



大数据科学与应用丛书

风控

大数据时代下的 信贷风险管理和实践

◎ 王军伟 著

大数据风控是互联网金融乃至传统金融风控的必然趋势，
它的发展给金融领域带来巨大的福音。



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>

 大数据科学与应用丛书

风控

大数据时代下的 信贷风险管理和实践

◎ 王军伟 著

大数据风控是互联网金融乃至传统金融风控的必然趋势，
它的发展给金融领域带来巨大的福音。



電子工業出版社

Publishing House of Electronics Industry

北京•BEIJING

内 容 简 介

本书对大数据时代下的信贷风险管理进行了介绍和剖析。首先,从经济学理论与实践应用上对信贷的产生和经济意义、信贷分析方法的变迁进行阐述;其次,对信贷整个生命周期中使用的 Cohort 分析、信贷业务开展、合同签订、风险监控预警、催收和不良资产处置、系统信息管理系统中报表等重要方法进行了深入讲解;最后,从财务数据、信用报告、交易流水等信贷角度方面分析借款者的还款能力和还款意愿,并提出了还款意愿的货币量化方法。同时,对传统信贷方法、IPC 信贷方法、巴塞尔协议方法、大数据风控进行优缺点分析,提出了基于 IPC 信贷、巴塞尔协议的大数据风控模式,并给出了不同情况下的具体实施方案,有助于信贷机构提高自身风险管理能力。

本书理论与实践相结合,适合银行、信用保证保险、消费金融、P2P、小贷公司、互联网金融、大数据风控等从业人员,以及有意从事金融工作的人员阅读与参考。

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

风控:大数据时代下的信贷风险管理和实践 / 王军伟著. —北京:电子工业出版社,2017.8
(大数据科学与应用丛书)

ISBN 978-7-121-31960-0

I. ①风… II. ①王… III. ①信贷管理—风险管理 IV. ①F830.51

中国版本图书馆 CIP 数据核字(2017)第 139680 号

策划编辑:李树林

责任编辑:赵 娜

印 刷:三河市鑫金马印装有限公司

装 订:三河市鑫金马印装有限公司

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:787×1 092 1/16 印张:16 字数:231 千字

版 次:2017 年 8 月第 1 版

印 次:2017 年 10 月第 2 次印刷

定 价:49.00 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) 88254463, lisl@phei.com.cn。

前言/Prologue

正如马云提出的“我们正处于 DT (Data Technology) 时代”，确实我们处在一个数据时代，未来 40 年是数据和量子时代。国家提出的“互联网+”其实质是数据，而互联网金融是“互联网+”的重要组成部分。互联网金融的本质是利用互联网技术来解决金融问题，而金融的本质是解决信息不对称。金融天生就是数据，所以互联网金融就是充分利用数据来推进金融的发展。

金融的核心是风险控制（简称风控），互联网金融也不例外，其核心是风控，只是互联网金融的风控更加依赖数据和数据样的，也就是我们提到的大数据风控。大数据风控的模式到底应该是什么样的？大数据风控与传统信贷风控模式、IPC 信贷风控模式、“信贷工厂”风控模式、巴塞尔协议风控模式是什么关系？大数据风控结构是怎样的？大数据风控如何实施？大数据风控如何评估？这是数据时代风控人需要解决的问题。

要理解风控，首先要对信贷有清晰的认识。本书主要从主流经济学的跨期选择、消费理论、厂商理论、帕累托效率或卡尔多-希克斯效率、马克思经济学等角度对信贷进行解读，这样将更加便于清晰理解信贷的意义。对于信贷业务则先从信贷机构的风控偏好来选择客户、刻画客户画像、设计产品、策划市场营销方案等方面做了简略的阐述；随后对风险管理涉及的 Vintage 分析，从应用和理论两方面进行了阐述；最后从信贷分析、信贷方法比较、审批决策和定价、借款合同和放款、风控预警、还款、逾期管理和 MIS 系统等方面进行阐述和探讨。本书第 2~4 章，对信贷的经济学基础、信贷分析秘密武器——Cohort 分析、信贷业务的开展进行了介绍。第 5 章从信贷分析涉及的财务分析、信用分析、银行流水分析、交易流水分析等方面分析客户的还款能力和还款意愿，并对各种软信息进行分析，从而校验还款能力和还款意愿，给出还款意愿的货币化量化方法，最后对

全面风险评估和压力测试方面进行了简要介绍。第 6 章主要对传统信贷方法、IPC 信贷方法、信贷工厂方法、巴塞尔协议方法、大数据风控进行优缺点分析，提出改进建议，最终给出可以操作的基于 IPC 信贷、巴塞尔协议的大数据风控模式。第 7 章从审批的角度阐述了大数据风控审批模式和实施经验及定价。第 8 章是关于合同的签订和放款。第 9 章给出风控预警相关模式和经验。第 10 章是还款阶段的简单阐述。第 11 章就催收和不良资产处理进行了阐述。第 12 章是关于风险管理的报表系统——MIS 系统的介绍，并给出了关键性报表。

本书基于作者对传统银行贷款、抵押质押或担保贷款、IPC 微贷分析技术、供应链贷款、评分卡贷款、巴塞尔协议下的贷款、大数据风控贷款的理解，结合自身的实践和一些具有丰富经验的人给予的分享和教诲而成，如哈尔滨银行互联网金融副总裁赖华、中银消费风控总监马新生、徙木金融策略总监吴淑玉、浙商银行资深风控专家辜小亮、资深风控专家殷少坤。赖华先生对作者而言既是兄长和朋友，更是风控方面的恩师，他对事物具有独特的看法和理解，给予了作者细心、无私的帮助和指导是完成本书写作的前提。殷少坤是 IPC 信贷技术方面的资深专家，他更是给予了作者很大的帮助。有他们无私的资料分享，以及作者自身在大数据分析和挖掘近 10 年的实践、Basel 协议 8 年的研究、近 4 年的风险实践，疯狂阅读吸收各种优秀金融书籍，在总结信贷风控各种技术的基础上，不断突破自我，形成此书。愿同有志于在这个行业发展的人一起为大数据风控发展尽绵薄之力。

感谢电子工业出版社的编辑李树林先生为本书的辛勤付出。感谢我的妻子周晓霞和两个孩子，为了尽快完成书稿，他们给了我充足的时间和鼓励。特别感谢我妻子周晓霞女士对我“疯狂”购书的包容，因为我每周要看 3 本书左右，也喜欢收藏一些书，她竟然想办法在家里专门辟出一块空间为我摆放图书。

因为个人经验和认知有限，书中难免存在不妥之处，恳请读者抱着“尽信书不如无书”的理念和质疑的态度来阅读此书，同时也希望读者就文中错误和不足之处指出并发电子邮件（wjw84221@aliyun.com）给我，在此表示感谢！

王军伟

2017 年 2 月 26 日

目录/Contents

第 1 章	导言 / 001
第 2 章	信贷的经济学基础 / 010
	2.1 信贷产生的经济学分析 011
	2.2 信贷分析方法随经济周期而发生变化 019
	2.3 信贷风控和策略的经济学分析 024
第 3 章	信贷分析秘密武器——Cohort 分析 / 027
	Cohort 分析的案例和模型 035
第 4 章	信贷业务的开展 / 040
	4.1 客户画像和产品设计 041
	4.2 市场开拓和营销 046
	4.3 申请调查 051
第 5 章	信贷分析 / 062
	5.1 硬信息分析 66
	5.2 软信息分析 113
	5.3 还款意愿量化方法 126
	5.4 全面风险管理 134
	5.5 压力测试——未来预期与敏感度分析 139

第 6 章	现有信贷方法的优缺点与改进建议 / 145
6.1	传统信贷的优缺点和改进建议 147
6.2	IPC 信贷的优缺点和改进建议 151
6.3	“信贷工厂”的优缺点和改进建议 154
6.4	巴塞尔协议模式的优缺点及改进建议 156
6.5	大数据风控模式的优缺点和改进建议 160
6.6	基于传统信贷、IPC 信贷、“信贷工厂”、巴塞尔协议 和大数据风控模式融合的展望 177

第 7 章	信贷的审批决策 / 180
7.1	信贷审批委员会决策模式 182
7.2	“信贷工厂”审批模式 188
7.3	大数据风控自动审批模式 189

第 8 章	信贷的合同签订及贷款发放 / 200
--------------	---------------------------

第 9 章	风险监控预警 / 207
--------------	---------------------

第 10 章	信贷的还款阶段 / 214
---------------	----------------------

第 11 章	逾期管理和不良资产处理 / 221
---------------	--------------------------

第 12 章	MIS 系统 / 234
---------------	---------------------

后记 / 243

参考文献 / 246

| 第1章 |
导 言

信贷是为产业发展服务的。

——经济学家 熊彼特

信贷是经济发展的发动机，没有信贷就没有经济增长，也不会有人类生活、生产的提高与改善。按照资本积累来说，仅靠自身积累是不足以支持经济发展的。信贷有着悠久的历史，据一些文献资料记载，信贷在公元前 2000 年前就已出现，到目前已有 4000 多年的历史了。

信贷的形式多样，先说说国家与政府层次的信贷。首先是国家层面的信贷有国家发行的债券，例如，美国发行的国债主要由私人资本、主权资本持有，美国国债达到其年 GDP 的 120%；日本发行的国债主要由其国民持有，国债发行率高达 220%；等等。其次是地方政府发行的地方债，如美国加州发行的地方债和中国不同省（直辖市、自治区）的地方债等。再次，国家及政府为了自身的经济发展进行的贷款，如从世界银行、亚洲基础设施投资银行、亚洲开发银行，或者具体国家的商业银行获得贷款等。这些债券或贷款募集的资金一般用于基础设施建设、战争、福利发放等，主要还款来源是税收，但国家债券或贷款也有违约的时候，比如希腊在欧洲债务危机中就发生了违约。

对于企业层次的信贷，主要是指企业借贷，是企业生产经营活动中为了支付人员工资、市场拓展、原材料费用等所需要的资金，而向信贷机构申请的贷款。这种借贷体现了企业家精神，形成了信贷机构和企业之间的权利和责任的时空性分离的契约。信贷机构对于企业放款方面表现了信贷机构对企业家精神的尊重和支持，另一方面借助具有企业家

精神的人或组织的创新，分享创新带来的利益，这成为促进经济发展的重要动力。

对于个人层次的信贷主要有个人投资和个人消费，其中，个人消费主要包括住房、汽车、结婚生子、美容整形、购物、上学等，这些都将产生借款需求。因为消费是生产的根本出发点和落脚点，是经济发展的根本动力，所以满足消费所产生的资金需求也是对经济发展的支持，体现在信贷机构对个人的消费进行支持，这不仅有助于提高个人消费者的消费能力，而且促进了企业的产品变现，同时也有利于经济的快速发展和提高人民的生活水平。

按照贷款的类型可以分为信用贷、抵押贷、质押贷、担保贷，也可以分为个人贷、企业贷、项目贷等。但无论表象如何，都是权利和义务分离，且权利和义务的实现存在不可忽视的时间差，进而使信用问题显得特别突出。比如各金融机构推出的消费贷、京东白条、阿里巴巴信用支付、阿里巴巴借呗、阿里巴巴花呗（其前身为阿里巴巴的虚拟信用卡），还有微众银行的微粒贷，实质都是约定好相关期限，在未来还本付息或者免息。这些都是信贷，并不需要用抵押、质押、担保来提供第二还款来源和增加借贷者的违约成本，但往往额度比有抵押的要低一些。信贷的还款来源主要是第一还款来源，而且无论如何都不能本末倒置。因此，第一还款来源是根本，需要放贷者衡量、评估第一还款来源的稳定性和持续性。

信贷的方式有很多，常见的有贷款、债券、赊账、应收应付、保理等。从借款者角度来看主要有个人、企业、政府，这些机构或个人借贷的目的一般是消费、资金周转、扩大生产等。整体而言，贷款都是贷给自然人和法人，自然人就是具有独立民事行为能力的人，法人则需要细分为个体户、小企业、大中企业、企业集团、事业单位，但不包含国家。同时，按照是否将经营与家庭分开和是否具有完善的公司治理架构来识

别个体户、小企业、大中企业。个体户或微企业一般是以个人或家庭为单位进行经营，家庭和企业之间无法分开，具体表现在私人收入和生意收入不区分，私人开支和生意开支不区分，没有完备的会计文本资料；小企业一般是法人、企业所有人的生意和家庭收支分开，具备相应的财报，但不具备完善的治理架构和制度，依靠公司所有人自己的兴趣、经验来管理；大中企业具有法人、财务报表，以及完善的公司治理架构和制度等。

在历史上，最开始时信贷基本上都以自己所认识和熟悉的人为基础，如亲朋好友之间的信贷关系，而后是以自身权势作为“威胁”来提高借款人的违约成本进行顺利放贷，如一些合法的地方小贷公司，或者非法私人放贷。对于在熟悉的朋友圈内开展的个人信贷业务，按照 Facebook 公布的信息，一般最多达到 200 人，因为我们真正熟悉的朋友不会太多，而朋友中有信贷需求的又少之又少，所以过去金融发展需要较多的从业人员。但这种熟人经济，严重地制约了经济发展，银行等中介机构充分利用对存款人或放贷人的担保和对借款人的充分了解，开展了非熟人之间的业务，促进了信贷行业的发展和经济发展。

银行信贷的出现是人类经济发展的重要里程碑，银行信贷满足了陌生借贷人的信贷需求。对于陌生借贷人，银行需要进行风险评估，如欺诈风险、信用风险等。对信贷风险而言，其核心就是信用风险。关于信用风险的评估，传统银行审批分析模式是：无论对于个人或企业贷款，传统银行贷款主要通过纸质申请，提交各种资料，如个人需要提供身份证、户口本、结婚证、工作收入证明等，企业则需要营业执照等，然后一一核实，结合银行自己或央行征信数据进行相关分析给出结论，一般需要 15 天左右完成相关贷款。除此之外，银行还需要客户提供抵押、质押和担保来缓释风险，但抵押、质押和担保被作为第二还款来源。目前，一些信贷机构过于强调抵押、质押和担保的作用而忽视自身的信贷系统、

风控体系、数据体系的建设，但这些机构对验证抵押、质押和担保的真实性和可靠性方面的经验却非常丰富，这也是其他采用 IPC 信贷、供应链技术、评分卡技术、大数据风控技术等信贷分析技术的机构值得学习的地方。

IPC 信贷分析技术主要是基于还款能力、还款意愿，立足实质重于形式，以“单人调查”、眼见为实的原则获得真实信息，通过交叉验证获得隐性信息的真实性，而后基于现金流来编制客户“资产负债表”“损益表”和“现金流量表”。供应链贷款基于产业链或供应链中核心企业对客户的控制能力进行贷款，如阿里巴巴依靠自身在互联网上强大的控制力来对用户进行放贷，海尔供应链金融也是如此。供应链金融根本在于其对信息真实性把控和对还款的处理手段，因为供应链信贷服务对象是供应链上下游客户，各种产品或原料具有相似的功能和作用，处于供应链中核心地位的企业在客户违约后对货物或原料的处置能力是非常强的。评分卡贷款是基于一些标准（如 Z-score 得分等），对优质客户进行快速审批的贷款方式，优质客户包括公务员、银行等机构管理层。巴塞尔协议下的贷款是银行利用征信数据和自身数据，计算出客户的违约率、违约损失率、期限、相关性，结合资本准备金和 RAROC 等进行的审批、授信、定价等。大数据信贷是基于征信数据或互联网上各种数据，如消费数据、运营商数据、申请行为数据等计算违约率，并进行风险定价的模式，以拍拍贷的“魔镜”系统为代表。对于企业贷款，传统银行贷款主要通过企业的性质、财务报表、规模和实地考察等来评估，在财务报表等放低要求的基础上主要以抵押、质押或担保贷款的抵押物来确定贷款。IPC 微贷技术需要将经营收入、家庭收入和支出进行整合而进行分析，供应链贷款主要核心企业对供应链上下游企业的控制来管控风险，评分卡贷款、巴塞尔协议下的贷款、大数据风控贷款下的企业和个人是类似的，不过增加了相应的财务报表等信息。

无论是对个人或企业的传统银行贷款、抵押质押或担保贷款，还是IPC 信贷贷款、供应链贷款、评分卡贷款、巴塞尔协议下的贷款、大数据风控贷款，根本原理都是衡量客户的还款能力和还款意愿，以及反欺诈。然而，在现实中很多小企业没有财务报表，家庭和生意不分离，信贷机构在这种情景中可以采用IPC 微贷分析技术构建客户的财物报表，这里对IPC 微贷分析技术进行比较简单的阐述。

小微企业贷款是以个人或家庭为单位的经营类贷款，这类贷款占整个小微企业贷款业务中的大部分，客户基本为各类个体工商户，以小贸易及小型服务生意为主，包括部分家庭作坊形式的小生产。一般来说，小微企业贷款会重视家庭特征，对所有权问题、所处行业特点关注及未来现金流的预测也是非常必要的，因为以家庭为单位的生意模式是微小贷款客户群体的主要特征。一般而言，这类客户对企业的治理不规范，如没有或具有不完备的会计制度及文档资料，同时私人收入、私人开支与其企业经营的收入、开支不区分。在中国，绝大部分这类客户都有工商管理总局授予的营业执照，而根据从事的行业不同，有些行业必须持有特殊执照，如开饭馆的应该有卫生许可证等。小企业贷款的客户群体主要是家庭特征相对弱化的小企业，这类企业拥有一定的报表，但是其产权及管理结构超出了家庭范围，因此根据不同情况，这类客户的家庭收入、开支与其业务收入、开支相对有一定的区分。

基于小微及小企业贷款的客户特征，信贷人员在实际工作中，不仅要分析客户的生意状况，还必须分析涵盖客户的整个社会经济状况。对于客户或客户家庭同时拥有多种生意，信贷人员的分析就必须包含所有的生意。由于小微及小企业贷款的客户特点，其社会信息也是贷款分析的重要部分，因为客户的经济行为是其社会角色的经济延伸。

通常，还款意愿主要是了解客户的社会角色，包括客户的个人情况、家庭情况、客户在其周边的名声等信息。这些信息能帮助我们了解客户

处于一种什么样的社会地位，客户至今是怎样在这样的社会环境中生活的，并且可以利用这些信息量化并勾勒出客户的特征。在贷款业务中，这部分信息可归为非财务信息或软信息，技术上通过运用软信息分析模型，如信息不对称偏差分析法，可以帮助信贷人员更好地了解客户在其社会经济环境中所处的位置，并根据客户与“标准”客户的偏差提出警示。依据这些软信息，信贷机构可以量化出客户的违约成本，从而合理对客户进行授信。

一般对于还款能力的分析，主要通过财务分析，以及对客户的经济状态进行静态及动态分析，估算出客户的还款能力。还款能力分析主要是对客户生意上的收入和非生意收入进行分析，分析思路基本上以现金流为核心。在实际操作中，由于大多数客户没有正规的会计报表，信贷机构必须获得必要的信息来编制客户的财务报表，通过编制财务报表来估算出客户的实际还款能力。对于客户提供的完整财务报表，信贷机构需要对报表中的数据进行逻辑交叉检验，一旦提供的报表存在逻辑异常，就需要重新构建相关报表。

还款意愿与还款能力还可以直接从客户的收入等财物信息、身份、地位等软信息进行分析获得，也需要结合客户的贷款目的进行分析，但对于客户的反欺诈则需要从客户产生借贷需求和为什么产生信贷等入手进行分析。

我们将从消费和供需理论、跨期消费理论、信息经济学、经济发展周期等理论来解读信贷产生的原因和带来的问题。按照斯蒂格勒和斯蒂格利茨等人的理论，信贷风险主要是信息不对称而引发了道德风险（如欺诈），以及逆向选择风险，或者因没有还款能力，从而产生欺诈、逾期等问题，信贷机构或组织不得不小心翼翼地对待信贷申请者的申请。在风险管理中有句话值得牢记——“能把钱放出去不算本事，把放的钱和利息全部收回才是真正的本事”。

信贷的业务流程包括确定目标客户群、产品设计、市场营销推广、信贷分析和审批、贷后管理等，在这个过程中有一种方法是自始至终都会用到——Cohort 分析，贷后分析称为 Vintage 分析。对于这种分析在这里我们将进行深入的阐述，便于信贷相关人员使用。

在本书中，对于还款能力的分析，需要结合“互联网+”的各种与收入和成本相关的要素进行解释，比如可以通过消费记录、广告等多个方面来估计还款能力。对于还款意愿分析，一般认为还款意愿无法通过大数据进行估计，而在实际操作中，我们发展出了还款意愿货币化方法来量化客户的还款意愿，从而很好地解决这个问题，比如当还款意愿货币化的值小于等于 0（或者信贷机构自身设置的一个阈值）就可以拒绝为该用户服务。

在对还款能力、还款意愿的分析中，存在多种信贷技术，不同的信贷技术具有自身的优缺点。针对不同方法的缺点，我们提出了相应的改进方法，并给出了大数据风控的架构图。根据我们对同一批数据上万模型结果的分析，方法不是最重要的，数据才是真正的关键。大数据风控核心是数据，在不断拓展数据的情况下，如何充分利用数据，以及不断通过阿里巴巴前首席数据官（CDO）车品觉先生提出的“联动”方式来不断让数据“动”起来，从而发挥出数据的价值。例如利用上网的 IP 找到对应的地址，并结合上网时间，判断该地址是用户的公司还是家庭住址，如果 IP 地址是公司地址，根据公司地址可以搜索公司名称从而判断客户的收入情况和风险所在，如果 IP 地址判断出是客户家庭地址，则可以结合房价或房租来判断客户的收入和支出。

当我们对各种数据真实性进行验证，量化出还款能力、还款意愿后，就需要决定是否进行放贷，这就是审批。在不同的信贷技术下，对应不同的审批模式，比如审批委员会模式、“信贷工厂”式审批和大数据风控的自动化审批模式，并对自动化审批模式内部如何决策进行了展示，同

时也就信贷定价进行了阐述。本书还针对贷后方面的风险预警、还款流畅和便利性、催收和不良资产处理、信息管理和报表系统（MIS）进行了简单阐述。虽然是简要介绍，但这并不代表着不重要，因为风险管理是全面的、全流程的风险管理，信贷业务流程中任何一环出问题都会造成逾期和不良率上升。

| 第2章 |

信贷的经济学基础

保险市场和信贷市场存在道德风险和逆向选择问题。

——诺贝尔经济学奖得主、经济学家 约瑟夫·斯蒂格利茨