



经管类专业虚拟仿真实验系列教材

大数据分析 与应用

基于IBM客户预测性智能平台

DASHUJU FENXI
YU YINGYONG

JIYU IBM KEHU
YUCEXING ZHINENG PINGTAI

蹇洁 主编

实验实训



Southwestern University of Finance & Economics Press
西南财经大学出版社



经管类专业虚拟仿真实验系列教材

大数据分析 与应用

基于IBM客户预测性智能平台

DASHUJU FENXI YU YINGYONG
JIYU IBM KEHU YUCEXING ZHINENG PINGTAI

蹇洁 主编

参编人员

廖显 余海燕 陈思祁 武建军
余若雪 杨梦丽 张英培 牛舒 张燕双



Southwestern University of Finance & Economics Press
西南财经大学出版社

图书在版编目(CIP)数据

大数据分析与应用:基于 IBM 客户预测性智能平台/蹇洁主编. —成都:西南财经大学出版社,2017. 8

ISBN 978 - 7 - 5504 - 3082 - 2

I. ①大… II. ①蹇… III. ①数据处理 IV. ①TP274

中国版本图书馆 CIP 数据核字(2017)第 141010 号

大数据分析与应用:基于 IBM 客户预测性智能平台

蹇洁 主编

责任编辑:高小田

封面设计:穆志坚

责任印制:封俊川

出版发行	西南财经大学出版社(四川省成都市光华村街 55 号)
网 址	http://www. bookcj. com
电子邮件	bookcj@foxmail. com
邮政编码	610074
电 话	028 - 87353785 87352368
照 排	四川胜翔数码印务设计有限公司
印 刷	郫县犀浦印刷厂
成品尺寸	185mm × 260mm
印 张	15
字 数	305 千字
版 次	2017 年 8 月第 1 版
印 次	2017 年 8 月第 1 次印刷
印 数	1—1500 册
书 号	ISBN 978 - 7 - 5504 - 3082 - 2
定 价	29.80 元

1. 版权所有,翻印必究。
2. 如有印刷、装订等差错,可向本社营销部调换。
3. 本书封底无本社数码防伪标识,不得销售。

经管类专业虚拟仿真实验系列教材

编 委 会

主 任：林金朝

副主任：万晓榆 卢安文 张 鹏 胡学刚 刘 进

委 员（以姓氏笔画为序）：

龙 伟 付德强 吕小宇 任志霞 刘雪艳

刘丽玲 杜茂康 李 艳 何建洪 何郑涛

张 洪 陈奇志 陈家佳 武建军 罗文龙

周玉敏 周 青 胡大权 胡 晓 姜 林

袁 野 黄蜀江 樊自甫 蹇 洁

总序

实践教学是高校实现人才培养目标的重要环节,对形成学生的专业素养、养成学生的创新习惯、提高学生的综合素质具有不可替代的重要作用。加强和改进实践教学环节是促进高等教育教学方式改革的内在要求,是培养适应社会经济发展需要的创新创业人才的重要举措,是提高本科教育教学质量的突破口。

信息通信技术(ICT)的融合和发展推动了知识社会以科学2.0、技术2.0和管理2.0三者相互作用为创新引擎的创新新业态(创新2.0)的形成。创新2.0以个性创新、开放创新、大众创新、协同创新为特征,不断深刻地影响和改变着社会形态以及人们的生活方式、学习模式、工作方法和组织形式。随着国家创新驱动发展战略的深入实施,高等学校的人才培养模式必须与之相适应,应主动将“创新创业教育”融入人才培养的全过程,应主动面向“互联网+”,不断丰富专业建设内涵、优化专业培养方案。

“双创教育”为经济管理类专业建设带来了新的机遇与挑战。一方面,经济管理类专业建设应使本专业培养的人才掌握系统的专门知识,具有良好的创新创业素质,具备较强的实际应用能力。另一方面,经济管理类专业建设还应主动服务于以“创新创业教育”为主要内容的相关专业的建设和发展。我们要更好地做好包括师资建设、课程建设、资源建设、实验条件建设等内容的教学体系建设,教学内容、资源、方式、手段的信息化为经济管理类专业建设提供了有力的支撑。《国家中长期教育改革和发展规划纲要(2010—2020年)》提出:“信息技术对教育发展具有革命性的影响,必须予以高度重视。”《教育信息化十年发展规划(2011—2020)》提出,推动信息技术和高等教育深度融合,建设优质数字化资源和共享环境,在2011—2020年建设1500套虚拟仿真实训实验系统。经济管理类专业的应用性和实践性很强,其实践教学具有系统性、综合性、开放性、情景性、体验性、自主性、创新性等特征,实践教学平台、资源、方式的信息化和虚拟化有利于促进实践教学模式改革,有利于提升实践教学在专业教育中的效能。但是,与理工类专业相比,经济管理类专业实践教学体系的信息化和虚拟化起步较晚,全国高校已建的300个国家级虚拟仿真实验教学中心主要集中在理工一类专业。因此,为了实现传统的验证式、演示式实践教学向体验式、互动式的实践教学转变,将虚拟仿真技术运用于经济管理类专业的实践教学显得十分必要。

重庆邮电大学经济管理类专业实验中心在长期的实践教学过程中,依托学校的信息通信技术学科优势,不断提高信息化水平,积极探索经济管理类专业实践教学的建设与改革,形成了“两维度、三层次”的实践教学体系。在通识经济管理类人才培养的基础上,将信息技术与经济管理知识两个维度有效融合,按照管

理基础能力、行业应用能力、综合创新能力三个层次，主要面向信息通信行业，培养具有较强信息技术能力的经济管理类高级人才。该中心 2011 年被评为“重庆市高等学校实验教学示范中心”，2012 年建成了重庆市高校第一个云教学实验平台——“商务智能与信息服务实验室”。2013 年以来，该中心积极配合学校按照教育部及重庆市建设国家级虚拟仿真实验教学中心的相关规划，加强虚拟仿真环境建设，自主开发了“电信运营商组织营销决策系统”“电信 BOSS 经营分析系统”“企业信息分析与业务外包系统”三套大型虚拟仿真系统，同时购置了“企业经营管理综合仿真系统”“商务智能系统”以及财会、金融、物流、人力资源、网络营销等专业的模拟仿真教学软件，搭建了功能完善的经济管理类专业虚拟化实践教学平台。

为了更好地发挥我校已建成的经济管理类专业虚拟实践教学平台在“创新创业教育”改革中的作用，在实践教学环节让学生在全仿真的企业环境中感受企业的生产运营过程，缩小课堂教学与实际应用的差距，需要一套系统、规范的实验教材与之配套。因此，我们组织长期工作在一线、具有丰富实践教学经验和企业从业经历的教学和管理团队精心编写了系列化实验教材，并在此基础上进一步开发虚拟化仿真实实践教学资源，以期形成完整的基于教育教学信息化的经济管理类专业的实践教学体系，使该体系在全面提升经济管理类专业学生的信息处理能力、决策支持能力和协同创新能力方面发挥更大的作用，同时更好地支持学校正实施的“以知识、能力、素质三位一体为人才培养目标，以创新创业教育改革为抓手，以全面教育教学信息化为支撑”的本科教学模式改革。各位参编人员广泛调研、认真研讨、严谨治学、勤勤恳恳，为该系列实验教材的出版付出了辛勤的努力，西南财经大学出版社为本系列实验教材的出版给予了鼎力支持，本系列实验教材的编写和出版获得了重庆市高校教学改革重点项目“面向信息行业的创新创业模拟实验区建设研究与实践（编号 132004）”的资助，在此一并致谢！但是，由于本系列实验教材的编写和出版是对虚拟化经济管理类专业实践教学模式的探索，经济管理类专业的实践教学内涵本身还在不断地丰富和发展，加之出版时间仓促，编写团队的认知和水平有限，本系列实验教材难免存在一些不足，恳请同行和读者批评指正！

林金朝

2016 年 8 月

前言

《大数据分析与应用——基于 IBM 客户预测性智能平台》是 IBM Predict Customer Intelligence 数据分析软件的指导教程，用于大数据分析与应用、数据挖掘等与数据分析相关的综合性课程。该指导书注重理论与实践相结合，把上机实验作为课程实践的重要环节，是教学过程中不可或缺的部分。实验课程与理论课程不同，要充分体现“以学生为中心”的模式，应以学生为主体，充分调动学生的积极性和能动性，重视学生自学能力与动手能力的培养。本书是数据分析相关课程的配套实验教材，编写这本书的目的是满足高校工商管理、电子商务、物流工程、信息管理与信息系统等专业学生学习之用。本书结合大数据的理论与实践，突出数据分析的应用分析，在实验中适当安排了具备认知性、操作性、验证性、综合性等特点的相关实验，以培养学生的动手能力、创新能力。

本教材实践环节通过对数据分析工具的详细介绍和对相关的行业案例的引用，加深学生对课堂教学内容的理解，增加其对大数据的感性认识，增强学生的实际动手能力，培养数据分析的意识与能力。本书共分为五章：第一章是 IBM 预测性客户智能介绍，引出大数据时代背景下的预测性客户分析；第二章是 IBM 预测分析平台系统介绍，涵盖了 IBM 预测性客户智能的整个框架、DB2 数据库、Data Studio、SPSS Modeler、Cognos 系列；第三章是预测模型，主要介绍了在实验部分所用的案例预测模型；第四章是预测性客户智能平台系统的基础操作，针对 IBM 预测性客户智能的 DB2、SPSS Modeler、Cognos 工具，介绍了数据库连接以及工具的基本操作步骤；第五章是 IBM 预测性客户智能平台系统的应用，教师可根据理论课上课需要结合本章行业案例进行相关知识的讲解。此外，本书所需的故障排除措施、术语解释及资料来源均在附录部分有所体现。

全书由蹇洁教授负责主编，廖显、余海燕、陈思祁、余若雪负责全书的统稿。具体参编人员分工如下：前言及第一章中 IBM 预测性客户智能介绍（蹇洁、牛舒），第二章 IBM 预测分析平台系统介绍（廖显、张燕双），第三章行业案例预测模型介绍（余海燕、张英培），第四章预测性客户智能平台系统的基础操作（武建军、杨梦丽），第五章预测性客户智能平台系统的应用（陈思祁、余若雪）。此外，刘路元、王金波、卢敬芝、李东云、罗才林等参与了本书所需资料的收集工作及相关文稿的翻译。陈粤、左逸飞、王健林、龙文思等参与了本书的实验

设计。

本书是重庆邮电大学电子商务特色专业建设的阶段性成果，教育部产学合作专业综合改革项目“大数据分析与应用”的成果。在本书的撰写过程中，得到了诸多同仁的帮助，在此对大家的辛勤工作表示诚挚的感谢！在撰写本书的过程中，编者参考和吸收了国内外相关领域的教学思想和教学内容，但由于数据分析所涉及的相关技术飞速发展，且鉴于我们的水平有限、时间仓促，书中难免有不妥之处，恳请读者与同行批评、指正！

编者

2016年8月

目 录

第一章 IBM 预测性客户智能简介 / 1

第一节 基于预测性客户分析的大数据时代到来 / 1

第二节 IBM 预测性客户智能平台方案简述 / 2

第三节 IBM 预测性客户智能方案的价值 / 3

第四节 IBM 预测性客户智能的业务优势 / 3

第二章 大数据预测性客户智能平台系统介绍 / 5

第一节 预测性客户智能框架介绍 / 5

第二节 DB2 数据库 / 6

一、DB2 介绍 / 6

二、Data Studio 工具介绍 / 6

第三节 SPSS Modeler 简介 / 7

一、SPSS Modeler 概述 / 7

二、SPSS Modeler 节点介绍 / 11

第四节 Cognos 系列简介 / 29

一、Cognos BI 概述 / 29

二、Cognos Framework Management 简介 / 31

第三章 预测模型 / 33

第一节 数据源 / 33

第二节 电信呼叫中心案例的预测模型 / 34

一、客户流失率模型 / 35

二、客户满意度模型 / 36

三、客户关联模型 / 37

四、客户回复倾向模型 / 37

五、分析决策管理中的电信模型 / 38

第三节 电信移动端的预测模型 / 38

一、用于移动端案例的聚合模型 / 38

二、预测流失模型	/ 39
三、呼叫中心预测模型	/ 39
四、建议接受倾向预测模型	/ 39
第四节 零售案例的预测模型	/ 39
一、数据准备为零售提供解决方案	/ 40
二、客户细分模型	/ 41
三、购物篮分析模型	/ 42
四、客户亲和模型	/ 43
五、响应日志分析模型	/ 43
六、库存建议模型	/ 44
七、零售案例中的部署模型	/ 45
八、使用零售案例模型分析 IBM 决策管理	/ 45
第五节 保险案例的预测模型	/ 46
一、保险案例中使用的数据	/ 47
二、客户分割模型	/ 47
三、客户流失预测模型	/ 48
四、客户终身价值模型 (CLTV)	/ 48
五、活动反馈模型	/ 50
六、人生阶段模型	/ 50
七、购买倾向模型	/ 50
八、保单推荐模型	/ 50
九、数据处理模型	/ 50
十、社群媒体分析模型	/ 51
十一、情绪评分模型	/ 51
十二、保险数据模型	/ 51
第六节 银行案例的预测模型	/ 53
一、亲和力分类模型	/ 54
二、客户流失率模型	/ 54
三、拖欠信用卡模型	/ 54
四、客户分类模型	/ 54

五、序列分析模型	/ 54
六、训练预测模型	/ 55
七、评估模型	/ 55
八、商务规则模型	/ 55
九、部署	/ 55

第四章 预测性客户智能平台系统的基础操作 / 56

第一节 数据库连接操作 / 56

一、实验目的	/ 56
二、实验原理	/ 56
三、实验内容	/ 58
四、实验步骤	/ 59

第二节 SPSS Modeler 中模型的建立 / 73

一、实验目的	/ 73
二、实验原理	/ 73
三、实验内容	/ 73
四、实验步骤	/ 74

第三节 Cognos Framework Management 创建元数据模型 / 94

一、实验目的	/ 94
二、实验原理	/ 94
三、实验内容	/ 94
四、实验步骤	/ 94

第四节 Cognos BI 制作可视化报表 / 112

一、实验目的	/ 112
二、实验原理	/ 112
三、实验内容	/ 113
四、实验步骤	/ 113

第五章 预测性客户智能平台系统的应用 / 118

第一节 电信行业案例 / 118

- 一、实验目的 / 118
- 二、实验原理 / 118
- 三、实验内容 / 118
- 四、实验步骤 / 118

第二节 保险行业案例 / 135

- 一、实验目的 / 135
- 二、实验原理 / 135
- 三、实验内容 / 135
- 四、实验步骤 / 135

第三节 零售行业案例 / 167

- 一、实验目的 / 167
- 二、实验原理 / 167
- 三、实验内容 / 167
- 四、实验步骤 / 167

第四节 银行行业案例 / 186

- 一、实验目的 / 186
- 二、实验原理 / 186
- 三、实验内容 / 186
- 四、实验步骤 / 186

附录 A 使用报表的配置 / 215

附录 B 故障排除问题 / 222

附录 C 术语解释 / 224

附录 D 资料来源 / 227

第一章 IBM 预测性客户智能简介

第一节 基于预测性客户分析的大数据时代到来

当前的企业拥有来源广泛的大量客户数据。尽管大多数企业认为这些数据可带来潜在的收益，但是，在将信息有效地转化为可行的洞察力方面，许多企业面临着困难。有效的客户分析战略有助于企业增加收入，避免不必要的成本支出，并且提高客户满意度。消费者和企业每天产生 2.5 Quintillion（1 Quintillion 相当于 10 的 18 次方）字节的数据。事实上，当前，全球 90% 的数据是在过去两年中创造的。这些数据来源多样：用于收集气候信息的传感器、社交媒体中发布的帖子、在线发布的数码照片和视频、销售点（POS）数据、在线购物交易记录、电子邮件内容和移动电话 GPS 信号等。由于上网设备和云服务的价格低廉，世界已经从现实互连向虚拟互连方式转变，与以前相比，产生了更多与客户相关的数据，并且能在更短的时间内完成数据的传输。如今，大多数企业高管都认识到了收集客户相关数据的价值。然而，许多人面临的挑战是如何从这些数据中获得洞察力，继而创造智慧的、主动的、与客户相关的交互通路。他们不确定如何有效地使用客户数据做决策，才能将洞察力转变为销售业绩的增长。采用业务分析的企业可以全面地利用数据、统计和定量分析、探索式和预防性建模以及基于事实的管理，从而在当前复杂的环境中做出更明智的决策。

IBM 预测性客户智能可以帮助人们了解目前状况以及下一步的目标，从依靠猜测进行决策转变为依靠预测进行决策。它可以帮助用户分析自己的结构化和非结构化数据中的趋势、模式和关系，运用这些内容来预测将来的事件，并采取行动以实现期望的成果。无论用户从事市场营销、客户服务、销售、财务、运营还是其他业务领域的工作，都可以随时运用 IBM 预测性客户智能软件中丰富的高级功能，包括在内部部署、在云中部署以及混合解决方案的形式。本产品服务组合将统计分析、预测建模、决策优化和计分、数据收集等功能结合在一起，为用户提供各种工具，解决组织所面临的所有数据难题，实现更好的成果。IBM 的预测性分析解决方案能够满足不同用户的各种需求，无论他们是刚刚入门的，还是经验丰富的分析人员。这些解决方案支持各种规模的企业（无论是中小企业还是大

型企业)利用预测性智能的强大能力,为战术性和战略性的决策指引方向。因此,大数据时代对人类的数据驾驭能力提出了新的挑战,也为人们获得更为深刻、全面的洞察能力提供了前所未有的空间与潜力。

第二节 IBM 预测性客户智能平台方案简述

IBM 预测性客户智能(Predictive Customer Intelligence, PCI)平台根据每个独特客户的购买行为、Web 活动、社交媒体参与情况等,提供与该客户最相关的建议,从而个性化客户体验。通过使用自动化,该集成软件解决方案将从多个内部和外部源收集客户信息,并对客户行为进行建模。然后,通过评分为用户提供可执行的定制行动,以在正确的时间向正确的客户提供正确的购物建议。

IBM 预测性客户智能平台包含以下功能,可帮助用户在与客户接洽的关键点自信地推荐个性化的相关建议:

1. 预测分析,可帮助用户预测每个客户的行为
 - (1) 将数据转变为洞察,帮助用户判断每个客户很可能需要的产品或接下来的行动,如接受购物建议、拖欠抵押贷款或取消保单。
 - (2) 通过预测技术指导一线客户交互和体验。
 - (3) 使用高级客户流失率模型预测并前瞻性地控制客户保留时间。
 - (4) 参与精准营销活动,前瞻性地识别客户服务问题。
2. 决策管理,可将预测模型评分转换为相应的行动
 - (1) 为每个具体的客户交互提供推荐的行动。
 - (2) 近乎实时地利用自动化且经过优化的交易决策。
 - (3) 通过灵活且直观的用户界面,针对每个客户交互开发和实施有针对性的配置和内容。
3. 实时评分,可随需生成和重新生成预测
 - (1) 对交易数据(例如大规模销售、客户服务和索赔交易)持续评分。
 - (2) 为客服人员、营销人员和业务分析员提供最新预测,而非预先计算好的静态历史记录。
 - (3) 支持一线人员在与客户交互时根据预测采取行动,并在了解到新信息后做出反应。
4. 跨营销活动进行优化,可识别针对每个客户来说最有利的决策
 - (1) 跨所有渠道,使营销活动的成果增加 20%。
 - (2) 扫描多个营销活动和业务约束,以查找最符合营销活动的客户。
 - (3) 将业务规则逻辑与通过预测建模获得的洞察相结合。
5. 客户生命周期价值细分,可对客户进行分类,并提供保留时间建议
 - (1) 根据生命周期价值的可能性,使用客户细分方法对客户进行分类。
 - (2) 接收自动生成的行动分配策略,该策略已针对长期期望的奖励进行了

优化。

(3) 根据客户生命周期价值细分,使用建议的行动留住客户。

(4) 使客户细分和建议结果可视化。

第三节 IBM 预测性客户智能方案的价值

1. 市场部门

(1) 对每个客户的交互数据、态度数据、描述数据以及行为数据进行整合,形成 360 度的独特客户视图,并且加以细分。同时能够基于数据流实时刷新客户视图,确保反映最真实的客户记录;

(2) 能够定义市场细分群体;

(3) 在实时接触当中,提供了交叉销售及向上销售方案;

(4) 形成有效的市场活动——正确的时间、正确的地点、正确的方案;

(5) 实时监测客户异动,主动提供客户挽留方案;

(6) 客户生命周期理论驱动;

(7) 提供 IBM Unica 即插即用的链接。

2. 客户服务部门

(1) 线上、线下的任意接触点,都能够为客户提供个性化的服务体验;

(2) 通过提供个性化或客户最有可能响应的方案,延长客户生命周期;

(3) 为每个客户匹配适合的客户代表,提高服务质量以及效率;

(4) 提供与客户接触的最优方式——线下沟通、智能手机、电子邮件、呼叫中心、社交网络等。

3. 销售部门

(1) 为销售部门提供一致的全面客户视图,以便销售人员全面洞察客户;

(2) 在与客户的接触当中,实时提供交叉销售/向上销售的专业洞察;

(3) 通过提供价格敏感度分析来提升利润和收入。

4. 客户体验及洞察部门

(1) 基于客户个体以及细分市场,均能提供一致的全面客户视图;

(2) 通过全品牌分析,识别影响效益的重要驱动因素,并改善整个企业的战略以及运营。

第四节 IBM 预测性客户智能的业务优势

IBM 预测性客户智能综合不同数据源数据,形成 360 度客户视图,获得客户洞察,并能根据客户的独特需求,提供个性化的客户体验,大大提高客户的满意度。

IBM PCI 解决方案是 IBM 全球智慧战略的重要组成部分,从研发、销售、服

务、支持等各个方面都得到了极大支持。

IBM 是全范围的高级分析提供商，提供业界顶尖的分析广度、深度。

IBM 预测性客户智能提供开箱即用的预置分析模板，汇集行业专家经验，尽享全球最佳实践。

IBM 预测性客户智能拥有灵活易用的操作界面，使整个企业都可以分享数据分析产生的价值。

IBM SPSS 西安研发实验室拥有超过 200 人的专业技术力量，为中国用户提供无与伦比的技术支持和响应。

第二章 大数据预测性客户智能平台系统介绍

第一节 预测性客户智能框架介绍

IBM 预测性客户智能的整体框架如图 2.1 所示，其中主要是由三个部分构成，分别是分析、实时配置以及可操作视图。而 Omni-channel 主要是整合业务流程和模型数据。

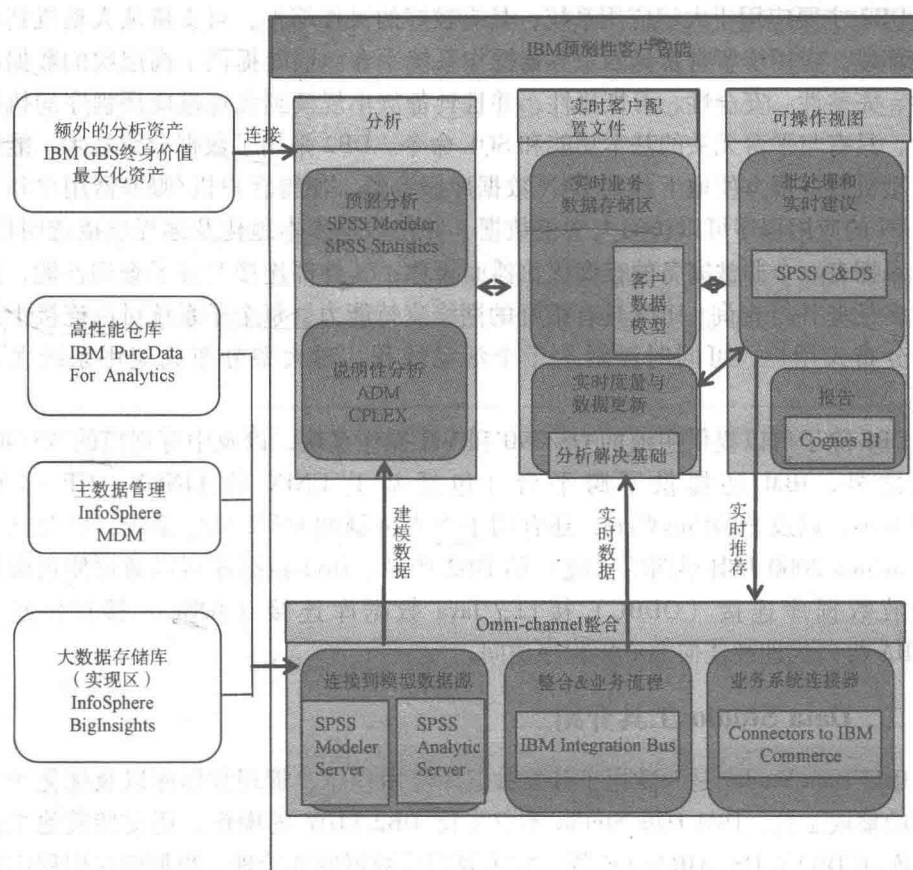


图 2.1 IBM PCI 框架介绍