

“十二五”国家重点图书出版规划项目

# 大数据技术与应用

丛书策划

上海大数据产业技术创新战略联盟(上海产业技术研究院)

上海市数据科学重点实验室(复旦大学)

丛书主编

朱扬勇 吴俊伟

# Big Data

Technology and Application Series

陈云  
主编

# 金融 大数据



上海科学技术出版社




大数据技术与应用

# 金融大数据

陈 云  
主编

上海科学技术出版社



**图书在版编目(CIP)数据**

金融大数据 / 陈云主编. —上海:上海科学技术出版社, 2015. 1(2015. 1 重印)  
(大数据技术与应用)  
ISBN 978-7-5478-2346-0

I. ①金... II. ①陈... III. ①金融—数据 IV.  
①F830. 41

中国版本图书馆 CIP 数据核字(2014)第 185843 号

**金融大数据**

陈 云 主 编

上海世纪出版股份有限公司 出版  
上海科学技术出版社  
(上海钦州南路 71 号 邮政编码 200235)

上海世纪出版股份有限公司发行中心发行  
200001 上海福建中路 193 号 www.ewen.co

苏州望电印刷有限公司印刷  
开本 787×1092 1/16 印张 14.75

字数: 330 千字

2015 年 1 月第 1 版 2015 年 1 月第 2 次印刷

ISBN 978-7-5478-2346-0/TP·31

定价: 60.00 元

# 内容提要

---

---

---

---

---

---

---

---

---

---

本书阐述了大数据技术在国内金融领域的发展和现状,从大数据视角加以思考和探索,为金融创新、金融大数据研究和应用提供有益支持。

全书共 15 章。第 1、2 章介绍金融大数据的应用现状、资源和技术等。第 3 至 14 章是本书重点,分别介绍了大数据技术在证券期货、银行、保险和互联网金融中的应用。第 3 至 5 章为证券期货应用,典型应用包括自动化交易策略设计、金融期货风险监管、证券期货舆情分析等。第 6 至 8 章为银行业应用,典型应用包括银行客户关系管理、银行风险管理、小微企业信贷等。第 9 至 11 章为保险业应用,典型应用包括保险精准营销、医疗保险业务优化、保险欺诈识别等。第 12 至 14 章为互联网金融应用,典型应用包括第三方支付、网络融资、网络基金等。第 15 章对金融业在大数据时代的应对战略、金融信息复合型人才培养提出了建议。

本书可供金融信息专业、计算机类数据科学等专业的高校师生使用,也可供金融行业从业人员、数据科学研究人员及政府决策和管理人员参考。

大数据技术与应用

## 学术顾问



中国工程院院士 邬江兴  
中国科学院院士 梅 宏  
中国科学院院士 金 力  
教授, 博士生导师 温孚江  
教授, 博士生导师 王晓阳  
教授, 博士生导师 管海兵  
教授, 博士生导师 顾君忠  
教授, 博士生导师 乐嘉锦  
研究员 史一兵

大数据技术与应用  
编撰委员会

---

**主 任**

朱扬勇 吴俊伟

**委 员**

(以姓氏笔画为序)

于广军 朱扬勇 刘振宇 孙景乐 李光亚 李光耀 杨 丽  
杨佳泓 吴俊伟 何 承 张鹏翥 陈 云 武 星 黄林鹏  
童维勤 蔡立志

# 本书编委会

---

## 主 编

上海财经大学 陈 云

## 编 委

(以姓氏笔画为序)

上海金融期货信息技术有限公司 史光伟

上海吉贝克信息技术有限公司 刘世平

上海财经大学 李艳红

上海财经大学 张 超

上海财经大学 郝晓玲

上海财经大学 钟 明

上海财经大学 俞 立

上海财经大学 姜 广

华院数据技术(上海)有限公司 宣晓华

上海财经大学 宫 剑

上海财经大学 徐晓萍

上海财经大学 曹志广

浙商期货有限公司 程 余

上海财经大学 谢 斐

# 前 言

---

---

---

---

---

---

---

---

---

---

随着信息技术的飞速发展,特别是云计算、大数据技术的应用,未来金融机构的核心竞争力很大程度上依赖于将数据转化为信息和知识的速度与能力,而这种转化速度和能力,取决于数据分析、挖掘和应用水平。随着移动支付、互联网金融等新兴金融业态的不断涌现,以“用户为中心”、强化用户体验的服务模式将成为未来金融业的重要发展方向。应用大数据技术可多方面掌握客户及业务的有效信息,全面综合分析客户的资产负债、流动性及客户行为,有助于金融机构进行产品创新、精准营销和风险管理,实现数据资产向战略资产和市场竞争力的转化。

本书力求把握金融大数据应用的最新动向,结合大数据技术在国内金融领域的发展和现状,以典型应用为支撑,从大数据视角进行分析,为金融创新与金融大数据研究和应用提供有益支持。

全书共 15 章。第 1、2 章介绍金融大数据概况,包括金融大数据应用现状、金融大数据资源、金融大数据处理技术等。第 3 至 5 章为证券期货应用,包括自动化交易策略设计、金融期货风险监管、证券期货舆情分析等。第 6 至 8 章为银行业应用,包括银行客户关系管理、银行风险管理、小微企业信贷等。第 9 至 11 章为保险业应用,包括保险精准营销、医疗保险业务优化、保险欺诈识别等。第 12 至 14 章为互联网金融应用,包括第三方支付、网络融资、网络基金等。其中第 3 至 14 章是本书的重点,分别介绍了大数据技术在证券期货、银行、保险、新兴互联网金融中的应用。第 15 章对金融业在大数据时代的应对战略、金融信息复合型人才培养提出了建议。

我国金融大数据应用刚刚起步,书中有些典型应用的数据规模可能不够大、采用的方法不够新,但新兴技术和产业的发展是“用”出来的,只有广泛应用才能发现需要突破的技术壁垒。遴选金融大数据典型应用,宣传和推广相关应用经验,为金融大数据提供



必要的发展基础是本书编写的重要目的。

本书的特色在于注重大数据技术与金融业务的结合,通过充分调研,并与多家业内有影响力的金融机构及数据公司合作研究,分析利用了大量第一手资料。书中内容着力于金融学与计算机科学的学科交叉,填补相关领域的资料空白,面向的读者包括金融从业人员、数据科学研究人员、金融信息专业学生等,力求通过抛砖引玉,引起业界人士和相关学者对金融大数据应用的重视,推动行业的发展。

本书由上海市金融信息技术研究重点实验室和上海国际金融中心研究院组织编撰,多位作者参与本书编写,其中张超编写了第3至5章证券期货业大数据应用,郝晓玲编写了第6至8章银行业大数据应用,李艳红和姜广编写了第9至11章保险业大数据应用,俞立编写了第12至14章新兴互联网金融大数据应用。徐晓萍、钟明、曹志广、谢斐在纲要和内容方面对本书提出了修改意见。另外,感谢合作单位提供了相关资料,对本书编写给予了大力支持,包括上海金融期货信息技术有限公司、友邦保险公司、上海大智慧股份有限公司、华院数据技术有限公司、平安财产保险公司上海分公司、吉贝克信息技术有限公司、交通银行、浙商期货等。

作者

# 目 录



<b>第1章 绪论</b>	<b>1</b>
• 1.1 大数据带来的变革	2
• 1.2 大数据在金融中的应用现状	3
1.2.1 证券期货业应用现状	3
1.2.2 银行业应用现状	6
1.2.3 保险业应用现状	8
1.2.4 跨行业互联网金融应用现状	10
• 1.3 大数据时代金融业的机遇和挑战	13
参考文献	16
<b>第2章 金融大数据资源</b>	<b>17</b>
• 2.1 国内的大数据资源	18
• 2.2 大数据处理流程	19
• 2.3 金融大数据资源及处理技术	21
2.3.1 金融大数据资源	21
2.3.2 金融大数据处理关键技术	23
参考文献	24

<b>第3章 大数据与自动化交易</b>	25
• 3.1 自动化交易在国内外的的发展	26
3.1.1 自动化交易的类型	26
3.1.2 自动化交易的国内外发展现状	29
• 3.2 典型自动化交易策略	30
3.2.1 期现套利策略	30
3.2.2 价差套利策略	33
3.2.3 流动性策略	35
• 3.3 基于数据挖掘的自动化交易策略	36
3.3.1 基于模式识别的策略研究	36
3.3.2 基于遗传算法的交易策略	37
参考文献	41
<b>第4章 大数据与金融期货市场风险监控</b>	43
• 4.1 金融期货及其特征	44
4.1.1 金融期货及其发展历程	44
4.1.2 金融期货种类和特点	45
4.1.3 金融期货在我国的发展	47
4.1.4 金融期货的风险	49
• 4.2 金融市场风险监控面临的挑战	49
• 4.3 基于大数据的金融期货市场风险监控	51
4.3.1 金融期货市场风险监控系统	51
4.3.2 大数据技术的应用	53
4.3.3 风险监控系統建设的意义	59
参考文献	60
<b>第5章 大数据与证券网络舆情</b>	61
• 5.1 网络舆情及其特征	62
• 5.2 网络舆情与资本市场	64

5.2.1	资本市场舆情管理	64
5.2.2	基于网络舆情的股票市场行情预测	67
• 5.3	基于大数据的投资者情绪分析	70
5.3.1	网页抓取技术	70
5.3.2	信息预处理技术	71
5.3.3	特征挖掘技术	72
5.3.4	情感极性分类技术	74
	参考文献	77
<b>第6章</b>	<b>大数据与银行客户关系管理</b>	<b>79</b>
• 6.1	银行客户关系管理	80
• 6.2	数据挖掘在客户关系管理中的应用	81
• 6.3	基于决策树方法的客户贷款风险预测	83
• 6.4	大数据在银行客户关系管理领域的应用案例	85
6.4.1	中信银行信用卡业务创新	85
6.4.2	交通银行呼叫中心精准营销案例	86
	参考文献	93
<b>第7章</b>	<b>大数据与银行风险管理</b>	<b>95</b>
• 7.1	银行风险管理体系	96
• 7.2	基于大数据的银行风险管理模式	99
• 7.3	基于大数据的银行风险管理案例	102
7.3.1	基于大数据处理的交易欺诈侦测	102
7.3.2	基于数据仓库的银行风险监管系统	104
7.3.3	银行开展全面风险管理的对策建议	109
	参考文献	110
<b>第8章</b>	<b>大数据与小微企业信贷</b>	<b>111</b>
• 8.1	小微企业信贷及其风险	112
• 8.2	基于大数据的小微企业信贷模式创新	113
• 8.3	基于数据挖掘的小微企业信用风险评估	118

8.3.1	算法设计	118
8.3.2	实验	119
<hr/>		
• 8.4	基于大数据的小微企业贷款案例	121
8.4.1	互联网金融公司 ZestFinance 基于大数据的信用评估体系	121
8.4.2	民生银行小微企业贷款实践	122
8.4.3	银行开展小微企业信贷的建议	124
	参考文献	125
<hr/>		
<b>第9章</b>	<b>大数据与保险精准营销</b>	<b>127</b>
<hr/>		
• 9.1	基于大数据的保险精准营销	128
9.1.1	大数据在保险营销中的应用	128
9.1.2	保险精准营销整体流程及系统架构	129
<hr/>		
• 9.2	保险精准营销核心技术	130
• 9.3	基于大数据的保险精准营销应用拓展	136
9.3.1	集成企业内部结构化数据	136
9.3.2	集成企业内部非结构化数据	137
	参考文献	139
<hr/>		
<b>第10章</b>	<b>大数据与医疗保险业务优化</b>	<b>141</b>
<hr/>		
• 10.1	医疗保险及其理赔流程	142
• 10.2	基于大数据的医疗保险产品定价优化	144
10.2.1	保险产品定价存在的问题	144
10.2.2	数据准备	144
10.2.3	聚类模型构建与评估	146
<hr/>		
• 10.3	基于大数据的医疗保险核保优化	147
10.3.1	医疗保险核保存在的问题	147
10.3.2	数据准备	148
10.3.3	客户风险级别判别模型的建立	148
	参考文献	149

<b>第11章 大数据与保险欺诈识别</b>	151
• 11.1 保险欺诈识别概述	152
• 11.2 行业车险信息平台及其反欺诈应用	153
11.2.1 车载数据的采集与共享	153
11.2.2 我国行业车险信息集中平台	154
11.2.3 行业车险信息平台在反欺诈中的应用	156
• 11.3 基于大数据的保险欺诈侦测方法	157
参考文献	161
<b>第12章 大数据与第三方支付</b>	163
• 12.1 第三方支付概述	164
12.1.1 第三方支付模式	164
12.1.2 第三方支付运营特点	169
• 12.2 第三方支付中的风险	170
12.2.1 业务风险分析	170
12.2.2 系统风险分析	171
• 12.3 基于大数据的第三方支付欺诈发现	172
12.3.1 第三方支付与网络欺诈	172
12.3.2 数据挖掘与欺诈发现	173
参考文献	177
<b>第13章 大数据与网络融资</b>	179
• 13.1 网络融资概述	180
• 13.2 网络借贷	181
13.2.1 P2P 模式	181
13.2.2 众筹模式	186
• 13.3 网络基金	189
• 13.4 基于大数据的 P2P 个性化推荐	190

13.4.1	P2P 网站中的个性化推荐	190
13.4.2	推荐系统	191
13.4.3	基于 VITA 系统的信贷产品匹配机制	193
	参考文献	195
<b>第14章 大数据与供应链融资</b>		197
• 14.1	供应链融资概述	198
14.1.1	供应链融资及其特征	198
14.1.2	供应链融资的意义	199
14.1.3	国内外供应链融资的发展	200
14.1.4	供应链融资中的风险	201
• 14.2	供应链融资产品与模式	203
• 14.3	基于大数据的企业信用评估	205
14.3.1	供应链中的企业信用问题	205
14.3.2	基于人工智能的信用评分模型	206
14.3.3	基于 PSO-BP 集成的企业信用评分	207
	参考文献	210
<b>第15章 展望</b>		213
• 15.1	大数据时代金融机构的战略	214
• 15.2	大数据时代的金融人才培养	215

# 第1章

A decorative horizontal line consisting of multiple thin parallel lines, with a slightly thicker vertical line intersecting it in the center.

## 绪论



随着信息技术的飞速发展,金融机构逐步实现数据集中、业务电子化和数字化。近年来,大数据技术在金融中的应用,进一步推进了金融服务模式创新、管理创新和产品创新,金融大数据时代已来临。大数据技术应用为金融创新提供了有效支持,为金融服务提供了准确的信息和知识,为消费者提供了个性化服务。金融大数据应用得到众多业界人士和学者的广泛关注。本章对金融大数据应用现状,以及金融业面临的机遇和挑战进行简述。

## 1.1 大数据带来的变革

近年,随着互联网应用爆发式增长,数据产生的速度和规模急剧提升。根据国际数据公司 IDC(Internet Data Center)的监测统计,2011 年全球数据总量已经达到 1.8 ZB(Zettabyte),且正以每两年增加一倍的速度快速增长,预计到 2020 年全球数据总量将达到 40 ZB,届时中国产生的数据量将占到全球总量的 21%。数据规模急剧增长的同时,数据类型也日益复杂,其中采用传统数据处理手段难以处理的非结构化数据已接近数据总量的 75%。随着互联网的快速发展,音频、文字、图片、视频等半结构化、非结构化数据大量涌现,而社交网络、物联网、云计算的广泛应用,使个人可以更加准确快捷地发布和获取数据。在科学研究、互联网应用、电子商务等诸多应用领域,数据规模、数据种类正在以极快的速度增长,大数据时代已悄然降临<sup>[1]</sup>。

迅速增长、庞大繁杂的数据资源给传统的数据分析、处理技术带来了巨大的挑战。为了应对这样的新任务,大数据技术、大数据工程、大数据科学和大数据应用等迅速成为信息科学领域的热点问题,得到了政府部门、经济以及科学等各个领域有关专家的广泛关注(表 1-1)。

表 1-1 对大数据的战略引导

国家/组织	时 间	计 划
美国	2012 年 3 月	《大数据研究和发展计划》大力推动大数据相关的收集、储存、保留、管理、分析和共享海量数据技术研究
中国	2012 年 3 月	《“十二五”国家科技计划信息技术领域 2013 年度备选项目征集指南》,明确提出“面向大数据的先进存储结构及关键技术”
日本	2012 年 7 月	《新 ICT 战略研究计划》重点关注“大数据应用”,提升国家竞争力
联合国	2012 年 7 月	《大数据政务白皮书》总结各国政府如何利用大数据提供更好的服务
英国	2013 年 1 月	注资 6 亿英镑发展 8 类高新技术,大数据独占 1.89 亿英镑