

绿色低碳机器学习研究与应用专题前言*

封举富¹, 俞扬², 刘淇³

¹(北京大学 智能学院, 北京 100871)

²(南京大学 人工智能学院, 江苏 南京 210023)

³(中国科学技术大学 计算机科学与技术学院, 安徽 合肥 230026)

通信作者: 封举富, E-mail: fengjf@pku.edu.cn; 俞扬, E-mail: yuy@nju.edu.cn; 刘淇, E-mail: qiliuq1@ustc.edu.cn

中文引用格式: 封举富, 俞扬, 刘淇. 绿色低碳机器学习研究与应用专题前言. 软件学报, 2024, 35(4): 1585–1586. <http://www.jos.org.cn/1000-9825/7020.htm>

近年来, 随着学术界与工业界在机器学习和人工智能领域投入越来越多的关注, 相关技术获得飞速发展, 机器学习已经被应用到社会生活的方方面面, 并产生巨大社会价值. 在“生成式人工智能”“元宇宙”等科技概念越来越受欢迎的今天, 机器学习面临的数据量不断增长、模型规模逐步扩大, 因而对计算能力的需求也在持续增长, 导致了机器学习系统能耗的上涨. 我国于 2020 年提出“双碳”目标, 逐步引导绿色技术创新, 实现经济增长和生态平衡的综合效益. 面向“双碳”目标, 如何高效使用数据和优化计算资源变得尤为关键. 因此, 亟需在机器学习领域内发展面向绿色低碳的新技术和新系统. 本专题将重点探讨在人工智能浪潮中, 机器学习所面临的机遇与挑战, 尤其关注低能耗机器学习的各种策略, 如模型量化、压缩、高效计算、结合逻辑推理的可信学习、小样本学习、知识融合等, 以及相关的计算平台和系统, 并探索机器学习在绿色低碳领域的实际应用.

本专题公开征文, 共收到投稿 35 篇. 论文均通过了形式审查, 其内容涵盖了机器学习在数据利用效率提升、算法使用效率优化和算力分布式优化等多个方面的相关理论、方法、系统和应用. 每篇投稿至少邀请 2 位专家进行评审. 稿件经初审、复审、终审 3 个阶段, 通过初审的论文还在第 19 届中国机器学习会议上进行了口头报告并与现场参会专家进行交流, 最终有 11 篇论文入选本专题, 一定程度上反映了我国在该专题下的研究水平. 根据主题, 这些论文可以分为 3 组.

(1) 数据利用效率提升

《基于多样真实任务生成的鲁棒小样本分类方法》提出了一种基于多样真实任务生成的鲁棒分类方法. 方法通过对有限任务进行混合来生成更多训练任务, 并通过约束任务的多样性和真实性, 有效提高了小样本分类的泛化能力.

《基于多模态对比学习的代码表征增强预训练方法》提出一种基于多模态对比学习的增强预训练模型, 旨在通过融合代码的多层次特征来实现更精准的语义建模, 有效提升数据利用效率并促进绿色低碳的发展.

《元强化学习研究综述》全面梳理了元强化学习领域的研究进展, 旨在通过元强化学习提高深度强化学习的样本效率和策略通用性, 以更小的数据量应对更多任务, 从而提高数据利用效率.

《局部一致性主动学习的源域无关开集域自适应》提出了主动学习的源域无关开集域自适应方法及局部一致性主动学习算法, 旨在有效利用少量专家标注的目标域样本, 在资源受限和隐私受保护的场景下实现鲁棒的迁移.

(2) 算法使用效率优化

《面向开集识别的稳健测试时适应方法》针对实际场景中类别协变量分布的变化提出了一种新的开集识别方法, 能够有效地适应协变量分布偏移, 确保机器学习模型在开放世界中的稳健部署与高效性能.

* 收稿时间: 2023-09-08; jos 在线出版时间: 2023-09-11



《神经网络结构搜索在脑数据分析领域的研究进展》引入神经网络结构搜索到脑数据分析,通过自动寻找最优的深层神经网络结构显著提高了在脑数据应用上的性能和效率。

《基于自适应权重的多源部分域适应》为了解决多源部分域适应中的累积误差和负迁移问题,提出了一种基于自适应权重的多源部分域适应方法,有效地利用源域的先验知识,同时量化不同源域的贡献度,优化了迁移效率和性能。

(3) 算力分布式优化

《分组并行的轻量化实时微观三维形貌重建方法》针对微观三维形貌重建中的计算复杂度问题,采用分组并行处理和轻量化网络结构,显著降低计算时间,实现了实时高精度的三维重建,且相较其他方法大幅减少了计算耗时。

《基于动态批量评估的绿色无梯度优化方法》提出了一种基于动态批量评估的绿色无梯度优化方法,通过自适应调整评估样本量来降低计算代价。

《多视角融合的时空动态 GCN 城市交通流量预测》提出了一种多视角融合的时空动态图卷积模型,整合了静态与动态、局部与全局的特性,从而全面捕获城市交通的多元空间关联,使模型在多种时间窗口下实现高效预测。

《基于强化联邦 GNN 的个性化公共安全突发事件检测》提出了一种强化联邦图神经网络方法,通过多方协作和梯度量化有效地进行突发事件检测。利用图采样减少数据偏差影响,并通过梯度量化与权重策略平衡模型性能与通信压力。

本专题主要面向机器学习领域的研究人员和工程技术人员,展现我国学者在绿色低碳机器学习各个子领域的最新研究成果。感谢《软件学报》编委会及编辑部、中国人工智能学会机器学习专委会对专题工作的指导和帮助,感谢专题全体评审专家及时、耐心、细致的评审工作,感谢踊跃投稿的所有作者。希望本专题能够对绿色低碳机器学习相关理论、方法、系统和应用领域的研究工作有所促进。



封举富(1967—),男,博士,北京大学教授,博士生导师,中国计算机学会人工智能与模式识别专业委员会副主任,中国人工智能学会机器学习专业委员会副主任,教育部新世纪优秀人才。主要研究领域为模式识别与机器学习,生物特征识别及其应用。在国内外期刊和国际会议上发表论文 100 多篇,主持研发的北大智能指掌纹自动识别系统已应用于国内多个省市。曾获 ACCV 1993 excellent paper, ACPR 2017 Best Poster Award。2000 年获中国高校科技进步二等奖,2012 年获公安部科学技术二等奖。



俞扬(1982—),男,博士,南京大学教授,博士生导师,中国计算机学会人工智能与模式识别专业委员会委员,中国人工智能学会机器学习专业委员会委员。主要研究领域为人工智能,机器学习,强化计算,相关成果获得过 4 项国际论文奖励和 3 项国际算法竞赛冠军。曾获 2020 CCF-IEEE 青年科学家奖,入选 2018 IEEE AI's 10 to Watch,获首届亚太数据挖掘青年成就奖,并受邀在国际人工智能联合会 IJCAI 2018 作“青年亮点报告”。入选国家级优秀青年人才项目,曾获全国优秀博士学位论文奖(2013 年)、CCF 优秀博士学位论文奖(2011 年)、江苏省杰出青年基金(2017 年)。



刘淇(1986—),男,博士,中国科学技术大学教授,博士生导师,CCF 高级会员,中国人工智能学会机器学习专业委员会委员,中国科学院青年创新促进会优秀会员。主要研究数据挖掘与知识发现、机器学习方法及其在智能教育等领域的应用,相关成果获得过 IEEE ICDM 2011 最佳研究论文奖、ACM KDD 2018 (Research Track) 最佳学生论文奖。还曾获中科院院长特别奖和优博、教育部自然科学一等奖(排名第 2)、吴文俊人工智能科技进步一等奖(排名第 3)、阿里巴巴达摩院青橙奖,入选了国家基金委优秀青年科学基金项目、中国科协“青年人才托举工程”。