

论司法裁判的人工智能化及其限度

宋旭光*

摘要:司法裁判人工智能化是指机器能代替法官独立完成某些司法决策。目前主要有两种进路:一是基于显式编码、封闭规则之算法的专家系统;二是基于机器学习算法的预测分析论。法律专家系统虽有多数积累,但限度已显。大数据算法虽方兴未艾,但也同样难以成功:第一,司法裁判从本质上就无法根据既往数据被预测;第二,机器裁判颠覆了许多司法的基本预设,消解了人的自主权,违背了人类发明人工智能的初衷,而诸如推动类案类判、限制自由裁量等辩护理由都是难以维系的;第三,司法大数据现实上难以支撑算法裁判在技术上的实现。总之,人工智能不能也不应当成为法官那样的决策主体,更为务实的方向是去发掘其作为辅助工具的价值。

关键词:司法裁判;人工智能;机器学习;大数据;法律预测论

一、导论

司法裁判的人工智能化并不是新近出现的问题,在起起伏伏的数次人工智能浪潮中,相关议题早已经被摆上了学者的研究日程。进入新世纪,尤其是2010年之后,随着大数据时代的到来,以及图像、语音、文本识别技术特别是算法、算力的发展,以机器学习(尤其是深度学习)为代表的人工智能获得了新的生机,诸如驾驶、翻译、对话等过去认为只能由人类智慧才能做出的事情,现在部分已经可以由机器完成,这些技术被广泛应用在商品推销、出行导航等诸多领域。当然,也有了许多法律智能系统问世,例如, Lex Machina法律分析工具, Ross智能系统, LawGeex系统, KMStandards、Seal等合同分析软件, COMPAS、CSOSA、LSI-R等犯罪预测工具,^[1]等等。在我国,也有不少法院推出了相应的智能辅助办案系统,如,北京市高级人民法院的“睿法官”智能研判系统,^[2]上海市高级人民法院的“206”刑事案件智能辅助办案系统,^[3]以及最高人民法院的“类案智能推送系统”,^[4]等等。这些产品都展现出人工智能

* 深圳大学法学院助理教授,法学博士。本文系中国法学会2019年度部级法学研究自选课题“论司法裁判人工智能化的可能与限度”(CLS2019D01)的研究成果。

[1] 关于这些系统的介绍,分别参见:<https://lexmachina.com>,最后访问日期:2020年3月9日;<https://rossintelligence.com>,最后访问日期:2020年3月9日;<https://www.lawgeex.com>,最后访问日期:2020年3月9日;<http://kmstandards.com>,最后访问日期:2020年3月9日;<https://www.seal-software.com>,最后访问日期:2020年3月9日;<https://www.equivant.com>,最后访问日期:2020年3月9日;<https://www.csosa.gov>,最后访问日期:2020年3月9日;<https://mhs.com>,最后访问日期:2020年3月9日。

[2] 参见刘子阳:《机器人法官“小睿”后台自动立案》,载《法制日报》2016年12月14日。

[3] 参见崔亚东:《人工智能与司法现代化》,上海人民出版社2019年版。

[4] 参见罗书臻:《认真学习贯彻党的十九大精神 深入推进智慧法院建设》,载《人民法院报》2018年1月6日。

技术强大的功能和潜力。于是,许多人开始憧憬人类进入司法裁判智能化时代的未来场景,^[5]但也有人认为这种时代所蕴含的危机(诸如法律人的失业问题)忧心忡忡。^[6]

在这种背景下,本文打算讨论一个前提性问题,即司法裁判的人工智能化是否可能实现,它又可能面临着哪些难题?为了回答这一问题,首先需要明确所谓司法裁判智能化以及智慧司法究竟指的是什么,具备何种“智能”的机器才能被认为实现了这种智能化。与人工智能算法和司法裁判理论相对应,下文将分两种进路进行分析:一种是显式编码、封闭规则的算法,即法律专家系统;另一种是机器学习算法,依靠大数据分析实现对判决的预测。^[7]前一种进路已经有了数十年的讨论,虽然成果颇丰,但终未显露胜利的迹象,目前也是疲态已显。后一种进路则方兴未艾,野心勃勃。本文讨论的重点就放在后一种进路之上,若要实现这种大数据算法裁判,法官的裁判必须是可被预测的,依靠学习算法可以实现这种预测,且算法裁判是正当的。为此,本文的论证将分三个层面:第一,从本质上看,司法裁判是否可以预测;第二,如果裁判可以被预测,那么,这种预测结果是否应当被应用于司法裁判;第三,如果司法裁判可以被预测且算法裁判具有正当性,那么,学习算法是否可以实现这种预测,它在现实中是否可行。本文的结论总体上是悲观的,即司法裁判的人工智能化是难以实现的,除非法学范式发生了根本性的改变。

二、司法裁判人工智能化的界定

为了讨论的有效性,这里首先要界定一个问题,即达至何种“智慧”才算人工智能化,或者说,至少在本文的讨论中,“司法裁判人工智能化”所指的究竟是什么?

一种较为常见的观点是,将人工智能技术的任何应用,例如,将那些可以通过电子眼识别车牌号码并做出放行动作的系统,都称为“人工智能”。这显然不合适。严格来讲,这种作为辅助工具的人工智能,与诸如多媒体、工业机器人等其他技术手段,并无本质不同,其至多只是一种工具的改进而已。例如,过去,制定法与案例往往以纸质媒介的形式存在,“人力找法”相对繁杂;当法律数据库建立之后,以搜索引擎或算法作为工具,可以辅助特定人(例如法官或律师)更加低成本、高效率、高精度地进行数据的检索、搜集和整理。^[8]

从字面上来看,人工智能应当是对人类自然智能的一种比拟,是机器(或计算机程序)模仿人类的一种能力。因此,在许多人看来,人工智能化指的是机器可以拥有像人类一样的思维和情感能力。通常,人类是理性的,能够通过自己独特的大脑和神经系统对外间世界作出精确反应或进行抽象思考,并在复杂的情境中作出审慎的判断;人类是有复杂情感的,有同情心和同感力,虽然有时候也会有偏见或冲动。但若真有一天,一种可以完全像人类一样(或超越人类)的机器成为现实的话,那人类就不再是独一无二的,这种“硅基人类”不是在司法裁判上,而是可能在整体上代替我们这些“碳基人类”,而且这样的替代似乎还具有极大的正当性(因为机器能比人类做得更好!)。在这种情境下,司法裁判人工智能化的问题便不再重要了,问题更多在于人类愿不愿意接受这种智能化,或者人类还有没有能力抵抗这种被机器替代的风险。

实际上,就目前来看,人类的这种理性和爱恨情仇依然被认为是独一无二的。因此,更值得关注的是一种比这更弱的人工智能,即机器能够独立完成某些原本需要依靠人类智慧才能完成的决定或任务

[5] 例如,有人断言,在司法裁判的许多方面,“AI司法都表现出超越人类司法的计算分析能力”(李飞:《人工智能与司法的裁判及解释》,载《法律科学》2018年第5期,第32页),我们进入了“算法日益代替人类进行各种决策尤其进行司法裁判的时代”[高学强:《人工智能时代的算法裁判及其规制》,载《陕西师范大学学报(哲学社会科学版)》2019年第3期,第167页]。

[6] 例如,有人认为,人工智能可能会造成法律行业的大裂变。See John O. McGinnis & Russell G. Pearce, *The Great Disruption: How Machine Intelligence Will Transform the Role of Lawyers in the Delivery of Legal Services*, 82 Fordham Law Review 3041 (2014).

[7] See Jesse Beatson, *AI-supported Adjudicators: Should Artificial Intelligence Have a Role in Tribunal Adjudication*, 31 Canadian Journal of Administrative Law and Practice 307, 307 (2018).

[8] See Benjamin Alarie, Anthony Niblett & Albert H. Yoon, *How Artificial Intelligence Will Affect the Practice of Law*, 68 (supplement 1) University of Toronto Law Journal 106, 112-115 (2018).

(例如,关于某一案件的法律推理、司法裁判等),这才是对人类智能的模拟。在这个意义上,“决策智能化”最简单的标志便是看算法程序运行的结果,究竟是找到了一系列依然需要人类进行阅读、分析、权衡、选择、决策的数据,还是它已经通过数据分析就相关问题的解答给出了结论,即使这些结论仅仅被当作是初步的建议。后者才算得上“智能”。

总之,本文讨论中的“司法裁判人工智能化”所指的主要不是机器拥有某种类似甚至超过法官的智能,也不是说机器可以辅助人类完成某项任务,而是说机器可以代替法官独立作出某些裁判决定。^[9]那么,根据这样的定义,目前人工智能技术在司法裁判中的应用(例如,应用在司法裁判的某一阶段或某一任务中,提高司法效率,或者设计人工智能系统,为法官决策提供协助)所涉及的都还是作为辅助工具的人工智能,而与司法裁判的智能化无关。

三、司法裁判人工智能化的两种路径

既然是司法裁判的人工智能化,首先当然与对司法裁判的认识相关。目前,司法裁判人工智能化的努力,依然没有完全逃脱出传统裁判理论框架的拘束,虽然也体现出许多差异。

(一) 司法裁判的性质与人工智能化

一般而言,有关司法裁判的过程,有形式主义与现实主义两种不同风格的说明:^[10]前者将司法裁判看作是一种以法律规范和案件事实为前提的推论过程,法官是依据逻辑推论规则推出判决的;后者则更多将司法裁判看作是经验的慎思过程,法官是在道德、政策、直觉(或者前理解)等现实因素的影响下作出决定的。前者认为,法学的任务在于建构一种完美的法律体系,借助这种体系,每个案件事实都对应着一个唯一的正确答案,裁判者只需依据形式逻辑法则便可以根据规范和事实推导出这个唯一的正确答案;后者认为,法学家应该更为关注法律的实效,重视经济、社会、政治、道德甚至裁判者的主观偏好等现实因素在法官裁判中的关键角色,并努力根据各种现实因素,实现对于法官行为的准确预测。

在法律人工智能领域也有两种相应的算法:第一,显式编码、封闭规则的算法,通过法律专家系统实现对人类法律推理的模拟并将之应用于司法裁判的决策;第二,机器学习算法,通过大数据分析训练,发现人类司法裁判的内在规律,并将之应用于对未来裁判的预测。后面的讨论将以“自动判决机”与“裁判学习器”的隐喻来象征这两种不同方向,无疑,前者代表了形式主义风格的理论,而后者带有很强的经验主义色彩,虽然这两种进路与形式主义裁判理论和现实主义裁判理论并不是严格对应的。

(二) 基于形式主义推演论的“自动判决机”

早在“人工智能”这一术语出现之前,形式主义法学(或机械法学)就已经酝酿了一种裁判智能化的美梦。^[11]如马克斯·韦伯(Max Weber)所讲,法学界早有这种想象存在:法官是一个诉讼机器,只需要将当事人的诉讼要求以及诉讼费用塞进这样的机器,便可以根据法典中推演出来的理由得到相关的答

[9] 有学者将这种强人工智能定义为“具备自主意识而进行刑事司法语境中的案件判决”(参见胡铭、张传玺:《人工智能裁判与审判中心主义的冲突及其消解》,载《东南学术》2020年第1期,第214页),可能会有不妥之处:第一,机器是否可以具有“自主意识”,依然严重存疑;第二,机器一旦可以具有“自主意识”,便成了前文所说的“硅基人类”,其与人类的关系可能会被重新定义,所谓裁判智能化将变得毫无意义。本文将使用“独立”或“自动”来刻画机器决策的特征,人工智能体不需要具有意识,也无需具有自主性,它所需要的是不依靠人类的指导或帮助就能给出决策的正确结果,而这些决策通常是人类需要“动脑子”(人类智能)才能解决的。

[10] 参见陈景辉:《人工智能的法律挑战:应该从哪里开始?》,载《比较法研究》2018年第5期,第140页;[美]理查德·波斯纳:《波斯纳法官司法反思录》,苏力译,北京大学出版社2014年版,第123-145页。

[11] 例如,高翔:《人工智能民事司法应用的法律知识图谱构建——以要件事实型民事裁判论为基础》,载《法制与社会发展》2018年第6期,第72页。

复。^[12]通常,这种“自动判决机”的运行原理是建立在“法律的公理体系”之上的;^[13]如果法律体系可以被建构得像公理体系那样,法学家便可以像数学家那样工作,法学问题也可以像数学问题一样通过“计算”予以解决。这种对法律推理的机械性说明,在很大程度上激发了人们对于“自动判决机”的探索,只要机器能够遵循相应的思维准则,依此自动推出相应的结论,即实现智能裁判,似乎也不是一个难题。^[14]

这便代表了司法裁判智能化的第一种思路:基于逻辑系统或计算机程序的建模,来模拟人的推理方式,其典型成果便是各种知识库系统或专家系统。简单来说,这种专家系统的建立至少需要三个条件:第一,提前建立一个储存在计算机之中包括相关素材(法律或案例等)的数据库;第二,对这些素材进行技术处理,建立索引使其可以通过关键词或关键要素进行检索;第三,建立一个搜索引擎,可以对输入的素材进行检索并输出想要的结果。^[15]根据这种思路,只要能够建构一个可靠、完全的数据库,依据提前设置的推论规则,便可以为该领域的个案提供解决方案。因此,许多辅助性的专家系统应运而生,它们主要包括两种基本的类型:其一是基于规则的论证模型,例如TAXMAN系统;其二是基于案例的论证模型,例如HYPO系统。^[16]

在司法裁判人工智能化的努力中,很长一段时间,形式主义路径都占据主导地位。必须看到,在原来那个储存成本和计算资源比较稀缺的时代,这种路径的简洁性显然更符合现实的要求。只是这种美梦,不仅没有随着法律体系的发展和科学技术的进步,而变得越来越指日可待,相反,越是到了科技进步日新月异、法律体系越来越复杂化的今天,这样的美梦就越显得遥远。人的推理包括法律推理过程太复杂了,模拟这种推理需要设计极其复杂的算法程序,以及极其惊人的计算能力和计算量,这是目前的理论水平和技术水平所不能承受的。例如,1984年开始的Cyc(来自英文单词“Encyclopaedia”)项目,试图通过将信息编码成计算机可识别的形式来表示人类常识,但直至今日,这个项目依然“路漫漫其修远”。^[17]

而且,法律推理的实质性和价值性也无法借用形式模型予以完全展现,法律推理的关键因素恰恰是计算机程序的短板。^[18]如今,法律体系变得越来越庞杂,立法更新越来越快,社会变迁、技术革新带来了太多的新型、疑难问题,对于这些问题的解答,法学家尚无法达成共识,遑论人工建构的知识系统。

总之,对于形式主义方案而言,关键之处在于建构一套可以模拟人类法律推理的模型,以法律规则或案例数据库为基础,借助提前植入的算法程序,推出相应结果。其难题在于,法律推理的复杂性和评价性是计算机程序难以模拟的。不过,如果将专家系统局限在某些特定的尤其是形式化程度较高的领域,它依然可能发挥作用,正如法律智能系统专家阿什利(Kevin D. Ashley)所言,“尽管法律专家系统不再是主流范式,但依然在许多语境中被广泛地使用”。^[19]当然,之于实现司法裁判人工智能化的目标而言,这样的现状显然是不能令人满意的。

(三) 基于大数据预测论的“裁判学习器”

随着算法的革新、计算机能力的增强以及神经科学等学科的进步,机器学习尤其是深度学习的兴起带来了一种新的希望,它推动着人类的技术范式从“大定律,小数据”向“大数据,小定律”转移。^[20]

[12] 参见[德] 马克斯·韦伯:《论经济与社会中的法律》,张乃根译,中国大百科全书1998年版,第355页。

[13] 参见舒国滢:《寻访法学的问题立场——兼论“论题学法学”的思考方式》,载《法学研究》2005年第3期,第3—4页。

[14] 参见张保生:《人工智能法律系统的法理学思考》,载《法学评论》2001年第5期,第11页。

[15] See Richard E. Susskind, *Expert Systems in Law: A Jurisprudential Approach to Artificial Intelligence and Legal Reasoning*, 49 The Modern Law Review 168, 169—170 (1986).

[16] 参见冯洁:《人工智能对司法裁判理论的挑战:回应及其限度》,载《华东政法大学学报》2018年第2期,第26页;梁庆寅、魏斌:《法律论证适用的人工智能模型》,载《中山大学学报(社会科学版)》2013年第5期,第118—128页;周元、熊明辉:《如何进行法律论证逻辑建模》,载《哲学动态》2015年第4期,第99—104页;

[17] 参见<https://www.britannica.com/topic/CYC>,最后访问日期:2020年4月10日。

[18] 参见冯洁:《人工智能对司法裁判理论的挑战:回应及其限度》,第27页。

[19] Kevin D. Ashley, *Artificial Intelligence and Legal Analytics: New Tools for Law Practice in the Digital Age* 10 (Cambridge University Press 2017).

[20] 参见王飞跃:《人工智能:第三轴心时代的来临》,载金观涛等:《赛先生的梦魇:新技术革命二十讲》,东方出版社2019年版,第118页。

如果说专家系统还是由人对知识进行总结,然后“教会”计算机的话,那么,根据机器学习模式,计算机则可以通过大数据跟着人类学习,从经验中汲取智慧。^[21]机器系统不再是按照人提前设置的路线模仿人类思维,而是“从样例中学习”、“在问题求解和规划中学习”、“通过观察和发现学习”、“从指令中学习”。^[22]尤其是互联网的大规模应用以及存储成本的降低带来了庞大低廉的数据,规模远超任何一个专家知识库,只要有一个足够强大的搜索引擎,似乎就可以从这些数据中搜寻到想要的内容。通过对数据的“学习”,机器便可以洞悉其中隐藏的“规律”或“关联”,从而实现对于特定事项或行为的预测。

司法裁判人工智能化的另一种路径也因此显现了出来,即通过学习算法,让学习器(learner)在大数据中自主学习如何“像法官那样”进行司法裁判。这种进路便不可能将司法裁判看作是一种形式主义的机械裁判过程,而更多与现实主义法学的许多想法相对应,尤其是后者有关经验方法和预测论的内容。但与现实主义法学不同,它并不必然要质疑法律规范的证成作用,这种“裁判学习器”必然要预设的是,裁判是有规律或关联可循的,而裁判证成的核心基础究竟是规则还是经验,并不关键。

从这种进路来看,司法裁判智能化的关注点就“从模拟法律推理的外在逻辑形式进一步转向探求法官的内在思维结构”,^[23]关键词从“规则”变成了“规律”,从“逻辑”变成了“概率”,从“推理”变成了“预测”,其核心的要素便在于建构一种能够实现对判决(或裁定、决定)进行预测的算法模型。^[24]这一点是与许多现实主义者的想法对应的:对于法律思维的模拟必须看到社会主流道德观念以及其他非理性内在因素的影响,对此,统计学数据分析和经验方法就可以派上用场了。美国法学家霍姆斯(Oliver W. Holmes)一百多年前就指明了这一方向:“对理性的法律研究来说,在法律的故纸堆里皓首穷经之人(the black-letter man)或许眼下大行其道,运用统计学之人以及经济学的行家里手则引领未来。”^[25]在霍姆斯的论断中隐含着某种预测论,他“所指的规律,正是对法院将会采取的实际举措作出的预测,而不是什么故作高深的东西”,^[26]在他看来,法律人的“学习目标就是预测,即预测公共强制力通过法院得到施展的概率”。^[27]

这和机器学习的原理是类似的:以历史数据预测未来行为。就此而言,机器学习技术可以提供有力的支撑,帮助法律人提高预测的质量。正如普通人基于经验数据对天气的预测一样,即使人们可能无法知道风暴或雨雪形成的物理机制,但却可以根据云图、风向或者动物的行为判断它们来临的几率。在这个意义上,对于法官内在思维的探究,必须依赖于经验数据的支撑,而深度学习算法使得机器可以从大数据中自主学习。^[28]如果我们能够以一种成熟的语言解析技术让裁判学习器“读懂”海量的判决书(或其他法律文件),再通过学习算法让学习器从这些经验数据中发现特定的规律或关系并产生模型,从而依据模型为新的个案提供(预测)相应的解决方案,那么,就不仅实现了裁判的人工智能化,而且有利于推动类案类判、限制自由裁量。这种方案看起来野心勃勃,但是,通过下文的分析,笔者依然将得出一个悲观的结论,并为这种消极的答案给出一种可能的辩护。

四、大数据预测论与司法裁判的不可预测

首先要讨论的问题是司法裁判的可预测性。如果想要通过大数据分析实现对司法裁判的预测,最基本的前提便是司法裁判应当是可被预测的,但这个前提在理论上似乎并没有得到足够的支持。

[21] 参见周志华:《机器学习》,清华大学出版社2016年版,第1页。

[22] 参见周志华:《机器学习》,第11页。

[23] 张保生:《人工智能法律系统的法理学思考》,第12页。

[24] 参见王禄生:《司法大数据与人工智能技术应用的风险及伦理规制》,载《法商研究》2019年第2期,第102页。

[25] [美]霍姆斯:《法律之道》,姚远译,载《厦门大学法律评论》(2015年下半年卷总第26辑),厦门大学出版社2015年版,第166页。

[26] [美]霍姆斯:《法律之道》,姚远译,第159-160页。

[27] [美]霍姆斯:《法律之道》,姚远译,第157页。

[28] 参见王飞跃:《人工智能:第三轴心时代的来临》,第118页。

（一）大数据预测论的隐含前提：司法裁判的可预测性

正如前述，司法裁判智能化两种路径的背后都有相应的理论基础，尤其是有关司法裁判之性质的预设。例如，专家系统的核心预设是，人类思维有某种固定的或可模仿的模式，通过刻画这些思维特征，便可以建构出某种自动推理模型，其中，规则模型预设了法律推理是一种规则指引的演绎推理，而案例模型则预设的是一种以案例为基础的推理模式。

机器学习系统的背后也有其特定的预设：人类智慧指引的决定或行为是有特定规律或特征可循的，通过大数据分析，学习算法可以发现这样的规律或者特征，从而实现对这些行为或决定的预测。例如，阿尔法狗（AlphaGo）之所以能在围棋领域所向披靡，一个关键原因在于其背后有清晰的规则，学习算法可以通过数据分析找到决胜的模式。同理，这样的算法模型是否有助于预测法官决策或者代替法官作出决策，关键之处在于法官的决策必须是有规律可循的，或者是受到某些可学得特定因素决定的。这种主张同样可以在法律现实主义那里找到联系，例如，卢埃林（Karl N. Llewellyn）便明确地表达：“关键在于观察法律官员做什么，他们如何处理纠纷或者其他任何事务，以及观察他们的所作所为，从中寻找某种独特的规律性——这种规律性使人们有可能对法律官员及其他官员今后的所作所为作出预测。”^[29]

总之，如果我们能够“按照既定目标，对判决书、案件卷宗、庭审视频等非结构化、半结构化数据进行探索和分析，揭示隐藏的、未知的规律性，并进一步将其模型化”，^[30]那司法裁判智能化便是可能的；但如果司法裁判活动并不存在这种规律或关系，即它是不可被预测的，那么，学习算法能否应用于司法裁判的预测，就成了一个伪问题。

（二）为什么司法裁判无法根据大数据分析被预测？

正如前述，基于大数据算法预测所依赖的是分析对象的规律性，例如，根据云图等信息能够预测天气，表面上是依靠经验观察的数据，但核心的内因还是在于天象规律的存在，符合什么样的条件，就会有何种天气。但是，司法裁判的作出却不是由经验数据支撑的，它是规范性的。即使经验数据可以用来解释影响司法裁判的某些因素，但裁判的正当性却并不能依赖于既往案件的规律性，除非存在一个更高的规范，即新的案件应该按照过去类似案件的裁判方案进行处理。于是有人自然想到了判例规则，但这个更高规范所要求的是严格地、机械地遵循先例，而现实法律系统中的“遵循先例”从来不是机械地遵循历史先例，而是由法官对它们作出区分并确定先例中的判决理由是什么，从而决定是否遵循它，核心依然在于法官对于先例的判断。^[31]

退一步讲，即使司法裁判是依赖于经验的，即使存在一个按照既往案件之规律进行裁判的更高规范，但这种规律性却也很难通过大数据体现出来，其中最大的困难在于裁判的经验与知识有许多都是默会的、隐含的，有些内容也许依据理性可以从实践中推导出来，但却不会体现在大数据之中。^[32]换句话说，案件裁判的依据和过程是未必体现在判决书或者其他材料之中的，法官们会隐藏自己内心的判案“逻辑”或“规律”，从大数据资料中未必能够分析出这些内在的思维结构。

更为重要的是，法治是规则之治，作为裁判正当依据的只能是规则而不能是规律。^[33]即使法律规则来自于生活世界的经验，即使某些依据特定规则作出的决定可能具有一定的规律性，但在司法裁判中，起关键证成作用的却必须是规则。当然，根据法律规则作出的裁判通常也应当具有可预测性，但这里所讲的是依据规则作出司法裁判的应然可预测性，而不是依据规律的经验可预测性。规范性不等于规律性，规范意义上的可预测性也不意味着司法裁判是可以实际被预测的。当规则发生变化时，所有的规律都变得不那么重要了。正如德国法学家基尔希曼（Julius Hermann von Kirchmann）所言，“立法者的三个

[29] [美] 卢埃林：《荆棘丛：关于法律与法学院经典演讲》，明辉译，北京大学出版社2017年版，第7页。

[30] 王禄生：《司法大数据与人工智能技术应用的风险及伦理规制》，第102页。

[31] 参见叶楦平：《遵循先例原则与英国法官的审判思维和方法》，载《比较法研究》2015年第1期，第16页。

[32] 参见潘庸鲁：《人工智能介入司法领域路径分析》，载《东方法学》2018年第3期，第114-115页。

[33] 参见王洪：《制定法框架下的判决论证模式》，载《比较法研究》2019年第2期，第146页。

更正词就可以使所有的文献成为废纸”。^[34]

那么,有人可能会反驳道,法官并不是或主要不是依据法律规则作出裁判的,他不可避免地会受到偏好、直觉或者道德观念的影响,正如许多现实主义者所主张的那样,那么,这种预测论是否能够成功呢?回答这一问题的关键便在于这种实质立场或非理性因素影响下的行为是否真的是有规律可循的。必须明确的是,非理性因素指引下的行为也未必是不可预测的,例如,人们的消费行为并不总是理性的,买卖选择往往会受各种偶然因素的影响,但大数据分析同样可以给出一种概率性的预测。但是,只要这种概率性的预测根据的是历史数据,就不能被强制作为现在决定的依据。道理显而易见,张三之前经常买法哲学的书,据此能够归纳出张三大概是喜欢法哲学书的,却无法预测出他现在想买它;音乐数据显示30岁以下的人群多选择听流行音乐,据此可以归纳出多数年轻人都喜欢听流行音乐,却无法据此推出16岁的李四也喜欢它。同样,司法决定往往是需要基于特殊的情境现时作出的,而预测所基于的数据却是历史数据,其中必然有某种推论的跳跃。但关键的困难在于,偶然因素影响下的司法决策情境是不断变化的,这种概率性的预测结果通常只能作为一种参考,而不能作为现在决策的正当化依据。“数学模型的本质是基于过去的的数据推测未来,其基本假设是:模式会重复。”^[35]如果我们不能确定司法裁判模式会不断重复,即使根据既往的数据可以做一些成功率较高的预测,数学模型也不能被用于司法裁判。

预测论难以成功的理由还在于,司法裁判具有评价性和价值性。大部分法学家都会承认,在司法裁判的过程中不可避免地会涉及到价值判断。例如,美国法学家德沃金(Ronald Dworkin)便明确表达了这一点:“法律推理意味着把特定的、个别的法律问题,置于法律推演诸原则或者政治之道德性的广大网络中来加以考察。实际上,除非你已经通过或者愿意通过一个巨大的、由诸多复杂的原则所构成的、贯通全局的理论系统来进行思考……否则你就没法对法律问题的正确答案进行思考。”^[36]因此,这里的问题又变成了实质的价值判断是否能够通过大数据分析予以预测。

价值判断关涉的主要是规范命题,预测关涉的是事实命题。^[37]一般来说,学习算法依靠训练数据集所学得的往往是经验上的相关关系或规律,它自身无法作出价值判断,除非可以将规范问题转化为某种经验问题予以处理。例如,以关于某种价值判断的既有数据作为训练数据集,让学习器从大数据中学得如何预测这种价值判断。但是,即使这种预测的准确率很高,但它依然会受困于前文所讲的问题,即规律仅仅是规律而不是规范,存在推不出合理,大概率未必真正确。由于评价性工作无法通过大数据分析予以解决,就此而言,司法裁判智能化再次遇到阻碍。正如美国学者桑斯坦(Cass R. Sustein)所言,“我相信,强的版本是错的,因为它没有抓住类比推理的核心特征——它不可避免地评价的、价值主导的特点”。^[38]在他看来,遵循先例的类比推理并不是形式性的,依据案例进行推理的关键点并不在于寻找更为相似的案例或作出更多的区分,也不在于找出相关的相似点与不同点,而是在于识别一个可以证成这种关于相似点或不同点之主张的原则,而寻找这一原则的任务是一项评价性工作,因此,结论显而易见,除非人工智能可以独立作出好的价值判断,否则它便不能从事类比法律推理。^[39]

总之,除非关于司法裁判(本质特征、功能定位等)的观念产生了根本性的改变,或者人工智能对于司法裁判的性质构成了根本性的挑战(例如出现了超级“硅基人类”),否则,司法裁判在本质上便是无法被智能化的。既然当前人工智能之于法律领域的应用依然没有超出传统裁判理论的范畴,那它必然要受困于传统裁判理论的那些难题。^[40]

[34] [德] J.H.冯·基尔希曼:《作为科学的法学的无价值性》,载《比较法研究》2004年第1期,第146页。

[35] [美] 凯西·奥尼尔:《算法霸权:数学杀伤性武器的威胁》,马青玲译,中信出版集团2018年版,第31-32页。

[36] [美] 罗纳德·德沃金:《身披法袍的正义》,周林刚、翟志勇译,北京大学出版社2009年版,第58页。

[37] 参见陈景辉:《人工智能的法律挑战:应该从哪里开始?》,第141页。

[38] Cass R. Sustein, *Of Artificial Intelligence and Legal Reasoning*, 8 University of Chicago Law School Roundtable 29, 31 (2001).

[39] See Cass R. Sunstein, *Of Artificial Intelligence and Legal Reasoning*, 8 University of Chicago Law School Roundtable 29, 32 (2001).

[40] 陈景辉:《人工智能的法律挑战:应该从哪里开始?》,第140页。

五、机器裁判的正当性难题

退一步讲,假设司法裁判是可以预测的,那么,这是否意味着将这种预测结果自动作为裁判结论就是正当化的?或者,更进一步来说,如果机器法官比人类法官做得“更好”,机器就真的能够代替人类裁判案件吗?

(一) 大数据算法应用的辩护理由之反驳

将大数据分析结果应用于司法裁判,最显明的一个优点可能在于,它可以保证判决的形式公平(同等情况同等对待,不同情况不同对待),“同案同判”或“类案类判”成了极佳的辩护理由:“当面对特定案件时,人工智能不但有能力识别出同类判决的海量数据,并且将它们进行合适的分类,然后在此基础上分析、总结出基本的模型,然后再将所获得的模型应用到待决案件中,最终实现类案类判的结果”;^[41]“同案同判的司法公正由此得以重塑”。^[42]但正如前述,法治是规则之治、理由之治,而不是规律之治,由于机器学习算法只能根据从大数据中学得的特征或规律建构模型并给出结果,那么,将这一结果进而当作裁判结论,就不是依法裁判,而是依照司法大数据中隐藏的规律进行裁判,但这些规律的合法性和正当性又是无法保证的。正如德沃金所言,一个正确的判决至少要满足两个维度的要求:符合要求,即它与过去的决定是相一致的;正当化要求,即它是通盘考量后的最佳答案。^[43]算法裁判最多只能做到当前判决与历史案例的一致性,却无法将正当化因素纳入决策过程。判决是个体化的、可能因人而异的,但是基于大数据的预测却是对于某一类人、某一类事的,它可能无法处理个案适用中极端不正义或不合目的的情况。而且,算法模型背后反映的是数据之间的相关关系,而非因果关系,它可能会以因果关系或者客观规律的面孔掩盖其背后的许多价值预设甚至偏见。^[44]因此,预测分析学即使可以保证类案类判(捍卫法的安定性),但也可能与实质的正义与道德要求相冲突。^[45]

当然,必须承认,如果只是将司法大数据作为数据库使用,通过特定算法搜索符合当前案件裁决的既往类案,推送给法官,正如最高人民法院的“类案智能推送系统”一样,当然可以提高司法效率。但这和“司法裁判人工智能化”并没有太大关系,它只是一种法官可以选择使用的工具而已。

与此相关,支持算法裁判的另一个理由是,算法裁判可以限制甚至消除人类法官的自由裁量空间,因为它依据的是既往数据中的客观规律。例如,美国许多州之所以对采用量刑辅助系统或犯罪风险评估系统持有某种积极态度,其中重要的理由就在于以算法限制并代替司法裁量。^[46]不难看出,这种想法的背后隐藏着对司法裁量的某种负面预设,即它是武断的或不受理性控制的,因此应当进行限制。但是,已经有许多法学研究表明,裁量的不确定性未必一定是负面的,有时候恰好能够借此在合法性空间中注入新鲜内容以实现法律内容的更新,协调法律条文的僵硬性与滞后性问题。^[47]而且,虽然裁量权的运用确实依赖于法官的意志,但它并非绝对自由的,它的行使必须符合法治原则,法官往往被要求依据道德价值或经验常识作出决定并给出相关理由。

况且,大数据分析是否真的能够消除这种裁量空间,也是一个未知数。裁量的行使要求立足于当前

[41] 陈景辉:《人工智能的法律挑战:应该从哪里开始?》,第140页。

[42] 李飞:《人工智能与司法的裁判及解释》,第33页。

[43] 参见[美] 罗纳德·德沃金:《法律帝国》,许杨勇译,上海三联出版社2016年版,第182-183页。

[44] See Emily Berman, *A Government of Laws and Not of Machines*, 98 Boston University Law Review 1278, 1283 (2018).

[45] 参见[德] 阿列克西:《法的安定性与正确性》,宋旭光译,雷磊校,载《东方法学》2017年第3期,第134-136页。

[46] See Pretrial Detention Reform Workgroup, *Pretrial Detention Reform: Recommendations to the Chief Justice*, Oct. 2017, https://newsroom.courts.ca.gov/internal_redirect/cms.ipressroom.com.s3.amazonaws.com/262/files/20179/PDRReport-FINAL%2010-23-17.pdf (accessed May 27, 2019).

[47] 有学者提出,应当根据自由裁量之于相关法领域的价值而作出区分,例如,在公权力机关享受大量的裁量权的领域,大数据预测工具是有价值的,但在那些自由裁量受到宪法或其他法律严格限制的地方,机器学习预测的应用可能会威胁到法治本身的运行。See Emily Berman, *A Government of Laws and Not of Machines*, 98 Boston University Law Review 1278, 1278 (2018).

的情境,而现时情境的复杂性和多变性可能是历史数据所不能涵盖的;它有时候需要的是一种创造力,一种实践理性,或者审慎的道德考量,但学习算法却只能处理经验数据的问题,即使它在表面上减少了自由裁量的空间,但这种裁量结果的合法性、正当性依然不能保证。更重要的是,依据算法进行裁量排除了人作为主体根据具体情境作出价值判断的权利,限制了人的主体地位。^[48]由此,将人类法官的裁量权交给机器行使,无论是否有实效,都是难以接受的,这一点尤其重要。

除此之外,算法裁判还隐藏着许多可能威胁法治的预设。例如,算法结果的正确性对大数据的完全性和正确性提出了极高的要求,但没有任何一种程序能够将所有的相关数据都纳入分析,因此完全性总是无法保证的;况且即使训练数据是完全的、精确的,也无法避免大数据分析给出的最终结论依然可能是不正确的、有偏见的,因为数据本身就隐藏着某种错误;而且,这些数据看起来越值得信赖,那么,这种带有偏见的决定似乎也就越容易被接受,并进一步加深这种错误。例如,对某移民城市犯罪数据的分析显示,来自X地的移民犯罪率远超来自其他地方的移民,监管机关很容易就依此作出推论,X地移民应当被严格监管,^[49]这就意味着X地犯罪行为被发现的几率会变得更,并进一步体现在数据上,以此导致恶性循环。在这个意义上,大数据决策不可避免地会面临且放大“多数决”的难题:一方面,对于庞大的大数据基数来讲,“少数异见”可能因为极小的占比和概率而被忽视,成为大数据决策中的“他者”;另一方面,当大数据决策被适用于所有人或所有相关行为时,即使它只有极小的误差(极小的占比或概率),也意味着大量的人或行为会因此而遭到不正确的对待。^[50]

综上所述,之所以反对算法裁决,并不在于它不能推动“类案类判”,而是在于它至多只能满足“同等对待”的形式要求,却忽视了判决的正当性要求;并不在于它不可以限制裁量,而是在于它可能会错误地限制裁量。总之,当机器不是在依据法律规则而是在依据规律进行裁判时,它会破坏法治的许多预设,人工智能给出的预测方案,最多只能作为一种说服论据,而不是权威论据,最多只能作为一种实践问题的建议,而不是理论上的正确答案。^[51]

(二) 机器代替人类裁判的正当性批判

假如运用大数据算法可以实现对裁判结果的预测,而且还具有种种优势,那是不是意味着机器就可以代替人类裁判案件呢?本文的回答依然是否定的。理由是多层面的,下文将指出,机器裁判至少违反了目前法学范式中的两个核心预设:其一,法学是实践智慧,司法裁判权是人(法官)的权力,而非机器的权力;^[52]其二,法治是理由之治,在裁判过程中法官应当保障当事人的参与权并给出有理由支持的决定,这是机器裁判做不到的。

首先,法学是一门实践智慧,司法是一种人事,司法决策依赖于人的实践理性,法律案件只能以特定的程序通过中立的第三人(法官)运用法律规范以及其他的人类知识(例如正义标准)来进行评价,只要这样的范式不发生变化,以机器代替法官作为决策的主体便是不恰当的。^[53]司法实践的知识和技能(特别是其中的默会知识)往往是在司法实践中逐渐学会的,这种学习过程难以被形式化为计算机语言输入电脑,也不能从大数据中分析得出。正如美国一位法官所强调的那样:“技术无法取代法律实施中检察官和辩护律师所拥有的司法知识、经验以及专业技术的深度。”^[54]法官裁判案件不仅仅需要法律知

[48] 参见陈景辉,《人工智能的法律挑战:应该从哪里开始?》,第141页。对这种观点的反驳,参见Aziz Z. Huq, *A Right to a Human Decision*, 106 Virginia Law Review (forthcoming 2020), <https://ssrn.com/abstract=3382521> (accessed April 13, 2020)。

[49] 实际上,“X地移民犯罪率高”是大数据分析揭示的规律,它推不出“X地移民应当作为被严格监管的对象”的结论,至少还应当再补充一个规则:“应当对犯罪率高的特定群体进行严格监管”。

[50] See Emily Berman, *A Government of Laws and Not of Machines*, 98 Boston University Law Review 1278, 1322-1323 (2018)。

[51] See Lawrence B. Solum, *Artificially Intelligent Law*, 1 BioLaw Journal - Rivista di BioDiritto 53, 55 (2019)。

[52] See Aernout Schmidt, *Ought Computers Adjudicate?* in *Liber Amicorum ter gelegenheid van de 60e verjaardag van prof. dr. Jaap van den Herik* 133, 135 (J. Donkers, L. Mommers, E. Postma & A. Schmidt eds., Maastricht: MICC - Universiteit Maastricht 2007)。

[53] 参见潘庸鲁,《人工智能介入司法领域路径分析》,第112页。

[54] Quentin L. Kopp, *Replacing Judges with Computers Is Risky*, Harvard Law Review Blog, (Feb. 18, 2018), <https://blog.harvardlawreview.org/replacing-judges-with-computers-is-risky/> (accessed May 27, 2019)。

识,还需要对社会效果的评价,甚至直觉、感觉、良知等。没有任何两个案件(或者事实要素)是完全相同的,对这些案件或事实进行区分所需要的往往是“人类的判断以及鲜活且自然的同感心”,^[55]这两项即使不是人类与生俱来、独一无二的力量,也是一种人工智能难以模拟的洞察力和情感。

但是,这种观点也许会遭遇一些反对意见:为什么在司法判决中人的判断和情感就是正当的,而机器所运用的学习算法就是不正当的?如果未来某一天机器也能模拟人的情感,那是不是意味着算法裁判就是正当的?这便涉及到一个更深层的问题,即人与技术的关系。传统上来讲,技术往往被当作人类为实现某一目标而使用的一种工具,但如果机器能在某些场合或某些领域中展现出人类智慧的话,那么坚持“人类中心主义”还是正当呢?就这一宏大问题,无法在此详细讨论。这里要说的是,反对机器裁判主要不是因为机器的裁判不如人的裁判更好,更重要的一个理由是,一旦它犯错,会导致无人负责甚至无人纠错的情况出现。况且机器决策具有标准化的特征,当它在某个具体情境中出现错误时,可能就意味着所有类似情境都会出错。而人的决策往往是个体化的,它不具有标准程序那样的扩散性。正如前述,司法裁判依赖于人的理性和意志,而理性的有限性和意志的薄弱性,一方面意味着,裁判者难免会在特定案件中遭遇疑难情况,甚至可能作出不正确的裁判,但另一方面也隐含着裁判者有着相应的应尽义务和责任,无论是纠错的义务,还是惩罚的责任,抑或仅仅是道义上被谴责的可能,都代表着拥有自主意志的人要对自己的决定负责。与人不同,机器是自动地而不是自主地作出决定,它是“无情的”,没有价值意涵,没有反思性,也无法知道自己已经“犯错”,更无力去纠正这种错误,甚至可以说,在它那里并没有错误一说,因为“错误”也必须用正确的代码表示才有可能被它识别。况且人们在多大程度上能够像容忍人类自身的错误那样容忍机器的毛病,这也是存疑的。因此,即使机器在某些领域能够代替人类作出更为精准的决定,即使我们可以给机器设置特定的“道德算法”,但算法程序还是可能会有缺陷,在它出现错误时,依然无法成为责任的承担者。^[56]

另一方面,算法裁判本身的证成理由只能有一个,那便是依据历史数据裁判是正确的,裁判模型会反复运行,但正如前述,这种正确性是无法保证的。即使它可以比人类做得更好(更正确),但由于机器裁判的过程并不是完全透明的、公开的,法官、律师和当事人都无法探知这一过程,机器也不会给出正当理由阐明裁判结果的正当性,这种正当性是无法被传递给人类的。这样的话,不仅司法过程的公开性、程序性、可说明性等传统价值遭到了破坏,甚至司法本身的正义价值(至少是程序正义)也被掩盖了。

而且,虽然司法应当具有谦抑性,但并不代表它排斥创新。如果没有法官通过裁判在司法实践中更新人们对于法律条文的认识,则很难想象诸如1789年《美国宪法》、1804年《法国民法典》在制定了一两百年后是如何跟上时代潮流的。只是指望立法者及时更新权威法律文本,是不现实的,法律的发展进步所依靠的更多是司法的创新。

更为重要的是,一旦形成对于算法程序的依赖,最终的结果可能是:理由不再重要,重要的是历史统计数据,论证不再重要,重要的是对过去的模仿。^[57]这种对于历史的模仿会腐蚀人类独立自主的思考能力,正如美国联邦最高法院首席大法官罗伯茨(John G. Roberts Jr.)所言,“我所担心的并不是机器开始像我们那样思考……我担心的是我们开始像机器那样思考”。^[58]一旦如此的话,人类实质上就以服从历史模型的方式臣服于机器,机器便以尊重历史经验的表象实际统治了人类,而裁判机器便成为这种统治模式最核心的一个“部件”。除非人类能忍受这种被统治的状况,否则机器裁判便是不可能被接受的。

总之,自主是人类的根本特征,面向某种未决状态,人类有自主思考、自主作出决定的权利,同时人

[55] Quentin L. Kopp, *Replacing Judges with Computers Is Risky*, Harvard Law Review Blog, (Feb. 18, 2018), <https://blog.harvardlawreview.org/replacing-judges-with-computers-is-risky/> (accessed May 27, 2019).

[56] 参见刘宪权:《人工智能时代机器人行为道德伦理与刑法规制》,载《比较法研究》2018年第4期,第42-43页。

[57] See Mireille Hildebrandt, *Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics*, 68 (supplement 1) University of Toronto Law Journal 12, 28 (2018).

[58] Deborah Cassen Weiss, 'Beware the Robots,' *Chief Justice Tells High School Graduates*, ABA J. (June 8, 2018), http://www.abajournal.com/news/article/beware_the_robots_chief_justice_tells_high_school_graduates (accessed June 9, 2019).

类还要因此肩负某种面向未来的责任,将人类自己作出选择、决定的权利和责任完全交给机器,显然并非人类发明人工智能的初衷。^[59]

六、大数据预测论的现实困境

还有一个问题需要回答,即如果司法裁判是可以被预测的,那么,这是不是意味着大数据算法就能够实现对司法裁判的预测呢?最后讨论这个技术难题,一方面是因为回答这个问题非常困难,技术是会不断(以指数级别的速度)进步的,即使今天不能以大数据算法实现对裁判结果的精准预测,不代表未来也不能,即使是人工智能专家,想必也无法对此作出准确的预测,况且笔者不是技术专家,无论作出何种讨论,都不会被当作权威的论述;另一方面,只要裁判是不可预测的,而机器裁判又不具有正当性,那技术问题的讨论实际上就变得不那么重要了。因此,在这里,只就实现这种预测分析论的现实困难展开一些基础分析。

一般来说,要想通过机器学习程序实现对裁判结果的准确预测,至少要满足三个条件:第一,人类的司法实践持续不断地被转化为可读数据储存起来,随着司法大数据越来越多,覆盖范围越来越广,质量越来越高,最终能满足学习算法对训练数据的需要;第二,深度学习所需要的算法与算力不断优化,文本解析等技术更加成熟,最终能够对海量的权威法律数据进行分析;第三,运用大数据算法预测裁判结果在法律共同体中获得认可(或具有可接受性),至少法律官员对这种运用要抱持一种积极的态度。^[60]但就目前而言,这三个条件几乎都没有被满足。

第一,司法领域难以提供学习算法所需要的训练数据。大数据之“大”不仅在于数量之庞大(volume),更是在于它的来源之广博(variety)、更新之迅速(velocity),或者变化之多端(variability)。^[61]即使司法实践中存在着一些在小规模数据中无法显现却在大数据中可能发现的“小定律”,但司法领域中也并没有如此庞大的、高质量的、多元化的数据支持这种学习算法的运行。与许多能够产生大量数据的领域不同,司法裁判领域的相关数据可能数量依然不够大,来源依然比较有限,多样性也未必足够。例如,国内大部分相关项目都是以“中国裁判文书网”(http://wenshu.court.gov.cn)公布的案例作为数据库基础的,截止到2020年4月11日16时45分,该网的文书总量已经达到了90700573篇,而且还在以极快的速度增加,不过,这些数据可能依然算不上“大”。更为关键的是,这些数据的质量是参差不齐的:有些类型的案件(特别是职务犯罪类、行政诉讼类案件)上网率不足,造成数据类型的结构性偏差;许多文书缺少说理部分,裁判的依据和推论过程等关键内容都没有展现;许多实质上决定裁判结果的数据无法体现。而且,因为它们是天生的缺陷,这些粗糙的数据还难以通过法律专家的清洗而得到质的改善。大家都知道,大数据算法更严格地受到“错进,错出”(Garbage in, garbage out)原则的限制,当不正确的、不充分的、过时的数据成为训练数据时,学习算法所发掘的便可能是不正确的特征,并因此获得不正确的结论。

除此之外,司法裁判有时候还会涉及到许多专业问题,诸如医疗事故、商业纠纷,如果要想通过大数据分析实现裁判的智能化,那么此类问题必须一并予以考虑,关于它们的判断也需要其所在领域预测分析学的支持。一方面,如何使用这些领域的数据库进行裁判是一个难题,另一方面,如何将这些数据与法律数据进行整合又将是另一个难题。

第二,司法裁判的数据并不容易被计算机“读懂”。众所周知,计算机所能阅读的仅仅是二进制的符号,算法是建立在这些符号的运算而非意义的生成之上的,因此,文本或语音中的以自然语言表述的内容,必须被转化为计算机语言,才能成为有用的数据。虽然目前借助文本解析技术,理论上计算机已

[59] 参见韩水法:《人工智能时代的人文主义》,载《中国社会科学》2019年第6期,第43页。

[60] See Benjamin Alarie, *The Path of the Law: Towards Legal Singularity*, 66 University of Toronto Law Journal 443, 455 (2016).

[61] See Dwight Steward & Roberto Cavazos, *Big Data Analytics in U.S. Courts: Uses, Challenges, and Implications* 2-3 (Palgrave Macmillan 2019).

经可以实现对文字的“阅读”,但“阅读”判决书以及其他法律文件却依然不是一件简单的事情。按照言语行为理论,某些话语(包括法律话语)并不仅仅是在表达意义,更是在做出某种行为,^[62]作为言语行为的法律话语是负载特定意图的,它们的意义往往是由其与其他话语之间的推论关系所决定的,其中还包含了许多只可意会不可言传的内容,诸如人的意识与反思以及话语共同体所共享的许多默会知识,都是计算机无法阅读或无法作为数据储存的。^[63]因此,只从语形上解析判决书可能是远远不够的,脱离了具体的语境,这些文字的意义是不能被理解的。^[64]

而且,与日常话语相比,法律话语还具有另外一些难以化解的难题。例如,虽然判决书有一定的结构性,但每一个法官使用的术语、写作风格都未必是相同的;法律话语中充满了诸如“公序良俗”、“善意第三人”等价值语词以及诸如“合理期待”、“可预期的”等高度模糊性的语词;严格来讲,每一个案件都是不同的,对任何一个案件之裁判结果的预测都可能会涉及到数以万计的类似案例,它们的判决依据、推论过程可能是各不相同的,将它们相互对比的组合数量及其产生的计算量非常大;一旦分析结果给出多个可能选项,在多个选项之间的选择往往还需要实质评价,而不仅仅是概率问题。这些都进一步加剧了司法数据的“阅读”难题。

第三,一般来说,只有在运用大数据算法进行裁判既是可行的又是可欲的时候,它才可能逐渐在法律官员以及社会大众中得到承认,最终成为司法判决实践的一部分。^[65]但这一切似乎都还没有发生。首先,深度学习等机器学习算法有许多不可探知或无法检验的内容,使其难以令人信任,例如,学习算法的运行有所谓“黑箱”存在,且可能出现诸如将黑人预测成猩猩等不可预知的“坏案例”(badcase);其次,出于对商业秘密或其他知识产权的保护,程序的源代码、训练数据与程序等信息可能无法公开,因此,在知识产权保护 and 算法透明性上我们面临着二选一的难题;再者,大部分法官和律师都不是程序员,基于自身知识的局限性,即使有公开的说明书,他们也很难读懂其背后的机制,正如美国波斯纳(Richard A. Posner)法官所言,法官对于科技几乎一无所知;^[66]最后,算法裁判的正当性和独立性依然可能为各种算法歧视问题所困,不仅学习算法的特征选择可能会受人之固有偏见的影响,而且表面客观的训练数据背后也可能隐藏着歧视因素。^[67]总之,虽然算法程序的设计者往往并非公共机关,但一旦涉及到司法裁判,机器也应当像政府机关一样保持信息的公开性和决策过程的透明性,^[68]但是算法的不透明性会让裁判过程变得无比神秘,法律人知识结构的缺陷也让算法裁判无法得到有效监督。基于这些理由,算法裁判可能依然很难获得法律共同体的认可,这也就意味着机器裁判的合法化过程依然任重道远,在这之前,包括法院在内的任何司法机关都不可能以算法代替人类智能进行决策,即使算法裁判在技术上是可行的。

最后,回到目前的技术现状来说,人工智能可以为法官裁判以及律师实务提供的帮助还是比较有限的,例如,虽然Ross智能系统已经可以在破产、知识产权、劳动等领域运用自然语言处理系统进行法学研究, LawGeex系统可以帮助公司法务对合同进行编辑和整理, Beagle系统可以帮助外行人对合同进行审查和管理,但是它们所涉及的几乎都是一些更为机械化、形式化的法律领域,所能处理的也多是简单的、无须复杂法律技能的事务。^[69]从本质上来看,人工智能系统在这些事务中展示出来的“智慧因素”并不多。

综上所述,由于司法领域的特殊性,司法数据的数量、质量以及它之于计算机的“可读性”都还不

[62] 诸如以言指事(locutionary act)、以言行事(illocutionary act)、以言成事(perlocutionary act)。See J.L. Austin, *How to Do Things with Words* (2d ed., Harvard University Press 1975); John R. Searle, *Speech Acts: An Essay in the Philosophy of Language* (Cambridge University Press 1969).

[63] See Mireille Hildebrandt, *Law as Computation in the Era of Artificial Legal Intelligence: Speaking Law to the Power of Statistics*, 68 (supplement 1) *University of Toronto Law Journal* 12, 26-27 (2018).

[64] See Robert B. Brandom, *Making It Explicit: Reasoning, Representing, and Discursive Commitment* xiii (Harvard University Press 1994).

[65] 参见郑戈:《算法的法律与法律的算法》,载《中国法律评论》2018年第2期,第78页。

[66] 参见[美]理查德·波斯纳,《波斯纳法官司法反思录》,苏力译,第92页。

[67] 参见郑智航、徐昭曦:《大数据时代算法歧视的法律规制与司法审查》,载《比较法研究》2019年第4期,第112-113页。

[68] See Alyssa M. Carlson, *The Need for Transparency in the Age of Predictive Sentencing Algorithms*, 103 *Iowa Law Review* 303, 303 (2017).

[69] See Sean Semmler & Zeeve Rose, *Artificial Intelligence: Application Today and Implications Tomorrow*, 16 *Duke Law & Technology Review* 85, 86-88 (2017).

够,即使经过专门的数据清洗和特征提取,可能也无法保证其作为训练数据集的质量,加上学习算法固有的不透明性等缺陷,更是降低了它的可接受性。总而言之,就现实情境以及可预见的未来情境而言,依靠预测分析学实现司法裁判的智能化,是“可望而不可即”的。

七、结语

上文的分析表明,无论是“自动判决机”,还是“裁判学习器”,可能都只是存在于科幻小说中的想象而已,实际上,它们目前也确实还只是一些“想象”而已。正如美国学者雷姆斯(Dana Remus)和利维(Frank Levy)所言,那些乐观断言人工智能将会代替律师、法官或者其他法律工作者的人,似乎都错误地理解了“计算机能够执行以及不能执行的法律任务的类型”。^[70]至少对于司法裁判工作而言,人工智能是不能胜任的。

不过,虽然本文对于司法裁判的人工智能化持一种消极性态度,但这并不表示笔者反对人工智能技术在司法领域的运用。相反,作为一种辅助工具,人工智能技术有着非常光明的应用前景,但这种应用应当立足于帮助法律工作者在特定任务或工作上实现更优的表现,而非代替他们进行决策。例如,出于“人少案多”的考虑,在司法裁判中引入各种技术手段减少审判人员的劳动量,提高司法效率,加强司法管理,这无疑是值得追求的,而且人工智能技术的进步也确实可以在这些方面改善司法裁判的品质,但如果由此认为人工智能可以代替人类裁判案件,则显然走得太远了。无论这种改善作用有多大,和其他的技术进步一样,人工智能依然只能是司法裁判的辅助工具,而不能成为决策的主体。除非技术的进步使得人类不再是唯一的理性存在者,机器也可以分享与人类一样甚至超越人类的智慧,除非法学的发展导致司法裁判的合法性、正确性预设被推翻,人类也乐于将自身事务的裁判权完全交给“无情的”、“不负责任的”机器,否则司法裁判人工智能化的消极前途都不会因为技术的进步而发生改变。

On Artificial Intelligentialization and Its Limitations in Judicial Adjudication

Song Xuguang

Abstract: Artificial intelligentialization in judicial adjudication means machines can replace human judges to make certain legal decisions independently. There are two approaches to intelligentializing legal decision making, the first approach is legal expert systems which run on explicitly coded, closed-rule algorithms, and the second one is predictive analytics which develop through machine learning algorithms. Expert systems have developed over the years and relevant limitations have appeared; while big data algorithms are making robust progress, but they are unlikely to be successful as well. Firstly, the adjudication results can not be predicted based on the regularity of big data in nature; secondly, machine adjudication destructs the traditional presumptions of administration of justice, dispels the human autonomy and disobeys the original intention of inventing artificial intelligence, and it is weaker to justify algorithm judgment for its promotion of similar cases being treated similarly or restriction of judicial discretion; lastly, as judicatory system has its particularity, it is difficult to satisfy the needs of learning algorithm technically. Therefore, artificial intelligence cannot and shall not be an adjudicator like a human judge, and what is more appropriate is to tap its potential as an assisting tool in judicatory.

Keywords: judicial adjudication; artificial intelligence; machine learning; big data; predictive theory of law

(责任编辑:倪鑫煜)

[70] Dana Remus & Frank Levy, *Can Robots Be Lawyers: Computers, Lawyers, and the Practice of Law*, 30 The Georgetown Journal of Legal Ethics 501, 504 (2017).