互联网金融系列丛书

倾力推荐

大数据金融与征信

何平平　车云月　编　著

大数据+金融落地实践之作
汇集大数据金融专家思想精华

大数据时代，深度剖析当下及未来征信行业发展机遇与路径

清华大学出版社

互联网金融系列丛书

大数据金融与征信

何平平　车云月　编著

清华大学出版社
北京

内 容 简 介

本书面向金融应用，系统地阐述了大数据金融与征信本身及其在现实生活中的应用，具有全面性、实用性和前瞻性等特色。全书共 8 章，第 1 章和第 2 章阐述大数据金融及大数据技术相关的基础知识问题，是后面章节的基础。第 3 章至第 6 章详细介绍大数据在银行业、证券业、保险业及互联网金融行业中的应用，是本书的主要内容。第 7 章重点阐述大数据在征信中的实际应用，是本书的另一重点问题，也是当代大数据研究的热点问题。第 8 章特别强调中国金融信息安全，这是大数据金融与征信的发展进程中不可避免的问题。本书力争把大数据与其实际应用糅合在一起介绍，力求活学活用。

本书可以作为高等学校互联网金融院系课程教材，也可供互联网金融研究者、从业者、管理人员参考所用。

图书在版编目(CIP)数据

大数据金融与征信/何平平，车云月编著. —北京：清华大学出版社，2017
(互联网金融系列丛书)
ISBN 978-7-302-48440-0

Ⅰ. ①大…　Ⅱ. ①何…　②车…　Ⅲ. ①金融—数据管理—研究　②互联网络—应用—金融—研究　Ⅳ. ①F830.49

中国版本图书馆 CIP 数据核字(2017)第 225246 号

责任编辑：杨作梅
封面设计：杨玉兰
责任校对：王明明
责任印制：沈　露
出版发行：清华大学出版社
网　　址：http://www.tup.com.cn, http://www.wqbook.com
地　　址：北京清华大学学研大厦 A 座　　邮　　编：100084
社 总 机：010-62770175　　邮　　购：010-62786544
投稿与读者服务：010-62776969, c-service@tup.tsinghua.edu.cn
质量反馈：010-62772015, zhiliang@tup.tsinghua.edu.cn
印 装 者：三河市铭诚印务有限公司
经　　销：全国新华书店
开　　本：185mm×260mm　　印　　张：18.25　　字　　数：440 千字
版　　次：2017 年 10 月第 1 版　　印　　次：2017 年 10 月第 1 次印刷
印　　数：1~3000
定　　价：42.00 元

产品编号：076698-01

前言

大数据金融是大数据在金融领域的重要应用。大数据金融市场前景广阔，预计未来 5 年到 10 年，金融大数据产业将迎来黄金增长期，大数据也将成为助推“大众创业、万众创新”浪潮的有力抓手。

本书为适应高等学校互联网金融专业人才培养的需要，从理论联系实际的原则出发，以大数据的实际运用为导向，对大数据在金融各行业的应用做了全面系统的介绍。

全书共分为 8 章，包括大数据金融概述、大数据相关技术、大数据在商业银行中的应用、大数据在证券行业中的应用、大数据在保险行业中的应用、大数据在互联网金融中的应用、大数据征信、大数据与中国金融信息安全。

由于大数据金融刚刚兴起，可供参考的资料不多，本书也仅仅是在这方面的一个探索，故全书整体框架以编者自己的思路进行呈现。本书以应用特别是金融领域前沿的应用为导向，以在各行业的实践为主线展开。本书内容新颖全面，论述问题极具现实意义。本书可以作为高等院校互联网金融专业相关课程的教材，也可供互联网金融研究者、从业者、管理人员参考。

全书主要有以下两大特点。

(1) 内容全面。

本书以大数据为出发点，结合国内外的发展现状及最新模式，系统地介绍了大数据在银行业、证券业、保险业、互联网金融行业及征信中的应用，并强调了在应用过程中，中国金融信息安全的重要性及保障机制。本书内容涵盖面极广，有效地为各行各业的读者提供了大数据金融与征信的宏观视图。

(2) 体例新颖。

本书秉承着注重实际运用的宗旨，编写体例上彰显了可读性和互动性。每章前有“本章目标”和“本章简介”，每章末有“本章总结”和“本章作业”。书中除了理论教学，还配有相关案例和解析，使理论与实践相结合，通俗易懂，开拓了学生的视野，可以更好地满足培养既懂专业知识又能运用所学知识解决实际问题的“复合型”经济人才需求。

本书由新迈尔(北京)特技有限公司组织研发，由何平平拟定大纲并进行统稿，湖南大学互联网金融研究所组织撰写。本书由何平平、车云月担任主编，以下研究生也参与了本书的编写：王杨毅彬、周春亚、张童、刘诗雨、刘晶宇。

本书编写过程中参考了大量的文献资料，有些已经在书后的参考文献中标注，而有些没有，在此一并表示感谢。囿于时间和个人能力，书中难免有疏漏和不妥之处，敬请读者批评指正。

何平平

目录

目录

目录

第1章

大数据金融概述

本章目标

- 掌握大数据的内涵与特征
- 了解大数据产生的背景
- 掌握大数据的类别
- 了解大数据的价值和应用领域
- 掌握大数据金融的内涵特点
- 掌握大数据金融相对于传统金融的优势
- 了解大数据给金融业带来的大变革
- 了解大数据给征信业带来的大变革
- 了解互联网大数据中的应用
- 掌握大数据金融的两种模式
- 了解大数据金融信息安全

本章简介

随着计算机技术和互联网的发展，大量的音频、图片、视频等结构化数据和半结构化数据不断涌现，传统的数据处理技术已经难以应对，因此大数据的概念应运而生。随着大数据技术的成熟，大数据已经广泛应用于商业、通信、医疗、金融等领域，给各行各业带来了巨大的价值。

近几年，大数据浪潮迅速席卷全球，数据成为企业重要的生产要素和战略资产，拥有大数据资产的企业将在竞争中占有优势。金融业本身就是基于数据与信息的产业，作为现代经济的核心，敏锐的金融行业正在积极拥抱大数据技术。大数据金融相对于传统金融有着无可比拟的优势，引起了金融行业广泛而深远的变革，包括银行业、保险业、证券业、征信业及互联网金融。

本章重点讲解大数据的内涵与特征、大数据的分类、大数据的处理流程以及大数据的价值和应用领域、大数据金融的内涵特点、大数据金融相对于传统金融的优势、大数据带来金融业和征信业大变革、互联网大数据的应用和大数据金融的两种模式。

1.1 大数据概述

在互联网中，大数据无处不在。无论是漫无目的的浏览网页、观看视频，还是发微博、聊微信，以及有目的性的搜索，基于每个用户都会产生数据，这些分散的数据汇集到网络中形成数据流，并最终聚集到网络服务提供商，形成大数据。

1.1.1 大数据的内涵与特征

1. 大数据与小数据

大数据(big data)是指在一定时间范围内无法用传统数据库软件进行采集、存储、管理和分析的数据集或数据群，需要通过新的处理模式才能体现出的具有高效率、高价值、海量、多样化特点的信息资产。利用数据挖掘分析技术可以使这些结构化、半结构化、非结构化的海量数据产生巨大的商业价值。小数据(small data)，或称个体资料，是以个体为中心，需要新的应用方式才能体现出的具有高价值、个体、高效率、个性化特点的信息资产。大数据和小数据有着本质的区别，虽然两者都是以创造数据价值为目的，但是在收集目的、数据结构、生命周期、分析方法及分析重点 5 个方面都存在着不同的定位。

1) 收集目的

小数据的目的性很强，往往是为了一个目标，制定规划进行收集、整理和分析，不会收集与其研究目的无关的数据。而大数据收集没有明确的目标，收集的数据范围更广，在数据采集阶段并不明确知道会产生什么结果。

2) 数据结构

小数据的数据基本来自相同的行业和领域，数据种类单一，结构统一，并采取一种有序排列的结构化方式。而大数据的数据来自不同的行业和领域，数据种类复杂，数据标准和格式有所不同，非结构化的数据居多，无法进行统一排序。

3) 生命周期

小数据的生命周期比较短，几乎只有几年的时间，待相关问题解决或相关项目结束之后，小数据一般会被删除。而大数据的工作主要是进行预测。只有基于完整的历史数据才能对未来进行相对准确的预测。因此，大数据的生命周期相对较长，大部分会被永久保留。

4) 分析方法

小数据采用一般的统计方法对收集的所有数据进行分析；而大数据因其复杂性一般通过分布式的方式进行分析，采用训练、学习、聚合、归一化、转化、可视化等多种不同的方法分析。

5) 分析重点

小数据是以个体行为数据为对象，主要是对个体数据信息进行全方位的精确的挖掘分析，重点在于深度；而大数据是以某个群体行为数据为对象，主要是对大范围大规模的数据处理分析，重点在于广度。

小数据不涉及大量的、急速的数据，或是繁多的信息种类，也没有隐含与大数据有关

的复杂化信息，并常以微观角度解释小型对象。而大数据则立于宏观角度，致力于表述宏观现象。简言之，用大数据得到规律，用小数据匹配个人。

2. 大数据的内涵

大数据的概念较为抽象。大数据中的“数据”是指广义的数据，不仅包括传统的结构化数据(即可以用二维表格表述的数据)，还包括非传统的非结构化数据(如视频、音频等)，大数据中的“大”既形容数据量多，也形容数据产生和变化的速度非常快。大数据的内涵主要体现在数据类型、技术方法和分析应用3个方面。

1) 数据类型方面

大数据不仅包括传统的结构化和半结构化的交易数据，还包括巨量的非结构化数据和交互数据，它是包括交易和交互数据集在内的所有数据集，如社交网站上的数据、在线金融交易数据、公司记录、气象监测数据、卫星数据和其他监控、研究和开发数据。

2) 技术方法方面

核心是从各种各样类型的数据中快速获取有价值信息的技术及其集成，依据大数据的生命周期的不同阶段可以将大数据处理技术分为大数据存储、大数据挖掘和大数据分析3个方面。大数据存储包括直接外挂存储(DAS)、网络附加存储(NAS)、存储域网络(SAN)等存储方式。大数据挖掘主要采用的是分布式挖掘和云计算技术。

3) 分析应用方面

重点是采用大数据技术对特定的数据集合进行分析，及时获得有价值的信息。常用数理统计方法进行数据分析，如可视化的数据分析工具。在数据分析过程中不仅需要计算机进行自动化的分析，还需要人工进行数据的选择和参数的设定。

3. 大数据的特征

大数据具有5个特征：大体量(Volume)、多样性(Variety)、时效性(Velocity)、准确性(Veracity)、价值性(Value)，如图1.1所示。

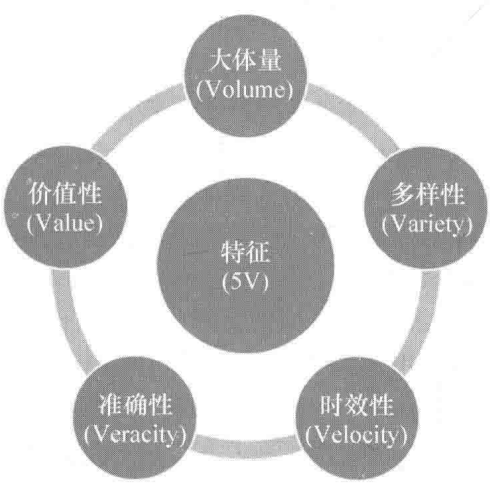

图1.1 大数据的特征

1) 大体量

大体量，即数据量大，是大数据的基本属性。大数据一般是指 10 TB(1 TB=1024 GB)规模以上的数据量，甚至可从数百 TB 到数十数百 PB、EB 的规模。资料显示，百度首页导航每天需要提供的数据超过 1.5PB(1PB=1024TB)。导致数据规模剧增的原因有：①传感器等各种仪器获取数据的能力大幅提高，越来越多的事物特征可以被感知，这些特征数据将会以数据的形式被存储下来。②互联网的普及，使数据的分享和获取越来越容易，无论是用户有意还是无意的分享或浏览网页都会产生大量数据。③集成电路价格的降低，使很多数据被保存下来。国际数据资讯(IDC)公司监测，全球数据量大约每两年翻一番，预计到 2020 年，全球将拥有约 35ZB 的数据量(见图 1.2)，并且 85%以上的数据以非结构化或半结构化的形式存在。

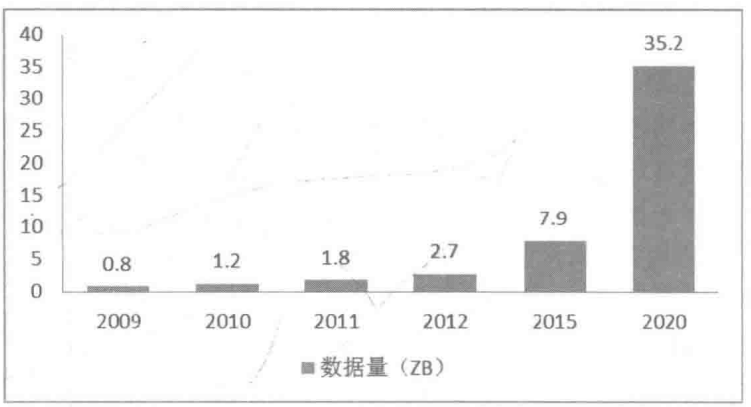

图 1.2 IDC 全球数据量使用情况及预测

2) 多样性

数据类型多样化是大数据的第二大特点。大数据包括各种格式和形态的数据。传统的数据大多是以二维表的形式存储在数据库中的文本类结构化数据。随着互联网的发展和传感器种类的增多，诸如网页、图片、音频、视频、微博类的未加工的半结构化和非结构化数据越来越多，以数量激增、类型繁多的非结构化数据为主。非结构化数据相对于结构化数据而言更加复杂，数据存储和处理的难度增大。目前，我国商业银行业务发展相关数据类型已从结构化数据扩展到非结构化数据。

3) 时效性

大数据的时效性是指在数据量特别大的情况下，能够在一定的时间和范围内得到及时处理，这是大数据区别于传统数据挖掘最显著的特征。大数据的流动速度快，当处理的数据从 PB 增加至 TB 时，超大规模的数据快速变化，使用传统的软件工具将难以处理。只有对大数据做到实时创建、实时存储、实时处理和实时分析，才能及时有效地获得高价值的信息。

4) 准确性

大数据的准确性是指保证处理的结果具有一定的准确性。结果的准确性涉及数据的可信度、偏差、噪声、异常等质量问题，原始数据的输入错误、缺失以及数据预处理系统的失效等会导致数据的不准确，进而分析得出一些错误的结论。因此，保证正确的数据格式

对大数据分析十分重要。

5) 价值性

大数据的价值性是指大数据包含很多深度的价值，对大数据的分析挖掘和利用将产生巨大的商业价值。数据量呈指数增长的同时，隐藏在海量数据中的有用信息却没有相应比例增长；相反，价值密度的高低常常与数据总量的大小成反比。这样反而使我们获取有用信息的难度加大。以商业银行监控视频为例，连续数小时的监控过程中可能有用的数据仅有几秒钟。

大数据的特征表明大数据不仅数据量巨大，种类繁多，对大数据的分析将更加复杂，更加追求速度，更注重时效性、准确性以及价值性。大数据不仅意味着数据总量的快速增长，其更大的意义在于：通过对大容量数据的交换、整合和分析，及时识别与发现新的知识，创造新的价值，带来“大知识”和“大发展”。作为一种重要的战略资产，大数据开启了一次全新的、重大的时代转型。

4. 大数据与传统数据的区别

大数据是以数量巨大、类型众多、结构复杂的数据集合以及基于云计算的数据处理和应用模式，通过数据的集成共享、交叉复用形成的智力资源和知识服务。大数据与传统数据在产生方式、存储方式、使用方式等方面都有所不同。

1) 产生方式

传统的数据是根据研究目的进行采集，采集的数据具有重要性。因为监管要求、业务逻辑或者技术便利，大数据具有“自产生”的特点，不需要特别的采集过程，比如搜索数据、交易数据等，尽管有些数据可能没有价值。

2) 存储方式

大数据的规模远远大于传统数据的规模。相对于传统数据库，量变引起质变，需要新的数据库技术来支持存储和访问。新型的大数据存储系统除了要具备高性能、高安全、高冗余等特征之外，还需具备虚拟化、模块化、弹性化、自动化等特征，才能满足具备大数据特征的应用需求。

3) 使用方式

传统数据是基于样本思维进行采集的，其分析方法主要是基于概率论理论和抽样理论。通常是通过这些样本数据推断总体，很难从这些数据中提炼出超出研究设计的知识。而大数据则是基于全体思维，所采集的数据基本能够代表整体，通过人工智能、神经网络等讲求高维和高效率的分析技术可以从这些详尽的数据中得出有价值的规律和知识。

5. 大数据的产生背景：计算机技术与互联网的发展

随着计算机的快速发展和互联网应用的成熟，数据量急剧增加，人类进入大数据时代。数据的采集、传输、存储、整合、管理、挖掘、分析等各项技术快速发展。

1) 计算机技术的发展

1946 年，第一台电子计算机的诞生开启了人类社会信息技术革命的序幕。截至目前，计算机技术的发展经历了大型主机、小型计算机、微型计算机、客户/ 服务器、互联网、云计算这六大阶段(见图 1.3)。

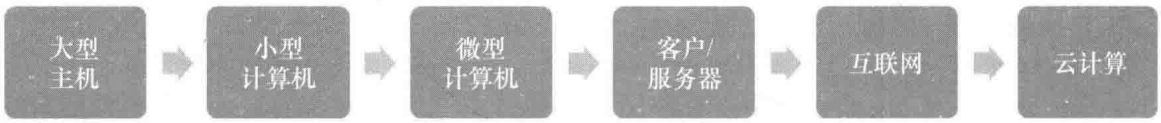

图 1.3　计算机技术经历的几个阶段

(1) 大型主机阶段(20 世纪 40—50 年代)。此阶段的计算机体型十分庞大，如第一台计算机由 18 800 个电子管组成，重量约 27 吨，占地约 150 平方米。在经历了电子管数字计算机、晶体管数字计算机、集成电路数字计算机和大规模集成电路数字计算机等发展历程后，计算机技术逐渐走向成熟。

(2) 小型计算机阶段(20 世纪 60—70 年代)。半导体和集成电路的改良使得大型主机经历了第一次缩小化，使用成本也因此降低，价格可被中小企业接受且能够满足中小企业的信息处理要求。现在很多企业使用的服务器都属于小型计算机，在体型上大于一般的个人计算机，小于大型主机。

(3) 微型计算机阶段(20 世纪 70—80 年代)。这个阶段是对小型计算机的缩小化，计算机已经缩小到可以放置在桌面上，因此被称为“微型计算机”或者“个人计算机”。1977 年美国苹果公司推出了 Apple 二代计算机，大获成功。1981 年 IBM 推出了 IBM – PC，经过不断的改良，功能不断加强，并占领了个人计算机市场，由此个人计算机得到了很大的普及。

(4) 客户机/服务器阶段。计算机的客户机/服务器结构起源于 20 世纪 60 年代，IBM 与美国公司建立了第一个全球联机订票系统，2000 多个订票终端被连在一起。在客户机/服务器结构中，网络的基础是客户机，核心是服务器，客户机通过服务器获得所需要的网络资源，其优点是能够充分发挥客户端的处理能力，减轻服务器的压力。

(5) 互联网阶段。1969 年，美国国防部研究计划署制定的协定将美国加利福尼亚大学洛杉矶分校、斯坦福大学研究学院、加利福尼亚大学和犹他州大学的 4 台主要的计算机连接起来，标志着计算机进入因特网阶段，即互联网阶段。此后，互联网经历了文本、图片、语音、视频阶段，带宽不断变快，功能越来越强大，这是人类迈向地球村坚实的一步。

(6) 云计算阶段。2008 年，“云计算”这个技术名词开始流行起来，它是一种基于互联网的计算方式，共享的软硬件资源和信息可以按照需求提供给计算机和其他设备。云计算阶段，计算机能力可以作为一种商品通过互联网进行流通。企业和个人不再需要购买昂贵的硬件，只需通过互联网来购买或者租赁计算能力，为所使用的计算功能付款。云计算囊括了开发、架构、负载平衡和商业模式等，是未来的软件业模式。

2) 互联网的发展

互联网不仅改变了传统的信息传播方式，也改变了人们的生活习惯。获取信息变得更加容易，足不出户便可了解世界新闻；沟通更加便捷，QQ、微信等网络工具将人们时刻联系在一起；购物消费更加容易，利用手机或电脑上网就可以快速实现商品交易。因此，互联网的发展不仅是一场信息革命，也是社会变革。根据第 38 次《中国互联网络发展状况统计报告》，截至 2016 年 6 月，中国网民规模达 7.10 亿人，其中手机网民规模达 6.56

亿人，占比 92.5%。网民行为因为互联网的发展更加多元化，文本、图片、音频、视频、地理位置等信息已经成为大数据增长最快的来源。

大数据与计算机技术和互联网的发展相辅相成。大体量的数据采集、存储、管理和挖掘因计算机和互联网技术的快速发展得以实现，数据的来源越来越丰富，形成信息流；大数据的信息流又通过社会生活和商业模式带动着资金流和物流的发展，进一步推动计算机与互联网技术的改进。大数据与计算机和互联网技术相互作用，相互促进，共同发展。

1.1.2 大数据的分类

大数据的种类很多，可以依照不同标准进行分类。

1. 按照大数据结构特征分类

按照大数据结构特征，可以将大数据分为结构化数据、非结构化数据和半结构化数据。

(1) 结构化数据。是指有结构的数据，也即行数据，在得到数据之前，其结构就是确定的。比如，传统的关系数据模型，可用二维结构表示。二维表中的数据就是典型的结构化数据，其结构事先通过数据模型的定义确定下来，在处理过程中不会改变。

(2) 非结构化数据。是指没有结构的数据，无法用数据库的二维逻辑结构来表现。包括所有格式的文档、文本、图片、视频、音频、各类报表以及标准通用标记语言下的子集 XML、HTML。它们通常没有数据模型，无法进行结构化处理。

(3) 半结构化数据。是指介于结构化数据和非结构化数据之间的数据。半结构化数据也是有结构的数据，与结构化数据不同的是，半结构化数据是先有数据，再有结构。半结构化数据一般是自描述的，数据的结构和内容混合在一起，没有明显的区分，其数据模型是数和图。常见的半结构化数据有 XML、HTML。

2. 按照大数据获取处理方式分类

按照大数据获取处理方式，可以将大数据分为批处理数据和流式计算数据。数据的批处理是指对数据进行批量的处理，如对数据进行成批的增加、修改、删除等操作。流式计算是指可以在实时处理的应用环境中，对大规模流动数据在不断变化的前提下进行持续计算、分析并能捕捉到有价值信息的分布式计算模式。流式数据具有实时性、易失性、突发性、无序性和无限性的特点。大数据的批处理和流式计算的区别如下表所示。

大数据批处理与流式计算的比较

性能指标	大数据流式计算	大数据批处理
计算方式	实时	批量
常驻空间	内存	硬盘
时效性	短	长
有序性	无	有
数据量	无限	有限
数据速率	突发	稳定

续表

性能指标	大数据流式计算	大数据批处理
是否可重现	难	易
数据精确度	较低	较高

3. 按照其他方式分类

按照大数据处理响应性能，可以将大数据分为实时数据、非实时数据和准实时数据；按照大数据关系，可以将大数据分为简单关系数据和复杂关系数据，如 Web 日志是简单关系数据，社会网络等具有复杂关系的图计算属于复杂关系数据。

1.1.3 大数据的价值

大数据最大的价值，是能够通过挖掘数据之间的相关性，把模糊的、隐含的、时滞性的问题，以可视化的、明确的、预演的方式展现出来，以便于决策和管理单元采取措施，改变所暴露的问题。这和传统的数据分析有着明显的不同。以往的数据分析或商业智能，更多的是面向过去已经发生的，而大数据是面向未来即将发生的。对金融行业来说，大数据主要有如下几点价值(见图 1.4)。

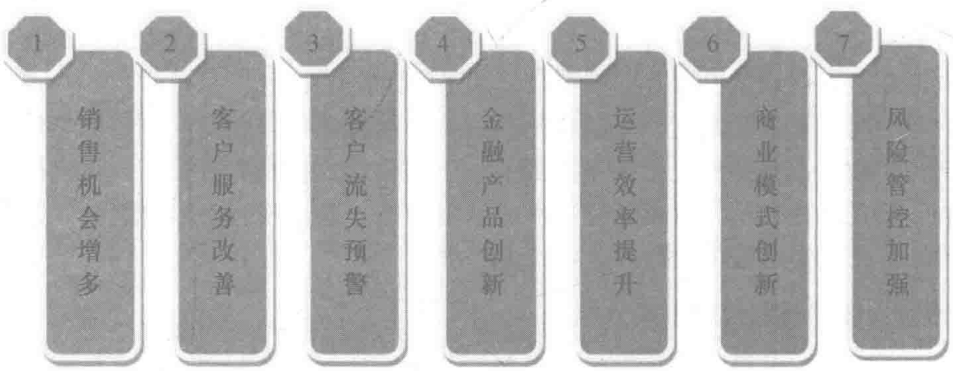

图 1.4 大数据在金融行业中的价值

1. 销售机会增多

金融企业掌握了海量的资金往来数据，再结合用户搜索行为、浏览行为、交易行为、评论历史、个人资料等数据，金融企业可以洞察消费者的整体需求，进而有针对性地进行产品生产、改进和营销。《纸牌屋》选择演员和剧情、百度基于用户喜好进行精准广告营销、阿里根据天猫用户特征包下生产线定制产品、亚马逊预测用户点击行为提前发货均是受益于互联网用户行为预测。

2. 客户服务改善

大数据的应用可以有效地改善客户服务。大数据不仅可以分析量化数据，还可以进行文本、语音分析。在客户体验方面，通过对交易数据、多渠道交互数据、社交媒体数据等的全面分析，帮助企业真正了解客户需求，并预测客户未来行为，从而为客户提供更好的

服务。在客户情感分析方面，通过对客服中心、社交媒体等数据的文本分析、语音分析，洞察客户情绪变化，分析客户的兴趣点、异常行为、意见、态度等，指导相关部门制定销售策略、市场策略等，并优化改进客户服务。

3. 客户流失预警

开发新客户往往比留住老客户要付出更高的成本。大数据技术的应用可以预警客户流失，减少客户流失率。利用大数据技术分析用户在整个相关产品里的使用行为的数据，识别可能流失的客户以及可能导致客户放弃的原因，如客户对产品不满意、对服务不满意、因为其他竞争对手等，以便企业及时采取策略，进行积极有效的改进。研究表明，客户在最终离开之前，很可能会持续关注或已经购买了竞争对手的产品，这些可以依据大数据进行探查。

4. 金融产品创新

大数据应用为金融行业突破传统金融产品带来了革新。高端数据分析系统和综合化数据分享平台能够有效地对接银行、保险、信托、基金等各类金融产品，使金融企业能够从其他领域借鉴并创造出新的金融产品。国内的数据挖掘最早基本也是基于授信所需要的分类挖掘算法而发展的。比如，金融贷款产品正在从抵押贷款向无抵押贷款演变，通过大数据应用建立信用评估机制，极大地提高了信用风险评级的及时性和准确性，抵押贷款模式正在逐步被信用贷款模式所取代。

5. 运营效率提升

在销售运营方面，金融机构能够通过现有客户的人际网络或业务网络，发现更多有价值的潜在客户，利用大数据的分析和预测模型，实现对客户消费模式和购买需求的分析，针对其个性需要展开精准营销，大大提升销售运营效率。在业务流程方面，通过大数据在存储和处理方面的优势，各种数据可被直接推送到需要这些信息的岗位，信息传递的中间环节被压缩，业务流程得到简化，从而带来巨大的效率提升空间。在资金需求预测方面，可以借助大数据构建资金需求预测模型，实现对资金需求的有效预算，帮助金融企业提高周转效率。

6. 商业模式创新

互联网金融和大数据技术正在对传统金融产生巨大冲击，大数据打破了信息不对称的局面，给金融商业模式带来了重大变化。一个很重要的表现形式是大数据的征信和网络贷款，可以根据企业行为数据计算出企业可能违约的概率，在这个基础上进行贷款，比如当前典型的阿里小贷。未来基于大数据的保险也是这样的，根据行为的数据进行保险差别的定价。比如，通过对人体的心率、体重、血脂、血糖、运动量、睡眠量等数据分析，预测客户的健康指数，帮助人身保险公司提高客户识别率，以此制定个性化的费率和承保方案。