

商业智能分析方法及案例研究

第一章 商业智能概述.....	3
1.1 商业智能的定义与发展.....	3
1.2 商业智能的关键技术.....	4
1.3 商业智能的应用领域.....	4
第二章 数据收集与预处理.....	5
2.1 数据收集方法.....	5
2.1.1 文档资料收集.....	5
2.1.2 问卷调查与访谈.....	5
2.1.3 网络爬虫.....	5
2.1.4 物联网技术.....	5
2.2 数据清洗与整合.....	5
2.2.1 数据清洗.....	5
2.2.2 数据整合.....	6
2.3 数据预处理技巧.....	6
2.3.1 数据降维.....	6
2.3.2 特征选择.....	6
2.3.3 数据归一化.....	6
2.3.4 异常值处理.....	6
2.3.5 数据加密.....	7
第三章 数据仓库构建.....	7
3.1 数据仓库的设计与规划.....	7
3.1.1 设计原则.....	7
3.1.2 设计步骤.....	7
3.2 数据仓库建模技术.....	7
3.2.1 星型模型.....	8
3.2.2 雪花模型.....	8
3.2.3 维度建模.....	8
3.3 数据仓库的功能优化.....	8
3.3.1 索引优化.....	8
3.3.2 数据分区.....	8
3.3.3 数据缓存.....	8
3.3.4 数据压缩.....	8
3.3.5 数据清洗与转换.....	8
3.3.6 数据维护与监控.....	8
第四章 数据挖掘与分析.....	9
4.1 数据挖掘的基本方法.....	9
4.2 数据挖掘在商业智能中的应用.....	9
4.3 商业智能分析工具与实践.....	9
第五章 联机分析处理（OLAP）.....	10
5.1 OLAP 的基本概念与分类.....	10
5.2 OLAP 多维数据模型.....	11

5.3 OLAP 在实际业务中的应用	11
第六章 数据可视化	12
6.1 数据可视化的原则与方法	12
6.1.1 数据可视化的原则	12
6.1.2 数据可视化的方法	12
6.2 数据可视化工具介绍	12
6.3 数据可视化在商业智能中的应用	13
6.3.1 销售数据分析	13
6.3.2 客户细分	13
6.3.3 财务分析	13
6.3.4 人力资源分析	13
6.3.5 市场趋势分析	13
6.3.6 供应链管理	13
第七章 商业智能报告撰写	13
7.1 报告撰写的基本原则	13
7.1.1 保证数据准确性	13
7.1.2 明确报告目的	14
7.1.3 结构清晰，层次分明	14
7.1.4 简洁明了，突出重点	14
7.2 商业智能报告的类型与结构	14
7.2.1 数据报告	14
7.2.2 分析报告	14
7.2.3 决策支持报告	14
7.2.4 报告封面	14
7.2.5 目录	14
7.2.6 正文	14
7.2.7 结论与建议	15
7.2.8 附录	15
7.3 报告撰写技巧与实践	15
7.3.1 选择合适的报告模板	15
7.3.2 使用图表和可视化工具	15
7.3.3 保持一致性和简洁性	15
7.3.4 强化逻辑性	15
7.3.5 注意报告的排版和格式	15
7.3.6 严格审查和修改	15
第八章 商业智能项目管理	15
8.1 商业智能项目规划与管理	15
8.2 项目风险管理	16
8.3 项目质量管理与评估	16
第九章 商业智能案例研究	17
9.1 零售行业案例	17
9.1.1 案例背景	17
9.1.2 商业智能应用	17
9.1.3 案例成果	17

9.2 金融行业案例.....	17
9.2.1 案例背景	17
9.2.2 商业智能应用.....	18
9.2.3 案例成果	18
9.3 制造行业案例.....	18
9.3.1 案例背景	18
9.3.2 商业智能应用.....	18
9.3.3 案例成果	18
第十章 商业智能发展趋势与展望.....	19
10.1 商业智能技术的发展趋势.....	19
10.1.1 大数据与云计算的深度融合	19
10.1.2 人工智能与商业智能的融合	19
10.1.3 增量式商业智能.....	19
10.1.4 个性化商业智能.....	19
10.2 商业智能在未来企业中的应用	19
10.2.1 企业战略规划与决策支持.....	19
10.2.2 优化企业运营管理.....	19
10.2.3 提升客户满意度.....	19
10.2.4 风险管理与预警.....	20
10.3 商业智能的发展前景与挑战.....	20
10.3.1 发展前景.....	20
10.3.2 挑战	20

第一章 商业智能概述

1.1 商业智能的定义与发展

商业智能（Business Intelligence，简称 BI）是指运用现代信息技术，对企业的数据资源进行整合、分析、挖掘和展示，为决策者提供及时、准确、全面的数据支持，从而提高企业运营效率、优化资源配置、增强竞争力的一种管理活动。商业智能起源于 20 世纪 80 年代，计算机技术和互联网的快速发展，商业智能逐渐成为企业信息化建设的重要组成部分。

商业智能的定义可以从以下几个方面进行阐述：

（1）商业智能是对企业内外部数据的整合与处理，包括数据的收集、清洗、转换、存储和分析等环节。

（2）商业智能的目标是为企业决策者提供有价值的信息，辅助决策者做出更明智的决策。

(3) 商业智能强调实时性、准确性和全面性，以满足企业对数据支持的需求。

商业智能的发展经历了以下阶段：

(1) 数据仓库阶段：20 世纪 80 年代，企业开始建立数据仓库，对大量数据进行整合和存储。

(2) 在线分析处理（OLAP）阶段：20 世纪 90 年代，企业开始使用 OLAP 技术对数据进行分析 and 处理。

(3) 数据挖掘阶段：21 世纪初，数据挖掘技术逐渐应用于商业智能领域，为企业提供更深层次的数据分析。

(4) 大数据与人工智能阶段：大数据和人工智能技术的发展，商业智能进入了新的发展阶段。

1.2 商业智能的关键技术

商业智能的关键技术主要包括以下几个方面：

(1) 数据集成技术：将企业内外部的数据源进行整合，形成统一的数据仓库。

(2) 数据清洗与转换技术：对数据进行清洗和转换，提高数据的质量和可用性。

(3) 数据存储技术：采用高效的数据存储方式，为数据分析和处理提供支持。

(4) 数据分析技术：包括统计分析、机器学习、数据挖掘等方法，用于发觉数据中的规律和趋势。

(5) 数据可视化技术：将数据分析结果以图形、图表等形式展示，便于用户理解和决策。

1.3 商业智能的应用领域

商业智能在众多领域都有广泛的应用，以下列举几个典型的应用领域：

(1) 市场营销：通过分析市场数据，了解客户需求、竞争对手状况和行业趋势，为企业制定有效的营销策略。

(2) 生产管理：通过实时监控生产线数据，优化生产流程、提高生产效率。

(3) 人力资源管理：分析员工数据，为企业提供招聘、培训、薪酬管理等

决策支持。

(4) 财务管理：分析财务数据，为企业提供投资、融资、成本控制等决策支持。

(5) 供应链管理：分析供应链数据，优化采购、库存、物流等环节，降低供应链成本。

(6) 客户关系管理：通过分析客户数据，提高客户满意度、降低客户流失率。

第二章 数据收集与预处理

2.1 数据收集方法

在商业智能分析中，数据收集是的一环。以下是几种常用的数据收集方法：

2.1.1 文档资料收集

企业内部和外部文档资料是数据收集的重要来源。这些资料包括企业报表、统计数据、市场调研报告、行业分析报告等。通过整理这些资料，可以获取关于企业运营、市场状况、竞争对手等方面的信息。

2.1.2 问卷调查与访谈

问卷调查和访谈是获取用户需求、市场反馈等信息的有效手段。企业可以根据研究目的设计问卷，通过线上或线下渠道发放，收集大量有价值的数据。针对特定对象进行深度访谈，也可以获取更为深入的信息。

2.1.3 网络爬虫

互联网的普及，网络数据成为商业智能分析的重要来源。利用网络爬虫技术，可以自动从网站上抓取大量数据，包括新闻、论坛、社交媒体等。这些数据可以帮助企业了解市场趋势、用户需求等信息。

2.1.4 物联网技术

物联网技术可以将各种设备连接到互联网，实现数据的实时收集。例如，通过智能设备收集的用户行为数据、企业生产过程中的实时数据等，都可以为商业智能分析提供丰富的数据来源。

2.2 数据清洗与整合

收集到的数据往往存在不完整、不一致、重复等问题，需要进行数据清洗和整合。

2.2.1 数据清洗

数据清洗主要包括以下几方面：

- (1) 去除重复数据：通过比对数据，删除重复的记录，保证数据唯一性。
- (2) 填补缺失数据：针对缺失的数据，可以根据其他相关数据进行填补，或采用插值、平均值等方法进行估算。
- (3) 数据类型转换：将不同类型的数据统一为统一的格式，便于后续分析。
- (4) 数据标准化：对数据进行标准化处理，消除数据之间的量纲影响。

2.2.2 数据整合

数据整合是将不同来源、格式和结构的数据进行整合，形成统一的数据集。数据整合主要包括以下步骤：

- (1) 数据源识别：确定需要整合的数据源，包括内部数据和外部的第三方数据。
- (2) 数据映射：建立不同数据源之间的映射关系，保证数据的一致性。
- (3) 数据合并：将不同数据源的数据合并为统一的格式和结构。
- (4) 数据校验：对整合后的数据进行校验，保证数据的准确性和完整性。

2.3 数据预处理技巧

为了提高数据分析的准确性和效率，以下几种数据预处理技巧值得借鉴：

2.3.1 数据降维

数据降维是指通过减少数据维度，降低数据复杂度的一种方法。常用的降维方法包括主成分分析（PCA）、因子分析等。数据降维有助于降低计算成本，提高模型训练速度。

2.3.2 特征选择

特征选择是指从原始数据中筛选出对分析目标有较大贡献的特征。通过特征选择，可以降低数据维度，提高模型功能。常用的特征选择方法包括相关性分析、信息增益等。

2.3.3 数据归一化

数据归一化是指将不同量纲的数据进行标准化处理，使其具有可比性。常用的归一化方法包括最小最大归一化、Zscore 归一化等。

2.3.4 异常值处理

异常值是指数据集中与正常数据相差较大的数据。异常值可能对数据分析结果产生负面影响，因此需要对其进行处理。常用的异常值处理方法包括删除异常值、替换异常值等。

2.3.5 数据加密

在商业智能分析过程中，数据安全。对敏感数据进行加密处理，可以有效保护数据隐私。常用的加密方法包括对称加密、非对称加密等。

第三章 数据仓库构建

3.1 数据仓库的设计与规划

3.1.1 设计原则

数据仓库的设计与规划是商业智能系统构建的核心环节。在数据仓库设计过程中，应遵循以下原则：

- （1）符合业务需求：数据仓库的设计应以满足企业业务需求为出发点，充分考虑业务发展的趋势和未来可能的需求变化。
- （2）数据一致性：保证数据仓库中数据的准确性、完整性和一致性，消除数据孤岛。
- （3）可扩展性：数据仓库应具备良好的可扩展性，能够适应数据量的增长和业务需求的变化。
- （4）安全性：保证数据仓库的安全性，防止数据泄露和非法访问。

3.1.2 设计步骤

- （1）需求分析：深入了解企业业务需求，明确数据仓库的目标和应用场景。
- （2）数据源选择：根据业务需求，选择合适的数据源，包括内部系统和外部数据。
- （3）数据模型设计：构建数据仓库的逻辑模型，包括事实表、维度表和关联表等。
- （4）数据集成：将选定的数据源进行清洗、转换和加载，形成统一的数据视图。
- （5）数据存储：选择合适的存储技术，如关系型数据库、NoSQL 数据库等。
- （6）数据仓库部署：将构建好的数据仓库部署到生产环境中，并进行功能优化。

3.2 数据仓库建模技术

3.2.1 星型模型

星型模型是数据仓库中最常用的建模方法，以事实表为中心，周围连接多个维度表。星型模型具有结构简单、易于理解、查询效率高等优点。

3.2.2 雪花模型

雪花模型是对星型模型的一种扩展，将部分维度表进一步拆分为多个子维度表，降低数据冗余。雪花模型在数据量较大时，可以提高查询效率。

3.2.3 维度建模

维度建模是一种以业务过程为驱动，构建数据仓库的方法。它强调从业务角度出发，将数据按照业务过程进行组织，提高数据的可用性和查询效率。

3.3 数据仓库的功能优化

3.3.1 索引优化

索引是提高数据查询速度的关键技术。在数据仓库中，应根据查询需求合理创建索引，提高查询效率。

3.3.2 数据分区

数据分区是将数据仓库中的数据按照一定规则划分到不同的存储区域。通过数据分区，可以提高数据查询的并发功能。

3.3.3 数据缓存

数据缓存是将频繁访问的数据存储在内存中，以提高查询速度。合理设置数据缓存，可以显著提高数据仓库的功能。

3.3.4 数据压缩

数据压缩技术可以减少数据仓库中数据的存储空间，提高数据传输速度。常用的数据压缩方法有行压缩、列压缩等。

3.3.5 数据清洗与转换

数据清洗和转换是提高数据质量的重要手段。在数据仓库构建过程中，应对数据进行清洗和转换，消除数据中的错误和重复记录。

3.3.6 数据维护与监控

数据仓库的维护与监控是保证系统稳定运行的关键。应定期检查数据仓库的功能指标，对异常情况进行排查和处理。同时建立数据备份和恢复机制，保证数据安全。

第四章 数据挖掘与分析

4.1 数据挖掘的基本方法

数据挖掘是商业智能分析的核心环节，它通过从大量数据中提取隐藏的、未知的、有价值的信息，为企业决策提供支持。以下是几种常见的数据挖掘基本方法：

（1）关联规则挖掘：关联规则挖掘旨在发觉数据集中各项之间的关联性，如频繁项集、关联规则等。该方法适用于购物篮分析、商品推荐等场景。

（2）分类与预测：分类方法根据已知数据的特征，将数据分为不同的类别。预测方法则根据已知数据预测未来数据的趋势。常见的分类与预测方法有决策树、支持向量机、神经网络等。

（3）聚类分析：聚类分析是将数据集划分为若干个类别，使得同一类别中的数据对象相似度较高，不同类别中的数据对象相似度较低。聚类分析在市场细分、客户分群等领域具有广泛应用。

（4）时序分析：时序分析是对时间序列数据进行分析，挖掘数据在不同时间段的规律和趋势。该方法在股票预测、销售预测等方面具有重要作用。

4.2 数据挖掘在商业智能中的应用

数据挖掘在商业智能中的应用广泛，以下列举几个典型场景：

（1）客户关系管理：通过数据挖掘技术，分析客户行为、消费习惯等特征，实现客户细分、客户价值评估、客户流失预警等功能。

（2）市场分析：运用数据挖掘方法，分析市场趋势、竞争对手情况，为企业制定市场战略提供依据。

（3）产品推荐：基于关联规则挖掘，为企业提供个性化推荐方案，提高用户满意度和转化率。

（4）风险控制：通过数据挖掘技术，识别潜在风险，为企业风险控制提供支持。

4.3 商业智能分析工具与实践

商业智能分析工具是数据挖掘与分析的重要载体，以下介绍几种常用的商业智能分析工具及其应用实践：

(1) Tableau: Tableau 是一款强大的数据可视化工具,用户可以通过拖拽方式实现数据图表的。在实际应用中,Tableau 可用于销售分析、财务分析等场景。

(2) SAP BusinessObjects: SAP BusinessObjects 是一款企业级商业智能平台,提供报表、分析、数据挖掘等功能。通过 SAP BusinessObjects,企业可实现对数据的深入挖掘和分析。

(3) Python: Python 是一种广泛应用于数据挖掘与分析的编程语言,拥有丰富的数据处理库(如 Pandas、NumPy 等)和机器学习库(如 Scikitlearn、TensorFlow 等)。在实际项目中,Python 可用于实现复杂的数据挖掘算法和模型。

(4) 实践案例:某电商平台运用数据挖掘技术,分析用户行为数据,实现用户细分和个性化推荐。通过实践,该平台提高了用户满意度,提升了转化率。

第五章 联机分析处理(OLAP)

5.1 OLAP 的基本概念与分类

联机分析处理(OLAP, OnLine Analytical Processing)是商业智能领域中的一种技术,它允许用户对大量数据进行复杂的数据分析,以支持决策制定过程。OLAP 系统与传统的联机事务处理(OLTP, OnLine Transaction Processing)系统不同,OLTP 主要用于日常事务的录入和处理,而 OLAP 则侧重于数据的查询和分析。

OLAP 的基本概念包括以下几个方面:

(1) 多维数据集: OLAP 系统中的数据通常以多维形式组织,这种形式也被称为立方体(Cube)。多维数据集允许用户从多个维度(如时间、地点、产品等)对数据进行切片、切块和钻取等操作。

(2) 切片与切块:切片是指在多维数据集中选择一个特定的维度的特定值,从而得到一个子集。切块则是选择多个维度的特定值,得到一个更小的数据子集。

(3) 钻取:钻取是 OLAP 中的一种操作,它允许用户从较高的聚合级别深入到较低的聚合级别,或者反之。例如,用户可以从年度数据钻取到季度、月份甚至天。

根据不同的技术实现方式,OLAP 可以分为以下几类:

以上内容仅为本文档的试下载部分，为可阅读页数的一半内容。

如要下载或阅读全文，请访问：

<https://d.book118.com/967134003120010051>