

大数据时代的金融

金融管理系统数据挖掘的
研究与效用

王雨霖 / 著



 復旦大學 出版社

大数据时代的金融

金融管理系统数据挖掘的研究与效用

王雨霖 / 著

 復旦大學 出版社

图书在版编目(CIP)数据

大数据时代的金融:金融管理系统数据挖掘的研究与效用/王雨霖著.
—上海:复旦大学出版社,2016.8
ISBN 978-7-309-12312-8

I. 大… II. 王… III. 金融-数据管理系统-研究 IV. F830.49

中国版本图书馆CIP数据核字(2016)第104803号

大数据时代的金融:金融管理系统数据挖掘的研究与效用
王雨霖 著
责任编辑/方毅超

复旦大学出版社有限公司出版发行
上海市国权路579号 邮编:200433
网址:fupnet@fudanpress.com http://www.fudanpress.com
门市零售:86-21-65642857 团体订购:86-21-65118853
外埠邮购:86-21-65109143
江苏凤凰数码印务有限公司

开本 890 × 1240 1/32 印张 8.625 字数 205 千
2016年8月第1版第1次印刷

ISBN 978-7-309-12312-8/F · 2269
定价:36.00元

如有印装质量问题,请向复旦大学出版社有限公司发行部调换。
版权所有 侵权必究

前 言

一、本书研究的科学意义

金融自从诞生以来,其发展历史就同数据管理与应用密不可分。从最早的纸质账簿到当今利用先进的信息系统完成交易过程登记、数据加工分析、风险计量、经营指标预测等工作,都是建立在数据的基础之上。作为全球较早使用信息化技术手段实现经营管理的行业,金融管理系统无论是从应用规模、还是应用水平上讲,在各行业中均处于领先地位。从国内情况看,特别是近几年以来我国金融从发展规模、市场化进程等方面发生了天翻地覆的变化,同时也正在加快金融管理系统信息化、数字化的发展步伐。在金融全球一体化、信息化的大背景下,综合分析国内外金融管理系统的未来发展趋势,概括起来有如下几个特点:

一是以客户为中心。金融管理系统对市场信息和客户信息要求充分掌握,深入挖掘分析客户需求与行为,通过为客户提供更好的产品与服务来维系良好的客户关系,提高客户满意度与忠诚度,从而提高客户的利润贡献度。

二是风险管理的量化与技术化。金融管理系统经过多年风险理论研究与管理实践,风险管理体系日趋完善,紧密围绕经济资本计量,以风险调整后收益率(RAROC)为核心指标。根据金融数据库中历年积累的业务信息资料,利用先进的信息技术手段进行风险分池,

精确计算违约率、违约下的损失率、违约暴露。

三是电子业务发展迅速。伴随着移动通信技术、互联网技术的飞速发展,客户交易行为偏好发生明显变化,更倾向于利用快捷、方便的电子渠道完成金融业务。此外,外汇买卖、贵金属买卖等即时、高频交易类中间业务需要先进的技术平台,对数据进行准确高效的加工处理。对于商业银行来说,由于网上银行、自助银行(ATM)、手机银行等电子渠道的交易成本仅是银行物理网点交易成本的十分之一,从节约成本的角度,更加有动力发展电子银行业务。

四是金融创新科学化。创新是一个企业实现可持续发展的源动力,但是只有有效的创新才能真正创造价值。如果像2008年金融危机导火索——金融衍生品、资产证券化过渡这类创新无序蔓延,所产生的危害是巨大的,危机会迅速从美国本土波及全球,再一次验证了“蝴蝶效应”,而且形成“W”型的持续周期。因此只有建立在数据分析基础上的理性且有价值的创新,科学利用金融工程理论工具的创新,才是金融机构实现差异化竞争的武器,成为金融发展的关键驱动因素。

总之,无论是在现在还是在未来,金融管理系统在客户服务以及经营管理中产生海量数据,如何有效管理好这些数据,利用它们提升客户服务水平,提高市场竞争能力,管理金融风险,服务金融创新,让数据创造更多的效用,是目前各国金融管理系统关注的重点问题。

二、学术价值及应用前景

世界著名统计管理学和质量管理学专家戴明经典言论“除非你是上帝,否则请用数据说话”,可谓家喻户晓。研究金融管理系统数据挖掘的运用及效用具有重要的理论及实践意义。当今信息化时代,网络经济与虚拟经济结合使得以银行为首的现代金融业经营管理模式发生重大变革,数据信息量增长迅猛,其商业价值备受关注,

作为信息载体的数据,已演变为金融机构一项特殊的无形资产。如何将数据效用融入传统效用理论,挖掘数据与金融价值创造的相关性,是对现有金融价值相关理论的重要补充和完善,具有重要的理论意义。

金融管理系统的数字效用体现在两个方面:一是未经加工处理的原始数据所体现的效用很小,数据经过整合、分类、清洗等加工处理后就演变成信息商品,对数据的进一步深入挖掘分析更是人类智慧的完美体现,产生更大的边际效用,因此数据信息具有微观经济学意义上的使用价值,即效用;二是有效的数据在金融管理系统内部各个业务领域应用广泛,将数据效用应用到金融的各项价值增值活动后,可以实现金融价值的增值。

而对于现代金融的运营管理实践而言,数据的重要性更是毋庸置疑。一方面,通过对相关数据的解读,分析金融资产结构、收入结构、客户结构、财务状况等,可以加强金融内部管理;另一方面,利用数据等相关信息资产,运用数据挖掘等相关数学工具,获取各种数据效用,金融管理系统有利于在产品研发、风险管控、优化管理流程、经营决策支持等活动中进行前瞻性分析。由此,强化金融管理系统数据效用分析具有重要的实践意义,尤其对中国金融管理系统来说,其意义更为明显。

三、本书研究的重点和难点

1. 金融管理系统数据挖掘的设计

在完成了系统需求分析和系统建设的相关技术的分析后,确定系统设计的可行性的基础上,进一步对系统进行相关的设计,本书主要从系统架构设计、系统功能模块设计、系统功能模块划分原则以及系统数据库和系统物理结构的设计上完成整个金融管理系统的详细设计。

2. 基于 Apriori 算法的数据分析

本书采用的数据挖掘技术是使用关联规则挖掘进行设计实现的,主要采用 Apriori 算法对金融管理系统数据进行分析,结合本书设计的金融管理系统的特点以及金融系统的数据特性,基于 Apriori 算法设计和实现对金融管理系统的相关的数据挖掘和分析。

四、本书的研究方法

本书的研究对象是金融管理系统的数据库,全书基于效用相关理论,运用数据挖掘方法,研究金融管理系统数据库效用,深入挖掘数据库及其管理对金融经营管理和业务发展的影响因素,理性评估数据库效用对金融管理价值的贡献,基于效用相关理论,运用数据挖掘方法,研究金融管理系统数据库效用。跨学科的研究过程中涉及经济学、管理学、信息技术科学等多个交叉学科。全书的主要研究方法包括:

(1) 运用定量分析与定性分析相结合的方法。定量分析方法通过量化影响金融管理系统数据库效用的各项因素,计量数据库效用对金融价值的影响度,对客观现象进行量化分析。定性分析对数据库效用的本质进行深入分析,发现其内在的客观规律。

(2) 运用经济理论与金融理论同金融管理系统实践相结合的方法,研究金融系统通过数据整合与挖掘创造价值的方法与途径。

(3) 采用管理学案例分析的方法。对本书中使用的案例展开分析,利用某银行企业级数据库中的真实数据,对比分析提高数据库有效性以及运用数据挖掘技术手段实现数据库价值增值的过程。

五、本书创新点

(1) 系统地将经济理论与金融理论同金融管理系统实践相结合

合,深入研究金融管理系统通过数据挖掘与运用创造价值的方法与途径。将金融系统管理、数据治理与信息技术应用有机整合,弥补国内这一领域研究的不足之处。

(2) 创造性地将数据作为金融管理系统专有重要战略资源,综合运用经济学领域、信息技术领域的效用理论与数据挖掘分析理论,结合金融管理系统理论与实践分析其在金融客户管理,风险管理,金融创新,构成竞争优势方面的价值贡献。

(3) 完成了金融管理系统的设计和实现,并且基于改进的Apriori算法实现了金融管理系统客户信贷数据的挖掘和数据分析,有效地提高了金融管理系统风险预测能力以及管理效率。

(4) 依据价值理论评估客户价值和客户关系价值,充分利用某全国性股份制银行企业级数据仓库中的数据,通过数据挖掘分析,合理对客户进行分类,深入分析客户资产情况与客户交易行为,探究客户数据质量方面存在的问题,提出数据质量改进建议,科学地评估金融管理系统在客户管理方面的成本与收益,通过技术手段实现金额管理系统客户价值最大化。

(5) 充分利用金融管理系统庞大的数据资产,基于数据挖掘分析的相关技术,研究金融管理系统通过产品、服务及流程方面的金融创新,创造金融价值和提升核心竞争力的途径。

目 录

第 1 章	绪论	1
1.1	开发背景	3
1.2	金融管理系统研究现状	4
1.3	本书主要工作及组织结构	6
第 2 章	数据挖掘理论	9
2.1	数据挖掘的定义和方法	11
2.2	数据挖掘的商业需求分析	14
2.3	数据挖掘研究和应用	16
2.4	关联规则挖掘	21
2.5	时间序列数据挖掘	27
2.6	动态数据挖掘	30
2.7	软件危机	33
2.8	本章小结	37
第 3 章	金融风险原理	39
3.1	金融风险的几个典型案例	42
3.2	金融风险的基本概念	56
3.3	金融风险的分类	57
3.4	金融风险管理	61

3.5	市场风险管理的 VaR 方法	66
3.6	VaR 的应用	72
3.7	信用风险度量方法	80
3.8	信用风险管理和监管	85
3.9	金融风险管理系统	88
3.10	本章小结	97
第 4 章	数据挖掘理论及应用的相关文献	99
4.1	数据挖掘理论与应用文献综述	101
4.2	效用、数据效用相关文献综述	108
4.3	现有研究不足	115
4.4	本章小结	117
第 5 章	金融管理系统的需求分析	119
5.1	系统可行性分析	121
5.2	系统需求分析	122
5.3	系统非功能需求	125
5.4	其他技术指标要求	131
5.5	相关技术与工具	132
5.6	本章小结	134
第 6 章	金融管理系统的设计	135
6.1	系统架构设计	137
6.2	系统具体功能设计	138
6.3	数据库设计	146
6.4	基于 Apriori 算法的数据分析	150
6.5	本章小结	156

第7章	金融管理系统的实现	159
7.1	金融管理系统实现环境	161
7.2	系统功能模块的实现	163
7.3	本章小结	172
第8章	金融管理系统的测试	173
8.1	测试环境	175
8.2	测试方法	175
8.3	测试用例	177
8.4	测试分析	178
8.5	测试结论	181
8.6	本章小结	182
第9章	数据挖掘在金融管理系统的效用研究	183
9.1	金融全面风险管理对数据的要求	185
9.2	数据挖掘在金融信用风险管理中的应用	197
9.3	数据挖掘在金融市场风险管理中的应用	207
9.4	数据挖掘在金融操作风险管理中的应用	212
9.5	金融管理系统的应用	221
9.6	本章小结	234
第10章	结论与展望	237
10.1	结论	239
10.2	展望	241
参考文献		247
致谢		263

图 目 录

图 2-1	数据挖掘是交叉科学	11
图 2-2	项集格先验原理的示意图	22
图 2-3	基于支持度的剪枝示意图	23
图 3-1	金融风险管理原理	41
图 3-2	2003 年 1 月至 2003 年 12 月欧元兑美元的汇率走势 ...	43
图 3-3	LME 3 月期铜价格走势	45
图 3-4	一年多来 LME 3 月、15 月、27 月远期价格	46
图 3-5	上海期铜 05120512 合约价格走势	46
图 3-6	0602 合约 77-1212	47
图 3-7	LME 库存 05.01.31-05.12.19	47
图 3-8	LME 库存 05.07-06.06	48
图 3-9	VaR 损失	66
图 3-10	VaR 在风险管理中的作用	68
图 3-11	IMA 方法资本保证金设置过程示意图	75
图 3-12	4 月 6 日的总资金风险状况	78
图 3-13	182.063 万元的最坏损失中的股票	79
图 4-1	信息价值链模型	114
图 6-1	系统整体架构设计图	137
图 6-2	系统功能结构图	139
图 6-3	安全管理模块功能结构图	140

图 6-4	系统用户角色分配及结构图	141
图 6-5	金融信息管理模块功能设计结构图	142
图 6-6	金融信息管理模块业务控制流程图	143
图 6-7	金融信用管理模块功能结构图	144
图 6-8	金融信用管理模块业务控制流程原理图	145
图 6-9	备份与日志管理模块功能设计结构图	145
图 6-10	金融管理系统主要实体概念模型	148
图 6-11	金融管理系统数据库物理结构模型	149
图 6-12	基于 Apriori 算法的金融信息数据挖掘模型	154
图 6-13	相同支持度不同事务数时三种算法执行时间对比分析	155
图 6-14	事务数固定不同最小支持度时三种算法执行时间	156
图 7-1	金融信息管理系统运行效果图	163
图 7-2	金融信息管理系统客户信息评估结果运行效果图	164
图 7-3	金融信用管理系统运行效果图	165
图 7-4	系统登录运行效果图	165
图 7-5	系统用户注册页面	166
图 7-6	系统管理员主页面	167
图 7-7	系统权限查询设置页面	167
图 7-8	权限详细设置实现运行效果图	168
图 7-9	基于数据挖掘的系统数据分析主界面运行效果图	169
图 7-10	客户数据预处理	169
图 7-11	采用 Apriori 算法的关联规则挖掘结果	170
图 9-1	全面风险管理信息系统目标架构	194
图 9-2	源数据库与数据挖掘库关系	198
图 9-3	市场风险系统接口架构	211
图 9-4	潜在欺诈交易流程图	221

表 目 录

表 2-1	数据集	24
表 2-2	频繁项集的支持度	24
表 3-1	一些重大金融风险事件给企业带来的损失统计	55
表 3-2	三大类主要的计算方法比较	69
表 3-3	VaR 数额与保证金系数	74
表 3-4	信用等级转换矩阵	82
表 3-5	债券利率期限结构	82
表 3-6	期末债券估价表	83
表 3-7	四种模型的比较	85
表 3-8	标准法风险计量	88
表 8-1	Bug 的状态定义	178
表 8-2	Bug 产生的原因及应对措施	180
表 9-1	源系统交换时间表	195
表 9-2	数据有效性对资本充足率影响分析	203
表 9-3	不同类型银行当前信息化建设工作	218
表 10-1	未来银行的发展趋势比较	243
表 10-2	动态数据仓库技术特点	245

第1章 绪论



1.1 开发背景

金融系统作为经济系统的重要组成部分,也是最核心的部分,是通过经济系统中各种相互关联的金融元素组成。金融管理系统是经济学领域中的一个新的概念,其提出主要是从信息管理角度出发,用于对金融相关的信息进行管理和分析。通常我们所说的金融管理系统的基本功能就是实现对金融系统相关联的信息进行管理,并实现金融资源的合理分配和利用,将复杂的金融系统相关的信息进行系统化、科学化的管理,进而为金融管理和金融决策者提供全面的金融信息处理和析的信息系统。通常我们构建的金融管理系统都是以管理功能为基础、金融信息数据分析处理为核心、金融综合业务数字化和网络化为手段,来进行构建的信息管理系统,通过金融管理系统,能够有效地提升管理效率和为决策者提供决策支持,对现代电子金融的发展具有非常重要的意义。

随着信息化技术的快速发展和应用,特别是在金融管理领域,目前对于构建金融信息应用系统的相关技术已经发展的非常成熟,并且各大金融机构,目前都在逐步的采用现代化的办公手段和业务管理手段替代传统的办公和金融业务管理方式,分布于银行的综合业务系统、信贷系统和个人征信系统等业务系统都已经建设得较为成熟,使整个银行业务化的范围大大突破了传统的地域和网点限制,金融客户也有了更大的选择,传统的以银行为主的金融单位的垄断能力也越来越弱。这些都进一步加大了各种金融机构之间的竞争,进而导致了金融业务的发展也由以内部核算为主的业务流程方式,转向了以金融客户为中心的经营服务。银行相关的服务产品以及相关