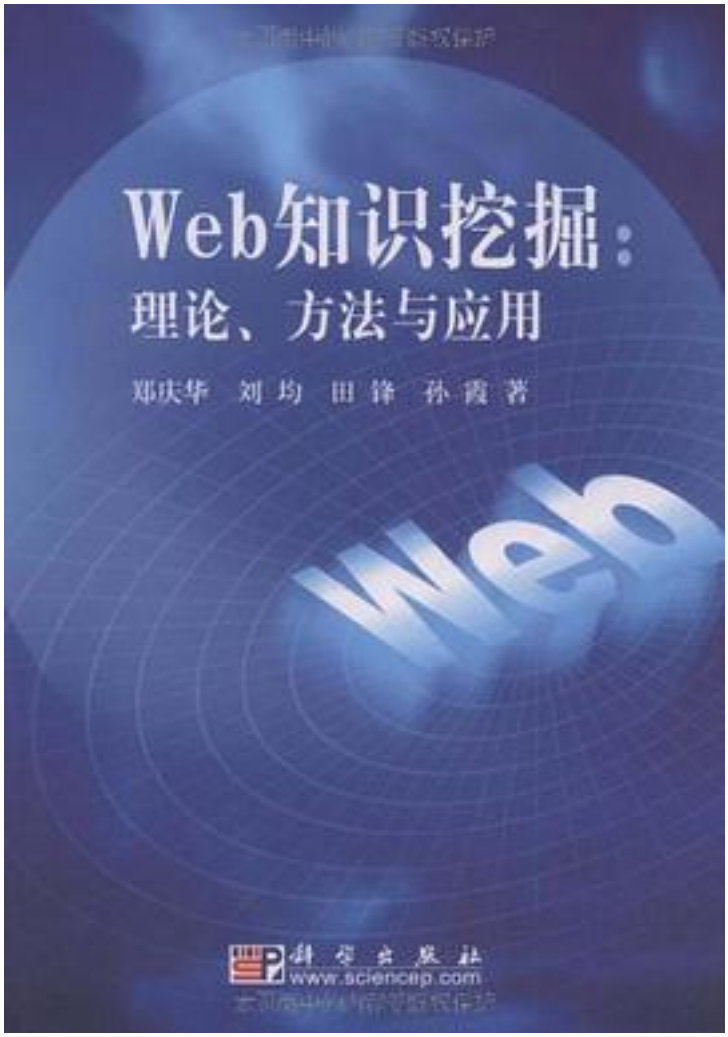
Web知识挖掘



Web知识挖掘_下载链接1_

著者:郑庆华

出版者:科学出版社

出版时间:2010-6

装帧:

isbn:9787030274991

《Web知识挖掘:理论、方法与应用》是一部关于Web知识挖掘的比较系统、完整，且

理论和实践相结合的著作，共含7章：第1章与第2章是Web知识挖掘概论，其中，第1章总体上对Web知识挖掘的现状、概念、典型方法、应用领域以及面临的挑战进行综述性说明；第2章介绍了Web知识挖掘的预备知识、分类体系、基本流程等内容。第3～6章是Web知识挖掘的理论与方法，分别论述了Web爬取、Web结构挖掘、内容挖掘、日志挖掘相关理论与方法，并系统总结了我们自己在元数据、概念、知识元等多个层次上的知识获取以及个性化知识服务等方面的工作。第7章是Web知识挖掘的实践与应用实例，以实例对Web结构挖掘、日志挖掘及内容挖掘的应用进行了说明。

《Web知识挖掘:理论、方法与应用》不仅系统地介绍了Web知识挖掘领域的基础理论与方法，也阐述了我们在该领域的创新性工作，因而适合不同类型与层次的研究人员及学生。

《Web知识挖掘:理论、方法与应用》可作为信息领域的科研与工程技术人员的参考书，也可作为计算机与相关专业的研究生和高年级本科生的教材或辅导书目。

作者介绍:

目录: 前言第1章 Web挖掘概述 1.1 Web发展历史与现状 1.1.1 Web技术发展 1.1.2 Web上的信息爆炸 1.2 Web挖掘的概念 1.2.1 典型的Web挖掘定义 1.2.2 Web挖掘与数据挖掘、信息检索、信息抽取的区别 1.3 Web挖掘面临的挑战 1.3.1 Web数据的高度复杂性 1.3.2 Web数据检索的局限性 1.4 Web挖掘的研究方向 1.5 小结第2章 Web挖掘的基础知识 2.1 Web挖掘的主要预备知识 2.1.1 数据挖掘 2.1.2 文本挖掘 2.1.3 信息检索 2.2 Web挖掘分类 2.2.1 Web数据的分类体系 2.2.2 Web挖掘分类 2.3 Web挖掘的主要应用 2.4 Web挖掘的基本流程 2.4.1 数据采集 2.4.2 数据预处理 2.4.3 模式挖掘 2.4.4 模式评估 2.5 Web挖掘领域的重要文献、国际期刊与会议、标准规范 2.5.1 Web挖掘领域的重要文献 2.5.2 Web挖掘相关的国际期刊与国际会议 2.5.3 Web挖掘相关的标准、规范及语言 2.6 小结第3章 Web爬取与页面组织管理 3.1 Web爬取概述 3.1.1 Web爬取的分类 3.1.2 Web爬取的基本原理 3.1.3 Web爬取面临的挑战 3.2 Web爬取中的主要技术问题 3.2.1 爬取次序 3.2.2爬取性能问题 3.2.3 爬取礼貌性问题 3.3 隐含Web爬取 3.3.1 隐含Web爬虫框架及工作机理 3.3.2 表单分析与提交 3.3.3 隐含Web爬虫实例HiWE 3.4 面向主题的Web爬取 3.4.1 主题相关度分析 3.4.2 确定下个访问URL 3.4.3 面向主题爬取的爬虫实例 3.5 爬取页面的存储与管理 3.5.1 爬取文档的特点 3.5.2 爬取文档的存储方法 3.5.3爬取文档的管理 3.6 小结第4章 Web结构挖掘 4.1 Web结构挖掘概述 4.1.1 Web结构挖掘的分类 4.1.2 Web结构挖掘的应用 4.2 PageRank算法 4.2.1 超链接分析的假设 4.2.2 随机冲浪(random surfing)模型 4.2.3 PageRank值的计算 4.2.4 PageRank算法的改进 4.2.5 PageRank算法在Google中的应用 4.3 HITS算法 4.3.1 HITS算法的基本思想 4.3.2 HITS算法具体过程 4.3.3 HITS算法与PageRank算法的对比 4.3.4 HITS算法改进 4.4 Hilltop算法 4.4.1 Hilltop算法基本思想 4.4.2 专家页面选取及分值计算 4.4.3 目标页面选取及分值计算 4.4.4 PageRank算法和Hilltop算法区别 4.4.5 Hilltop算法的缺陷 4.5 Web宏观结构特性分析 4.5.1 Web的无尺度特性 4.5.2 Web的小世界(small world)特性 4.5.3 “蝴蝶结”和“日冕”现象 4.5.4 Web宏观结构特性的主要应用 4.6 小结第5章 Web内容挖掘 5.1 Web页面的特征表示 5.1.1 特征表示的基本原理 5.1.2特征的离散化 5.1.3 Web页面特征分析 5.1.4页面文本建模 5.2 Web页面分类 5.2.1 分类方法综述 5.2.2 基于内容的网页分类 5.3 Web页面聚类 5.3.1 聚类方法综述 5.3.2 基于内容的页面聚类 5.4 面向Web的信息抽取 5.4.1 信息抽取概述 5.4.2命名实体识别 5.4.3 实体关系检测 5.4.4 页面元数据抽取 5.5 面向Web的本体学习 5.5.1 面向文本的本体学习概述 5.5.2 概念获取 5.5.3 概念关系获取 5.5.4 试验结果与分析 5.6 面向Web的知识元及其关联抽取 5.6.1 知识元及其关联抽取概述 5.6.2知识元抽取 5.6.3 知识元前序关系抽取 5.7 多媒体数据挖掘 5.7.1 图像数据的挖掘 5.7.2 视频数据的挖掘 5.7.3 音频数据的挖掘 5.8 Web内容挖掘的未来研究方向 5.9 小结第6章 Web日志挖掘 6.1 Web日志挖掘概述 6.1.1

Web日志挖掘的分类 6.1.2 Web日志挖掘的典型应用 6.1.3 Web日志挖掘的流程 6.2 Web日志预处理 6.2.1 Web日志数据的格式 6.2.2 Web日志数据清洗 6.2.3 用户识别和会话识别 6.2.4 访问路径填充 6.2.5 事务识别 6.3 序列模式挖掘 6.3.1 序列模式的定义 6.3.2 GSP算法 6.3.3 PrefixSpan算法 6.4 Web用户行为模式挖掘 6.4.1 研究现状 6.4.2 相关概念 6.4.3 用户行为模式挖掘工作机理 6.5 Web用户个性挖掘 6.5.1 个性挖掘的基本概念 6.5.2 个性属性归并 6.S.3 用户个性聚类 6.5.4 个性特征与行为的关联规则分析 6.5.5 个性特征的获取 6.5.6 实例 6.6 Web用户兴趣感知 6.6.1 研究现状 6.6.2 基于建构主义的学习兴趣感知 6.6.3 用户兴趣模型的表示和更新 6.6.4 用户兴趣感知举例 6.7 Web日志挖掘的未来研究方向 6.8 小结第7章 Web挖掘的应用实例 7.1 应用1：面向网络学习的学习者个性挖掘 7.1.1 学习者模型和数据收集 7.1.2 学习者个性挖掘机理 7.1.3 PELDIS工作流程 7.1.4 个性挖掘实例 7.2 应用2：海量Web资源中的知识处理与服务 7.2.1 体系结构与工作机理 7.2.2 基于主题图的Web资源组织与管理 7.2.3 主题图的自动生成 7.2.4 多维关联索引构建与检索结果的个性化排序 7.2.5 个性化资源推荐与导航 7.2.6 基于SOA的Yotta系统实现 7.3 小结参考文献

・・・・・・(收起)

Web知识挖掘_下载链接1_

标签

算法

评论

Web知识挖掘_下载链接1_

书评

Web知识挖掘_下载链接1_