

机器学习如何赋能 管理学研究？*

——国内外前沿综述和未来展望

刘景江 郑畅然 洪永淼

摘要:机器学习正在深刻改变管理学的研究范式与方法。如何运用机器学习更好地赋能管理学研究已经成为学术界关注的前沿热点议题。然而,机器学习在中国管理学研究中的应用仍处于初级阶段。本文基于1999~2021年发表在工商管理 and 会计财务两大研究领域的国内外顶级期刊的学术文献,识别了学术界借助机器学习开展管理学实证研究的4种核心途径:变量测量、事件预测(包括事件分类)、因果推断和理论构建;梳理了每个途径的代表性文献的研究主题、研究问题、数据集、机器学习算法和研究结论;提出了使用机器学习赋能管理学研究的主要策略,并讨论了中国学者运用机器学习开展中国特色管理理论研究的未来机会。本文显示:将机器学习与传统计量经济学相结合有助于做出更加精准的因果推断;机器学习能够在模式发现这一理论构建的关键步骤中发挥重要作用;将机器学习与多案例分析相结合有助于富有成效地开展理论构建。本文为如何采用机器学习提升管理学研究质量、推进管理学研究范式变革和构建中国特色管理理论提供了方法论指引和方向性启示。

关键词:机器学习 文献计量 因果推断 理论构建 未来展望

DOI:10.19744/j.cnki.11-1235/f.2023.0106

一、引言

机器学习(Machine Learning)作为人工智能的一个重要分支(贝尔,2020),在医疗健康、金融服务、自动驾驶、市场营销等领域的应用均表现出了令人兴奋的巨大优势(乔丹、米切尔,2015),并在实践界和学术界都取得了令人惊叹的快速发展。机器学习催生的以数据、算法和算力为基本要素的计算社会科学(Computational Social Science)这一新型研究范式和新兴交叉学科(陈,2018;梅森等,2014),正在极大地改变社会科学的传统研究范式和方法(郭峰、陶旭辉,2023;胡安宁等,2021;罗家德、樊瑛,2018)。因而,机器学习技术日益受到社会科学研究者的广泛关注。正如瓦里安(2014)指出,机器学习技术有助于用更有效的方式构建复杂关系。阿西和因本斯(2019)强调指出,“机器学习工具正在成为跨学科的标准”;研究者有必要摆脱对主流的“数据建模文化”(Data Modeling Culture)的排他性依赖,灵活采用基于机器学习的“算法建模文化”(Algorithmic Modeling Culture)(陈强,2021;布莱曼,2001b),并同时保留传统计量经济学的优势^①。洪永淼和汪寿阳(2021c)进一步提出,“大数据与机器学习计量经济学”这一门新兴学科正在茁壮成长。机器学习技术与传统计量经济学方法的深度融合将是未来管理研究的新趋势(阿西、因本斯,2019;陈冬梅等,2020;陈强,2021)。

机器学习技术、方法与工具具有一些独特优势,如从复杂高维数据环境中提取有价值的信息以更好地帮助变量测量(Variable Measurement)(洪永淼、汪寿阳,2021a,2021b),直接基于数据灵活选择函数形式和构建算法以更好地帮助样本外的事件预测(Event Prediction)(洪永淼、汪寿阳,2021c;穆来纳森、斯皮斯,2017),改进传统计量经济学方法以更好地帮助因果推断(Causal Inference)(阿西、因本斯,2019),发现稳健且具有可解释性的模式以更好地帮助理论构建(Theory Building)(蒂德尔、艾森哈特,2020;乔杜里等,2021;施雷斯塔等,

*本项研究得到国家自然科学基金基础科学中心项目“计量建模与经济政策研究”(基金号:71988101)、国家自然科学基金项目“绩效反馈、制度逻辑与科技企业创业成长研究”(基金号:71872164)和国家自然科学基金项目“患者偏差行为的形成机制与医患关系管理策略研究”(基金号:72074189)的资助。感谢匿名审稿专家提出的宝贵修改意见。洪永淼为本文通讯作者。

2021)等。因此,它可以弥补管理学研究所存在的短板:第一,管理学研究传统上未能从海量、复杂、高维、非结构化、多模态数据中提取有用信息(洪永森等,2023;洪永森、汪寿阳,2021a),也未能克服社会经济系统的非实验性、不可逆性和时变性,以及社会经济数据的种种缺陷(洪永森,2007)。第二,传统上,管理学定性研究常常采用少数甚至单个案例分析来构建理论,而没有使用大样本数据(施雷斯塔等,2021)。受制于有限样本和人工经验,这些研究很难从数据中识别新颖、稳健、具有预测性和相对可解释性的模式(蒂德尔、艾森哈特,2020;乔杜里等,2021)。并且,这些研究所构建的理论具有较弱的泛化能力。第三,基于观测数据(Observational Data)的管理学定量研究通常采用传统计量经济学方法,而这一方法存在诸多明显的局限性:(1)它基于模型驱动思想,通常事先假定一系列苛刻的假设条件和一个小的线性低维参数模型(洪永森、汪寿阳,2021c),无法捕获高维动态大数据中呈现的时变性、非线性和非平稳性特征(洪永森、汪寿阳,2023;汪寿阳等,2019),并存在模型误设偏差(洪永森、汪寿阳,2021c)。(2)它以统计显著性和样本内拟合优度为模型优劣主要评估标准,模型的样本外预测能力较弱,泛化能力较差(陈强,2021;蒂芬,2019;洪永森,2021)。(3)它只能刻画主要因素的影响,无法对一因多果、同因异果、一果多因、同果异因、多因多果和复合因果且有滞后性的事件做出可靠的因果推断(陈等,2023;洪永森,2007)。

已有许多国外文献采用机器学习方法,研究资产定价(冯等,2020;顾等,2020;卡洛米里斯、马梅斯基,2019;马内拉、莫雷拉,2017)、并购与公司治理(马尔霍特拉等,2018;乔杜里等,2019;游,2018)、公司业绩(蒂德尔、艾森哈特,2020)、知识产权(乔杜里等,2020)、公司雇佣(栾等,2019)、审计(谢等,2020)、非对称和私人信息(刘等,2020)等热点主题。国内的管理学研究逐渐开始重视运用机器学习方法。例如,国内期刊《管理世界》发表的多篇论文运用机器学习方法研究了信息和市场效率(宫晓莉等,2020;马黎珺等,2019;沈艳、王靖一,2021;张宗新、吴钊颖,2021)、企业融资与治理(黄益平、邱晗,2021;沈艺峰等,2017)、企业行为(胡楠等,2021)等主题。

然而,机器学习在我国管理学研究中的应用仍处于初级阶段。特别是,我国顶级管理学期刊在工商管理研究领域和会计财务研究领域^②的实证文献极少运用机器学习进行因果推断和理论构建。造成这一现象的原因可能是,在中国管理学研究领域目前很少有文献系统梳理机器学习方法给传统管理学研究范式与方法带来的挑战。极少有文献对实质性使用机器学习方法研究管理问题的最新成果进行系统综述并由此提出未来研究机会。虽然已有少数国内文献梳理了机器学习对经济学研究的影响(黄乃静、于明哲,2018;蒋锋、张文雅,2022;王芳等,2020),但是缺乏从变量测量、事件预测(包括事件分类)、因果推断和理论构建等4个方面系统而具体地阐述采用机器学习方法赋能经济学研究的途径和策略。

中国具有独特的政治、经济、文化和社会制度,其中蕴含的中国故事具有明显异于西方世界的特性(王永贵等,2021)。尽管许多学者强调了这些中国情境的独特性,因所采用的研究范式和研究方法还不够完善,特别是缺少既符合科学研究范式又适合于中国情景的实证研究方法的创新与应用(柴国俊、孙若宸,2022),目前中国管理学的学术影响力较弱(洪永森、汪寿阳,2020,2021b;王永贵等,2021)。显然,这些不利于提升中国管理学在国际学术界的“话语权”(贾良定等,2015;王永贵等,2021),也不利于更好地为世界管理学贡献中国智慧(王永贵等,2021)。简而言之,方法与思想同等重要(洪永森、汪寿阳,2020)。因此,这进一步说明,很有必要综述国内外顶级期刊发表的运用机器学习方法赋能管理学研究的前沿文献,并据此评估现有学术进展和发现未来研究机会。

本研究具有3个主要贡献。第一,以国际上得到广泛认可的UTD-24期刊^③和“2021中国最具国际影响力学术期刊(人文社会科学)”前20名中的管理学期刊^④于1999~2021年正式发表的文献为数据,采用编码方法,识别了工商管理和会计财务两大研究领域的国内外学者应用机器学习进行管理学实证研究的4种核心途径:变量测量、事件预测、因果推断和理论构建。第二,从研究主题、研究问题、数据集、机器学习算法和研究结论等多个方面系统总结了这些途径的代表性文献,从而描绘了一幅机器学习赋能管理学研究的图景。第三,在分析以往研究方法不足的基础上,深入阐述了一系列采用机器学习开展管理学研究的主要策略和未来机会。我们特别强调,未来研究应更加关注机器学习在因果推断和理论构建这两大方面的重要作用。总之,本研究

将有助于我国学者采用合适的最新机器学习技术、方法与工具,聚焦于中国情境、中国问题和中国数据,变革中国管理学研究范式,创新中国管理学研究方法,讲好中国管理故事,构建具有国际影响力的中国特色管理理论(黄群慧,2018;贾良定等,2015;李宝元等,2017;王永贵等,2021)。

二、机器学习发展概况

(一)机器学习定义和本质

塞缪尔(1959)最早将机器学习定义为,“在没有被显性编程的情况下,使计算机具有学习能力的研究领域”(杰隆,2019)。但该定义过于宽泛且仅为定性描述,因此米切尔(1997)通过引入3个概念:经验、任务和性能,给出了更加具体且严谨的定义,“机器学习是一个计算机程序。如果该计算机程序在任务上的性能随着经验提高,则称它从与某些任务和性能指标有关的经验中学习”(杰隆,2019)。

本质上,机器学习是“一个数学优化问题与实现该优化问题的计算机算法问题”(洪永森、汪寿阳,2021a)。它使计算机具备从大量数据中学习的能力(陈强,2021;洪永森、汪寿阳,2020,2021a)。机器学习是“从训练数据中寻找一个优化算法,使预测测试数据的损失函数加上惩罚项最小化,以达到最优样本外预测效果”(洪永森、汪寿阳,2021a)。面对实时发生、数量和体量大、类型多和信息密度低的大数据,非常适合运用机器学习方法对复杂高维大数据进行降维,实现更加精准的预测(洪永森,2021;洪永森、汪寿阳,2021b),更好地帮助识别因果关系和构建新的理论,从而变革管理学研究范式,创新管理学研究方法。

(二)机器学习发展简史

根据不同时期的研究内容和目标,本文将机器学习的发展历程划分为5个阶段(周志华,2016)。第一阶段是20世纪50年代初到60年代中叶奠定基础的热烈时期,主要研究“没有知识”的学习,代表性研究是1952年塞缪尔设计的跳棋程序,罗森布拉特(1958)基于神经网络“连接主义”(Connectionism)提出的感知机(Perceptron)。第二阶段是20世纪60年代中叶到70年代中叶停滞不前的冷静时期,主要研究将各个领域的知识植入到学习系统里,出现了基于逻辑表示的“符号主义”(Symbolism)学习技术,如亨特等(1966)提出的“概念学习系统”。第三阶段是20世纪70年代中叶到80年代中叶重拾希望的复兴时期,主要聚焦于从学习单个概念扩展到学习多个概念,探索不同的学习策略和学习方法。

第四阶段是20世纪80年代中叶到21世纪初现代机器学习的成型时期。此时,机器学习已成为一门独立的学科(卡博内尔,1990)。20世纪80年代中期,布莱曼等(1984)提出的CART算法(Classification and Regression Trees)和昆兰(1986)开发的ID3算法(Iterative Dichotomiser 3)使得以决策树(Decision Tree)为代表的“符号主义学习”占据主流。同年,鲁梅尔哈特等(1986)提出的反向传播算法推动了神经网络“连接主义学习”发展的第二次高潮。20世纪90年代中期,科尔特斯和瓦普尼克(1995)提出的支持向量机(Support Vector Machine, SVM),作为“统计学习”(Statistical Learning)的代表大放异彩。与此同时,“集成学习”(Ensemble Learning)也成为机器学习的重要延伸,最著名的是提升法(Boosting)(弗罗因德、夏皮尔,1996,1997;夏皮尔,1990)、装袋法(布莱曼,1996)和随机森林(Random Forest)(布莱曼,2001a)等算法。

第五阶段是21世纪初至今大放光芒的蓬勃发展时期。欣顿和萨拉胡季诺夫(2006)掀起了以“深度学习”(Deep Learning)为名的热潮,代表性算法有深度信念网络(欣顿等,2006)和生成对抗网络(古德费洛等,2014)等。可以预见的是,在未来,强化学习、对偶学习、迁移学习、分布式学习以及元学习等多种技术将成为机器学习重要的前沿方向(普拉莫德等,2021),不同流派也会通力合作,将各自的方法与人类知识相结合(乔丹、米切尔,2015),进入“人机协同”时代,从而更好地帮助人们解决实际应用中的复杂问题。

(三)机器学习应用场景

机器学习拥有强大的数据处理能力、数据挖掘能力和信息洞察能力,能够有效提高生产力和工作效率,增强企业核心竞争力,目前已经被广泛应用于各行各业(贝尔,2020;杜尔哈尔等,2020;普拉莫德等,2021)。在医疗健康领域,将机器学习技术与大数据相结合,借助可穿戴设备和传感器传输患者的体重、心率、脉搏、血压

理论述评

和血糖水平等数据,实时监测和评估患者健康状况,有利于医生根据历史信息分析患者病情,制定个性化治疗方案。在金融服务领域,一是利用机器学习技术建立财务欺诈预测模型,有效识别高风险客户和财务欺诈行为,实现风险预警和反欺诈;二是利用机器学习技术建立信用评分模型,准确客观地评估中小微企业和个人信贷风险,提升普惠金融服务精准度;三是利用机器学习技术进行算法交易,确定最佳交易时间和投资机会,辅助和改善投资决策。

在市场营销领域,零售商广泛运用机器学习技术分析历史销售数据,帮助制定营销活动、客户商品规划和价格优化策略,还结合自然语言处理和自然语言理解技术,配备虚拟助手和聊天机器人,为顾客提供定制化和个性化购物体验。在交通运输领域,一方面运用监督学习、非监督学习和强化学习等方法,以大数据为原料实现自动驾驶智能感知和智能决策;另一方面综合利用云计算、人工智能、机器学习和物联网等技术,搭建数据处理中心和资源整合平台,建立高效、便捷、安全、经济的城市智能交通运输体系,以数字力量驱动智慧城市构建。

三、研究设计与方法

(一)数据收集

本文使用标准抽样方法(巴顿,1990)确定文献范围,研究国内外管理学顶级期刊发表的、实质性使用机器学习方法的文章。首先,本文选取UTD-24期刊,以“machine learning”、“decision tree”、“support vector machine”、“random forest”、“artificial neural network”和“deep learning”等为关键词,对目标期刊的所有在库文章进行全篇检索,把正式发表时间限定到2021年12月末,得到一张包含1258篇文献的初步文献清单。其中,会计领域52篇,财务领域72篇,信息系统领域322篇,营销领域208篇,管理科学领域522篇,工商管理领域82篇。考虑到篇幅有限和用途梳理的全面性,本文只关注工商管理和会计财务两大研究领域^⑤。因最早实质性运用机器学习方法的文章发表于1999年,本文聚焦于这两大领域中正式发表于1999~2021年且实质性使用机器学习方法进行实证研究的文章。我们逐一阅读这两大领域中的206篇文章,剔除没有实际使用机器学习方法进行数据分析的文章,也剔除评论、书评和会议征稿等论文,最终筛选出符合条件的62篇文章。其中,工商管理领域18篇,会计财务领域44篇。如图1所示,从这些国际管理学期刊的文章来看,机器学习方法被实质性使用的频率呈现逐年上升的趋势^⑥。特别是,2018年开始出现了快速增长的态势。

接着,类似地,本文选取“2021中国最具国际影响力学术期刊(人文社会科学)”前20名中的管理学期刊,以“机器学习”、“决策树”、“支持向量机”、“随机森林”、“人工神经网络”和“深度学习”等为关键词对上述中文期刊的文章进行全文检索,以“machine learning”、“decision tree”、“support vector machine”、“random forest”、“artificial neural network”和“deep learning”等为关键词对上述英文期刊的文章进行全文检索,同样限定正式发表时间为2021年12月末,得到一张包含153篇文献的清单。其中,工商管理领域54篇,会计财务领域52篇,其他领域47篇。因国内最早实质性使用机器学习方法的文章发表于2004年,本研究聚焦于工商管理和会计财务两大领域中正式发表于2004~2021年且运用机器学习方法进行实证研究的文章。我们逐一阅读这两大领域中的106篇文章,最终筛选出符合条件的43篇文章。其中,工商管理15篇,会计财务28篇。图2显示了这些国内管理学期刊历年发表实质性使用机器学习方法进行实证研究的文章的数量^⑦。

(二)数据编码与文献分析

在确定研究目标后,我们按照以下3个步骤对数据进行编码和文献分析。第一步,根据以往理论和实证研究(阿西、因本斯,2019;蒂德尔、艾森哈特,2020;格里默等,2021;洪永森,2021;洪永森、汪寿阳,2021a,2021b;黄乃静、于明哲,2018;

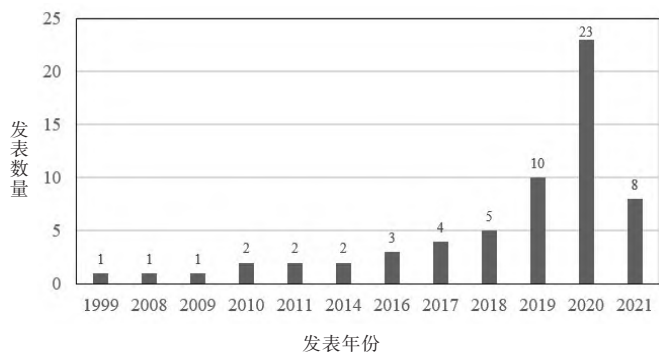


图1 两大研究领域国际顶级管理学期刊上实质性运用机器学习方法的文章发表趋势

乔杜里等,2020;施雷斯塔等,2021),我们总结出机器学习方法在管理学实证研究中的4种核心用途:变量测量、事件预测、因果推断和理论构建,如图3所示。

变量测量是根据一种规则,用数量的方法描述研究对象所具备的某种特征或行为,其目标是对变量之间的关系进行量化推断(陈晓萍等,2008)。事件预测是使用已掌握的经验或知识,预先推知和判断事物未来发展状况(阿西,2019),其目标是预料来自不同观测总体的样本已经或将要在未来实现的结果(格里默等,2021)。因果推断是借助理论和对制度细节的深入了解,估计事件和选择对给定结果的影响(坎宁安,2021),其目标是比较在同一干预措施下不同反事实(Counterfactual)结果之间的差异(格里默等,2021)。理论构建是构建概念及其相互关系,以展示一种现象是如何和为什么发生的过程(焦亚、皮特雷,1990;科利、焦亚,2011;克里斯蒂安森、钱丹,2017),其目标是建立稳健且具有可解释性的理论。

变量测量、事件预测、因果推断和理论构建是管理学实证研究的4项关键任务。它们既相互区别又紧密关联。理论构建在管理学实证研究中占据着核心地位(班伯格,2018)。管理学顶级期刊格外强调文章的理论贡献(科利、焦亚,2011)。实证研究的核心目标是理论构建。衡量一个“好”的实证研究的首要标准是它能够建立稳健且具有可解释性的理论。因果推断是理论构建的先决条件。事件预测是因果推断的必要前提。变量测量是开展管理学实证研究的根基。总之,这4个途径相辅相成,构成目的与手段的关系,前1个途径是后面1个、2个或3个途径的基础。

机器学习具有一系列独特优势,如自动学习数据节省研究时间(乔杜里等,2019),改进传统的计量经济学方法(阿西、因本斯,2019),从复杂高维的数据环境中提取有意义的信息(洪永森、汪寿阳,2021a,2021b),运用灵活的函数形式构建精度更高的模型(洪永森、汪寿阳,2021a,2021b,2021c;穆来纳森、斯皮斯,2017),使用正则化和交叉验证方法提高模型的泛化能力(蒂德尔、艾森哈特,2020;蒂芬,2019;乔杜里等,2021;瓦里安,2014)等。因此,它特别适用于更好地进行变量测量、事件预测、因果推断和理论构建。

具体地,采用机器学习从复杂高维数据环境中提取有用信息进行变量测量,可以看作是给机器学习系统布置任务以预测分配标签。从这个角度来看,变量测量问题被机器学习转化成为事件预测问题(格里默等,2021)。虽然预测和因果推断有着很大的不同,但是因果推断的本质是比较同一组样本的不同反事实,估计反事实的实质是预测(洪永森、汪寿阳,2021a)。机器学习的精准预测能力可以对不可能真实发生的反事实进行合理估计,从而助力因果关系的识别(郭峰、陶旭辉,2023;瓦里安,2014)。理论构建的核心是发现模式和理论解释(施雷斯塔等,2021)。虽然机器学习自身不能提供理论解释(施雷斯塔等,2021),但是它可以采用其强大的算法,依靠多次迭代比较来寻找出最优预测模型(蒂德尔、艾森哈特,2020),从大数据中挖掘稳健模式,从而助力理论构建(施雷斯塔等,2021)。

第二步,本文根据美国经济学会《经济文献杂志》(JEL)所创立的主题分类系统对文献进行编码。若文章包含分类号,则根据已经标记好的分类号进行编码;若不

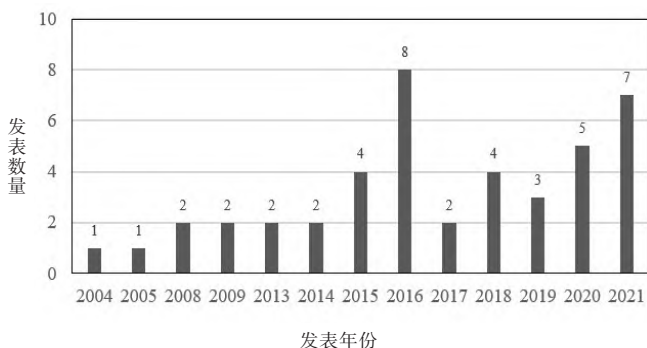


图2 两大研究领域国内顶级管理学期刊上实质性运用机器学习方法的文章发表趋势

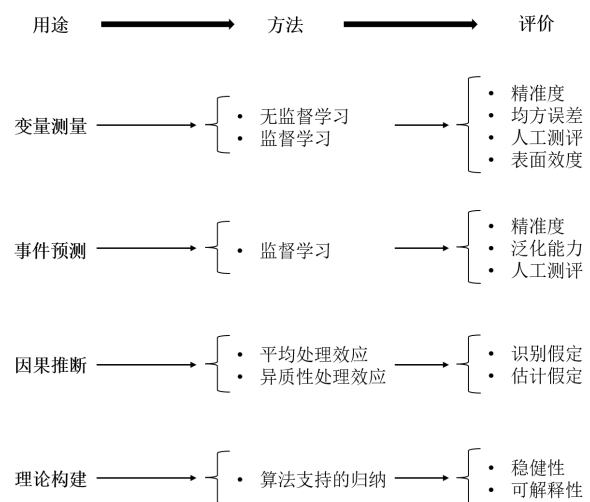


图3 机器学习方法在管理学研究中的4种用途

理论述评

包含分类号,则仔细阅读文章的摘要、引言和正文,根据文献主要探讨的主题,人工选择一个最适合的分类号。为了保证编码的客观性,本研究编码人员针对编码不一致的情况及时讨论、力求统一。若仍存在难以形成统一意见的文章,则向相关领域的专家寻求建议,最大程度避免由单一编码者的主观臆断对结果产生的偏差。

第三步,我们精心阅读变量测量、事件预测、因果推断和理论构建等4种核心途径的代表性文献,细致梳理这些文献的研究主题、研究问题、数据集、机器学习算法和研究结论;深入系统地分析以往研究方法的优势与不足,进而从上述4种途径阐述运用机器学习赋能管理学研究的策略和未来机会。

四、机器学习在管理学研究中的应用综述

工商管理 and 会计财务作为管理学的两大核心研究领域,包含大量来自个人、企业和政府的文本、图像、音频、视频等极具信息价值的非结构化数据。传统方法无法对这些非结构化数据进行量化分析,只能进行定性分析。借助机器学习方法,学者们可以从这些非结构化数据中挖掘、提取和构建诸如高管人格特质、管理者自恋、公司文化、媒体文章语调和投资者情绪等有意义的变量(洪永森、汪寿阳,2021a,2021b),运用灵活的函数形式和降维技术来实现更精准的预测(洪永森、汪寿阳,2021b,2021c),利用正则化和交叉验证方法提高模型泛化能力以帮助因果推断和理论构建(蒂德尔、艾森哈特,2020;蒂芬,2019;乔杜里等,2021;瓦里安,2014),从而更好地开展这两大领域中关键问题的实证研究。因此,本部分以这两大研究领域为例,以机器学习赋能管理学研究的4种核心用途为主线,全面回顾和系统梳理UTD-24期刊和国内顶级管理学期刊于1999~2021年正式发表的文章。具体来说,本文遵循重点性原则和典型性原则,按照这些领域和用途,总结归纳了代表性文献的研究主题、研究问题、数据集、机器学习算法和研究结论^⑧。

(一)工商管理

1. 变量测量

现实生活存在海量非结构化数据。工商管理研究领域的学者采用多种机器学习方法,可以从这些非结构化数据中挖掘具有管理意义的信息来进行变量测量,进而开展相关研究。代表性文献如表1所示。

以往研究从非结构化数据中,挖掘与CEO人格特质相关的信息,测量CEO的人格,进而估计不同人格特质对并购与公司治理决策的影响。例如,马尔霍特拉等(2018)运用SVM对2002~2012年Execucomp数据库中S&P 1500公司的2381名CEO的口语文本(CEO与财务分析师季度财报电话会议中问答环节所说的文字)进行人格评分,测量CEO的外倾性。研究发现,具有外倾性的CEO更有可能参与并购。乔杜里等(2019)通过对哈佛商学院“创造新兴市场”项目中的CEO进行访谈,结合无监督学习中的隐含狄利克雷分布(Latent Dirichlet Allocation, LDA)和情感分析,对CEO的口头沟通文本进行内容和价值提取。同时,应用监督机器学习中的卷积神经网络(Convolutional Neural Networks, CNN),对CEO访谈视频中的面部表情进行编码和情绪识别,构建出5种不同沟通风格(兴奋、严厉、戏剧性、漫不经心和忧郁)的测量,进而得到在表达方式上戏剧化的CEO不太可能监督重大收购的结论。

在企业行为研究中,哈里森等(2019)采用1996~2014年S&P 1500公司的3573名CEO的财报电话会议记录,

表1 工商管理研究领域采用机器学习进行变量测量的代表性文献

研究主题	作者(年份)	研究问题	数据集	机器学习算法	主要结论
并购与公司治理(G34)*	马尔霍特拉等(2018)	CEO的外倾性如何影响公司并购行为?	2002~2012年Execucomp数据库中S&P 1500公司CEO的口语文本	·SVM	外倾性CEO更有可能参与并购,更频繁地开展并购,并进行更大规模的并购。
	乔杜里等(2019)	如何测量CEO的沟通风格? CEO沟通风格如何影响收购?	2012~2018年哈佛商学院“创造新兴市场”项目中CEO的访谈文本和视频	·LDA ·CNN	CEO有5种沟通风格;在表达方式上戏剧化的CEO不太可能监督重大收购。
企业行为(D21)	哈里森等(2019)	CEO的五大人格特质如何影响公司战略变革?	1996~2014年S&P 1500公司CEO的口语文本	·GBM	开放性和宜人性的CEO对战略变革有直接影响,另外3种人格特质没有直接影响。
	胡楠等(2021)	管理者短视主义如何影响企业长期投资?	2007~2018年A股公司年度财务报告的MD&A部分	·Word2Vec	管理者短视导致企业减少资本支出和研发支出。

注: *G34为JEL主题代码,下同。

运用梯度提升机(Gradient Boosting Machine, GBM)算法,开发一个开放式语言工具(Open-Language Tool)来测量CEO的五大人格特质。结果发现,开放性的CEO对战略变革有正向直接影响,宜人性的CEO对战略变革有负向直接影响,其他3种没有直接影响。胡楠等(2021)选取沪深两市2007~2018年所有A股公司年度财务报告为初始研究样本,提取其中管理讨论和分析部分(Management Discussion and Analysis, MD&A),利用文本分析、Word2Vec机器学习算法和词典法构建管理者短视主义指标,得出管理者短视主义特质会导致企业减少资本支出和研发支出的结论。

2. 事件预测

机器学习可以更有效地探索变量之间的相关性,进而对事件做出较为精准的预测。在工商管理研究领域,运用机器学习技术进行预测的文章主要聚焦于技术变迁和企业管理等主题。代表性文献如表2所示。

学者们在知识产权管理研究中采用机器学习方法预测专利申请人策略性地省略相关信息或增加无关信息的行为。例如,乔杜里等(2020)将机器学习算法应用于美国专利商标局对专利的审查程序。他们发现,当代理人为了从有偏的预测结果中获益而策略性地改变算法输入时(即存在输入不完全性),机器学习技术得出了比布尔技术(Boolean Technology)更差的预测结果。他们进一步发现,当将特定领域专业知识和特定年份技能这两个关键人力资本属性完全纳入用于生成预测的训练数据集和算法中时,运用机器学习技术进行预测的精准度得到显著提升。因此,这些研究结果表明,只有将机器学习和人力资本相结合,才能够更好地减轻机器学习的预测偏差,有效地提升专利审查效率。

也有学者在公司雇佣决策研究中采用机器学习方法预测求职者的未来表现。比如,栾等(2019)选用航空公司236名求职者数据,分别运用套索(Least Absolute Shrinkage and Selection Operator, LASSO)、随机森林和SVM三种流行的机器学习算法预测求职者的未来表现,并与启发式推断策略和逻辑回归的预测精准度进行比较。结果发现,当样本量较大时,LASSO和随机森林算法的预测表现与逻辑回归相似,但稍微落后于启发式策略,SVM的表现落后于上述两种算法。当样本量较小时,LASSO和随机森林算法比逻辑回归更差,远落后于启发式策略。虽然SVM的预测精准度优于其他两种算法,但仍未能超越启发式策略。

还有学者在电子商务研究中运用机器学习方法对顾客在线评论和企业管理反馈策略进行分类。比如,郭晓妹等(2021)利用SVM机器学习方法,对国内某旅行网站99027名顾客二次入住酒店后发表的在线评论,进行积极、偏积极、偏消极和消极的评论情绪分类,并根据酒店的两次回复内容对道歉承诺类管理反馈策略和非

表2 工商管理研究领域采用机器学习进行事件预测和理论构建的代表性文献

用途	研究主题	作者(年份)	研究问题	数据集	机器学习算法	主要结论
事件预测	知识产权(O34)	乔杜里等(2020)	存在输入不完全性偏差时,人力资本如何补充机器学习算法以提高机器学习的预测精准度?	美国专利商标局专利数据	·机器学习算法	存在输入不完全性偏差时,机器学习的预测效果较差;两种关键人力资本可以在输入不完全的情况下对机器学习算法进行补充,以提高机器学习预测精准度。
	公司雇佣(M51)	栾等(2019)	生态理性视角下,启发式策略在管理决策中的表现会怎样?	航空公司236名求职者数据	·LASSO ·随机森林 ·SVM	在管理决策常见的条件下,启发式策略比复杂的理性策略在预测效果上表现更好。
	电子商务(L81)	郭晓妹等(2021)	在消费者情绪多样性的情况下,道歉承诺对顾客二次满意度有怎样影响?	国内某旅行网站99027名顾客二次入住酒店数据	·SVM	道歉承诺类管理反馈策略对提高顾客二次满意度有积极作用;顾客的矛盾情绪和负面情绪对顾客二次满意度有负向影响;在顾客消极情绪下,道歉承诺能够进一步提升满意度。
理论构建	公司业绩(L25)	蒂德尔和艾森哈特(2020)	什么是收益模式的最优选择?	2015年11月苹果公司App Store上的移动产品数据	·惩罚性多分类逻辑回归 ·决策树 ·随机森林	成功的商业模式需要价值获取与价值创造相契合。
	争议解决(J52)	何等(2020)	如何有效解决社区治理纠纷?	GitHub中包含“许可”问题的项目数据	·LASSO ·装袋法 ·ANN ·RR	当社区中的治理纠纷带有群体问题解决过程而非多方讨价还价过程的特征时,更可能被解决;群体过程和群体属性是解决社区治理纠纷的两个必要条件。
	模型构建与估计(C51)	施雷斯塔等(2021)	如何利用机器学习技术从数据中构建理论?	模拟数据集	·随机森林 ·LASSO ·GBRT ·神经网络 ·RR	在组织管理领域中,可以利用机器学习算法从大规模数据中归纳出具有稳健性和可解释性的模式,从而帮助理论构建。

理论述评

道歉承诺类管理反馈策略进行分类。研究发现,适当的道歉承诺管理反馈策略能够提高顾客的二次满意度,但需要根据顾客的不同情绪,采取相应的反馈策略,才能进一步提升满意度。

3. 理论构建

理论构建是机器学习技术的一个最新用途。虽然机器学习自身不能构建理论,但是它能够帮助研究者增强理论构建的规模和精准度(蒂德尔、艾森哈特,2020;格里默等,2021;施雷斯塔等,2021)。一方面,机器学习能够在定量数据中寻找模式,使用算法系统地估计和比较多种替代模型,选择最具有预测性的稳健模型(阿西,2019)。另一方面,机器学习能够采用正则化和交叉验证等技术来限制过拟合和过度复杂模型,从而提高预测和泛化能力(蒂德尔、艾森哈特,2020;乔杜里等,2021;瓦里安,2014)。工商管理研究领域的学者日益运用机器学习赋能理论构建。代表性文章如表2所示。

例如,蒂德尔和艾森哈特(2020)综合运用机器学习和多案例理论构建等方法,提出了一个高性能收益模式—活动系统配置的框架。具体地,他们对2015年11月苹果公司App Store上的全部移动产品进行抽样,得到了66652个产品样本,随机从中选择了400个受欢迎的产品和400个不受欢迎的产品。为了确保训练和交叉验证的数据平衡,他们进一步收集并预留了100个受欢迎和100个不受欢迎的产品作为样本外的测试数据,得到了1000个产品作为机器学习的样本。3种互补的机器学习方法(惩罚性多分类逻辑回归、决策树和随机森林)帮助他们获得了稳健且相对可解释的结果,而不是以可解释性为代价所产生的具有最佳预测的复杂输出。通过整合探索性数据分析所得到的趋势和案例理论抽样的路线图、多案例理论构建所确定的理论概念和机制、机器学习在案例研究中采用更精准的效应大小、等值路径、非线性和配置所提供的大规模证实和扩展,他们得出成功商业模式的核心是价值获取与价值创造相结合的结论。

何等(2020)采用机器学习构建了解决社区治理纠纷的理论。具体地,他们使用LASSO、装袋法和人工神经网络(Artificial Neural Network, ANN)3种机器学习算法与岭回归(Ridge Regression, RR),对183个GitHub项目中包含的数据进行正则化与特征选择,以获得更容易解释且稳健的模式,并通过反复分割数据,以获得样本外预测精准度更高的模型。机器学习帮助他们发现:在较大的讨论组和没有许可证项目(相比于有许可证的项目)中,社区治理纠纷更可能得到解决。进而,他们随机抽取61个案例进行编码,探寻特征变量之间的因果关系,以加强理论机制的解释。案例的文本分析帮助他们发现:群体过程和群体属性是解决社区治理纠纷的两个必要条件。基于这些发现,他们开发了一个整合群体过程和群体属性的理论模型,以解释与在线社区的核心治理原则有关纠纷的解决。

施雷斯塔等(2021)提出了一套利用多种机器学习算法助力理论构建的程序。基于一个模拟数据集,他们综合使用随机森林、LASSO、梯度提升回归树(Gradient Boosting Regression Tree, GBRT)、神经网络和RR等多种算法确定关键预测因子,识别出具有稳健性和可解释性的关联作为溯因理论构建的输入,有效助力了理论构建工作的开展。研究表明,在组织管理领域,机器学习可以作为从大规模数据中构建理论的工具。

(二) 会计财务

1. 变量测量

会计财务研究领域已有许多学者运用机器学习方法,将非结构化数据转变为定量数据,对变量进行测量。代表性文献如表3所示。

从会计和公司治理等研究主题来看,学者们应用机器学习对媒体文章、财务报表、分析师报告和推特推文等非结构化文本中的变量进行测量。例如,朴素贝叶斯算法(Naive Bayesian Algorithm)常被用来测量媒体文章语调(游等,2018)、10-K与10-Q报表中MD&A的语调(亨利、利昂,2016)、分析师报告中的情绪(黄等,2014)和个人推文总体意见(巴托夫等,2018)。LDA也被用来测量10-K报表中的主题(布朗等,2020;戴尔等,2017)。它达到与朴素贝叶斯方法异曲同工的效果。

从政府政策和监管等研究主题来看,以往研究运用多种机器学习对非结构化文本中的变量进行测量。比如,劳里等(2020)单独使用LDA识别和测量美国证券交易委员会意见函中关注的主题及其相对重要性,综合

运用LDA和KL散度(KL Divergence)识别美国证券交易委员会意见函的边际信息价值。此研究发现,企业在IPO之前会在招股说明书中增加对美国证券交易委员会关注主题的披露。

2. 事件预测

机器学习能够采用灵活的函数形式,在数量庞大且高度相关的数据中选择预测因子并实施降维处理(格里默等,2021;洪永森、汪寿阳,2021a,2021b),同时使用正则化和交叉验证缓解过拟合风险和限制过度复杂模型(蒂德尔、艾森哈特,2020;施雷斯塔等,2021)。因而,它在事件预测上具有很大的优势。会计财务研究领域的学者日益运用机器学习开展事件预测。代表性文章如表4所示。

表3 会计财务研究领域采用机器学习进行变量测量的代表性文献

研究主题	作者(年份)	研究问题	数据集	机器学习算法	主要结论
会计(M41)	亨利和利昂(2016)	哪种方法最适合评估财务披露语调?	2004~2012年8-K报表、1995~2007年10-K和10-Q报表	·朴素贝叶斯	在财务披露和资本市场的背景下,等权词频测量法产生了强大且可复制的结果。
	黄等(2014)	分析师报告文本是否具有信息价值?	1995~2008年S&P 500指数公司发布的分析师报告	·朴素贝叶斯	分析师报告文本对未来五年的收益增长有预测价值。
	巴托夫等(2018)	个人推特推文能否帮助预测公司收益和股票回报?	2009~2012年Russell 3000指数公司的推特推文	·朴素贝叶斯	个人推文的总体意见可以预测公司即将到来的季度收益和公告回报。
	布朗等(2020)	财务报表披露的主题在预测故意误报时是否具有渐进的信息量?	1994~2012年10-K报表、会计和审计执行公告、财务重述报告和10-K/As报表	·LDA	主题测量标准比常用的财务报表和文本风格测量标准提供了显著的增量预测能力。
	戴尔等(2017)	10-K报表披露的变化趋势是怎样的?	1996~2013年10-K报表	·LDA	10-K信息披露的趋势是长度、模板、粘性和冗余度增加,具体性、可读性和硬信息减少。
并购与公司治理(G34)	游等(2018)	政府控制如何影响媒体的信息中介和公司治理作用?	CSMAR中2005~2014年上海和深圳证券交易所的上市公司信息、2005~2014年在外国证券交易所双重上市的中国公司信息以及2004~2014年中国八家商业报纸文章	·朴素贝叶斯	市场导向的媒体文章比国家控制的媒体文章更具批判性、更准确、更全面、更及时;市场导向的媒体文章对公司治理具有显著影响。
政府政策和监管(G38)	劳里等(2020)	IPO之前,美国证券交易委员会和公司如何沟通?	2005~2016年美国证券交易委员会给952家公司IPO的意见函	·LDA ·KL散度	公司会在美国证券交易委员会关注的确切主题上增加招股说明书披露。

表4 会计财务研究领域采用机器学习进行事件预测的代表性文献

研究主题	作者(年份)	研究问题	数据集	机器学习算法	主要结论
会计(M41)	佩罗尔斯等(2017)	如何改进财务报表欺诈模型的预测能力?	51家欺诈公司和15934家非欺诈公司年度数据	·SVM ·逻辑回归 ·装袋法	OU和PVU可提高欺诈预测模型的精准度。
	鲍等(2020)	如何构建精准的欺诈预测模型?	1991~2008年美国上市公司财务报表数据	·RUSBoost ·逻辑回归 ·SVM	采用集成学习方法构建的预测模型,在精准度、泛化能力和稳健性等方面优于传统方法。
资产定价(G12)	马内拉和莫雷拉(2017)	新闻隐含波动率能否解释股票市场风险溢价?	1890~2009年《华尔街日报》头版文章隐含的股票市场波动率数据和1896~2009年资产定价数据	·SVR	新闻隐含波动率很好地预测了样本外的期权隐含波动率和实际波动率。
	卡洛米里斯和马梅斯基(2019)	新闻文章内容如何预测股票市场风险和收益?	1996~2015年新闻文章数据、国家层面股票市场指数数据和宏观数据	·Louvain社区发现算法 ·LDA ·弹性网	新闻文章所包括的特定主题情绪、频率和熵,可以预测股票市场的回报率、波动率和跌幅。
	顾等(2020)	如何运用机器学习方法预测股票风险溢价?	1957~2016年纽约证券交易所、美国证券交易所和纳斯达克上市公司月度股票回报数据和国债利率数据	·弹性网 ·降维技术中的PCR ·降维技术中的PLS ·BT ·随机森林 ·ANN	机器学习方法在预测股票风险溢价方面优于传统方法;神经网络和回归树是最好的预测方法。
信息和市场效率(G14)	陈等(2019)	金融科技创新有什么价值?	2003~2017年批量数据存储系统的专利申请数据和股票价格数据	·朴素贝叶斯 ·KNN ·随机森林 ·梯度提升 ·SVM ·神经网络	金融科技创新对创新者和整个金融行业具有价值。
企业融资与治理(G30)	黄益平和邱晗(2021)	大科技信贷的新信用风险管理框架是什么及其有何影响?	中国某头部金融科技公司2017年6月~9月抽样逐笔贷款数据	·随机森林	大科技信贷新的风险管理框架主要依靠大科技生态系统和大数据风控模型等两大工具;大科技信贷有利于提高金融普惠性。

从会计等研究主题来看,学者们采用机器学习预测上市公司的财务报表欺诈行为。例如,佩罗尔斯等(2017)选取51家欺诈公司和15934家非欺诈公司年度数据,采用SVM建立了财务报表欺诈的预测模型,并提出了3种不同的改进方法。这一研究发现,多子集观测欠采样(OU)和按欺诈类型划分的多子集变量欠采样(PVU)方法可以有效提升欺诈模型的预测精准度;多子集变量欠采样(VU)方法对欺诈预测模型精准度的改善取决于各子集中选择的具体变量。鲍等(2020)以1991~2008年美国证券交易委员会会计和审计执行公告中披露的重大会计错报作为会计舞弊样本,运用集成学习方法中的随机欠采样自适应提升法(Random Under-Sampling with AdaBoost, RUSBoost),预测欺诈公司与非欺诈公司。进而,他们使用逻辑回归和SVM建立了两个基准欺诈预测模型,与集成学习模型的预测结果进行比较后发现,采用集成学习方法构建的会计欺诈预测模型,不仅大大提升了传统欺诈预测方法的精准度,而且改善了模型的稳健性和泛化能力。

从资产定价等研究主题来看,越来越多的学者使用多种机器学习方法预测股票市场的风险溢价。比如,马内拉和莫雷拉(2017)选取1890~2009年《华尔街日报》头版文章隐含的股票市场波动率数据和资产定价数据,运用比普通最小二乘法(Ordinary Least Squares, OLS)能更好地处理高维特征空间的支持向量回归(Support Vector Regression, SVR),预测期权价格与词频的关系。这一研究发现,新闻隐含波动率很好地预测了样本外的期权隐含波动率和实际波动率。卡洛米里斯和马梅斯基(2019)使用Louvain社区发现算法和LDA,识别与分类1996~2015年所有汤森路透社新闻文章主题,预测新闻文章中特定主题的词流对股票回报率、波动率和跌幅的影响。进而,他们运用弹性网进行模型选择和系数缩减,探讨系数动态变化并进行样本外预测。这一研究发现,新闻文章所包括的特定主题情绪、频率和熵等信息,可以预测股票市场的风险和回报。顾等(2020)采用在纽约证券交易所、美国证券交易所和纳斯达克上市的所有公司的每月个人股票总回报等数据,运用惩罚线性回归中的弹性网、降维技术中的主成分回归(Principal Components Regression, PCR)和偏最小二乘法(Partial Least Squares, PLS)、回归树中的提升树(Boosted Tree, BT)和随机森林、ANN等多种机器学习方法,预测股票风险溢价。这一研究表明,使用机器学习算法进行预测的投资者能够获得巨大经济收益,并识别了回归树和神经网络是预测性能最好的机器学习算法。

从信息和市场效率研究主题来看,以往研究运用机器学习对非结构化文本中的事件进行分类。例如,陈等(2019)基于2003~2017年批量数据存储系统的专利申请数据,综合运用朴素贝叶斯、K-最近邻法(K-Nearest Neighbor, KNN)、随机森林、梯度提升、线性支持向量机(Linear SVM)、高斯支持向量机(Gaussian SVM)和神经网络等多种机器学习算法,识别和分类金融科技创新。这一研究发现,金融科技创新对创新者和整个金融行业具有价值。

从企业融资与治理研究主题来看,黄益平和邱晗(2021)基于中国某头部金融科技公司2017年6月~9月的抽样逐笔贷款数据,利用随机森林等算法,预测小微企业的信用贷款违约行为。结果发现,机器学习模型在传统数据和所有数据的情况下均优于逻辑回归模型,且当样本量增大和信息时效性增强时,机器学习模型比逻辑回归模型增益更大。由此可见,运用机器学习算法的大数据风控模型能够比传统风控模型更加精准地预测违约行为,大科技信贷能够利用该模型提供普惠金融服务。

3. 因果推断

在人工智能时代,计量经济学与机器学习结合得越来越紧密。它们如何有效结合成为了最有价值的前沿研究议题(瓦里安,2014)。特别是,计量经济学与机器学习的有机结合将会是今后解决因果推断问题的新方法(阿西,2019)。机器学习可以帮助改进基于传统计量经济学方法的因果推断(阿巴迪、卡塔尼奥,2018;阿西、因本斯,2019)。一方面,机器学习算法可以通过筛选重要的协变量,帮助估计平均处理效应和异质性处理效应(阿西、因本斯,2019)。另一方面,机器学习技术可以通过正则化,在海量数据中进行变量选择,并可以考虑更加复杂的变量关系,构建系统而全面的模型,进行有效的反事实推断(瓦里安,2014)。在会计财务研究领域,采用机器学习进行因果推断的研究日益增加。代表性文献如表5所示。

从信息和市场效率、抵押贷款等研究主题来看,以往研究运用机器学习有效估计处理效应,从而改进因果

表5 会计财务研究领域采用机器学习进行因果推断的代表性文献

研究主题	作者(年份)	研究问题	数据集	机器学习算法	主要结论
信息和市场效率(G14)	雷塞等(2019)	不确定性如何影响市场流动性?	美国股权不动产投资信托数据、每日交易数据和日内交易数据	·PDS LASSO	因受桑迪飓风的不确定性影响,房地产投资信托基金的交易量相对较少,买卖价差更大。
抵押贷款(G21)	科斯特洛等(2020)	是否和如何采用额外的自由裁量权来增强机器对贷款结果的影响?	2018~2019年月度贷款数据	·神经网络 ·因果树	当预测没有社交媒体的私人借款人的未来信用风险时,加入人的自由裁量权有助于增强机器的预测效果。
非对称和私人信息(D82)	刘等(2020)	房地产市场是否存在委托代理问题?	2000~2016年来自佐治亚州亚特兰大和亚利桑那州凤凰城的多重上市服务数据	·DS LASSO	房地产市场不存在代理问题。
资产定价(G12)	冯等(2020)	新因子对资产定价有何贡献?	1976~2017年的150个因子	·DS LASSO ·弹性网 ·PCA ·前向逐步回归	几个新因子在解释资产价格方面具有显著的增量定价能力。

推断。比如,雷塞等(2019)遵循贝洛尼等(2014b)的建议,使用机器学习文献中广泛使用的后向双选 LASSO (Post-Double-Selection LASSO, PDS LASSO) 算法,自动寻找并控制重要的协变量进行平均处理效应的估计,以消除主观选择协变量的担忧,提供了不确定性对市场流动性具有不利影响的证据。科斯特洛等(2020)首先运用神经网络对延期和及时还款的结果进行分类,然后运用阿西和因本斯(2016)提出的因果树(Causal Tree)方法,创建一个协变量空间的分区,并在分区的每个元素中估计处理效应,识别相关横截面分割,计算条件平均处理效应。他们得到了与运用横截面测试一致的结果,即当借款人没有社交媒体账户且处于竞争激烈的市场时,贷款人调整基于机器的推荐对贷款结果的影响更加明显。

从非对称和私人信息、资产定价等研究主题来看,采用双选 LASSO (Double-Selection LASSO, DS LASSO) 选择因子并解决遗漏变量偏差问题以改进因果推断的做法,已经成为学者们的共识。双选 LASSO 分两个阶段进行,第一阶段的 LASSO 以选择一个精简模型为目标,第二阶段的 LASSO 旨在确定最有可能导致遗漏变量偏差的因子。这样,既可以得到一个稳健的模型,又可以减轻潜在模型选择错误导致的遗漏变量偏差问题。例如,刘等(2020)在高维非结构化数据中使用双选 LASSO 程序进行变量选择,有效地解决了模型遗漏变量偏差问题,得出了房地产经纪人并没有使用他们的信息优势损害客户利益的结论。冯等(2020)运用双选 LASSO 方法分析新因子对资产定价的贡献,并采用弹性网、主成分分析(Principal Component Analysis, PCA)和前向逐步回归(Forward Stepwise Regression)进行稳健性检验。这一研究证实:双选 LASSO 通过克服模型选择错误能够做出正确推断;近年提出的几个新因子在解释资产价格方面具有显著的增量定价能力。

五、未来研究展望

本部分从变量测量、事件预测、因果推断、理论构建等4个方面,系统阐述采用机器学习助力管理学研究的策略,进而重点讨论中国学者未来运用机器学习,植根中国情境,挖掘中国数据,变革研究范式和创新研究方法,开展中国特色管理理论研究的主题方向、重要问题、实施策略和主要建议。

(一)变量测量

变量测量可以帮助研究者对变量之间的关系进行量化推断,是社会科学的核心任务之一。管理学作为一门交叉性的社会科学,长期以来受到数据稀缺、数据结构单一的困扰,只能在少量数据的基础上,使用计算能力较弱的统计技术进行分析(格里默等,2021)。大数据所具有的海量性、高速性、多样性、真实性等四大特征(洪永森、汪寿阳,2021c;迈尔-舍恩伯格、库克耶,2013),一方面极大地丰富了管理学研究的数据资源,另一方面给传统统计学与计量经济学的分析方法带来了前所未有的挑战。仅仅依靠传统方法已无法对容量大、维度高、变化快、种类多、噪音高的大数据进行有效的变量测量和数据分析(洪永森、汪寿阳,2021a,2021b),亟需借助新的技术、方法与工具。

机器学习可以帮助学者打破管理学研究主要采用结构化数据进行变量测量的传统,从海量、高维的非结构化数据中提取有价值的信息进行变量测量,将原本只能进行定性分析的数据转变为可以进行定量分析的数

理论述评

据,从而扩大了变量的可得性,丰富了定量研究数据库。因此,机器学习日益成为助力国内外管理学研究领域学者开展变量测量的重要方法。目前,正如前文所述,朴素贝叶斯机器学习技术,已被广泛应用于从文本数据中挖掘语义信息(巴托夫等,2018)、判断语音语调(德佛朗哥等,2011;亨利、利昂,2016;游等,2018)和捕获情绪差异(黄等,2014)。此外,已有研究运用机器学习算法,从音频和视频等数据中提取感兴趣的信息,测量人格特质(哈里森等,2019;马尔霍特拉等,2018;乔杜里等,2019)。

然而,在管理学研究领域,以往运用机器学习进行变量测量的文献存在两个主要方面的不足:通常使用单一的机器学习算法;采用单一来源的数据。例如,在会计财务研究领域广泛使用的朴素贝叶斯算法,存在结果难以复制的弱点(洛克伦、麦克唐纳,2016)。在工商管理研究领域常用来测量CEO五大人格特质的开放式语言工具,只能捕捉人们在口语或书面语言中表现出来的特征,难以测量诸如个人魅力这类需要从视频数据中提取信息的变量(哈里森等,2019)。因此,我们认为,当进行变量测量时,未来的研究可以采用以下两个策略,以提高测量效度。第一,尝试运用多种机器学习算法或融合机器学习算法与其他方法,以弥补单一方法的固有局限性。例如,哈里森等(2020)指出,综合运用基于机器学习算法的开放式语言工具、基于预定义字典的封闭语言工具和视频测量方法,能够帮助研究者从不同方面测量高管人格特质。第二,采用函数数据、矩阵数据、区间数据、符号数据等新型结构化数据(洪永森、汪寿阳,2021a,2021b,2021c),或者融合新型结构化数据与其他数据,选择合适的机器学习算法,构建高效度的变量测量,尽可能减少测量误差,从而最大限度规避研究的内生性问题。

在中国情境下,本文认为可以积极运用多种机器学习方法,对不同类型数据进行深入挖掘,构造独特的变量及其测量。举例来说,企业文化是一个企业的灵魂,是一个企业核心能力的重要组成部分。正如前文所述,中国具有独特的制度和因素,其源远流长的历史文化深刻地影响着企业高层管理者的认知和行为,从而在很大程度上塑造了企业文化。因此,中国学者可以着手研究“中国企业文化具有哪些关键维度?”、“中国企业文化是如何被塑造的?”和“中国企业文化及其每个维度如何影响企业创新行为?”等3个问题,围绕企业文化的3个层次(物质文化、制度文化和精神文化),企业的使命、愿景和价值观,企业文化的功能和塑造,以及企业创新行为等内容,综合使用诸如基于数据降维的LDA算法、基于数理统计的贝叶斯分类算法、基于高低维转换的SVM算法和基于深度学习的CNN算法等多种机器学习方法,更好地从文本、图像、音频、视频等非结构化数据和量化的结构化数据中收集和挖掘有价值的信息,构建企业文化等核心构念,进而做出具有中国特色的高质量研究。

(二)事件预测

如何采用以往的数据和经验,根据变量之间的关系建立统计模型,对感兴趣的事件进行有效预测,是管理学研究领域的一项非常重要的议题(格里默等,2021)。传统计量经济学的主要目标是因果推断,即推断 x_i 对 y_i 的因果效应。为了有效识别和解释因果关系,研究者往往需要事先对 $f(x_i; \beta)$ 的函数形式做很强的假定。因此,它关注的重点是 β ,强调模型的可解释性(陈强,2021),并侧重于构建容易解释的简约模型(蒂芬,2019;格里默等,2021;施雷斯塔等,2021)。以统计显著性和样本内拟合优度为模型优劣主要评估标准的传统计量经济学模型的预测能力较弱(陈强,2021;蒂芬,2019)。机器学习的主要目标不是进行样本内预测,而是在新的、未知的数据集上根据 x_i 预测 y_i (格里默等,2021),关注的重点是 \hat{y}_i 。为实现此目标,它不对模型进行任何事先假定,而是想方设法使用任何可能的函数,甚至难以解释的黑箱方法(陈强,2021)。因此,它强调模型的准确性而非可解释性(蒂芬,2019)。以样本外预测精准度为衡量标准的机器学习模型(蒂芬,2019),可以提供良好的预测能力(洪永森、汪寿阳,2021a),从而帮助研究者做出更精准的预测(格里默等,2021)。因此,机器学习日益成为赋能国内外管理学研究领域学者开展事件预测的重要方法。正如前文所述,在工商管理研究领域,已有学者在人力资源管理决策(栾等,2019)、专利技术识别(乔杜里等,2020)和管理反馈策略(郭晓姝等,2021)中运用机器学习进行事件预测。在会计财务研究领域,已有许多学者使用机器学习技术对会计欺诈(鲍等,2020;佩罗尔斯等,2017)、股票市场风险溢价(顾等,2020;卡洛米里斯、马梅斯基,2019;马内拉、莫雷拉,

2017)和企业违约贷款比例(黄益平、邱晗,2021)进行预测。

然而,已有学者指出,虽然基于机器学习算法可以构建精准度更高的模型,但其预测过程可能不易被我们理解(约曼斯等,2019)。因此,我们认为,当进行事件预测时,未来的研究可以采用以下两个策略,以提高预测效度和预测结果的可解释性。第一,将机器学习算法与人的判断和专业相结合(乔丹、米切尔,2015),减轻机器学习算法可能产生的潜在偏差(乔杜里等,2020),从而提高模型的预测精准度、样本外泛化能力和可解释性。第二,灵活采用大数据集和小数据集,比较机器学习模型的预测精准度。学者们采用机器学习进行事件预测时不应局限于大数据集。当前,管理学研究领域的一些数据的样本量难以达到让机器学习获得精准预测的要求。因此,在小数据集上训练机器学习算法以提高模型预测精准度,可能更具有实际意义(栾等,2019)。

本文认为,在中国情境下研究资本市场投资者情绪问题,是运用机器学习方法进行事件预测的重要研究方向之一。中国资本市场经历30余年探索、改革和发展,已成为全球第二大资本市场。中国特有的政治、经济、制度、文化和资本市场等因素和投资者的心理因素共同影响投资者的情绪。基于这些独特情境,中国学者可以着手研究“如何测度和分类中国资本市场投资者情绪?”和“中国资本市场投资者情绪如何影响其投资行为?”等2个问题,围绕资本市场、股价波动、投资者心理、投资者情绪、投资者偏好和投资行为等内容,收集和挖掘海量的非结构化数据和结构化数据(包括高频甚至实时结构化数据),综合使用诸如SVM、随机森林和提升算法等多种机器学习方法,对投资者情绪进行分类或预测(如积极、消极和混合等多种复杂情绪),探索不同情绪对投资行为的影响,进而为引导投资者理性投资做出重要贡献。

(三)因果推断

如何高效地开展因果推断是计量经济学研究领域的核心议题(埃尔南、罗宾斯,2020;安格里斯特、皮施克,2009,2014;亨廷顿-克莱因,2022;坎宁安,2021;因本斯、鲁宾,2015)。以往学者指出,采用观测数据来估计一个事件或政策干预的因果效应是有问题的(蒂芬,2019),因为这样会遇到因果推断的基本问题(霍兰,1986),即只能观测到某一组样本接受了处理的结果,无法观测这一组样本没有接受处理的状态(即反事实)。因此,无法在实施某一干预后,直接比较同一组样本的两种不同状态,以估计因果效应(陈强,2014;洪永森、汪寿阳,2021a)。传统计量经济学常用样本匹配、断点回归设计、双重差分、合成控制和工具变量等方法进行因果推断。虽然机器学习不能直接揭示因果关系,但是它可以采用灵活的算法实施降维(格里默等,2021;洪永森、汪寿阳,2021a,2021b),解决高维问题或“维数灾难”,通过准确估计反事实和处理效应,帮助精准识别与测度因果效应,从而助力因果推断(洪永森、汪寿阳,2021a,2021b,2021c)。因此,学者们日益将机器学习与传统计量经济学相结合进行因果推断(阿西、因本斯,2016;贝洛尼等,2014a;郭峰、陶旭辉,2023;李超、求文星,2021;钱浩祺等,2021;韦杰、阿西,2018)。本文主要从平均处理效应、异质性处理效应和多种方法融合等3个方面,阐述机器学习赋能管理学研究中因果推断的未来机会。

1. 平均处理效应

因果推断领域的一个重要分支是在无混淆性(Unconfoundedness)假设下,估计平均处理效应(阿西、因本斯,2019;罗森鲍姆、鲁宾,1983;因本斯、鲁宾,2015)。这一假设要求在控制了协变量后,潜在结果独立于处理分配(Treatment Assignment),换句话说就是处理分配与随机分配(Random Assignment)结果一致(阿西,2019)。在此假设下,因果效应可以简单地从处理组与控制组的平均差异之中求得(埃尔南、罗宾斯,2020;安格里斯特、皮施克,2009)。但是,对于管理学研究中的大多数问题,实施随机对照试验设计往往不切实际,甚至根本不可能(蒂芬,2019)。因此,大部分实证研究工作只能依赖观测数据(陈强,2014)。然而,真实世界会存在大量潜在的协变量,如果研究者无法有效识别和控制重要的协变量,则可能估计出有偏的处理效应(黄乃静、于明哲,2018)。机器学习可以发挥其强大的降维和预测能力(格里默等,2021;洪永森、汪寿阳,2021a,2021b),准确估计反事实和平均处理效应,从而帮助研究者做出高质量的因果推断(洪永森,2021;洪永森、汪寿阳,2021a;瓦里安,2014)。在本研究中,我们重点从样本匹配、断点回归设计、双重差分(包括合成控制)和

工具变量等4个方面阐述运用机器学习赋能传统计量经济学开展因果推断的策略。

(1) 样本匹配。样本匹配法的主要思路是,针对处理组个体,在控制组中寻找特征尽可能相似的个体,与处理组个体进行匹配,运用控制组个体的结果估计处理组个体的反事实结果(陈强,2014;因本斯、鲁宾,2015)。传统的样本匹配法包括协变量匹配法和倾向得分匹配法(阿巴迪、卡塔尼奥,2018)。协变量匹配法,顾名思义就是直接根据协变量,在处理组和控制组中进行一对一或一对多近邻匹配(阿巴迪、卡塔尼奥,2018;陈强,2014)。

然而,这种方法仅适用于协变量较少的情况。如果协变量较多,则可能会在高维空间遇到数据稀疏问题,很难进行匹配(陈强,2014;李超、求文星,2021)。因此,罗森鲍姆和鲁宾(1983)提出的倾向得分匹配法是指将高维协变量整合为一个倾向得分后再进行匹配。这样,无须对所有协变量进行匹配,极大降低了偏误。但是,倾向得分匹配法也不是完美的,因为这种方法是采用已知且相对于样本量较小的协变量,无法适用于协变量数量远大于样本数量的情形(阿巴迪、卡塔尼奥,2018;钱浩祺等,2021)。

机器学习有助于解决上述传统匹配法在高维空间上遇到的问题(韦斯特赖希等,2010)。已有学者运用神经网络(濑户口等,2008)、SVM(斯韦雷多斯基、鲍尔迪,2009)、分类与回归树(Classification and Regression Tree, CART)(怀斯等,2014;濑户口等,2008)、提升法(哈德等,2006;麦卡弗里等,2004)等机器学习算法,在高维协变量数据情形下估计倾向得分,并将其作为传统逻辑回归的替代(韦斯特赖希等,2010)。例如,林登和亚诺尔德(2016)直接使用最优判别分析(Optimal Discriminant Analysis)的机器学习算法,按照协变量分布区分研究小组。该算法不仅可以确定个人如何根据协变量的最优切点进行自我选择,而且对偏斜数据和异常值不敏感。它克服了传统方法的诸多缺陷,可以作为评估匹配研究中协变量平衡的传统方法的有力补充或替代方法(林登、亚诺尔德,2016)。卡里姆等(2018)比较了高维倾向得分算法、机器学习算法(随机森林、LASSO和弹性网)和混合方法(前两者的结合)在高维协变量空间中识别和选择混淆因素的能力。这一研究发现,在大多数情况下机器学习算法与高维倾向得分算法的表现一致,甚至更好,而混合方法则表现最好。

(2) 断点回归设计。断点回归设计思想起源于心理学文献(西斯尔思韦特、坎贝尔,1960),由哈恩等(2001)奠定了其在计量经济学领域的理论基础。这一方法经常被用来估计选举结果和政策干预等因果效应(阿纳斯塔索普洛斯,2019)。与大多数因果推断方法相比,断点回归设计需要的假设较少,能够估计临界值附近的局部平均处理效应(李文钊、徐文,2022;纳拉亚南、卡利亚南,2020),与随机实验最为相似(赫兰兹等,2018;李、勒米厄萨,2010)。当处理是基于潜在的连续性强制变量是否超过临界值时,可以采用断点回归设计来估计这种处理效应;只有当一个强制变量高于临界值时才会被处理,否则被分配到控制组(安格里斯特、皮施克,2009;陈强,2014;坎宁安,2021;因本斯、勒米厄,2008)。

但是,断点回归设计仍有不足之处。例如,它没有提供预处理协变量选择的具体方法,往往使用少量观测值进行处理效应估计,这会导致处理效应的大小和精准度因所选强制变量而出现大幅波动。这样,研究者在选择协变量时,往往以处理效应统计显著性为目标,而不是以处理效应估计精准度最大化为目标(阿纳斯塔索普洛斯,2019)。

目前,已有学者将擅长处理高维数据和数据拟合的机器学习与断点回归设计相结合,以有效弥补上述缺陷。例如,赫兰兹等(2018)开发了首个能够用于自动发现断点回归设计的统计机器学习方法(Statistical Machine Learning Approach)。该方法能在任意维度数据中识别可解释的局部断点回归设计,并且可以在没有专家监督的情况下无缝计算处理效应,具有严格的统计和计量经济学有效性。他们将该方法应用于各种合成和真实数据集,展示了它在诸如未观测到的变量、大量噪声和模型设定错误等不利条件下的稳健性能。在选择具体函数形式方面,布兰森等(2019)将在机器学习领域广泛使用的高斯回归应用于断点回归设计,提出的高斯过程回归方法(Gaussian Process Regression, GPR)不仅在一定程度上摆脱了传统方法对函数形式事先设定的依赖,而且在覆盖率、区间长度和均方误差等特性比标准的局部线性回归(Local Linear Regression)和最先进的局部线性回归等方法表现更好。在模型假设条件方面,纳拉亚南和卡利亚南(2020)运用复杂机器学习算法

计算大量高维消费者的购买倾向得分,得到具有连续性的强制变量。这一做法很好地满足了处理状态在断点处发生跳跃式变化、强制变量具有连续性和强制变量无法被精准操纵等苛刻的假设条件。

值得一提的是,阿纳斯塔索普洛斯(2019)将自适应LASSO、常用于机器学习的正则化方法和断点回归设计估计相整合,提出了断点回归设计中局部平均处理效应的自适应LASSO估计方法。这一方法分为4个步骤:第一步是研究者预处理协变量选择;第二步是自适应LASSO正则化;第三步是自动模型选择;第四步是采用卡洛尼科等(2014)提出的方法执行最终模型的正则化稳健估计。这一方法特别适用于小数据集。它允许研究者将实质性知识与自动协变量选择算法相结合,不仅有助于最大限度地提高局部平均处理效应估计的精度,而且能够最大限度地避免研究者盲目追求局部平均处理效应的统计显著性。

(3)双重差分和合成控制。双重差分法,因其原理和模型设置简单、容易理解与运用,自20世纪90年代初以来一直是实证研究人员的重要工具(卡德,1990;卡德、克鲁格,1994)。然而,它存在缺陷。例如,当处理组和控制组中随时间变化的协变量受到处理的影响时,使用它来恢复无偏因果效应很困难(泽尔多、哈特菲尔德,2021)。处理组和控制组的组成因处理而改变(安格里斯特、皮施克,2009,2014)。合成控制法(阿巴迪、加迪扎巴尔,2003;阿巴迪等,2010,2015),通过使用一组协变量的加权平均值,而不是单一的协变量或者协变量的简单平均值,构建处理组的反事实,从而明显改进了双重差分法的估计效果。

但是,无论是双重差分法还是合成控制法均无法很好地选择控制因素,包含太少就会造成偏差,包含过多则导致低精度。因此,不适合在高维空间中进行反事实估计(金恩,2018)。为此,已有学者进行了相关尝试,运用机器学习方法替代传统方法,采用负权重为重构处理单元提供了更大的灵活性,从而有效地改进了传统方法对处理效应的估计。例如,杜琴科和因本斯(2016)发现带有惩罚项的弹性网,可以作为合成控制法计算权重的替代方法,并且在有大量潜在协变量的环境中表现良好。本-迈克尔等(2021)将RR和合成控制法相结合,提出了岭增强合成控制法。这一方法允许负权重情况存在,采用外推法改善拟合结果,同时运用正则化减少过拟合风险,从而极大地提高了反事实估计的能力。

特别是,金恩(2018)为比较一个被处理单元的时间序列与由大量控制单元构建的反事实提供了一个通用框架,旨在通过最小化欠拟合和过拟合之间的权衡来预测反事实。这一研究将两种传统方法与PCR、LASSO和贝叶斯结构时间序列模型(Bayesian Structural Time Series)等机器学习算法进行对比,发现当潜在的控制因素很多或者被处理单元与控制因素有很大不同时,机器学习方法的表现优于传统方法。当应用于三个宏观经济案例,该方法再次证明了灵活的机器学习方法能够更大程度地捕捉数据生成过程,进行更精准的反事实估计,从而助力因果推断。

(4)工具变量。以解决内生性为目标的工具变量法是计量经济学领域进行因果推断的有力武器(安格里斯特、皮施克,2009;坎宁安,2021;因本斯、鲁宾,2015)。但是,在现实世界想要找到一个“完美”的工具变量是非常困难的,要么可能无法完全满足外生性,与某些随机扰动项相关(陈云松,2012;胡安宁,2015),要么可能与自变量关系非常微弱,仅为弱工具变量,因而不仅无法显著提高有效性,而且会增强偏差性(邦德等,1995;邱嘉平,2020)。

目前,已有文献将机器学习算法与工具变量相结合,不仅可以帮助研究者摆脱长久以来对工具变量的线性模型假设,还可以利用机器学习强大的预测能力帮助研究者获得最佳工具变量(郭峰、陶旭辉,2023),从而大幅提高因果推断的精度。例如,辛格和孙(2019)运用去偏机器学习方法(Debiased Machine Learning),估计带有高维数据的编译器(Complier)参数。这种方法无需对工具变量识别的假定进行任何函数形式的限制,并且采用这种方法得到的估计量具有一致性、渐进正态性和半参数有效性(辛格、孙,2019)。麦卡洛克等(2021)将贝叶斯叠加回归树(Bayesian Additive Regression Trees, BART)与工具变量相结合提出的基于贝叶斯叠加回归树的工具变量算法,可以处理非线性和具有非正态误差的结构方程,从而突破了长期以来对工具变量的线性模型假设,让工具变量方法的运用更加灵活。

本文以辛格等(2020)提出的机器学习工具变量(Machine-Learned Instrumental Variables, MLIV)算法为例

理论述评

进行详细说明。工具变量法导致的有偏或不精准的因果效应估计受到多个学科的批评,当前的解决方法主要是从可获得的外生数据中构造近似的最佳工具变量。但是,这一方法需要对最佳工具变量结构进行明确的理论假设,而这些假设的相关性会随着实际应用而变化。为此,他们运用LASSO回归、RR、弹性网、随机森林、极端梯度提升(eXtreme Gradient Boosting, XGBoost)和神经网络等多种机器学习方法,提出了一种直接从数据中学习得到强工具变量的MLIV算法。这一算法利用可获得的外生信息从算法上优化效率准则,直接利用算法构造最佳工具变量,允许从样本数据中同时进行工具变量训练和因果推断,并且不需要依赖任何关于模型或最佳工具变量结构的明确假定,减轻了弱工具变量的诅咒,因而是一个很有前途的工具。此外,他们还提出了正式的渐进理论,展示了机器学习工具变量的半参数效率,证明了该算法能够大幅增强从观测数据中进行因果推断的性能。

2. 异质性处理效应

因果推断领域的另一个重要分支是研究处理效应在不同样本中的异质性(阿西,2019;阿西、因本斯,2019)。传统方法常常采用在回归模型中加入协变量与处理变量的交互项来估计异质性处理效应。然而,该方法只能在协变量数量有限的情况下得到异质性处理效应的一致估计(阿西、因本斯,2015;黄乃静、于明哲,2018)。若协变量数量很多,则一方面无法判断到底哪些是重要的协变量,另一方面不能在模型中无限添加交互项。这样,交互项的设定非常困难,很可能具有很强的主观性,甚至随意性(胡安宁等,2021)。

运用“数据驱动”的机器学习算法,有利于在高维协变量空间解决上述问题(钱浩祺等,2021)。例如,阿西和因本斯(2016)根据回归树方法提出因果树方法,估计异质性处理效应,从而进行因果推断。因果树方法首先对协变量空间进行划分,然后在子空间内计算处理组与控制组的平均差异求得平均处理效应,最后比较不同子空间的平均处理效应得到异质性处理效应。虽然这一方法可以通过交叉验证避免过拟合问题,但因果树的结构存在一定随意性,可能产生有偏估计(阿西、因本斯,2019)。因此,韦杰和阿西(2018)提出了一种基于随机森林的因果森林(Causal Forest)方法。该方法首先构建一个包含许多因果树的因果森林,然后在每一颗树上根据阿西和因本斯(2016)的方法进行计算,最后对这些因果树的计算结果进行平均,即可估计出异质性处理效应。这一研究展示,因果森林对于真实处理效应具有逐点一致性,并且具有渐近高斯和中心抽样分布等性质。值得一提的是,莱希纳(2018)基于韦杰和阿西(2018)提出的修正因果森林(Modified Causal Forest),已被灵活运用于会计财务领域的最新研究之中(奥德里诺等,2022)。因此,我们认为,采用基于随机森林的因果森林、修正因果森林(莱希纳,2018)和广义随机森林(阿西等,2019)来估计异质性处理效应是机器学习帮助研究者有效进行因果推断的重要策略。

3. 多种方法融合

除了上述机器学习赋能因果推断的重要策略,我们认为,融合传统计量经济学的因果推断与基于机器学习算法的异质性处理效应估计等多种方法也是一种非常重要的策略。已有研究采用多种方法融合进行因果推断。比如,巴尔加利斯托菲和格内科(2020)将因果树和工具变量相结合,提出的带有工具变量的因果树(Causal Tree with Instrumental Variable, CT-IV),可以在复杂高维数据空间中估计处理效应的异质性。接着,他们对CT-IV进行改进,提出了带有工具变量的诚实因果树(Honest Causal Tree with Instrumental Variable, HCT-IV)。后者比前者能更可靠地估计异质性因果效应。王等(2021)将工具变量法纳入因果树,提出的工具变量树(Instrumental Variable Tree),不仅可以纠正观测数据中可能存在的内生性偏差,还可以获得比传统因果树方法更好的覆盖率和更小的均方误差。此外,他们进一步提出的工具变量森林(Instrumental Variable Forest),得到了比广义随机森林更高的准确性和层次性(Stratification)。可以肯定的是,未来机器学习将会在因果推断中发挥越来越大的重要作用(阿巴迪、卡塔尼奥,2018;阿西,2019;阿西、因本斯,2019)。

如何运用机器学习算法精准构建反事实结果以做出有效的政策评价,是因果推断领域需要重点关注的研究议题(郭峰、陶旭辉,2023)。中国特有的文化基因提倡基于道德规范与社会舆论的监管模式,更加要求上市公司注重和满足外部利益相关者的诉求(王永贵等,2021)。中国具有独特的资本市场,政府在经济运行中发

挥着重要的宏观调控作用。国有企业,作为中国资本市场价值创造的重要载体,存在双重委托代理关系。这些诸多因素影响上市公司的治理目标、治理结构和治理模式。基于这些中国情境,中国学者可以围绕上市公司的治理目标、治理结构、治理能力、治理模式、治理效果、创新行为和创新绩效等内容,构建大数据库,尝试研究“中国上市公司如何构建高质量的治理结构?”、“中国上市公司治理结构如何影响企业创新行为?”和“中国上市公司治理模式对不同规模企业的创新绩效有何影响?”等3个问题。需要指出的是,这些研究问题通常涉及海量、高维、高频和非结构化的数据,传统方法无法从中提取有效信息进行高质量运用;传统方法也常因假设条件过于苛刻而失效,无法较好回答某些重要的研究问题。因此,亟需在管理学研究中引入机器学习算法,拓展传统方法适用边界,增强样本匹配随机性和提升反事实结果预测精准度(郭峰、陶旭辉,2023;钱浩祺等,2021)。我们建议学者将树型算法、惩罚回归算法和神经网络算法等多种机器学习算法,与样本匹配、断点回归、双重差分、合成控制和工具变量等计量经济学方法相结合,在高维空间匹配样本、识别断点和构造强工具变量,精准估计反事实、平均处理效应、异质性处理效应和分位数处理效应,进而有效识别因果关系,提升因果推断性能。可以肯定的是,秉承开放思维,深度融合实证模型、机器学习和随机实验等多种方法综合提升研究的内部效度和外部效度,寻求具有更高准确性、稳健性和泛化能力的因果效应结论是因果推断的未来研究趋势(郭峰、陶旭辉,2023;李文钊,2018a,2018b;李文钊、徐文,2022)。

(四)理论构建

在管理学的传统中,理论构建仅仅是少数案例研究者的特权(施雷斯塔等,2021)。大家普遍认为从定性数据中构建的理论才具有可靠性(洛克,2015;沙阿、科利,2006)。目前,大部分研究者尚未意识到机器学习为实现从定量数据中构建理论所能提供的价值。造成这一现象的可能原因主要是两个方面:第一,使用机器学习技术不能产生统计上一致的系数估计或可靠的标准误(乔杜里等,2021),从而无法直接应用于传统的演绎性假设检验(穆来纳森、斯皮斯,2017)。第二,运用复杂机器学习算法所产生的预测缺乏可解释性(蒂德尔、艾森哈特,2020;施雷斯塔等,2021)。

然而,正因为机器学习具有这些特性,它能够在模式发现这一理论构建的关键步骤中发挥重要作用。借助机器学习,研究者可以从数据中发现复杂和稳健的模式,并把这些模式作为溯因理论创建的输入,同时在不带来过拟合风险的情况下创建一个单独的保留样本来检验理论,从而有助于从数据中创建理论(施雷斯塔等,2021)。施雷斯塔等(2021)提出了一个基于机器学习的理论构建流程。该流程包含4个步骤:样本拆分、模式发现、理论解释和理论检验。样本拆分是指将数据随机分为算法支持的归纳样本(即用于模式发现)和大小相等的保留样本(即用于假设检验);模式发现是指采用机器学习算法在归纳样本中发现具有稳健性和可解释性的模式;理论解释是指输出一组样本外和模式外的假设;理论检验是指在保留样本中进行假设检验(施雷斯塔等,2021)。其中,理论解释阶段和理论检验阶段需要人的专业知识和判断力,不能交给机器学习算法来完成(乔杜里等,2021;施雷斯塔等,2021)。因此,我们认为,当进行理论构建时,未来的研究可以采用这一流程。特别是,未来的研究可以借鉴以往文献的具体做法(蒂德尔、艾森哈特,2020;何等,2020),综合使用机器学习和多案例分析这两种高度互补的方法进行理论构建。

基于中国情境,运用机器学习方法探究关系国计民生的重大经济社会议题,是未来理论构建的主要方向。中国在脱贫攻坚实践中,积极借鉴国际经验,紧密结合中国实际,创造性地提出并实施精准扶贫方略,创造了中国减贫史乃至人类减贫史上的伟大奇迹^⑨。企业是精准扶贫的重要主体。针对这个特定的中国情境,中国学者可以着手研究“为什么中国企业能够在精准扶贫过程中发挥举足轻重的作用?”和“中国企业助力脱贫攻坚的成功模式是什么?”等2个问题。首先,围绕企业精准扶贫的多元动机、行为模式和经济结果等内容,收集和挖掘海量、复杂、高维的农村贫困数据和企业精准扶贫数据,利用机器学习善于处理非结构化数据的优势,针对文本数据运用朴素贝叶斯、LDA和SVM等具有良好表现的算法,针对高维稀疏特征数据运用LASSO回归、RR和GPR等具有良好预测效果的算法,测量贫困家庭和贫困人口的收入、生活质量、健康状况、劳动能力和教育水平以及企业扶贫政策的制定、执行和监督等变量,识别和分类这些贫困家庭、贫困人口和企业扶贫

理论述评

政策,预测贫困对象的发展能力。接着,利用机器学习的强大降维和预测能力,准确估计反事实,全面科学评估企业精准扶贫政策的实施效果,建立因果关系。最后,利用机器学习算法擅长从数据中学习并归纳总结规律的优势,发现新的、具有稳健性和可解释性的企业扶贫方式、体制和机制,并综合运用机器学习(例如,随机森林、提升法、惩罚回归、神经网络等算法)和多案例分析等高度互补的方法,构建中国特色的企业精准扶贫理论。

六、结论与讨论

以弥补管理学研究传统上所存在的短板为目标,本研究采用1999~2021年发表在工商管理 and 会计财务两大研究领域的国内外顶级期刊的学术文献,识别了学术界借助机器学习赋能管理学实证研究的核心途径;从多个角度系统梳理了这些途径的代表性文献;详细阐述了机器学习赋能管理学研究的主要策略,并重点讨论了中国学者运用机器学习开展中国特色管理理论研究的未来机会(主题方向、重要问题、实施策略和主要建议)。本研究得出如表6所示的主要结论。可以预见的是,在未来,变量测量、事件预测、因果推断、理论构建等4种核心途径的融合将日益紧密。它们的融合为机器学习赋能管理学研究提供了更加具有深度和广度的未来机会。例如,事件预测可以用来揭示数据中难以假设的复杂和未知关系,开发新的理论构念及其测量,或者按照预测的相对精准度比较竞争理论(克鲁帕、米努蒂-梅扎,2022),从而更好地进行理论构建(蒂德尔、艾

表6 本研究主要结论

管理学研究短板	赋能途径	主要策略	算法举例	研究机会举例
<ul style="list-style-type: none"> ·无法充分利用海量、复杂、高维、非结构化、多模态数据中的信息。 ·受制于社会经济系统的非实验性、不可逆性和时变性,以及社会经济数据的缺陷。 ·定性研究基于少数个案,未使用大样本数据;受限于人工经验,难以识别新颖、稳健、具有预测性和可解释性的模式;构建的理论具有较弱的泛化能力。 ·基于观测数据的定量研究常常采用传统计量经济学方法,而这一方法存在诸多明显的局限性。例如,无法捕获高维动态大数据中呈现的时变性、非线性和非平稳性特征;存在模型误设偏差;模型的泛化能力较差;无法对多因多果且有滞后性的事件做出精准的因果推断。 	变量测量	<ul style="list-style-type: none"> ·利用机器学习算法充分挖掘文本、图像、音频、视频、电子化记录信息等非结构化数据中的有用信息。 ·运用多种机器学习算法或融合机器学习算法与其他方法,弥补单一方法的固有缺陷。 ·融合信息含量丰富、数据形式多层次化的函数数据、矩阵数据、区间数据、符号数据等新型结构化数据与其他数据,选择合适的机器学习算法,构建高效度的变量测量。 	<ul style="list-style-type: none"> ·朴素贝叶斯 ·LDA ·CNN ·SVM ·GBM 	<ul style="list-style-type: none"> ·构建具有中国特色的企业文化等核心构念
	事件预测	<ul style="list-style-type: none"> ·利用机器学习算法直接从超高维动态复杂数据中挖掘潜在数据证据和函数关系,灵活选择模型和函数形式,同时将机器学习算法与人的判断和专业知有机相结合,提高模型的可解释性,减轻模型误设偏差以及因机器学习算法产生的潜在偏差。 ·运用交叉验证和正则化等方法,防止模型过拟合,提高模型的样本外预测精准度和泛化能力。 ·灵活采用大数据集和小数据集,训练和提高机器学习模型的预测精准度。 	<ul style="list-style-type: none"> ·SVM ·逻辑回归 ·随机森林 ·LASSO ·ANN ·KNN 	<ul style="list-style-type: none"> ·分类或预测中国资本市场投资者情绪等关键事件
	因果推断	<ul style="list-style-type: none"> ·利用机器学习的强大降维和预测能力,准确估计不可能真实发生的反事实和平均处理效应,以做出高质量的因果推断。比如,引入机器学习算法改进高维空间的协变量匹配;利用机器学习算法自动识别断点;运用机器学习算法合理选择控制因素;采用机器学习算法寻找最佳工具变量。 ·采用修正因果森林和广义随机森林等机器学习算法来估计异质性处理效应。 ·充分融合传统计量经济学的因果推断与机器学习算法,更好地估计异质性处理效应。 ·利用机器学习在海量、高维、高频数据中高效识别复杂关系的优势,精准预测反事实结果,同时结合LASSO等机器学习降维技术剔除次要或滞后影响因素。 	<ul style="list-style-type: none"> ·自适应LASSO断点回归法 ·岭增强合成控制法 ·去偏机器学习 ·贝叶斯叠加回归树 ·MLIV ·修正因果森林 ·广义随机森林 ·工具变量树 ·工具变量森林 	<ul style="list-style-type: none"> ·识别中国上市公司治理结构影响企业创新行为的因果关系 ·识别中国上市公司治理模式影响企业创新绩效的因果关系
	理论构建	<ul style="list-style-type: none"> ·平均处理效应 ·样本匹配 ·断点回归设计 ·双重差分和合成控制 ·工具变量 ·异质性处理效应 ·多种方法融合 	<ul style="list-style-type: none"> ·采用基于机器学习的理论构建的4步流程:样本拆分、模式发现、理论解释和理论检验。在模式发现这一关键步骤中,利用机器学习算法在归纳样本中精准识别稳健且具有相对可解释性的模式。同时,结合人的专业知识和判断力进行理论解释和理论检验。 ·综合运用机器学习和多案例分析等互补方法进行理论构建。 	<ul style="list-style-type: none"> ·随机森林 ·LASSO ·RR ·ANN ·GBRT

森哈特,2020)。

中国独特的政治、经济、文化和社会制度,蕴含明显异于西方世界的特性(王永贵等,2021)。特别是,我国改革开放以来建立的将“看不见的手”与“看得见的手”结合的中国特色社会主义市场经济体制驱动中国经济取得了惊人的发展(洪永森、汪寿阳,2021b)。这些情境、事实和数据给构建中国特色管理理论提供了肥沃的土壤(黄群慧,2018)。中国管理故事需要构建和使用中国本土化理论才能进行有效阐释。况且,中国管理学界已经形成普遍共识,即中国管理学的根本任务是构建中国管理学原创性理论(洪永森、汪寿阳,2021b;贾良定等,2015;王永贵等,2021)。这给构建中国特色管理理论提出了富有深远意义的迫切要求(王永贵等,2021)。中国学者在运用机器学习方法构建中国特色管理理论时,一定要聚焦于党在经济领域、政治领域、社会领域、文化领域和生态文明领域百年奋斗所取得的重大成就和历史经验;一定要围绕国家重大战略需求,把握时世发展变迁的关键特征(柴国俊、孙若宸,2022);一定要植根中国特有的制度和因素,深入挖掘中国情境的内涵,深度透视中国问题的本质;一定要权衡理论的普适性、科学性与中国特色(王永贵等,2021)。

同时,学者还应该清醒地认识到机器学习不是万能的,它也存在固有的不足和缺陷。第一,机器学习的黑箱特质无法清晰、透明地展示思考决策的过程和逻辑,造成模型可解释性较低等问题。对此,研究者们可以通过理论思考和实证检验对黑箱进行有效合理地拆解,从而收获新视角和新洞见(陈云松等,2020)。第二,机器学习容易出现过拟合现象,导致模型泛化能力不足和样本外预测精度不高等问题。通过实施正则化、样本分割技术、使用更多数据、删除无用特征等方法,能够有效提升模型的可迁移性(钱浩祺等,2021)。第三,因机器学习算法歧视带来的公平性和伦理性问题不容忽略。为此,我们需要在描述语言、建模技术和求解方法上引入因果理论(李家宁等,2023)。

因此,中国学者在运用机器学习方法进行管理学研究时,需要尽可能构建多来源的非结构化和结构化数据集;权变地联合使用多种互补的机器学习算法,将机器学习方法与传统计量经济学方法有机结合进行因果推断;遵循基于机器学习构建理论的基本流程:样本拆分、模式发现、理论解释和理论检验,综合运用机器学习、人的判断和专业知识和基于个人、团队和组织的多案例分析等高度互补的方法进行理论构建。对于大数据特别是胖大数据,需要更加关注模型选择而不是解释变量的统计显著性,这样更有助于改进对数据的拟合或预测效果(洪永森、汪寿阳,2021a)。此外,在研究范式和研究方法方面,需要借鉴国际同行所使用的范式和方方法,使用国际同行能够理解的学术语言,讲述中国管理故事。

综上所述,机器学习正在深刻改变管理学的研究范式与方法。我们希望越来越多的中国学者能够基于中国情境、中国问题和中国数据,大胆运用国际前沿机器学习技术、方法与工具推动中国管理学研究的范式变革和方法创新,构建原创性的中国特色管理理论,持续为我国经济建设、社会发展和文化繁荣提供精准指导,并提升中国管理学的国际学术影响力,更好地为世界贡献中国管理智慧^①。

(作者单位:刘景江、郑畅然,浙江大学管理学院;洪永森,中国科学院大学经济与管理学院、中国科学院大学数字经济监测预测预警与政策仿真教育部哲学社会科学实验室、中国科学院数学与系统科学研究院)

注释

- ①关于机器学习与传统的统计学以及计量经济学的区别和联系,请参见陈强(2021)的第4~7页和洪永森、汪寿阳(2021a)。
- ②因为会计与财务关系紧密,本文参照何佳讯等(2021)的做法把这两个领域合并为一个领域,统称为会计财务领域。
- ③UTD-24期刊是指美国德克萨斯大学达拉斯分校界定的24本国际管理学顶级期刊。它覆盖6个研究领域。本文采用何佳讯等(2021)对这6个领域的中文译名。每个领域包含的期刊如下。

会计领域: The Accounting Review、Journal of Accounting and Economics 和 Journal of Accounting Research。

财务领域: Journal of Finance、Journal of Financial Economics 和 The Review of Financial Studies。

信息系统领域: Information Systems Research、Informs Journal on Computing 和 MIS Quarterly。

营销领域: Journal of Consumer Research、Journal of Marketing、Journal of Marketing Research 和 Marketing Science。

管理科学领域: Management Science、Operations Research、Journal of Operations Management、Manufacturing and Service Operations Management 和 Production and Operations Management。

工商管理领域: Academy of Management Journal、Academy of Management Review、Administrative Science Quarterly、Organization Science、Journal of International Business Studies 和 Strategic Management Journal。

④“2021 中国最具国际影响力学术期刊(人文社会科学)”前 20 名中有 7 本管理学期刊,它们是:管理世界、Frontiers of Business Research in China、International Journal of Innovation Studies、中国管理科学、Journal of Management Science and Engineering、中国社会科学和 Fudan Journal of the Humanities and Social Sciences。

⑤通过系统阅读管理学领域中实质性运用机器学习方法进行实证研究的文章,我们发现在工商管理领域,学者们大多运用机器学习进行变量测量、事件预测和理论构建;在会计财务领域,学者们大多运用机器学习进行变量测量、事件预测和因果推断;而在其他领域,学者们大多运用机器学习进行变量测量和事件预测。因此,本文聚焦于工商管理和会计财务这两大积极探索机器学习不同用途的核心领域。

⑥图 1 中未显示的年份的文献数为 0。

⑦图 2 中未显示的年份的文献数为 0。

⑧根据我们的数据,在国际顶级管理学期刊上,工商管理研究领域没有运用机器学习赋能因果推断的文献,会计财务研究领域没有运用机器学习赋能理论构建的文献。国内顶级管理学期刊在工商管理研究领域和会计财务研究领域均没有运用机器学习赋能因果推断和理论构建的文献。因此,本部分未能对这些文献进行综述。

⑨参见:中华人民共和国国务院新闻办公室:《人类减贫的中国实践》,《人民日报》,2021 年 4 月 7 日。

⑩中外文人名(机构名)对照:贝尔(Bell);乔丹(Jordan);米切尔(Mitchell);陈(Chen);梅森(Mason);瓦里安(Varian);阿西(Athey);因本斯(Imbens);布莱曼(Breiman);穆来纳森(Mullainathan);斯皮斯(Spiess);蒂德尔(Tidhar);艾森哈特(Eisenhardt);乔杜里(Choudhury);施雷斯塔(Shrestha);蒂芬(Tiffin);冯(Feng);顾(Gu);卡洛米里斯(Calomiris);马梅斯基(Mamaysky);马内拉(Manella);莫雷拉(Moreira);马尔霍特拉(Malhotra);游(You);栾(Luan);谢(Hsieh);刘(Liu);塞缪尔(Samuel);杰隆(Géron);罗森布拉特(Rosenblatt);亨特(Hunt);卡博内尔(Carbonell);昆兰(Quinlan);鲁梅尔哈特(Rumelhart);科尔特斯(Cortes);瓦普尼克(Vapnik);弗罗因德(Freund);夏皮尔(Schapiro);欣顿(Hinton);萨拉胡季诺夫(Salakhutdinov);古德费洛(Goodfellow);普拉莫德(Pramod);杜尔哈尔(Dulhare);巴顿(Patton);格里默(Grimmer);坎宁安(Cunningham);焦亚(Gioia);皮特雷(Pitre);科利(Corley);克里斯蒂安森(Christiansen);钱丹(Chandan);班伯格(Bamberger);哈里森(Harrison);何(He);亨利(Henry);利昂(Leone);黄(Huang);巴托夫(Bartov);布朗(Brown);戴尔(Dyer);劳里(Lowry);佩罗尔斯(Perols);鲍(Bao);阿巴迪(Abadie);卡塔尼奥(Cattaneo);雷塞(Rehse);贝洛尼(Belloni);科斯特洛(Costello);迈尔-舍恩伯格(Mayer-Schönberger);库克耶(Cukier);德弗朗哥(De Franco);洛克伦(Loughran);麦克唐纳(McDonald);约曼斯(Yeomans);埃尔南(Hernán);罗宾斯(Robins);安格里斯特(Angrist);皮施克(Pischke);亨廷顿-克莱因(Huntington-Klein);鲁宾(Rubin);霍兰(Holland);韦杰(Wager);罗森鲍姆(Rosenbaum);韦斯特赖希(Westreich);濑户口(Setoguchi);斯韦雷多斯基(Sweredoski);鲍尔迪(Baldi);怀斯(Wyss);哈德(Harder);麦卡弗里(McCaffrey);林登(Linden);亚诺尔德(Yarnold);卡里姆(Karim);西斯尔思韦特(Thistlethwaite);坎贝尔(Campbell);哈恩(Hahn);阿纳斯塔索普洛斯(Anastasopoulos);纳拉亚南(Narayanan);卡利亚南(Kalyanam);赫兰兹(Herlands);李(Lee);勒米厄萨(Lemieux);勒米厄(Lemieux);布兰森(Branson);卡洛尼科(Calonic);卡德(Card);克鲁格(Krueger);泽尔多(Zeldow);哈特菲尔德(Hatfield);加迪扎巴尔(Gardeazabal);金恩(Kinn);杜琴科(Doudchenko);本-迈克尔(Ben-Michael);邦德(Bound);辛格(Singh);孙(Sun);麦卡洛克(McCulloch);莱希纳(Lechner);奥德里诺(Audrino);巴尔加利斯托菲(Bargagli Stoffi);格内科(Gnecco);王(Wang);洛克(Locke);沙阿(Shah);克鲁帕(Krupa);米努蒂-梅扎(Minutti-Meza);美国专利商标局(U.S. Patent & Trademark Office);美国证券交易委员会(Securities and Exchange Commission);会计和审计执行公告(Accounting and Auditing Enforcement Releases);纽约证券交易所(New York Stock Exchange);美国证券交易所(American Stock Exchange);纳斯达克(National Association of Securities Dealers Automated Quotations);批量数据存储系统(Bulk Data Storage System);不动产投资信托(Real Estate Investment Trusts);多重上市服务(Multiple Listing Service)。

参考文献

- (1)柴国俊、孙若宸:《微观计量经济学应用中的若干问题探析》,《河北经贸大学学报》,2022 年第 6 期。
- (2)陈冬梅、王俐珍、陈安霓:《数字化与战略管理理论——回顾、挑战与展望》,《管理世界》,2020 年第 5 期。
- (3)陈强:《高级计量经济学及 Stata 应用》,高等教育出版社,2014 年。
- (4)陈强:《机器学习及 Python 应用》,高等教育出版社,2021 年。
- (5)陈晓萍、徐淑英、樊景立:《组织与管理研究的实证方法》,北京大学出版社,2008 年。
- (6)陈云松:《逻辑、想象和诠释:工具变量在社会科学因果推断中的应用》,《社会学研究》,2012 年第 6 期。
- (7)陈云松、吴晓刚、胡安宁、贺光桦、句国栋:《社会预测:基于机器学习的研究新范式》,《社会学研究》,2020 年第 3 期。
- (8)宫晓莉、熊熊、张维:《我国金融机构系统性风险度量与外溢效应研究》,《管理世界》,2020 年第 8 期。
- (9)郭峰、陶旭辉:《机器学习与社会科学中的因果关系:一个文献综述》,《经济学(季刊)》,2023 年第 1 期。
- (10)郭晓姝、张焱、徐健:《道歉承诺类管理反馈策略对顾客二次满意度的影响——基于顾客多样情绪调节效应》,《中国管理科学》,2021 年第 2 期。
- (11)何佳讯、葛佳焯、张凡:《中国学者管理学研究的世界贡献:国际合作、前沿热点与贡献路径——基于世界千种管理学英文期刊论文(2013~2019 年)的定量分析》,《管理世界》,2021 年第 9 期。
- (12)洪永森:《计量经济学的地位、作用和局限》,《经济研究》,2007 年第 5 期。
- (13)洪永森:《理解现代计量经济学》,《计量经济学报》,2021 年第 2 期。
- (14)洪永森、刘俸奇、薛涧坡:《政府与市场心理因素的经济影响及其测度》,《管理世界》,2023 年第 3 期。
- (15)洪永森、汪寿阳:《数学、模型与经济思想》,《管理世界》,2020 年第 10 期。
- (16)洪永森、汪寿阳:《大数据、机器学习与统计学:挑战与机遇》,《计量经济学报》,2021 年 a 第 1 期。
- (17)洪永森、汪寿阳:《大数据革命和经济学研究范式与研究方法》,《财经智库》,2021 年 b 第 1 期。
- (18)洪永森、汪寿阳:《大数据如何改变经济学研究范式?》,《管理世界》,2021 年 c 第 10 期。

- (19)洪永森、汪寿阳:《人工智能新近发展及其对经济学研究范式的影响》,《中国科学院院刊》,2023年第3期。
- (20)胡安宁:《社会科学因果推断的理论基础》,社会科学文献出版社,2015年。
- (21)胡安宁、吴晓刚、陈云松:《处理效应异质性分析——机器学习方法带来的机遇与挑战》,《社会学研究》,2021年第1期。
- (22)胡楠、薛付婧、王昊楠:《管理者短视主义影响企业长期投资吗?——基于文本分析和机器学习》,《管理世界》,2021年第5期。
- (23)黄乃静、于明哲:《机器学习对经济学研究的影响研究进展》,《经济学动态》,2018年第7期。
- (24)黄群慧:《改革开放四十年中国企业管理学的发展——情境、历程、经验与使命》,《管理世界》,2018年第10期。
- (25)黄益平、邱晗:《大科技信贷:一个新的信用风险管理框架》,《管理世界》,2021年第2期。
- (26)贾良定、尤树洋、刘德鹏、郑祎、李珏兴:《构建中国管理学理论自信之路——从个体、团队到学术社区的跨层次对话过程理论》,《管理世界》,2015年第1期。
- (27)蒋锋、张文雅:《机器学习方法在经济研究中的应用》,《统计与决策》,2022年第4期。
- (28)李宝元、董青、仇勇:《中国管理学研究:大历史跨越中的逻辑困局——相关文献的一个整合性评论》,《管理世界》,2017年第7期。
- (29)李超、求文星:《基于机器学习的因果推断方法研究进展》,《统计与决策》,2021年第11期。
- (30)李家宁、熊睿彬、兰艳艳、庞亮、郭嘉丰、程学旗:《因果机器学习的前沿进展综述》,《计算机研究与发展》,2023年第1期。
- (31)李文钊:《因果推理中的科学模型——反事实、选择性偏差与赫克曼结构计量经济学模型》,《实证社会科学》,2018年a第2期。
- (32)李文钊:《因果推理中的潜在结果模型:起源、逻辑与意蕴》,《公共行政评论》,2018年b第1期。
- (33)李文钊、徐文:《基于因果推理的政策评估:一个实验与准实验设计的统一框架》,《管理世界》,2022年第12期。
- (34)罗家德、樊瑛:《计算社会科学如何扎根真相——以如何找出中国风险投资的“产业领袖”为例》,《探索与争鸣》,2018年第7期。
- (35)马黎珺、伊志宏、张澈:《廉价交谈还是言之有据?——分析师报告文本的信息含量研究》,《管理世界》,2019年第7期。
- (36)钱浩祺、龚嫣然、吴力波:《更精确的因果效应识别:基于机器学习的视角》,《计量经济学报》,2021年第4期。
- (37)邱嘉平:《因果推断实用计量方法》,上海财经大学出版社,2020年。
- (38)沈艳、王靖一:《媒体报道与未成熟金融市场信息透明度——中国网络借贷市场视角》,《管理世界》,2021年第2期。
- (39)沈艺峰、王夫乐、黄娟娟、纪荣嵘:《高管之“人”的先天特征在IPO市场中起作用吗?》,《管理世界》,2017年第9期。
- (40)汪寿阳、洪永森、霍红、方颖、陈海强:《大数据时代下计量经济学若干重要发展方向》,《中国科学基金》,2019年第4期。
- (41)王芳、王宣艺、陈硕:《经济学研究中的机器学习:回顾与展望》,《数量经济技术经济研究》,2020年第4期。
- (42)王永贵、汪寿阳、吴照云、吴晓波、毛基业、戚聿东、张维、苏宗伟、朱旭峰、杨开峰、杨立华、席西民、李新春、张玉利、徐向艺、刘志阳、闫妍、魏江、郜亮亮:《深入贯彻落实习近平总书记在哲学社会科学工作座谈会上的重要讲话精神 加快构建中国特色管理学体系》,《管理世界》,2021年第6期。
- (43)张宗新、吴钊颖:《媒体情绪传染与分析师乐观偏差——基于机器学习文本分析方法的经验证据》,《管理世界》,2021年第1期。
- (44)周志华:《机器学习》,清华大学出版社,2016年。
- (45)Abadie, A. and Cattaneo, M. D., 2018, “Econometric Methods for Program Evaluation”, *Annual Review of Economics*, vol.10(1), pp.465~503.
- (46)Abadie, A., Diamond, A. and Hainmueller, J., 2010, “Synthetic Control Methods for Comparative Case Studies: Estimating the Effect of California’s Tobacco Control Program”, *Journal of the American Statistical Association*, vol.105(490), pp.493~505.
- (47)Abadie, A., Diamond, A. and Hainmueller, J., 2015, “Comparative Politics and the Synthetic Control Method”, *American Journal of Political Science*, vol.59(2), pp.495~510.
- (48)Abadie, A. and Gardeazabal, J., 2003, “The Economic Costs of Conflict: A Case Study of the Basque Country”, *American Economic Review*, vol.93(1), pp.113~132.
- (49)Anastasopoulos, L. J., 2019, “Principled Estimation of Regression Discontinuity Designs”, *arXiv*: 1910.06381.
- (50)Angrist, J. D. and Pischke, J. S., 2014, *Mastering Metrics: The Path from Cause to Effect*, New Jersey: Princeton University Press.
- (51)Angrist, J. D. and Pischke, J. S., 2009, *Mostly Harmless Econometrics: An Empiricist’s Companion*, New Jersey: Princeton University Press.
- (52)Athey, S., 2019, “The Impact of Machine Learning on Economics”, in Agrawal, A., J. Gans, and A. Goldfarb, eds: *The Economics of Artificial Intelligence: An Agenda*, University of Chicago Press, Chicago.
- (53)Athey, S. and Imbens, G. W., 2015, “A Measure of Robustness to Misspecification”, *American Economic Review: Papers & Proceedings*, vol.105(5), pp.476~480.
- (54)Athey, S. and Imbens, G. W., 2016, “Recursive Partitioning for Heterogeneous Causal Effects”, *Proceedings of the National Academy of Sciences*, vol.113(27), pp.7353~7360.
- (55)Athey, S. and Imbens, G. W., 2019, “Machine Learning Methods That Economists Should Know About”, *Annual Review of Economics*, vol.11(1), pp.685~725.
- (56)Athey, S., Tibshirani, J. and Wager, S., 2019, “Generalized Random Forests”, *The Annals of Statistics*, vol.47(2), pp.1148~1178.
- (57)Audrino, F., Chassot, J., Huang, C., Knaus, M., Lechner, M. and Ortega, J. P., 2022, “How Does Post-Earnings Announcement Sen-

timent Affect Firms' Dynamics? New Evidence from Causal Machine Learning”, *Journal of Financial Econometrics*, <https://doi.org/10.1093/jjfinc/nbac018>.

(58) Bamberger, P. A., 2018, “AMD—Clarifying What We Are about and Where We Are Going”, *Academy of Management Discoveries*, vol.4(1), pp.1~10.

(59) Bao, Y., Ke, B., Li, B., Yu, Y. J. and Zhang, J., 2020, “Detecting Accounting Fraud in Publicly Traded U.S. Firms Using a Machine Learning Approach”, *Journal of Accounting Research*, vol.58(1), pp.199~235.

(60) Bargagli Stoffi, F. J. and Gnecco, G., 2020, “Causal Tree with Instrumental Variable: An Extension of the Causal Tree Framework to Irregular Assignment Mechanisms”, *International Journal of Data Science and Analytics*, vol.9(3), pp.315~337.

(61) Bartov, E., Faurel, L. and Mohanram, P. S., 2018, “Can Twitter Help Predict Firm-Level Earnings and Stock Returns?”, *The Accounting Review*, vol.93(3), pp.25~57.

(62) Bell, J., 2020, *Machine Learning: Hands-on for Developers and Technical Professionals*, Indianapolis: John Wiley & Sons.

(63) Belloni, A., Chernozhukov, V. and Hansen, C., 2014a, “High-Dimensional Methods and Inference on Structural and Treatment Effects”, *Journal of Economic Perspectives*, vol.28(2), pp.29~50.

(64) Belloni, A., Chernozhukov, V. and Hansen, C., 2014b, “Inference on Treatment Effects after Selection among High-Dimensional Controls”, *Review of Economic Studies*, vol.81(2), pp.608~650.

(65) Ben-Michael, E., Feller, A. and Rothstein, J., 2021, “The Augmented Synthetic Control Method”, *Journal of the American Statistical Association*, vol.116(536), pp.1789~1803.

(66) Bound, J., Jaeger, D. A. and Baker, R. M., 1995, “Problems with Instrumental Variables Estimation When the Correlation Between the Instruments and the Endogenous Explanatory Variable is Weak”, *Journal of the American Statistical Association*, vol.90(430), pp.443~450.

(67) Branson, Z., Rischard, M., Bornn, L. and Miratrix, L. W., 2019, “A Nonparametric Bayesian Methodology for Regression Discontinuity Designs”, *Journal of Statistical Planning and Inference*, vol.202, pp.14~30.

(68) Breiman, L., 1996, “Bagging Predictors”, *Machine Learning*, vol.24(2), pp.123~140.

(69) Breiman, L., 2001a, “Random Forests”, *Machine Learning*, vol.45(1), pp.5~32.

(70) Breiman, L., 2001b, “Statistical Modeling: The Two Cultures”, *Statistical Science*, vol.16(3), pp.199~215.

(71) Breiman, L., Friedman, J. H., Olshen, R. A. and Stone, C. J., 1984, *Classification and Regression Trees*, New York: Wadsworth International Group.

(72) Brown, N. C., Crowley, R. M. and Elliott, W. B., 2020, “What Are You Saying? Using topic to Detect Financial Misreporting”, *Journal of Accounting Research*, vol.58(1), pp.237~291.

(73) Calomiris, C. W. and Mamaysky, H., 2019, “How News and Its Context Drive Risk and Returns around the World”, *Journal of Financial Economics*, vol.133(2), pp.299~336.

(74) Calonico, S., Cattaneo, M. D. and Titiunik, R., 2014, “Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs”, *Econometrica*, vol.82(6), pp.2295~2326.

(75) Carbonell, J. G., 1990, *Machine Learning: Paradigms and Methods*, Cambridge: MIT Press.

(76) Card, D., 1990, “The Impact of the Mariel Boatlift on the Miami Labor Market”, *Industrial and Labor Relations Review*, vol.43(2), pp.245~257.

(77) Card, D. and Krueger, A. B., 1994, “Minimum Wages and Employment: A Case Study of the Fast-Food Industry in New Jersey and Pennsylvania”, *American Economic Review*, vol.84(4), pp.772~793.

(78) Chen, L., Pelger, M., and Zhu, J., 2023, “Deep Learning in Asset Pricing”, *Management Science*, in-press.

(79) Chen, M. A., Wu, Q. and Yang, B., 2019, “How Valuable Is FinTech Innovation?”, *Review of Financial Studies*, vol.32(5), pp.2062~2106.

(80) Chen, S. H., 2018, *Big Data in Computational Social Science and Humanities*, Cham: Springer.

(81) Choudhury, P., Allen, R. T. and Endres, M. G., 2021, “Machine Learning for Pattern Discovery in Management Research”, *Strategic Management Journal*, vol.42(1), pp.30~57.

(82) Choudhury, P., Starr, E. and Agarwal, R., 2020, “Machine Learning and Human Capital Complementarities: Experimental Evidence on Bias Mitigation”, *Strategic Management Journal*, vol.41(8), pp.1381~1411.

(83) Choudhury, P., Wang, D., Carlson, N. A. and Khanna, T., 2019, “Machine Learning Approaches to Facial and Text Analysis: Discovering CEO Oral Communication Styles”, *Strategic Management Journal*, vol.40(11), pp.1705~1732.

(84) Christiansen, B. and Chandan, H. C., 2017, *Handbook of Research on Organizational Culture and Diversity in the Modern Workforce*, Hershey: IGI Global.

(85) Corley, K. G. and Gioia, D. A., 2011, “Building Theory about Theory Building: What Constitutes a Theoretical Contribution?”, *Academy of Management Review*, vol.36(1), pp.12~32.

(86) Cortes, C. and Vapnik, V., 1995, “Support-Vector Networks”, *Machine Learning*, vol.20(3), pp.273~297.

(87) Costello, A. M., Down, A. K. and Mehta, M. N., 2020, “Machine+Man: A Field Experiment on the Role of Discretion in Augmenting AI-Based Lending Models”, *Journal of Accounting and Economics*, vol.70(2~3), No.101360.

(88) Cunningham, S., 2021, *Causal Inference: The Mixtape*, New Haven & London: Yale University Press.

- (89) De Franco, G., Wong, M. H. F. and Zhou, Y., 2011, "Accounting Adjustments and the Valuation of Financial Statement Note Information in 10-K Filings", *The Accounting Review*, vol.86(5), pp.1577~1604.
- (90) Doudchenko, N. and Imbens, G. W., 2016, "Balancing, Regression, Difference-in-Differences and Synthetic Control Methods: A Synthesis", NBER Working Paper, No.22791.
- (91) Dulhare, U. N., Ahmad, K. and Ahmad, K. A. B., 2020, *Machine Learning and Big Data: Concepts, Algorithms, Tools and Applications*, Hoboken: John Wiley & Sons.
- (92) Dyer, T., Lang, M. and Stice-Lawrence, L., 2017, "The Evolution of 10-K Textual Disclosure: Evidence from Latent Dirichlet Allocation", *Journal of Accounting and Economics*, vol.64(2~3), pp.221~245.
- (93) Feng, G., Giglio, S. and Xiu, D., 2020, "Taming the Factor Zoo: A Test of New Factors", *Journal of Finance*, vol.75(3), pp.1327~1370.
- (94) Freund, Y. and Schapire, R. E., 1996, "Experiments with a New Boosting Algorithm", In Proceedings of the 13th International Conference on Machine Learning.
- (95) Freund, Y. and Schapire, R. E., 1997, "A Decision-Theoretic Generalization of On-Line Learning and an Application to Boosting", *Journal of Computer and System Sciences*, vol.55(1), pp.119~139.
- (96) Géron, A., 2019, *Hands-on Machine Learning with Scikit-Learn, Keras, and TensorFlow: Concepts, Tools, and Techniques to Build Intelligent Systems*, Sebastopol: O'Reilly Media.
- (97) Gioia, D. A. and Pitre, E., 1990, "Multiparadigm Perspectives on Theory Building", *Academy of Management Review*, vol.15(4), pp.584~602.
- (98) Goodfellow, I. J., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., Courville, A. and Bengio, Y., 2014, "Generative Adversarial Nets", In Proceedings of the 27th International Conference on Advances in Neural Information Processing Systems.
- (99) Grimmer, J., Roberts, M. E. and Stewart, B. M., 2021, "Machine Learning for Social Science: An Agnostic Approach", *Annual Review of Political Science*, vol.24(1), pp.395~419.
- (100) Gu, S., Kelly, B. and Xiu, D., 2020, "Empirical Asset Pricing via Machine Learning", *Review of Financial Studies*, vol.33(5), pp.2223~2273.
- (101) Hahn, J., Todd, P. and Van der Klaauw, W., 2001, "Identification and Estimation of Treatment Effects with a Regression-Discontinuity Design", *Econometrica*, vol.69(1), pp.201~209.
- (102) Harder, V. S., Morral, A. R. and Arkes, J., 2006, "Marijuana Use and Depression Among Adults: Testing for Causal Associations", *Addiction*, vol.101(10), pp.1463~1472.
- (103) Harrison, J. S., Thurgood, G. R., Boivie, S. and Pfarrer, M. D., 2019, "Measuring CEO Personality: Developing, Validating, and Testing a Linguistic Tool", *Strategic Management Journal*, vol.40(8), pp.1316~1330.
- (104) Harrison, J. S., Thurgood, G. R., Boivie, S. and Pfarrer, M. D., 2020, "Perception Is Reality: How CEOs' Observed Personality Influences Market Perceptions of Firm Risk and Shareholder Returns", *Academy of Management Journal*, vol.63(4), pp.1166~1195.
- (105) He, V. F., Puranam, P., Shrestha, Y. R. and von Krogh, G., 2020, "Resolving Governance Disputes in Communities: A Study of Software License Decisions", *Strategic Management Journal*, vol.41(10), pp.1837~1868.
- (106) Henry, E. and Leone, A. J., 2016, "Measuring Qualitative Information in Capital Markets Research: Comparison of Alternative Methodologies to Measure Disclosure Tone", *The Accounting Review*, vol.91(1), pp.153~178.
- (107) Herlands, W., McFowland III, E., Wilson, A. G. and Neill, D. B., 2018, "Automated Local Regression Discontinuity Design Discovery", In Proceedings of the 24th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining.
- (108) Hernán, M. A. and Robins, J. M., 2020, *Causal Inference: What If*, Boca Raton: CRC Press.
- (109) Hinton, G. E., Osindero, S. and Teh, Y. W., 2006, "A Fast Learning Algorithm for Deep Belief Nets", *Neural Computation*, vol.18(7), pp.1527~1554.
- (110) Hinton, G. E. and Salakhutdinov, R. R., 2006, "Reducing the Dimensionality of Data with Neural Networks", *Science*, vol.313(5786), pp.504~507.
- (111) Holland, P. W., 1986, "Statistics and Causal Inference", *Journal of the American Statistical Association*, vol.81(396), pp.945~960.
- (112) Hsieh, T. S., Kim, J. B., Wang, R. R. and Wang, Z., 2020, "Seeing Is Believing? Executives' Facial Trustworthiness, Auditor Tenure, and Audit Fees", *Journal of Accounting and Economics*, vol.69(1), No.101260.
- (113) Huang, A. H., Zang, A. Y. and Zheng, R., 2014, "Evidence on the Information Content of Text in Analyst Reports", *The Accounting Review*, vol.89(6), pp.2151~2180.
- (114) Hunt, E. B., Marin, J. and Stone, P. J., 1966, *Experiments in Induction*, New York: Academic Press.
- (115) Huntington-Klein, N., 2022, *The Effect: An Introduction to Research Design and Causality*, Boca Raton: CRC Press.
- (116) Imbens, G. W. and Rubin, D. B., 2015, *Causal Inference for Statistics, Social, and Biomedical Sciences: An Introduction*, New York: Cambridge University Press.
- (117) Imbens, G. W. and Lemieux, T., 2008, "Regression Discontinuity Designs: A Guide to Practice", *Journal of Econometrics*, vol.142(2), pp.615~635.
- (118) Jordan, M. I. and Mitchell, T. M., 2015, "Machine Learning: Trends, Perspectives, and Prospects", *Science*, vol.349(6245),

pp.255~260.

(119) Karim, M. E., Pang, M. and Platt, R. W., 2018, "Can We Train Machine Learning Methods to Outperform the High-Dimensional Propensity Score Algorithm?", *Epidemiology*, vol.29(2), pp.191~198.

(120) Kinn, D., 2018, "Synthetic Control Methods and Big Data", *arXiv*:1803.00096.

(121) Krupa, J. and Minutti-Meza, M., 2022, "Regression and Machine Learning Methods to Predict Discrete Outcomes in Accounting Research", *Journal of Financial Reporting*, vol.7(2), pp.131~178.

(122) Lechner, M., 2018, "Modified Causal Forests for Estimating Heterogeneous Causal Effects", *arXiv*:1812.09487.

(123) Lee, D. S. and Lemieux, T., 2010, "Regression Discontinuity Designs in Economics", *Journal of Economic Literature*, vol.48(2), pp.281~355.

(124) Linden, A. and Yarnold, P. R., 2016, "Using Machine Learning to Assess Covariate Balance in Matching Studies", *Journal of Evaluation in Clinical Practice*, vol.22(6), pp.848~854.

(125) Liu, C. H., Nowak, A. D. and Smith, P. S., 2020, "Asymmetric or Incomplete Information about Asset Values?", *Review of Financial Studies*, vol.33(7), pp.2898~2936.

(126) Locke, K., 2015, "Pragmatic Reflections on a Conversation about Grounded Theory in Management and Organization Studies", *Organizational Research Methods*, vol.18(4), pp.612~619.

(127) Loughran, T. and McDonald, B., 2016, "Textual Analysis in Accounting and Finance: A Survey", *Journal of Accounting Research*, vol.54(4), pp.1187~1230.

(128) Lowry, M., Michaely, R. and Volkova, E., 2020, "Information Revealed through the Regulatory Process: Interactions between the SEC and Companies ahead of Their IPO", *Review of Financial Studies*, vol.33(12), pp.5510~5554.

(129) Luan, S., Reb, J. and Gigerenzer, G., 2019, "Ecological Rationality: Fast-and-Frugal Heuristics for Managerial Decision Making under Uncertainty", *Academy of Management Journal*, vol.62(6), pp.1735~1759.

(130) Malhotra, S., Reus, T. H., Zhu, P. and Roelofsen, E. M., 2018, "The Acquisitive Nature of Extraverted CEOs", *Administrative Science Quarterly*, vol.63(2), pp.370~408.

(131) Manela, A. and Moreira, A., 2017, "News Implied Volatility and Disaster Concerns", *Journal of Financial Economics*, vol.123(1), pp.137~162.

(132) Mason, W., Vaughan, J. W. and Wallach, H., 2014, "Computational Social Science and Social Computing", *Machine Learning*, vol.95(3), pp.257~260.

(133) Mayer-Schönberger, V. and Cukier, K., 2013, *Big Data: A Revolution That Will Transform How We Live, Work, and Think*, Boston: Houghton Mifflin Harcourt.

(134) McCaffrey, D. F., Ridgeway, G. and Morral, A. R., 2004, "Propensity Score Estimation with Boosted Regression for Evaluating Causal Effects in Observational Studies", *Psychological Methods*, vol.9(4), pp.403~425.

(135) McCulloch, R. E., Sparapani, R. A., Logan, B. R. and Laud, P. W., 2021, "Causal Inference with the Instrumental Variable Approach and Bayesian Nonparametric Machine Learning", *arXiv*:2102.01199.

(136) Mitchell, T. M., 1997, *Machine Learning*, New York: McGraw-Hill.

(137) Mullainathan, S. and Spiess, J., 2017, "Machine Learning: An Applied Econometric Approach", *Journal of Economic Perspectives*, vol.31(2), pp.87~106.

(138) Narayanan, S. and Kalyanam, K., 2020, "Behavioral Targeting, Machine Learning and Regression Discontinuity Designs", Working Paper.

(139) Patton, M. Q., 1990, *Qualitative Evaluation and Research Methods (2nd edition)*, Newbury Park: Sage.

(140) Perols, J. L., Bowen, R. M., Zimmermann, C. and Samba, B., 2017, "Finding Needles in a Haystack: Using Data Analytics to Improve Fraud Prediction", *The Accounting Review*, vol.92(2), pp.221~245.

(141) Pramod, A., Naicker, H. S. and Tyagi, A. K., 2021, "Machine Learning and Deep Learning: Open Issues and Future Research Directions for the Next 10 Years", in Tyagi, A. K., eds: *Computational Analysis and Deep Learning for Medical Care: Principles, Methods, and Applications*, John Wiley & Sons, Hoboken.

(142) Quinlan, J. R., 1986, "Induction of Decision Trees", *Machine Learning*, vol.1(1), pp.81~106.

(143) Rehse, D., Riordan, R., Rottke, N. and Zietz, J., 2019, "The Effects of Uncertainty on Market Liquidity: Evidence from Hurricane Sandy", *Journal of Financial Economics*, vol.134(2), pp.318~332.

(144) Rosenbaum, P. R. and Rubin, D. B., 1983, "The Central Role of the Propensity Score in Observational Studies for Causal Effects", *Biometrika*, vol.70(1), pp.41~55.

(145) Rosenblatt, F., 1958, "The Perceptron: A Probabilistic Model for Information Storage and Organization in the Brain", *Psychological Review*, vol.65(6), pp.386~408.

(146) Rumelhart, D. E., Hinton, G. E. and Williams, R. J., 1986, "Learning Representations by Back-Propagating Errors", *Nature*, vol.323(6088), pp.533~536.

(147) Samuel, A. L., 1959, "Some Studies in Machine Learning Using the Game of Checkers", *IBM Journal of Research and Development*, vol.3(3), pp.210~229.

(148) Schapire, R. E., 1990, "The Strength of Weak Learnability", *Machine Learning*, vol.5(2), pp.197~227.

- (149) Setoguchi, S., Schneeweiss, S., Brookhart, M. A., Glynn, R. J. and Cook, E. F., 2008, "Evaluating Uses of Data Mining Techniques in Propensity Score Estimation: A Simulation Study", *Pharmacoepidemiology and Drug Safety*, vol.17(6), pp.546-555.
- (150) Shah, S. K. and Corley, K. G., 2006, "Building Better Theory by Bridging the Quantitative-Qualitative Divide", *Journal of Management Studies*, vol.43(8), pp.1821-1835.
- (151) Shrestha, Y. R., He, V. F., Puranam, P. and von Krogh, G., 2021, "Algorithm Supported Induction for Building Theory: How Can We Use Prediction Models to Theorize?", *Organization Science*, vol.32(3), pp.856-880.
- (152) Singh, A., Hosanagar, K. and Gandhi, A., 2020, "Machine Learning Instrument Variables for Causal Inference", In Proceedings of the 21st ACM Conference on Economics and Computation.
- (153) Singh, R. and Sun, L., 2019, "Debiased Machine Learning in Instrumental Variable Models for Treatment Effects", *arXiv*: 1909.05244.
- (154) Sweredoski, M. J. and Baldi, P., 2009, "COBEpro: A Novel System for Predicting Continuous B-Cell Epitopes", *Protein Engineering, Design and Selection*, vol.22(3), pp.113-120.
- (155) Thistlethwaite, D. L. and Campbell, D. T., 1960, "Regression-Discontinuity Analysis: An Alternative to the Ex Post Facto Experiment", *Journal of Educational Psychology*, vol.51(6), pp.309-317.
- (156) Tidhar, R. and Eisenhardt, K. M., 2020, "Get Rich or Die Trying... Finding Revenue Model Fit Using Machine Learning and Multiple Cases", *Strategic Management Journal*, vol.41(7), pp.1245-1273.
- (157) Tiffin, M. A. J., 2019, "Machine Learning and Causality: The Impact of Financial Crises on Growth", IMF Working Paper, No. 19/228.
- (158) Varian, H. R., 2014, "Big Data: New Tricks for Econometrics", *Journal of Economic Perspectives*, vol.28(2), pp.3-28.
- (159) Wager, S. and Athey, S., 2018, "Estimation and Inference of Heterogeneous Treatment Effects Using Random Forests", *Journal of the American Statistical Association*, vol.113(523), pp.1228-1242.
- (160) Wang, G., Li, J. and Hopp, W. J., 2021, "An Instrumental Variable Forest Approach for Detecting Heterogeneous Treatment Effects in Observational Studies", *Management Science*, vol.68(5), pp.3399-3418.
- (161) Westreich, D., Lessler, J. and Funk, M. J., 2010, "Propensity Score Estimation: Neural Networks, Support Vector Machines, Decision Trees (CART), and Meta-Classifiers as Alternatives to Logistic Regression", *Journal of Clinical Epidemiology*, vol.63(8), pp.826-833.
- (162) Wyss, R., Ellis, A. R., Brookhart, M. A., Girman, C. J., Funk, M. J., LoCasale, R. and Stürmer, T., 2014, "The Role of Prediction Modeling in Propensity Score Estimation: An Evaluation of Logistic Regression, bCART, and the Covariate-Balancing Propensity Score", *American Journal of Epidemiology*, vol.180(6), pp.645-655.
- (163) Yeomans, M., Shah, A., Mullainathan, S. and Kleinberg, J., 2019, "Making Sense of Recommendations", *Journal of Behavioral Decision Making*, vol.32(4), pp.403-414.
- (164) You, J., Zhang, B. and Zhang, L., 2018, "Who Captures the Power of the Pen?", *Review of Financial Studies*, vol.31(1), pp.43-96.
- (165) Zeldow, B. and Hatfield, L. A., 2021, "Confounding and Regression Adjustment in Difference-in-differences Studies", *Health Services Research*, vol.56(5), pp.932-941.

How Can Machine Learning Empower Management Research? A Domestic-Foreign Frontier Review and Future Prospects

Liu Jingjiang^a, Zheng Changran^a and Hong Yongmiao^{b,c,d}

(a. School of Management, Zhejiang University; b. School of Economics and Management, University of Chinese Academy of Science;

c. MOE Social Science Laboratory of Digital Economic Forecasts and Policy Simulation, University of Chinese Academy of Science;

d. Academy of Mathematics and System Science, Chinese Academy of Sciences)

Abstract: Machine learning is profoundly changing the research paradigm and method of management. How to use machine learning to empower management research better has become a cutting-edge hot topic of academic attention. However, the application of machine learning in Chinese management research is still in its early stages. Based on academic literature published in top domestic and international journals in the fields of business management, accounting, and finance from 1999 to 2021, this paper identifies four core approaches for the academic community to conduct empirical research in management through machine learning: variable measurement, event prediction (including event classification), causal inference, and theory building. It then sorts out the research topics, questions, data sets, machine learning algorithms, and research conclusions of the representative literature for each approach. It also puts forward the main strategies for empowering management research using machine learning. Finally, it discusses future opportunities for Chinese scholars to use machine learning to study management theory with Chinese characteristics. This paper demonstrates that: Combining machine learning with traditional econometrics can help make more accurate causal inferences; Machine learning can play an important role in pattern discovery, a critical step in theory building; Combining machine learning with multi-case analysis enables us to carry out theory building effectively. Overall, it provides methodological guidance and directional inspiration for how to use machine learning to enhance the quality of management research, promote the paradigm change of management research, and build management theory with Chinese characteristics.

Keywords: machine learning; bibliometrics; causal inference; theory building; future prospect

How Can Machine Learning Empower Management Research? A Domestic–Foreign Frontier Review and Future Prospects

Liu Jingjiang^a, Zheng Changran^a and Hong Yongmiao^{b,c,d}

(a. School of Management, Zhejiang University; b. School of Economics and Management, University of Chinese Academy of Science; c. MOE Social Science Laboratory of Digital Economic Forecasts and Policy Simulation, University of Chinese Academy of Science; d. Academy of Mathematics and System Science, Chinese Academy of Sciences)

Summary: Machine learning is profoundly changing the research paradigm and method of management. Using machine learning can remedy the shortcomings of three aspects of management research. Firstly, management research has traditionally failed to extract useful information from massive, complex, high-dimensional, unstructured, and multimodal data. It has also failed to overcome the non-experimental, irreversible, and time-varying nature of socioeconomic systems and various defects of socioeconomic data. Secondly, traditionally, qualitative management research often uses a few or even a single case analysis to build a theory rather than using large-sample data. Restricted by a limited number of cases and experience, these studies have difficulties in identifying novel, robust, predictive, and relatively interpretable patterns from the data. The generalization abilities of the theories built by these studies are also weak. Thirdly, quantitative research in management based on observational data typically adopts traditional econometric methods, which have many apparent limitations. Therefore, how to use machine learning to empower management research better has become a hot topic of academic frontiers.

Machine learning has been employed to study many management problems by foreigners. However, the application of machine learning in Chinese management research is still in its early stages. Especially, the empirical studies of top management journals in China in the research fields of business management, accounting, and finance have rarely used machine learning for causal inference and theory building. Meanwhile, very little literature systematically reviews the latest achievements in the substantive use of machine learning to study management problems or proposes future research opportunities.

Based on academic literature published in top domestic and international journals in the fields of business management, accounting, and finance from 1999 to 2021, this paper identifies four core approaches for the academic community to conduct empirical research in management through machine learning: variable measurement, event prediction (including event classification), causal inference, and theory building. It then sorts out the research topics, questions, data sets, machine learning algorithms, and research conclusions of the representative literature for each approach. It also puts forward the main strategies for empowering management research using machine learning. Particularly, specific strategies for using machine learning to make causal inferences are elaborated from average treatment effects, heterogeneous treatment effects, and a multi-method fusion. Finally, it discusses future opportunities for Chinese scholars to use machine learning to study management theory with Chinese characteristics.

This paper demonstrates that: (1) Combining machine learning with traditional econometrics can help make more accurate causal inferences. (2) Machine learning can play an important role in pattern discovery, a critical step in theory building. (3) Combining machine learning with multi-case analysis enables us to carry out theory building effectively. We emphasize that future research should pay more attention to the vital role of machine learning in causal inference and theory building. Overall, it provides directional inspiration for how to use machine learning to improve the quality of management research, supplies new insights for promoting the paradigm change of management research, and offers methodological guidance for building management theory with Chinese characteristics.

Keywords: machine learning; bibliometrics; causal inference; theory building; future prospect

JEL Classification: C1, C8, G3, M2