

人工智能在学术不端行为风险控制中的应用

刘 鑫,谭彩霞

(金陵科技学院学报编辑部,江苏 南京 211169)

摘 要:人工智能(AI)技术在学术不端行为检测中具有较强适用性和广阔应用前景。当前学术不端行为检测系统存在数据库涵盖范围有限、图表和公式检测结果不理想、人为修改以规避检测等问题。人工智能的大数据智能理论,以自然语言处理和图形、图像识别为核心的精密算法,深度学习和类脑智能计算能力,能够弥补检测系统的不足,从技术上有效控制学术不端行为风险。

关键词:人工智能;学术不端行为;检测系统;自然语言处理;深度学习;大数据

中图分类号:G23

文献标识码:A

文章编号:1673-131X(2018)03-0080-05

The Application of Artificial Intelligence in the Risk Control of Academic Misconduct

LIU Xin, TAN Cai-xia

(Jinling Institute of Technology, Nanjing 211169, China)

Abstract: Artificial intelligence(AI) technology has strong applicability and broad application prospect in the detection of academic misconduct. The current detection system has such problems as limited database coverage, the unsatisfactory test results of the charts and formulas, and artificial modification to avoid detection problems prominent. The application of AI theory and technology such as the large data intelligence theory of artificial intelligence, the precision algorithm based on the natural language processing and image graphics recognition as the core, deep learning and brain-like intelligence computing capabilities and so on can make up for the deficiency of the detection system, and technically effective control the risk of academic misconduct.

Key words: artificial intelligence; academic misconduct; detection system; natural language processing; deep learning; big data

2017 年 4 月 20 日,著名学术出版机构施普林格·自然出版集团以涉嫌同行评议造假为由,宣布撤销旗下《肿瘤生物学》期刊 2012—2016 年发表的 107 篇论文。此次事件将学术不端问题再次推向舆论的风口浪尖。在职称评定“唯 SCI”“唯核心”的大背景下^[1],我国科研诚信情况并不乐观,学术

不端行为呈现多样性和复杂性特征。然而,单纯依靠投稿作者道德约束和行业内的声明很难对学术不端行为风险进行有效的控制。此外,很多学术不端行为都是在文章刊载后才被发现的,此时不良影响已经造成。因此,在学术出版规范和学术不端行为界定等相关标准尚未完善的情况下,利用人工智

收稿日期:2018-08-07

基金项目:江苏省现代教育技术研究 2016 年度立项课题(2016-R-47146);江苏高校哲学社会科学研究基金项目(2018SJA0489)

作者简介:刘鑫(1984-),女,吉林蛟河人,副编审,硕士,主要从事智媒体、数字出版、云出版、学术期刊影响力和传播力研究。

能、大数据等新兴信息技术对论文进行检测,有助于识别部分学术不端行为,减少不良学术影响。目前,各学术期刊出版单位主要利用学术不端行为检测系统来控制学术不端行为风险。因此,本文在对国内主要学术不端行为检测系统进行介绍和对比分析的基础上,探索人工智能在学术不端行为检测中的可行性和应用前景,以期为学术不端行为的风险控制提供新思路。

一、国内主要学术不端行为检测系统对比及问题分析

(一)发展概况

1. 自有数据库系统。中国知网是国内最早涉猎学术不端行为检测的数据库运营商,2008 年其学术不端文献检测系统(AMLC)正式上线。AMLC 以《中国学术文献网络出版总库》为全文比对数据库,同时支持与互联网资源及自建资源库进行比对,是期刊出版单位使用最多的检测系统。AMLC 支持从词、句子到段落的数字指纹定义,并可对图、表等特殊检测对象进行基于标题、上下文、图表等内容的相似性检测处理^[2]。中国知网推出 AMLC 后,万方数据和维普资讯也相继推出 WFSD(万方检测)、WPCS(维普论文检测),运营方式与 AMLC 相似。

2. 指纹数据库系统。目前,答辩前对毕业论文进行检测已成为大部分高校学生毕业工作的一个

重要环节。PaperPass 就是主要面向广大毕业生的检测系统。PaperPass 的比对指纹数据库由超过 9 000 万种的学术期刊和学位论文以及一个超过 10 亿数量的互联网网页数据库组成^[3]。但由于没有自己的数据库资源,该指纹库提取的仅是文章关键性特征信息,并非全文。考虑到每家检测系统的数据库有差异,PaperPass 支持检测者建立自建库对比源,提交检测的论文会与自建库、系统的本地库以及网络库进行对比检测。此外,PaperPass 营销手段新颖独特,例如通过扫码赢得免费检测的活动深受广大学生的欢迎。与 PaperPass 类似的检测软件还有百度论文检测平台、PaperRight 等,这些系统都是借助其他数据库资源进行检测的,虽然“借东风”的运营方式成本较低,但是容易出现比对资源不稳定的问题。

(二)检测系统的对比分析

本文借鉴张旻浩等的研究方法^[4],从网络模式、后台核心数据库、检测文件支持类型、功能特点和是否支持个人使用 5 个维度对 AMLC、WFSD、WPCS 和 PaperPass 进行对比分析(表 1)。通过对比可以看出,基于自有数据库建立的检测系统在功能和形式上大同小异,支持的文件类型也非常相似,区别在于后台核心数据库的差异。而 PaperPass 因没有自有数据库,所以容易出现资源不稳定的问题,但其页面质感强,操作便捷,因而深受大学生喜爱。

表 1 学术不端行为检测系统的对比

系统名称	网络模式	后台核心数据库	检测文件支持类型	功能特点	是否支持个人使用
AMLC	B/S	CNKI 数据库	caj、doc、pdf、txt	检测速度快、准确率高、召回率较高、抗干扰性强	否
WFSD	B/S 或 C/S	万方数据库	pdf、doc、txt、rtf	算法精准科学、报告详实全面	是
WPCS	B/S	四维对比源(维普数据库、用户自建库、TONDA 共享数据库、WEB 资源)	txt	检测资源丰富、有创新性检测指标	是
PaperPass	B/S	动态指纹数据库	doc	页面质感强、操作简单、促销方式新颖、算法更新快	是

(三)问题分析——以 AMLC 为例

目前,学术期刊出版单位使用最多的检测系统是 AMLC,因而本文以 AMLC 为例分析学术不端行为检测系统存在的问题。

1. 不同格式文本检测结果差异较大。根据笔

者所在单位工作流程要求,所有录用的稿件均需进行 3 次学术不端行为检测:审稿前、网络优先出版前、正式刊印出版前。一般审稿前检测为 Word 版本,优先出版和正式刊印出版前检测为 PDF 版本。但在实际操作中发现,这两种文本格式的论文检测

结果相差很大,最大误差达到20%。

2. 数据库资源有限,对跨库资源和图书资源检测受限。AMLC是以《中国学术文献网络出版总库》为全文比对数据库,包括《中国学术期刊网络出版总库》等自有数据库、互联网资源(包含贴吧等论坛资源)和个人比对库等资源,数据资源比较丰富。但由于各大数据库运营商出于利益考虑,会与一些出版单位签订独家代理合同,因而数据库以外的文献和部分网络文献则不能被检测到。

3. 语种限制。AMLC的数据库以中文文献为主,虽然也包含英文数据库(涵盖期刊、博硕、会议的英文数据以及德国Springer、英国Taylor & Francis期刊数据库等),但一方面英文数据库不全,另一方面没有小语种数据库,对于一些将外文文献直接翻译成中文后完全或部分抄袭的行为,目前难以进行检测。

4. 适用性不强。AMLC对文字复制的检测准确率较高,但对图片、表格和公式尚无法完全辨别检测,对以同义词、近义词、改变句式结构等手法规避查重的手段也无法识别。

实际上,无论是基于自有数据库的AMLC、WFSD、WPCS还是基于指纹数据库的PaperPass,都未解决基准数据库有限、图表检测失效、数据造假难辨、语义修改降重识别难等问题。

二、智媒时代学术不端行为检测的新技术可能

自2014年8月中央《关于推动传统媒体和新兴媒体融合发展的指导意见》发布以来,国内媒体融合在继续呈现出平台化、分众化、流体化特征的同时,智能化、创意化等智媒特征也在融入其中。人工智能、物联网、AR/VR等新兴技术赋予了媒体智能基因,使信息传播更自动、精准、远影响——智媒时代已经开启^[5]。2017年,百度、腾讯等互联网产业巨头在全球范围内掀起了人工智能商业化浪潮,随着深度学习、图像(语音)识别、机器视觉、数据挖掘等智能技术的创新和突破,人工智能与文化内容产业的融合不断加深,文化产业进入人工智能新纪元,也为人工智能技术在学术不端行为检测中的适用提供了可能。

(一)人工智能在文化产业中的应用

随着科技的迭代更新速度加快,人工智能不断

与金融、医疗、交通等多个行业深度融合,成为影响世界产业格局的核心科技。我国也把发展人工智能上升到国家战略层面。国务院于2017年7月发布《新一代人工智能发展规划》,对我国人工智能发展的态势、总体要求、重点任务、资源配置和保障措施进行了系统介绍和部署;党的十九大报告进一步指出,要推动互联网、大数据、人工智能与实体经济深度融合。在国家政策引导和扶持下,人工智能在向纵深方向发展并与文化产业深度融合。

1. 内容推荐算法。人工智能内容推荐已不是新鲜话题,其在新闻、音乐、视频等APP中广泛使用。利用人工智能算法,可以对平台用户的多维度数据进行统计分析,据此估算用户的阅读(收听、观看)喜好和倾向,然后进行有针对性的推荐,打造内容一站式管家。人工智能内容推荐算法能够快速实现用户与信息的精确匹配,降低用户获取信息的成本,满足其多元化、个性化的内容需求。但在利用人工智能内容推荐算法过程中,应尽量避免“信息茧房”带来的负面效应。

2. 人工智能写作。一是智能新闻写作。2010年,美国思科公司工程师与美联社等新闻机构开始利用Wordsmith人工智能平台撰写新闻稿件。2013年,人工智能撰写的新闻稿件数量达到3亿篇^[6]。2017年由中国地震台网“地震信息播报机器人”自动编写的《四川阿坝州九寨沟县发生7.0级地震》稿件,从创作到发布仅用时25秒。人工智能新闻写作在节省大量劳动力的同时,也在重大突发事件面前充分展现了人工智能的速度优势。二是文学作品写作。2017年5月,微软人工智能“小冰”创作并发布了诗歌集《阳光失了玻璃窗》,这是人类历史上第一部百分之百由人工智能创作的诗集,是人工智能在人文领域的一大突破。但是,目前人工智能仍是程序控制下的被动思维,很难具有人的情感和主动创造性。因此,与程式化的新闻写作相比,具有思想特质的文学作品的人工智能化道路还很漫长。

3. 人工智能翻译。一直以来,自然语言处理是机器翻译的最大缺陷。随着人工智能深度神经网络的发展,机器翻译开启人工智能新纪元。近两年,神经网络机器翻译(Neural Machine Translation, NMT)技术异军突起,翻译质量大幅跃升。在文化全球化的推动下,高度成熟的翻译技术或将掀起第二次全球化的浪潮^[7]。2018年3月,微软

研究团队就研发出首个在新闻报道的翻译质量和准确率上媲美人类专业译者的翻译系统。该系统突破此前 NMT 的局限,实现了机器翻译与人类自然语言翻译的一致性和协调性,是神经语言程序学(NLP)里程碑式的突破^[8]。人工智能在机器翻译中的应用,将使翻译结果更加“信、达、雅”。

(二)人工智能应用于学术不端行为检测的可行性

人工智能是在计算机科学、控制论、信息论、心理学、语言学等多种学科相互渗透的基础上发展起来的一门新兴边缘学科,应用领域涵盖符号计算、模式识别、专家系统、机器翻译等诸多方面。国务院发布的《新一代人工智能发展规划》指出,我国在人工智能领域取得了重要进展,国际科技论文发表量和发明专利授权量已居世界第二,自适应自主学习、直觉感知等初步具备跨越发展的能力,中文信息处理、智能监控、生物特征识别等逐步进入实际应用阶段^[9]。这些都为人工智能技术在学术不端行为检测中的应用提供了理论和技术可能。

目前,在主流学术不端行为检测系统中,检测结果不尽如人意的主要表现有:一是数据库涵盖范围有限,跨库检索尤其是跨语言检索困难;二是图表、公式检测结果不理想,容易引发人为二次加工抄袭;三是根据查重结果进行语义修改以规避检测问题突出。利用人工智能大数据智能理论,能够将数据驱动与知识引导相结合,利用大数据深链打破对比数据库的局限;以自然语言处理和图形、图像识别为核心的精密算法,使图形、图像识别更为精确智能;高级机器的深度强化学习和类脑智能计算,能够使对自然语言的处理从“机器化”向“类人化”转变,从而解决跨语言识别和语义识别问题。

三、人工智能在学术不端行为检测中的应用前景

(一)基于深度学习的 NLP 在语义检测中的应用

自然语言处理(Natural Language Processing, NLP)是指利用计算机对人类自然语言信息进行处理和加工,最终实现人机对话的理论和方法^[10]。作为人工智能的重要领域,NLP 的目标是使计算机(机器)具有与人类一样的语言理解能力。NLP 是学术不端行为检测的重要技术手段,但语言的任意性、可变性、无限性和二义性等特征使 NLP 在单词边界界定、语义歧义消除、不规范输入识别等方

面尚难以突破。

目前,学界公认的学术不端行为主要包括:抄袭、篡改和伪造。在现有技术水平下,检测系统主要对文字抄袭进行对比检测,各检测系统的检测原理基本相同,即首先对检测文献按照全文、段落、句子三个层级分层处理,然后根据层级分别创建指纹并与比对资源进行比对。由于检测系统最小指纹粒度为句子,所以原则上只要检测文献与比对文献存在一个相同的句子,就能被检测系统发现(表 2)。NLP 的句子级分析分为词法分析、句法分析和语义分析三个层面^[11]。例如,在对某一句文本进行分析时,首先要进行分词和指派词性(词法分析),然后利用句法规则识别句法成分并判断句法功能(句法分析),最后结合语义角色(语义分析)得到该句的逻辑语义表达。由于汉语是表意文字,因此在对汉语文献进行自然语言处理(即中文信息处理)时,面临分词、词性标注、深层文法句法和语义分析等问题。

表 2 AMLC 抄袭界定标准

类型	重合文字条件	总文字复制比例
轻度句子抄袭	连续重合文字均<200 字	<10%
句子抄袭	连续重合文字均<200 字	≥10%
轻度段落抄袭	连续重合文字≥200 字	<30%
段落抄袭	连续重合文字≥200 字	≥30% 且 <50%
整体抄袭	重合文字≥(总字符数/2)	≥50%

近年来,基于神经网络(Neural Networks, NNS)的深度学习(Deep Learning)在自然语言处理中的应用给以上问题的解决带来了契机。与其他机器学习方法类似,人工智能领域的深度学习也是利用数学模型对特定问题进行建模以解决该领域内相似问题的过程,即将计算机需要学习的数据导入复杂的、包含多个层级的数据处理网络(深度神经网络)中,然后检查经过这个网络处理得到的数据是否符合要求。如符合就保留该目标模型,不符合则继续调整网络参数,直至输出符合要求^{[6]77-88}。NLP 是以比较抽象的符号来表达概念的,但也存在一些问题,比如两个词的含义相近但词形不匹配,计算机在进行信息处理时就会认为它们是两个词,这就为人为语义修改以规避检测提供了可能。传统的 NLP 主要采用人工定义词、词性、类别标签等原子特征,而深度学习则把这些原子特

征进行量化,再利用多层神经元网络提取特征,计算不同层次的语言单元之间的相似度^[11]。此外,随着图片、音频、视频等多媒体形态数据在信息处理中的应用日益广泛,信息处理由单媒体信息处理逐渐向文本、图像、音视频等多模态信息处理转变。基于深度学习的多模态信息处理能够通过多模态信息建模进行语义的度量、分析和识别,为检测系统识别图表与文字转换抄袭行为提供依据。

(二)大数据深链、区块链在数据库扩展中的应用

数据库是学术不端行为检测系统的核心。无论是精密算法还是 NLP,覆盖面广的数据库资源都是其应用的基础。而中国知网的 AMLC 之所以为大多数期刊出版单位所认可,主要原因是其具有强大的比对数据库优势。而这正是 PaperPass 这类基于指纹数据库系统的劣势。但在大数据时代,无论自有数据库有多大,都无法承载爆发式增长的数据。“互联网+”时代,信息交换、存储和处理中产生了大量的数据。1986—2017 年,信息存储能力每 3 年翻一番,2017 年全球每天通过互联网通道交换的信息数量是 1986 年的 217 倍^{[6]89}。大数据的介入使最初的关系型数据库在处理超大规模和高并发的 SNS 型 Web2.0 纯动态网站时力不从心,数据库向能够对海量级数据进行管理和分析的非关系型数据库转变。非关系型数据库具有高并发性和可拓展性,数据查找速度快,更容易进行分布式扩展,并且非关系型数据库在处理文档数据时不需要预先定义表结构,可以对图片进行结构算法分析,能够解决现有检测系统数据库无法进行海量信息处理的难题。但实际上,随着数据量的无限极增长,无论是关系型数据库还是非关系型数据库,其中心化的云存储构建服务是昂贵和低效的^{[12]116},数据中心化成为检测系统数据库扩展的瓶颈。而区块链作为一种新型的去中心化数据库技术,能够解决这一问题。

区块链技术(Blockchain Technology, BT)是一种互联网数据库技术,其特点是去中心化、公开透明、信息不可篡改,让每个人均可参与数据库信息记录^{[12]31}。近年来,随着比特币等数字货币的出现,其核心技术——区块链逐渐进入公众视野。虽然区块链是随着数字货币的兴起而被广泛关注的,但实际上,区块链并非仅与金融相关,其在人工智

能的诸多领域都有所建树,当下产业风向正从“互联网+”走向“区块链+”,而区块链技术将成为下一代数据库架构技术。区块链数据库是继关系型数据库和非关系型数据库后的第三种数据库类型,其运行在基于区块链的新型云计算平台上,无需架设任何服务器。例如, Sia 是基于区块链数据的云存储平台,平台中个人和用户的数据被分散存储在众多节点中,并可被自动化智能合约追踪^{[12]117},海量信息存储和信息安全得到了双重保证。

参考文献:

- [1] 刘鑫,刘琳. TOP 学术期刊跨文化传播与国际影响力研究述评[J]. 中国报业, 2016(18): 14 - 15
- [2] 潘萍. 论“学术不端文献检测系统”在学报编辑工作中的应用[J]. 云南大学学报(自然科学版), 2011(S2): 82 - 84
- [3] PaperPass 官网. PaperPass 介绍[EB/OL]. [2018 - 03 - 20]. http://www.paperpass.com/?utm_source=bdpz
- [4] 张旻浩,高国龙,钱俊龙. 国内外学术不端文献检测系统平台的比较研究[J]. 中国科技期刊研究, 2011(4): 514 - 521
- [5] 林竹鸣. 未来之路:2018 传媒业内容生产趋势[J]. 中国记者, 2018(1): 11 - 13
- [6] 李开复,王咏刚. 人工智能[M]. 北京:文化发展出版社, 2017
- [7] 搜狐网. NMT 机器翻译技术或将开启第二次全球化浪潮[EB/OL]. (2016 - 12 - 22)[2018 - 03 - 20]. http://www.sohu.com/a/122262870_116441
- [8] 新智元: AI-era 微信公众号. 人工智能再下一城: 机器翻译提前 7 年达到人类专业翻译水平[EB/OL]. (2018 - 03 - 15)[2018 - 03 - 25]. http://www.thepaper.cn/newsDetail_forward_2029654
- [9] 国务院. 新一代人工智能发展规划[EB/OL]. (2017 - 07 - 20)[2018 - 03 - 29]. http://www.gov.cn/zhengce/content/2017-07/20/content_5211996.htm
- [10] 刘金凤. 面向自然语言处理的汉语句子语义知识库构建研究[D]. 烟台:鲁东大学, 2009
- [11] 中国中文信息协会. 中文信息处理发展报告(2016)[R/OL]. (2016 - 12 - 24)[2018 - 04 - 05]. <http://cips-upload.bj.beebos.com/cips2016.pdf>
- [12] 黄步添,蔡亮. 区块链揭秘:构建基于信用的下一代互联网[M]. 北京:清华大学出版社, 2016

(责任编辑:唐银辉)