

数据挖掘在图书馆工作中的应用

Application of Data Mining in Library Work

范茜

Qian Fan

云南省图书馆 中国·云南昆明 650031

Yunnan Library, Kunming, Yunnan, 650031, China

摘要: 我们正处于信息化时代, 计算机日益普及, 网络信息服务质量不断提高。作为一种最新的计算机信息技术, 数据挖掘技术被广泛适用于各个行业中。对于图书馆而言, 文献信息规模较大, 怎样管与利用这些有价值的文献信息, 为广大读者提供个性化服务, 这需要数据挖掘技术的支撑。对此, 本研究以数据挖掘技术为切入点, 重点分析了数据挖掘技术在图书管理中的运用方法, 以此来更好地指导图书管理工作。

Abstract: We are in the information age, the computer is increasingly popular, network information service quality is improving. As a new computer information technology, data mining technology is widely used in various industries. For the library, the literature information scale is large, how to manage and use these valuable literature information, to provide personalized services for the readers, this needs the support of data mining technology. In this regard, this study takes data mining technology as the entry point, and focuses on analyzing the application method of data mining technology in the library management, so as to better guide the library management.

关键词: 图书馆; 图书管理; 数据挖掘技术

Keywords: library; library management; data mining technology

DOI: 10.12346/sde.v3i9.4354

1 引言

数据爆炸和知识管理系统时代性查逐渐成为了目前图书馆管理中存在的一系列问题, 特别是近年来, 随着中国经济的高速发展以及图书登记业务的持续性增加, 使得图书馆登记方面的数据也呈现出海量增长的趋势, 而原有的数据库管理方式已经无法适应全新的数据容量, 会出现检索能力不足等一系列问题。而为了进一步优化图书资源结构, 充分发挥出图书资源的价值, 则务必要引进先进的技术, 其中, 数据挖掘与可视化, 作为可以有效提高国土规划数字档案管理的重要技术水平, 可以提高其规划的文档信息服务质量, 并有效改善用户体验效果。

2 数据挖掘技术特征

数据挖掘实质上是一个发现知识的过程。一般而言, 数据挖掘是统计、数据库、模糊数学、神经网络等学科为一体的技术, 从不完善、模糊、大规模的随机数据中准确识别有价值、有潜能的数据处理过程四。其主要特征表现为如下几点。

2.1 数据的分类

数据的分类, 则是细致、准确地分析数据库中的各数据,

并对不同类别进行准确描述, 或者确定分类模型, 或者对分类规则进行明确, 然后根据分类规则分类与记录其他类型^[1]。同时, 积极预测已经分类的模型, 并结合实际情况, 深入剖析相关数据及相关特点, 并基于这些所获取的数据信息对新收集的数据进行深入分析。

2.2 数据的关联分析

对于数据挖掘而言, 关联规则是重点技术日。在数据库寻求的过程中, 各数据项目中必然存在一定的关联规则。对数据关联进行分析, 可找到大多数数据间的密切关系。挖掘频繁出现的数据, 以此来提升关联规则挖掘的速度与质量。另外, 数据管理分析的重点在于寻找到各数据中所隐藏的关联情况, 经过对消费者购买的不同商品, 联系分析出消费者购买习惯, 进而为零售商营销方案提供可靠数据。

2.3 聚类

对于数据分析而言, 聚类是重要分析工具, 综合分析大规模的数据点, 促使数据的类别之间尽量地相似, 不同类别的数据则进行有效区分。因此, 数据的聚类过程是一种无监督分类, 这便说明了聚类实质上是把某一个基于某种相似性的特殊性进行类别划分, 然而聚类未如实记录分类, 也不

【作者简介】范茜(1982-), 女, 中国云南昆明人, 本科, 中级馆员, 从事图书管理研究。

了解具体分类情况,根据差异性的差别,对其分为独特的积累,并合理确定本身实际归属^[2]。

2.4 浅析预测

预测,则是基于没有个体对象自身属性的观察值,准确地预测对象属性将来的素质。在各种较大数据库中,数据挖掘能够积极寻求信息。例如,积极预测市场,数据挖掘以前的各种数据,对以后的投资提供可靠的数据支持。

3 数据挖掘技术在图书管理中的应用

3.1 图书管理中数据的特征

现代社会中,图书收藏量持续增长,图书馆管理规模日益扩大,因此图书管理中的数据呈现出如下特征。

3.1.1 信息隐晦性

信息隐晦性也被叫做潜在性。图书馆在管理过程中出现了大规模的数据,应对这些信息进行有效组织,而且应挖掘其中所蕴含的价值,对此需借助一定的手段才可获取有价值的信息。

3.1.2 关联性

对于图书而言,借阅与查询是信息产生的主要环节,且相关信息之间存在千丝万缕的关系。

3.1.3 海量性

随着图书馆经营时间的增长及馆藏量的增加,图书管理过程中必然产生了大量的数据,海量趋势日益明显。

3.2 图书管理中数据挖掘的需求

近年来图书馆发展中推出了多项管理业务,然而服务上有一定的局限性,大多是简单让读者获得图书方面的咨询,并提供图书的借阅与归还等内容,根据信息实施管理,不具有主动性,也未为读者提供个性化服务。所以,为了能够实现智能化图书管理,管理者应积极摒弃传统服务理念,要提高对读者实际需求的重视程度,从以前简单的检索信息方面进行创新,让图书馆的服务更加多元化。从图书日常管理中工作来看,检索与借阅图书的各个环节中产生大规模的数据。要想对读者阅读习惯及爱好等情况进行准确把握,只是分析这些信息还是不够的^[3]。针对这种情况,图书管理业务的开展要与数据挖掘技术相结合,加大对数据的深层次挖掘力度,这样图书馆才能全面掌握读者的需求与爱好,便于有针对性地开展服务,让读者需求得到满足。

3.2.1 借阅行为模式与读者群分析

①孤点分析:从读者群实际情况出发,关注其借阅特点与个性化需求,对借阅历史进行研究,评价图书的受欢迎度;②聚类统计:从读者情况、借阅记录数据等出发,确定读者借阅行为和群体特点,发挥聚类数据的作用,进一步挖掘其行为规则和模式;③叙述统计:对读者、图书借阅率等进行分析,准确划分读者群体,为推出相关业务服务创造条件。

3.2.2 个性化服务工作

①关联规则分析:对相关信息之间的关联进行深入分析,归纳总结出读者借阅习惯,判断读者实际需求,有利于从各

种读者群体特点出发,保证图书服务更具针对性”。

②时间序列分析:全面掌握读者各种图书借阅的时间、顺序等信息,这样有利于了解其图书借阅习惯,便于在适当时间精准推介相应图书。此外,在时间分析后,也便于对图书的安排更具科学合理。

3.2.3 文献排架管理方面

为提高图书管理水平,需要提高对文献排架的重视程度。以数据挖掘技术为支撑,通过一系列数据了解图书借阅、群体阅读及受欢迎的图书等情况,从而对图书资源进行整合,确保图书资源充分发挥自身价值。

3.3 功能介绍

第一,图书文件的管理,目前图书馆数字图书管理中,无论是数据挖掘或是可视化,都会应用到先进的信息技术,能够对PDF文件进行图像识别中文字的提取和处理,也能够有效地修正一些错误的内容,并且提高识别的精度和效率。更能够为后续的深入学习和文本挖掘提供坚实的基础保障,其中文档文件管理中会包括档案,PDF上传和文档的查询加工等功能。解锁列表中也会给出所需要查询文档的文档编号中文名文档内容等一系列内容,通过点击可以查看文档的元数据和相关的关联数据内容。第二,图书馆的术语管理,这一功能模块中可包括增加删减改动查询等一系列的功能^[4]。第三,图书馆图书流程关联,这一功能更加侧重于对于图书馆图书流程特征建设的开发,根据业务流程,管理逻辑的关系来建立流程的规则库和流程管理项目,实现图书馆图书流程可视化。

4 结语

总而言之,在当今开放且多元的网络环境下,图书馆数字图书资源的数据挖掘和可视化技术的改革创新已经成为必然选择,由于图书馆规划数字图书极具现实意义,图书资源也会与许多图书馆规划工作密切相关,因此更是需要在海量数据趋势下,加深对于图书、档案等资源的深度开发,并有效实现知识管理来充分发挥出数字图书的最大价值。因此,今后也需要充分提升图书馆数字图书部门的参与力度。还要与智慧城市相结合,打造具有生态化系统的整体开放性图书管理模式,更要真正实现。图书馆数字图书馆的多样化功能,真正推动图书管理又好又快的发展。

参考文献

- [1] 王阳.新业态环境下高校图书馆专业馆员职业能力研究[J].林区教学,2021(8):56-59.
- [2] 吕勇,陈文勇.高校图书馆员开展嵌入式服务的新机会:嵌入知识合成研究团队[J].情报探索,2021(8):81-86.
- [3] 罗艳丽.现代信息化视域下高校图书馆管理服务模式创新[J].华东纸业,2021,51(4):5-8.
- [4] 金龙.从俄罗斯帝国公共图书馆的开创探究奥列宁的图书馆实践与思想[J].图书馆研究与工作,2021(8):74-82.