是我，对此我表示歉意并将竭力修正。而我却又不是本书的唯一贡献者，出版社的编辑、公司的同事、身边的朋友以及读者朋友们的每一条建议都帮助过我进步，使这本书更加精细。

欢迎通过我的微信 [ustelixin](#) 与我取得联系，我会尊重大家的每一条意见和建议，并愿意与大家一起探讨有关数据产品的任何话题。

无以名状的感谢

尽管我是此书的唯一作者，但成书并非我一人所能为。儿时囫圇吞枣习得的典籍，现在不求甚解读下的篇章，是家庭给予我最宝贵的精神财富。

数据思维的启迪要归功于中科大的陈恩红教授，同时，实验室为我提供了徜徉在数据海洋的一叶扁舟。

科大讯飞公司中的一些同事参与了本书早期内部培训版本的审校，头条号与PMCAFF社区中的网友也向我提出了极具思考性的问题，在此不详细列举他们的姓名，一并谢过。

尽管成书早期也接触了其他出版社和编辑，但最终还是选择了博文视点，选择了给我提出建议最多，需要我动笔修改最多的孙奇俏老师，因为我相信她的认真与执着会让我更加专业。

关于内容，我如履薄冰，无论是鼓励，还是建议，抑或是批评，我都一并感谢并将铭记，且承诺会认真修正。

最后，再次深深感谢家人！

李鑫

2018年2月于合肥

读者服务

轻松注册成为博文视点社区用户（www.broadview.com.cn），扫码直达本书页面。

- **提交勘误：**您对书中内容的修改意见可在 **提交勘误** 处提交，若被采纳，将获赠博文视点社区积分（在您购买电子书时，积分可用来抵扣相应金额）。
- **交流互动：**在页面下方 **读者评论** 处留下您的疑问或观点，与我们和其他读者一同学习交流。

页面入口：<http://www.broadview.com.cn/33695>



目 录

Contents

第一部分 产品经理的前世今生

第 1 章 产品经理的前世

1.1 产品经理究竟是什么	4
1.1.1 咬文嚼字说产品经理	4
1.1.2 产品经理的历史溯源	5
1.2 泛产品经理与产品经理	6
1.2.1 产品经理的专业取向	7
1.2.2 产品经理的泛化	8
1.3 互联网产品经理的规定动作	12
1.3.1 需求调研	12
1.3.2 竞品分析	14
1.3.3 原型设计	16

第 2 章 产品经理的今生

2.1 卖家秀：自我提升的几项技能	20
2.1.1 从需求文档到动机文档	20

2.1.2	从竞品分析到广义竞品分析	22
2.1.3	从原型设计到交互设计	24
2.2	买家秀：弄垮团队的若干“要领”	28
2.2.1	越过产品雷池	28
2.2.2	踏入团队雷池	29
2.2.3	迈向公司雷池	30

第3章 产品经理的入行

3.1	入行做产品的几种可能	34
3.1.1	源自技术岗	34
3.1.2	源自业务岗	35
3.1.3	源自应届生	36
3.2	上岗后的第一件事	37
3.2.1	产品全图	38
3.2.2	行业全图	39
3.2.3	产业全图	40
3.3	工作中如何学习	41

第二部分 古往今来的数据思维

第4章 历史中的数据思维

4.1	人口普查：最早的数据埋点策略	46
4.1.1	埋点的技术视角	46
4.1.2	埋点的时机与策略	48
4.2	命令与征服：可视化最早的用意	49
4.2.1	可视化大家说	50