

AI对未来会计变革研究——基于技术视角

黄亮

(集美大学 财经学院,福建 厦门 361021)

摘要:人工智能目前是全社会关注的热点,人工智能技术丰富多样,发展迅速,有必要结合会计领域自身的特点,对影响会计发展的人工智能关键基础技术进行梳理,从技术角度探讨AI对未来会计的影响。介绍了OCR、NLP和大数据等人工智能基础技术,探讨这些技术对未来会计发展的影响路径,最后系统分析了变革中可能遇到的瓶颈问题和解决方案。

关键词:人工智能;光学字符识别;自然语言处理;数据挖掘;大数据

中图分类号:F230 **文献标志码:**A **文章编号:**1008-5092(2021)04-0053-06

人工智能(Artificial Intelligence,简称AI)是研究、开发用于模拟、延伸和扩展人的智能的理论、方法、技术及应用系统的一门新的技术科学。AI是引领未来的战略性技术,为了在新一轮国际科技竞争中掌握主导权,2017年7月8日,国务院印发并实施了《新一代人工智能发展规划》,从国家层面制定了未来AI的战略部署,这显示了政府高层对AI发展的高度重视。

AI在各行各业都有着广阔的应用前景,以制造业为例,德勤在2019年发布的《人工智能制造业应用调查报告》中指出,93%的受访企业认同AI将成为全球制造业增长和创新的关键技术,83%的企业认为AI已经或将在未来五年内对企业产生实际可见的影响。在AI的冲击下,会计相关领域的变革也难以避免。

一、AI在会计领域中的关键技术

AI目前是计算机领域的研究热点,相关AI技术丰富多样,不断更新,有必要结合会计自身的特点,对其中将会影响会计领域发展的关键技术进行梳理,这样就可以从技术视角揭示AI环境下会计相关领域未来的发展路径。

1. 光学字符识别(Optical Character Recognition,简称OCR)

OCR是计算机视觉研究领域的分支之一。OCR技术起源于德国,这个概念是由德国的科学家Tausheck于1929年首先提出的。它是模拟人类的视觉,把图像中的文本内容识别提取出来,并给予数字化的技术。

电子数据是计算机处理加工和网络传输的“原料”,在实际应用中,除了电子数据外,还充斥着数量庞大的印刷体文本、手写文本、相关纸质单据和图片等,这些有用的信息只有进入计算机,成为计算机可以识别的数据,相关AI算法才能对其加以利用。因此,OCR为代表的识别技术是AI算法应用的基础。

以财务共享系统为例,为了节约企业的运营成本,越来越多的大中型企业采用财务共享模式,对于远程分公司的常用单据发票等通过扫描方式生成图像,然后通过网络,传送到远在千里外的财务共享服务中心。共享服务中心接收图像,首先采用OCR扫描识别技术,快速对图像的真伪进行鉴别,并自动识别图像中的会计要素,自动对票据建立索引并归档。^[1]据统计,与传统的会计人工

收稿日期:2021-04-26

基金项目:福建省财政厅资助项目(SCZ201902)。

作者简介:黄亮(1978—),女,福建厦门人,讲师,硕士,研究方向:会计理论、大数据下的内部控制。

录入数据方案相比,发票 OCR 扫描识别理论上可以减少 70% ~ 80% 以上的工作量,单据发票信息的采集效率产生质的飞跃。而且,由于面向特定行业的专业型 OCR 系统信息格式较为固定,识别的字符集相对小,随着 OCR 技术的不断进步,对于常用单据发票的识别准确率已经与人工不相上下。

2. 自然语言处理(Natural Language Processing, 简称 NLP)

OCR 识别技术仅仅是 AI 应用的初步,AI 还需要结合上下文和相关语境,理解其中的准确含义,由此 NLP 技术应运而生。NLP 是对自然语言进行处理和理解的系统,它的主要功能是计算机能处理和理解输入的自然语言,并返回响应的结果。NLP 是人类和计算机交流的关键技术,当计算机可以识别并理解使用者录入的文字和语音,就可以采用数据挖掘、大数据分析、财务机器人(Robotic Process Automation, 简称 RPA)、专家系统等 AI 算法对这些电子数据进行后续加工处理。由此,诸如阅读和分析财务报告、报销单据、阅读法律文本、识别风险等烦琐专业的工作将变得简

单、高效和客观,而且处理的边际成本几乎为零。

值得指出的是,目前在财务报告领域应用广泛的 XBRL 技术之所以能引起重视,关键在于通过唯一含义的、标准化的标签,使财务报告中的项目可被计算机识别和理解,从而便于计算机处理。据统计,有了 XBRL 的支持,计算机可以在一分钟内快速准确地阅读上百份的财务报告,并实时提交分析结果。由于 XBRL 格式有其规范性,因此,计算机理解 XBRL 格式的文档语义远比自然语言容易得多,XBRL 取得的巨大成功可以印证计算机识别和理解语言文字的重要性和巨大的潜力。

NLP 技术的发展任重道远,在一些应用环境相对单一的领域已经得到了较好的应用,如单据报销领域,以“人人差旅”为例,在 NLP 和财务机器人的支持下,风险的识别、补助标准的计算、财务的审核以及支付和核算都几乎可以自动实现(如表 1),有效地缓解了商旅人士对公司差旅报销诟病最多的报销耗时长、收集整理发票难等问题。华为等大公司纷纷采用了类似的报销系统,极大地缓解了财务人员和报销人员的负担,节约了成本,效率也随之大为提高。

表 1 基于 NLP 和 RPA 技术的人人差旅报销系统

角色	员工	管理者	员工	员工	审核	出纳	会计
场景顺序	发出出差申请	管理者审批	出行方案选择	差旅报销	财务审核	财务付款	财务核算
NLP + RPA (财务机器人)	NLP 出差申请单生成	NLP 内审风险池数据分析,疑似风险点控制	NLP 智能推荐出行方案	根据规则设定, NLP 分析报销差旅,并进行补助标准计算	NLP 实现模糊算法,连号、重复、相近号、相近金额查验	RPA 技术自动操作异构系统网银支付	RPA 智能核算,后台事务自动处理

资料来源:刘丹彤. 自然语言处理、知识图谱及其在财务中的应用[R]. 金蝶软件研究报告

3. 数据挖掘(Data Mining)和大数据(Big Data)

随着高速互联网的出现和单位存储成本的急剧降低,电子数据呈几何级的爆炸增长, IDC 预计到 2020 年,全球将总共拥有 35 ZB 的数据量。每时每刻都有大量的电子数据被创造,凯文·凯利指出,“一切都被记录,一切都被数字化。不管你现在做什么行业,未来都是数据生意”。2019 年 10 月,中国信息通信研究院发布了《全球数字经济新图景(2019 年)》显示,我国 2018 年数字经济规模为 4.73 万亿美元,保持全球第二大数字经济体地位。海量的电子数据并不是孤立存在的,也不

应该孤立,它们之间总会存在着千丝万缕的联系,人们迫切希望从中挖掘出有价值的信息。基于此,以数据挖掘和大数据为代表的 AI 技术引起了人们的广泛重视。

数据挖掘可以在海量数据中利用各种分析工具发现模型与数据间的关系,发现被隐藏的、被忽略的因素。大数据则是数据挖掘概念的再升级,是大量、快速和多样的信息资产。它的提出是和目前爆炸式“实时动态”增长的电子数据相适应的。相关资料显示,在数据爆炸的社会,交通数据日均增长量为 1 PB,金融数据日均增长量为 0.5 PB,食品安全日均增长量为 0.03 PB,公安安防数

据日均增长量为 5 PB。面对目前爆炸式实时动态增长的电子数据,需要数据处理速度遵循“1 秒定律”,即一般要在秒级时间范围内给出分析结果,否则就无法适应激烈的社会竞争节奏,这就需要采用大数据中的流计算(Stream Computing)模式,实现动态数据的实时分析,^[2]这是大数据和传统的数据挖掘技术最大的区别。

二、AI 的关键技术对未来会计发展的影响路径

随着 AI 应用的深入,未来的会计相关领域已经发生或将要发生天翻地覆的变化,这是时代赋予会计的机遇和挑战。本文认为,根据相关会计工作的特点,AI 对未来会计的影响主要遵循两条路径(图 1):

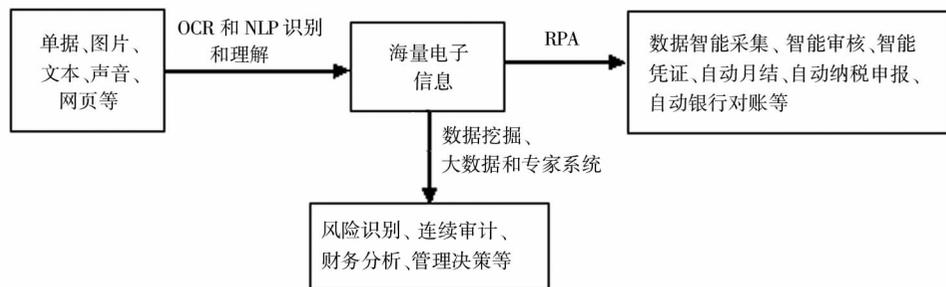


图 1 AI 对未来会计的影响路径

1. 规则导向的标准化工作被财务机器人取代

在日常会计工作中,充斥了大量的可重复性、规则导向性、可标准化、可流程化的核算事项,比如数据智能采集、智能审核、智能凭证、自动月结、自动纳税申报、自动银行对账等工作,^[3]这些工作由于有明确的流程和判定标准,便于计算机模拟,完全可以交给财务机器人(RPA)完成。2017年10月,德勤 RPA 的问世是财务数字化领域的标志性事件,RPA 随之迅速得到了大量的应用,如中化国际引进普华永道机器人实现了财务流程自动化,把财务人员从银行对账、月末入款提醒、进销项差额提醒、增值税验证这些枯燥重复的工作中解放出来。据中国强财财务软件公司人员调查,RPA 可将开票的效率提高大概 70% 之多。^[4]RPA 的广泛应用不仅极大缓解了会计人员的工作强度,而且其凭借客观、准确和及时的特点也很大程度上加强了会计信息的相关性和可靠性。

2. 数据挖掘和大数据在更高级决策中的应用

在 RPA 取得成功之后,使用者还希望 AI 能在风险识别、连续审计、财务分析、管理决策等更复杂场景的工作中发挥作用。有了 OCR 和 NLP 等技术的支持,就有了计算机处理所需要的海量基础数据,这样就可以通过数据挖掘、大数据和专家系统等,让计算机拥有相关会计领域专家的智慧和经验。

由于这些问题往往都属于半结构化问题或非结构化问题,数据挖掘、大数据和专家系统的应用并不能提供给我们确切的解决方案,但其及时提供的提示和建议却可以让决策人员事半功倍。以 2008 年德意志国家银行的乌龙事件为例,在各位高管几乎都知悉雷曼兄弟可能倒闭的情况下,该银行却在雷曼倒闭的 10 分钟后按照外汇掉期协议,通过计算机自动付款系统,向雷曼兄弟公司的银行账户转入 3 亿欧元,这笔巨款立刻随着雷曼兄弟的轰然倒塌化为乌有。这种愚蠢的错误发生的原因在于公司风控系统的客户信息库没有对雷曼兄弟的相关资料进行及时的更新。客户信息库如果单纯依靠手工收集和判断,很容易出现信息的滞后,而 AI 却完全可以通过爬虫技术从主流财经网站、主流媒体、政府官方网站等实时获取客户的信息,并运用相关 AI 算法对数据进行加工处理,从而对客户风险进行判断,更新客户信息库,这样在自动付款系统转账前,系统会自动查询与该客户相关的资料,向决策部门自动提交预警信息,供决策部门确认,否则无法执行。这种高效、客观、及时的内控方式将可以避免这类乌龙事件的发生。

三、AI 变革中存在的瓶颈问题

随着 AI 在各行各业的广泛应用,会计理论界

和实务界均表现出乐观的情绪,认为 AI 将对未来的会计产生革命性的影响。很多高校也相继开设了“智能会计”专业方向,设置了很多相关的课程。本文虽赞同 AI 对未来会计的巨大影响,但也认为不能过于乐观,其中有不少瓶颈问题有待解决。

1. OCR 和 NLP 等关键技术还有待改进

随着计算机性能的提高和相关算法的改进,无论是印刷体英文字母,还是规范的手写英文,很多软件的英文字母识别准确率可以接近 100%,但与字母文字相比,汉字的识别要困难得多,主要原因在于:(1)汉字的特点。汉字数量庞大,并且汉字的词汇无法使用空白符分隔,也加大了识别难度。(2)汉字书写不规范。随着研究和设备的不断进步,目前的技术对印刷体汉字的识别率已经非常高了,如文通 OCR 对于版面整洁的印刷字符的识别,正确率达 99.96%,但由于每个人的书写习惯和字体的不同,手写汉字的识别率却不高,即使对于规范手写汉字,识别率一般也仅仅达到 85% 左右。(3)环境的干扰。在单一场景的基础上,目前的识别算法比较理想。如 OCR 发票识别技术对常见的发票(如增值税专用发票、增值税普通发票)的识别率几乎可以达到 100%,因为它可以根据地税、国税发票的特征、共性特别定制识别技术。但如果要检测和识别自然场景中的所有图像,还存在很大的困难,这是因为图像中汉字的大小、字体、角度、颜色等存在差异,图像本身还可能存在扭曲、脏污、遮挡、定向模糊、背景杂乱等问题,这都导致了文字识别的困难。

微软创始人比尔·盖茨曾经表示,“语言理解是人工智能领域皇冠上的明珠”。相比于性能趋于饱和的计算机视觉和语音识别技术,NLP 因技术难度太大、应用场景太复杂,目前的发展还差强人意,这一直是制约 AI 发展的瓶颈之一。目前的 NLP 的应用主要包括如下几种:(1)“机器翻译”,克服不同语言的障碍。(2)“情感分析”,判断表达观点和态度的正负面性。(3)“智能问答”,让机器人回答客户提出的各种问题。(4)“信息提取”,能提取文本中包含的关键信息并标准化。(5)“文本分类”,对文档资料进行归类整理。(6)“语言生成”,理解人类语言,自动生成文章摘要。^[5]目前 NLP 比较成熟的应用还主要停留在对情感分析、关键信息提取、文本分类等方面,还无法非常准确地实现机器翻译、智能问答和

语言生成,这主要是因为自然语言有上下文的语境、词义消歧、模棱两可的表达方式等困难,^[6]距离通过图灵所设定的“图灵测试”仍相差甚远。

2. 数据挖掘和大数据技术所需数据质量欠缺以及“信息孤岛”问题

首先,目前企业内外的数据有结构化数据(如财务数据、供应商和销售商资料等),也有非结构化数据(如企业官网及百科类数据、财经资讯数据、新闻网站、社交媒体、舆情数据、各类声音图像等),由于结构化数据便于处理,大多数公司将他们的分析集中于结构化的数据。但高德纳咨询公司的研究表明,80%~90%的企业数据都是非结构化数据,并且以每年 55%~65% 的速度增长,缺少了这些非结构化数据的支持,收集的数据就会碎片化。因此,如何把游离在决策数据链外的非结构化数据充分利用起来,是亟须解决的重点和难点。其次,互联网上充斥了各种各样的海量电子数据,互相矛盾,噪音严重,甚至会存在很多“脏”数据。中科院院士鄂维南指出,“数据质量非常差,数据里面噪音非常多,我国情况尤其严重,这是我国大数据发展的难题”。这些噪音数据不仅对决策无益,还会干扰计算机的判断,AI 如何对这些数据进行甄别是亟须解决的难题。最后,“信息孤岛”也是一个严重的问题。“信息孤岛”可以分为逻辑性和物理性两种,逻辑性是指不同部门由于理解的不同和沟通的缺乏,对同样的数据存储的内容、格式、长度、精度也不同。物理性是指由于部门壁垒或数据的隐私问题,数据在不同部门相互独立存储,独立维护,彼此间相互孤立,从而导致了数据的碎片化。以上这些问题都导致了数据挖掘和大数据技术所需数据质量欠缺。

3. AI 的综合技能和情商的缺失

我们也看到,当会计人员在欣喜于 AI 提供便利的同时,也会隐隐担心工作会不会被 AI 所取代。事实上,在 AI 的冲击下,被 AI 淘汰的只能是只懂得从事简单核算的会计人员,经济越发展,越需要复合型的、懂管理、擅沟通的会计人才,会计人员的转型和提升刻不容缓。正如缙世泽等学者阐述的那样,“高级的财务人员要结合宏观经济、整个行业的情况、尽可能站在最高的视野用更长远的眼光去分析企业的财务运营,以实现资源的最合理配置。”^[7]而 AI 很难有这种宏观的视野和各行各业庞杂的知识,如一些貌似和决策问题无

直接关系的经验、经历、性格,乃至宏观和微观环境的变化都可能影响最终的决策,毕竟经营决策本身是非常复杂的,但AI在算法设计时是很难把这些复杂因素考虑进去的,而且AI毕竟只是冷冰冰的机器,没有情商可言,缺少与他人沟通的能力和技巧,如AI可以战胜围棋高手,但它不会为此而发出来自内心的欢呼。因此,综合技能和情商的缺失决定了AI只能做中高端会计人员的助手。

四、AI变革中瓶颈问题的解决方案

虽然目前AI在未来会计变革中还存在很多困难,但其解决并非不可能,我们尝试着提出以下解决方案:

1. 强大的需求推动技术瓶颈的解决

纵观人类科技的发展,几乎都是由需求推动的。经济的发展和人口老龄化的矛盾促使各行各业对AI有着强烈的需求,这也是AI技术在近期蓬勃发展,关键技术瓶颈问题接连被突破的主要动因。根据MarketsandMarkets的研究,仅自然语言处理(NLP)市场规模就预计将从2019年的102亿美元增长到2024年的264亿美元。这充分反映了市场对AI技术的强大需求。我们相信,OCR和NLP等关键技术遇到的困难也能在这种强大需求下迎刃而解。

2. 通过标准化和建立信息平台等方式解决数据质量问题

一方面,廉价的存储成本使海量的数据存储成为可能,但如果数据的质量堪忧,其带来的作用可能适得其反。对于海量的非结构数据问题,需要用技术手段解决。虽然对如何管理、使用、快速价值化非结构化数据,当前并没有很好的解决方案,但在强烈的需求推动下,中国电子技术标准化研究院推出了《非结构化数据管理解决方案白皮书》,该白皮书在理论层面对解决非结构化问题具有标志性意义。白皮书立足于非结构化数据管理应用实践,结合DCMM国家标准体系框架,从宏观层面,提出了非结构化数据管理体系及能力成熟度模型。该白皮书为如何解决非结构数据问题提供了非常好的思路。至于网上充斥的海量“脏数据”问题,目前解决起来还比较棘手,我们建议采用循序渐进的方式解决,如目前可以考虑

从国家层面进行统一规划,在保证数据安全的前提下,建立具有各行业公信力的信息平台(已有的平台也要根据需求进行细化和改进),从而保证数据的可信度。

另一方面,为了解决逻辑性的“信息孤岛”问题,我们建议采用制定数据规范、定义数据标准、建设维护元数据的方式,从而减少数据的不兼容问题。如可以考虑采用以XBRL为代表的跨平台、可扩展的统一数据格式。由于格式的统一和XBRL自身的特性,使数据的无缝传送和计算机自动处理成为可能,处理效率得到了极大提高。至于物理性的“信息孤岛”问题,主要源于小农经济的部门封闭、信息私有的狭隘观念,可以考虑从行业层面,在保证信息安全的前提下,建立信息强制披露机制。

3. 加强对会计人才综合技能和情商的培养

AI不能完全替代人来完成会计工作是会计领域的共识,这就需要在培养会计人才时,除了培养学生的专业知识外,还应对学生的知识结构和视野格局进行拓展,如数据整合和数据处理能力,金融、税务、财政等相关外围知识的培养。同时,还应加强学生人际沟通和商业谈判等软知识的培养,只有这样,才能突出人相对于AI的不可替代性,让AI成为辅助会计人员决策的不可或缺的工具,而不是替代者。

五、结论

AI在会计相关领域的应用已经不可逆转,这主要是由两者各自特点所决定的:(1)AI中的大数据本身包含了会计领域相关数据。(2)AI更适应基于多变量的可描述规则,而企业会计准则、风险监控、财务分析等是可描述的,只要数据可以被识别,融合专家知识的各类算法都可以利用这些规则赋予计算机以智能。(3)财务会计中大量重复的、可标准化的工作,非常适合AI来模拟人工的判断和操作,^[7]这也决定了AI对未来会计的影响路径。其中,关键的AI技术主要有OCR、NLP、数据挖掘和大数据技术,虽然这些技术遇到了发展中的瓶颈问题,但在强大需求的推动下,我们依旧坚信,随着AI技术的不断发展,未来会计的相关领域可能会产生革命性的变化。

参考文献:

- [1] 林枫春. 财务共享服务中心发票扫描识别录入管理解决方案[EB/OL]. (2017-06-05) [2021-03-26]. https://www.sohu.com/a/146206956_397987.
- [2] 佚名. 大数据和数据挖掘之间, 主要有什么关系? [EB/OL] (2019-03-23) [2021-03-26]. <https://blog.csdn.net/duozhishidai/article/details/87877072>.
- [3] 韩向东, 余红燕. 智能财务的探索与实践[J]. 财务与会计, 2018(17): 11-13.
- [4] 彭启发, 王慧秋, 王海兵, 等. 会计人工智能存在的风险与对策研究[J]. 会计之友, 2019(5): 114-119.
- [5] 德勤. 智慧审计的七种武器之自然语言处理[EB/OL]. (2019-04-03) [2021-03-26]. <http://news.esnai.com/2019/0403/187968.shtml>.
- [6] 章成志. 财务报告的智能分析[R]. 合肥: MPACC 大数据+会计研讨会, 2018.
- [7] 缙世泽. 会计人的终结者还是开拓者——财务机器人对会计人的影响分析[J]. 财会研究, 2018(8): 77-81.

Study on Revolutionary Impact of Artificial Intelligence on Future Accounting —— From a Technical Perspective

HUANG Liang

(School of Finance and Economics, Jimei University, Xiamen Fujian 361021, China)

Abstract: Artificial Intelligence (AI) is currently a hot spot in computer sciences. Diversity and fast development of AI technologies make it essential to study which critical basic technologies are most likely to impact accounting and what revolutionary impact will be imposed on future accounting by AI. The study presents basic AI technologies such as OCR, NLP and Big Data on accounting, reveals their impacting paths on future accounting, and discusses possible bottlenecks encountered during the change and their solutions.

Keywords: AI; OCR; NLP; data mining; big data

(责任编辑:沈建新)