

# 机器学习在图书情报领域的应用研究

李达人

湖北省图书馆 湖北 武汉 430000

**[摘要]**目前,伴随互联网时代的高速发展,计算机技术以及相关产品,正在朝着智能化的方向不断转换,而机器学习属于计算机朝着智能化发展的一个主要渠道,与此同时也属于人工智能技术的核心架构,目前在诸多领域已得到大范围应用。本文主要针对机器学习目前在图书情报领域的应用进行研究,然后基于此,分析了一系列应用现状,以供参考。

**[关键词]**机器学习;图书情报领域;应用

**[DOI]** 10.12252/j.issn.2096-6288.2021.09.064

## 前言

目前,大数据时代背景下,时常出现信息迷航、信息过载等诸多问题,基于此种现状,应用机器自动学习,以计算机设备的图像识别、信息检索、语音识别等诸多功能为基础,可以从体量庞大且繁琐冗杂的数据信息中,自动化的提取有关图书情报的数据信息,从而提高数据获取环节的效率。但目前,我国在机器学习层面针对图书情报问题的研究相对较少,所以有关人员需要对此予以高度重视,深层次的研究图书情报领域对于机器学习的应用,在提升机器学习技术水平的同时,推动其在图书情报领域获取更加长远的发展。

## 1 图书情报领域中机器学习的应用研究

### 1.1 自动化文本分类

对于图书受众有关信息,根据一定的规范对其实行自动化的文本分类,并且针对分类体系中的所有内容实行有效的标识,不仅有助于用户对有关内容实现精准迅速的定位,提高检索环节的速度、便捷性,又可以弥补传统形式存在的周期长、效率低、费用高等诸多缺陷,降低对财力、物力、人力资源造成的浪费。自动化文本在分类层面的规范相对较多,应用传统的人工方式很难真正意义上的实现分类,然而应用人工学习的方式,可以针对用户在信息层面进行有效的研究以及分析,结合用户喜爱的各种类型,确定最为合理的分类方法,以用户需求为主要目标,可以个性化对文本信息展开分类,从而为用户提供更大程度的便利。目前,在自动化文本分类中,情绪识别属于应用最为常见的一种形式,可以针对用户信息有关词汇进行有效分析,然后以此为基础,针对文本信息实现自动化的分离,从而向用户提供有关数据信息。在图书自动化文本分类中,对于机器学习进行科学合理的应用,并建立书目层次划分的类别系统模型,以用户环节的图书检索内容为参考,通过科学合理的论证以及分析,针对有关信息内容进行类别划分,可以按照用户搜索的词汇,针对图书情报进行有效的服务。

### 1.2 智能化信息检索

智能化的信息检索顾名思义是指,为用户在信息查询、搜索等层面,提供个性化的需求以及服务。在信息层面实现智能化检索,需要以机器学习技术为基础,通过机器学习技术针对智能化检索终端,收集到的所有用户在搜索、浏览层面的数据

信息进行整理以及分析,由此可以确定用户具体需求,突出信息的实际种类,用户在输入检索关键词之后,系统就会自动弹出有关信息,伴随用户输入信息内容的不断增加,弹出的信息也可以实现自动化的调节以及转变。在此环节,输入内容需要以用户信息的具体要求为参考依据,尽最大可能弹出价值最高的数据信息,有利于用户最快速度的查找自己所需要的图书数据信息<sup>[1]</sup>。

### 1.3 个性化信息的推荐服务

个性化的信息推荐服务顾名思义是指,针对信息过滤技术实现深层次的优化以及完善,以机器学习功能为基础的智能化终端技术,针对用户信息在阅读、浏览、检索等层面的有关记录进行识别以及分析,通过科学合理的方式对用户进行推荐,由此可以最大程度的满足用户对有关信息的浏览需求。在图书情报的个性化信息推荐服务中,对机器学习技术进行应用,能够针对用户近一段时间内浏览的所有图书信息,有关的数据信息、读书偏好等进行有针对性的分析、整理,在用户对信息进行检索期间,不仅可以按照用户的具体需求将有关图书进行显示,同时也可以按照分析结果,将主题相关的所有数据信息进行罗列。当前,已经有多数业内人士,开始针对个性化的数据信息服务展开深层次的研发,以机器学习为基础对其进行现代化、科学化的设计,能够针对信息实现迅速的分析以及反馈,就时效性层面而言拥有极大程度的优势,不仅可以在图书情报领域,为有关用户在服务层面提供更加个性化的推荐。

## 2 图书情报领域中机器学习在发展层面面临的机遇以及挑战

### 2.1 机器学习面临的机遇

#### 2.1.1 新算法效应为图书情报领域带来积极向上的影响

近些年以来,机器学习中的概率学习、增强学习、深度学习等崭新算法,如雨后春笋般不断涌现,为人工智能的高速发展营造了有利条件,与此同时也为机器学习的深层次发展,注入了崭新的原动力。因为机器学习属于人工智能研发制作的核心,上述算法并不是单纯的在某个独立领域进行应用,而是在人工智能的诸多领域作为分支对其进行应用,由此在人工智能层面形成了一系列的新型现代化技术,比如机器人技术、语音识别技术、计算机视觉技术等,极大程度的推动了人工智能技

术的高速发展,以及优化了其他形式的人工技能技术,同时又可以反过来推动机器学习技术的深层次发展,机器学习技术与其他形式的人工智能技术,相互之间属于相辅相成、协同发展的关系,从而在人工智能行业形成一种新算法效应。比如,智能化的机器人聊天技术,目前在图书馆中进行大范围的应用,不仅可以让用户在服务层面实现自助化的互动目标,同时可以强化用户和图书馆相互之间的交流互动,提高用户在服务层面的诸多体验,极大程度的节约图书馆内部的人力资源,为用户提供咨询、参考、图书检索等诸多现代化、信息化的服务,极大程度的解决图书管理人员的工作量,就本质角度而言,是以机器学习为基础,构建一个以用户问答数据库为基础的答疑、检索的系统。目前,清华大学、合肥工业大学等诸多大学的图书馆,已经在具体实践环节应用到了智能化的聊天机器人,此项技术在应用环节明显让上述学校的图书馆,在智能化、自动化服务层面得到了极大程度的提升,并且获取了学生的一致好评<sup>[2]</sup>。

### 2.1.2 新一代科学信息技术在图书情报领域得到大范围运用

在机器学习层面一项成功的应用,并不是拥有最好的算法,而是拥有相对最多的数据,机器学习主要是将人工智能行业中,主要彰显的数据为主的思想发挥到了目前的极致。对于机器学习目前的核心技术而言,深度学习主要从体量庞大的数据搜集对象中,针对潜在或者有关特征进行收集以及整理,从而完成有关任务,强化学习主要是应用海量数据信息来进行试错,由此在问题解决层面找到最优化的方案。但是,不管在深度学习还是强化学习层面想要实现目标,前提都是需要有体量庞大的数据信息对其作为支撑。而体量庞大的数据信息从何处获取,这一问题一直是长期以来,制约机器学习深层次发展的一大阻碍。新一代现代化科学信息技术,所提供的体量庞大的数据信息,可以更好的契合机器学习在发展层面的诸多需求,更好的解放机器学习数据有关的收集、获取问题,对于机器学习在发展层面而言迎来了极大的机遇,所以就某种程度而言,可以将其视为机器学习实现快速发展的推动器。新一代科学信息技术形成的体量庞大的数据信息,可以帮助机器展开次数更多的试验以及试错,从而持续提高机器学习的能力,而机器学习能力在提升层面,又可以极大程度的帮助公众获取更加精准的数据信息。新一代的科学信息技术与机器学习,相互之间实现有机结合的交互应用,能够深层次的让数据飞轮效应得到进一步扩展,也有利于机器学习乃至整体人工智能行业的高速发展。另外,新一代的信息技术深层次的发展就某种程度而言,为人类提供了可靠性更强的合作方式,来辅助机器学习的深层次发展。比如,百度可以应用用户的反馈以及无偿贡

献等,诸多隐性的合作方式来优化以及完善系统,在翻译层面的精准性以及可靠性,也可以通过雇用体量庞大的专业翻译人员,与其相互之间进行沟通联系等诸多方式,来提高翻译工作的质量以及效率<sup>[3]</sup>。

### 2.2 机器学习面临的挑战

数据问题长期以来,都是机器学习在图书情报领域,很难实现深层次发展的一个主要原因。在机器学习领域流传一句话,“得数据者即得天下”,然而机器学习如果没有体量庞大的数据知识,就很难发挥出自身最大化的作用,数据开放性相对较低,就会导致机器学习在数据来源渠道层面受到诸多阻碍,从而影响精细学习自身作用的最大化发挥。数据开放性相对较低,不仅包含线上、线下数据无法实现共享以及同步,同时也包含线下实体相互之间的有关信息,无法进行有效的沟通以及联系。目前,以高校内部的图书馆阅读人员的阅读信息为具体案例,高校图书馆虽然可以掌控馆内读者有关的图书借阅信息,但是由于诸多因素的限制,无法掌控读者在豆瓣、当当等网络上的购买、阅读情况,而大多数读者通常情况下的阅读行为,又只是在线上开展,由此导致机器学习无法精准掌控,读者在阅读层面的喜好以及需求,从而无法展开精准化、个性化的推荐服务,最终对用户的服务体验造成不利影响<sup>[4]</sup>。

### 结束语

综上所述,在现代化计算机科学领域中,机器学习是实现人工智能的基础所在,以机器学习为前提,能够为用户在服务层面提供质量更好的体验。在图书领域中针对机器学习进行应用,不仅可以提高新型现代化信息技术以及算法,针对数据开展的整合以及分析,为用户在提取图书情报层面,提供更加便利、智能化、个性化的信息检索以及服务,同时为自动化文本分类提供更加便利性的服务。目前,有关机器学习在图书情报领域中的应用研究,相对而言内容较少,所以需要针对其进行深层次的研究以及探索。

### 参考文献

- [1]肖玥,李念祖,谢鹏.机器学习在智慧图书馆的应用研究[J].山东图书馆学刊,2021(04):63.
- [2]勇美菁,钟永恒等.国内图书情报领域人工智能研究演化分析[J].科技管理研究,2020,40(11):155.
- [3]王晰巍,贾若男等.图书情报领域人工智能的研究热点及发展趋势研究[J].图书情报工作,2019,63(01):70.
- [4]张坤,王文韬,谢阳群.机器学习在图书情报领域的应用研究[J].图书馆学研究,2018(01):47.

### 作者简介:

李达人,1991.5.6男,汉族,湖北荆州,馆员(中级职称),本科,湖北省图书馆,图书情报。