

顶会观察

NeurIPS 2022

河北工业大学 许铮铨

国际神经信息处理系统大会(Neural Information Processing Systems, 简称 NeurIPS)是机器学习和计算神经科学领域的顶级国际会议,是人工智能领域水平最高,录用难度最大,论文影响力最强的会议之一,与 ICML、ICLR 并列为全球机器学习三大顶会。NeurIPS 不仅属于中国计算机学会推荐的人工智能领域 A 类国际学术会议,而且是 Core Conference Ranking 推荐 A*类会议, H5-index 高达 198, Impact Score 达到 33.49。本届 NeurIPS 大会于 2022 年 11 月 28 日至 2022 年 12 月 9 日举行,为期两周,其中第一周为线下会议,在美国新奥尔良 Ernest N. Morial 会议中心召开,第二周则为线上会议,以在线交流、录制视频和异步聊天等多种方式为无法到线下会场的与会者提供线上讨论与互动的机会。参会人员身份被划分为 Full-time Student、Academic 和 Non-academic,其中学生身份的参会人员会获得注册等费用上减免。

一、会议亮点

多样的会议展现: 1、线下海报展示: 为了方便线下的快速交流,本次会议优先考虑以海报会议为中心的面对面互动讨论的方式。2、线上论文会议: 除了直接与作者交流了解他们工作外,每篇论文在 NeurIPS 网站上都有一个单独页面,线上与会人员可以在网页中找到一个 5 分钟的视频和一个专门聊天频道来异步讨论该论文相关工作。3、期刊展示: 今年 NeurIPS 引入了从期刊到会议的轨道,与会者可以在其中了解被 NeurIPS 相关期刊接受的论文的工作。本次会议共有 2,672 篇会议论文被接收,此外还有 41 篇来自 JMLR 期刊和 33 篇来自 ReScience 的期刊论文在大会期刊轨道进行了展示。

严格的委员会评审: 本届 NeurIPS 2022 大会由 Sanmi Koyejo(斯坦福大学)和 Shakir Mohamed(DeepMind)担任 General Chair,由 Alekh Agarwal(谷歌研究院)、Alice Oh(韩国科学技术院)、Danielle Belgrave(DeepMind)和 Kyunghyun Cho(纽约大学)担任 Program Chair。同时,各个分区的主席均来自著名的大学和科研机构,如纽约大学、牛津大学和普林斯顿大学,或者主流的人工智能企业,如 DeepMind、微软和谷歌等。本届 NeurIPS 的绝大部分论文都经过不少于四名专家的评审,并且通过 rebuttal 的形式让作者与评审专家能够对审稿意见中的相关问题进行充分讨论,以最大程度解决审稿意见中的分歧。但从大会组织方发布的相关公告中仍发现,虽然经过了严格的评审和充分的讨论,论文作者和审稿人之间仍存在一定的分歧;因此大会组委会还专门对审稿分歧这一问题展开调查,通过分析 NeurIPS 2022 中记录的关于论文质量的分歧(包括共同作者之间、作者和审稿人之间以及审稿人委员会成员之间的分歧)。组委会的调查结果表明论文质量的评估是一项从本质上具有挑战性的复杂任务,即使通过严格的评审流程,依然没有客观正确的答案,因此,组委会也呼吁研究人员应对论文提交结果持保留态度。

伦理审查程序: 为了激发整个领域的伦理研究、实践和反思,本次会议将伦理审查程序作为提高 NeurIPS 2022 作者和审稿人伦理意识和参与度的一个重要环节。通过结合前几版伦理审查流程的成功经验,组委会在本次会议中着重讨论了新版伦理审查流程可靠性和适用性,从而实现进一步巩固伦理审查政策,建立一个向前

推进式的审核流程。组委会表示伦理审查目的不是惩罚性或排他性的；相反，伦理审查旨在告知、教育和阐明道德问题，以便作者可以通过公开讨论的方式来解决冲突问题。今年，组委会还发布 NeurIPS 临时道德准则初稿，旨在为社区提供更全面的道德准则和会议的期望。

二、录用情况

NeurIPS 2022 收到的有效投稿和录用数量都有显著提高，大会共收到了 10411 篇投稿，最终接收了 2672 篇论文，接收率约为 25.6%。相较于 2021 年，今年 NeurIPS 2022 的投稿量提升了 12%(1289 篇)，整体录用论文数量也在去年的 2334 篇的基础上提升了 338 篇，增长率约为 14%。在接收的论文中，有 199 篇论文录用于 Oral Presentations，比去年增加了 143 篇，Oral 率约为 7.4%(比去年提升 3%)。今年 Spotlight 文章数量为 523 篇，相较于去年提升了 263 篇。NeurIPS2022 会议涵盖的方向包括：识别(检测、分类与检索)、3D 视觉、图像与视频的生成、深度学习与表示学习、视觉与语言、迁移学习、图神经网络、场景分析和理解、无监督学习、数据集与评估等方向。值得关注的是，在本次会议中，标题出现词汇最多的前五名依次为 Multi-Training、Self、Optimal 和 Learning，摘要中出现最多的五个词汇依次为 Learning、Model、Data、Models 和 Methods。来自中国的学术机构在本次大会取得了相当不俗的成绩。据大会公开的文章接受列表统计(以第一作者单位为准)，清华大学以 85 篇论文位列榜首，北京大学排名全球第六名，上海交大和中科大分列全球第九和第十二名。此外斯坦福大学和谷歌在本次会议中仍有不俗表现，分别有 72 篇和 65 篇论文被录用。此外，华人学者也占据了个人录用情况统计排名的前列，悉尼大学刘同亮(Tongliang Liu)教授，香港浸会大学韩波(Bo Han)教授、上海交通大学严骏驰(Junchi Yan)教授和伊利诺伊大学厄巴纳-香槟分校李博(Bo Li)教授均以 11 篇论文被录用的优异成绩共同位列排行榜的第三位。

三、主题报告

本届 NeurIPS 2022 会议邀请了七位主题报告讲者(Keynote Speakers)，报告主题内容涵盖了包括语言模

型、脑启发研究、扩散模型、图神经网络等方面，其中具有较高讨论度和影响力的报告简介如下。

The Forward-Forward Algorithm for Training Deep Neural Networks. 图灵奖得主、深度学习先驱多伦多大学 Geoffrey Hinton 教授带来了关于一种新的神经网络学习方法——前向-前向算法(Forward-Forward Algorithm, FF)的介绍。该方法不同于反向传播算法，它包含两个前向传递过程，其中一个使用正数据，另一个使用网络自身生成的负数据。Hinton 认为，FF 算法的优点在于：它更好地模拟了大脑皮层的学习过程，并且能够在使用低功耗硬件模拟的同时节省大量计算资源。他主张放弃计算机软硬件分离的形式，而是将未来的计算机设计为“非永生的”(mortal)，以提高计算资源的利用率。因此，FF 算法是最佳的学习方法之一，能够高效地运行在这种硬件中。

Are Large Language Models Sentient? 纽约大学神经科学家 David Chalmers 教授带来了 GPT-3、LaMDA 2 和相关的大型语言模型是否具有感知能力的报告。近年来，随着自然语言处理领域的蓬勃发展，关于大型语言模型是否具有轻微意识的问题呈现了不同的讨论。Chalmers 从生物学，感官系统角度出发，结合全局模型的理念对 Large Language Model(LLM)进行多角度的分析，他认为关于 AI 意识的相关问题不能简单的作出判断，并预测在十年内，即使我们没有人类水平的 AGI，我们也可能拥有非常智能化意识的系统。

Conformal Prediction in 2022. 来自斯坦福大学的 Barnum-Simons 数学与统计学讲席教授 Emmanuel Candes 为与会者报告了共型预测在 2022 年的发展现状和主要的研究工作。共型预测近年来在学术界和工业界引起了广泛的讨论和研究，大有流行之势。共型预测主要为未来的数值观测提供了准确的预测区间，并且无需进行任何分布假设。本次演讲回顾了共型预测的基本原理，并回顾了过去 2-3 年内的一些主要贡献。Candes 还讨论了适用于定量和分类标签增强一致性的方式，并尝试将共型预测应用于金融经济学，市场行为等领域。最后，Candes 借用预测 COVID19 病例轨迹的示例进一步说明共型预测的具体应用方向。

Interaction-Centric AI. 来自韩国科学技术院计算学院的 Juho Kim 教授介绍并讨论了以交互为中心的人工智能。虽然现有的模型算法在很多人工智能领域都取得了卓越的性能,但这些模型上的进步很少转化为对现实世界用户的影响。当前,很多模型由于忽视了人类与人工智能交互的动态性和复杂性,从而进一步限制人工智能在实际环境中的应用。因此,报告者认为人机交互应该被视为设计人工智能应用程序的首要对象。在本次演讲中, Kim 展示了一些使用 AI 来支持复杂现实生活任务的新型交互系统,并讨论了人机交互设计中的复杂难题和解决方案,分享模型设计经验。最后, Kim 呼吁要为“以交互为中心的 AI”建立稳健的区块——一种设计和工程化人机交互的系统方法,进而补充并克服以模型和数据为中心的人工智能模型的局限性。

The Data-Centric Era: How ML is Becoming an Experimental Science. 来自谷歌大脑的 Isabelle Guyon 教授带来了《以数据为中心的时代: ML 如何成为一门实验科学》的主题报告。Guyon 表示,最近的分析显示,使用更多的数据和更深层的网络并不是一种可持续的进步方式。同时,其他指标表明机器学习越来越依赖于良好的数据和基准。这不仅可以训练更强大/更紧凑的模型,还可以合理评估新想法并强调测试模型的可靠性、公平性和安全性。在 2021 年, NeurIPS 启动了数据集和基准测试轨道,并诞生了以数据为中心的 AI 计划,开启了“以数据为中心的时代”。与设计 and 训练 ML 模型相比,数据科学家将花费更多时间在理解问题、设计实验和选取工程数据集上。Guyon 通过挑选一些引起争议并引发人深省的热门话题进行讨论,并在报告结尾强调当前需解决的一些紧迫的问题。

四、 热点论文

本届 NeurIPS 2022 共有 13 篇论文获得杰出论文奖, 1 篇论文获得时间检验奖, 2 篇论文获颁杰出数据集和基准论文奖。获奖论文涵盖了梯度评估, 最优学习, 梯度下降, 文档检索等方面。

时间检验奖: ImageNet Classification with Deep Convolutional Neural Networks. AlexNet 因首次将卷积网络(CNN)应用于 ImageNet 训练挑战赛, 并获得

远超当时的其他方法的优越表现而获得了该奖项。 AlexNet 的成功对机器学习社区产生了巨大的影响, 标志着卷积神经网络(CNN)成为图像分类的核心模型, 随后 ImageNet 冠军的获胜者都采用了卷积神经网络结构, 使得深度学习迎来了新的高潮。随着卷积网络在分类任务上的成功, CNN 网络随后也在目标检测任务, 图像分割任务等一系列深度学习领域中继续发扬光大。

杰出论文: ProcTHOR: Large-Scale Embodied AI Using Procedural Generation. 海量数据集和高容量模型推动了计算机视觉和自然语言理解的许多最新进展, 来自华盛顿大学的 AI2 PRIOR 团队受此启发提出了一种基于 AI 环境的生成框架 ProcTHOR。该框架能够对各种交互式、可定制式虚拟环境的任意大数据集进行采样, 在交互和操纵任务中训练和评估模型。团队通过使用 10000 套 AI 生成的房屋样本和一个简单的神经网络模型, 展示了 ProcTHOR 的威力和潜能。在 ProcTHOR 框架下使用 RGB 图像进行训练获得的模型, 在没有明确的映射和人工任务监督的前提下, 能够在 6 个 AI 基准任务上取得最先进成果, 证明 ProcTHOR 框架强大。

杰出论文: High-dimensional limit theorems for SGD: Effective dynamics and critical scaling. 来自纽约大学的 Courant 团队研究了在高维区域中具有恒定步长的随机梯度下降(SGD)的尺度极限并且证明了有限维函数的轨迹在维数达到无穷大时的极限定理。该方法允许选择跟踪的有限维函数的轨迹和步长, 并产生了 ODE 和 SDE 极限。在从多峰时间尺度收敛, 以及从随机初始化收敛到概率远离零的次优解等方面展示了惊人的效果提升。

杰出数据集和基准论文奖: MINEDOJO: Building Open-Ended Embodied Agents with Internet-Scale Knowledge. 受人类在开放世界中不断学习和适应的启发, 来自 NVIDIA 的团队提出了构建通用 Agent 的三位一体要素: (1)支持多种任务和目标的环境; (2)大规模多模态知识数据库; (3)灵活可扩展的 Agent 体系结构。该团队在此基础上提出了 MINEDOJO, 这是一个基于流行的 Minecraft 游戏构建下的新框架, 它具有一个模拟套件, 其中包含数千种不同的开放式任务, 以及一个包含 Minecraft 视频教程、维基页面和论坛讨论的

互联网规模知识库。该方法能够解决各种以自由形式指定的开放式任务，而无需任何手动设计的密集奖励。

杰出论文：Is Out-of-Distribution Detection Learnable? 在 out-of-distribution (OOD) 领域，来自悉尼科技大学澳大利亚人工智能研究所的团队提出了 OOD 检测的一种新的学习理论：probably approximately correct (PAC)。该团队首先找到 OOD 检测可学习性的必要条件；然后借助该条件，证明了在某些场景下 OOD 检测的可学习性的几个不可能性定理。基于这一观察，团队进而给出几个充分必要条件来表征 OOD 检测在一些实际场景中的可学习性，从而为监督学习提供新的优化思路。

杰出论文：Gradient Estimation with Discrete Stein Operators. 梯度估计是对分布参数的期望梯度进行近似，也是当前解决许多机器学习问题的核心。然而，当分布是离散的时，最常见的梯度估计器会出现方差过大问题。为了提高梯度估计的质量，来自斯坦福大学和清华大学的团队引入了一种基于 Stein 算子的离散分布方差缩减技术。通过使用这种技术为 REINFORCE leave-one-out 构建灵活的 control variates。并且该团队提出的 control variates 可以在线调整以获得最小化方差，并且不需要对目标函数进行额外评估。

其余获奖文章还涉及到的领域有图像文本模型，扩散生成模型，文档检索网络，梯度优化和数据修剪等方面的创新成果。其中加州大学伯克利分校的论文 LAION-5B: An open large-scale dataset for training next generation image-text models 提出了一个由 58.5 亿个 CLIP 过滤图像文本对组成的数据集 LAION-5B，使得大规模多模态模型的研究更加民主化。谷歌的论文 Photorealistic Text-to-Image Diffusion Models with Deep Language Understanding 提出了一种文本到图像的扩散模型 Imagen；该模型建立在大型 transformer 语言模型理解文本的能力上，并依赖于高保真图像生成中扩散模型的强度，达到了前所未有的照片真实性和深度的语言理解。剑桥大学的获奖论文 Gradient Descent: The Ultimate Optimizer 展示了如何通过对反向传播进行简单修改来自动计算超参数，并能够轻松地将该方法应用于其他优化器和超参数。

五、 讲习班，研讨会和竞赛

NeurIPS 2022 还开展了 13 场讲习班(Tutorials), 62 场研讨会 (Workshops) 和 25 项专业竞赛 (Competitions)。讲习班主要聚焦机器学习及其相关新型领域的主题。部分关注度和影响力较高的讲习班如下。

- The Role of Meta-learning for Few-shot Learning
- Foundational Robustness of Foundation Models
- Advances in NLP and their Applications to Healthcare
- Incentive-Aware Machine Learning: A Tale of Robustness, Fairness, Improvement, and Performativity
- Theory and Practice of Efficient and Accurate Dataset Construction

研讨会主要是对正在进行的前沿工作和未来发展方向进行提供了一个讨论平台。今年的研讨会主要关注大型语言模型，深度学习与化学生物学科交叉等方面的问题。部分关注度和影响力较高的研讨会列举如下。

- Second Workshop on Efficient Natural Language and Speech Processing (ENLSP-II)
- Progress and Challenges in Building Trustworthy Embodied AI
- AI for Science: Progress and Promises
- Synthetic Data for Empowering ML Research
- AI for Accelerated Materials Design (AI4Mat)
- Order up! The Benefits of Higher-Order Optimization in Machine Learning

在竞赛方面，今年的竞赛同样聚焦于 Computer Vision, NLP 和 Multi-modal 等主流任务。从机器学习方法的角度出发，今年强化学习相关的竞赛显著增加。在应用方面，生物医学和健康领域的竞赛数目最多最高；与过去几年相比，竞赛数量和重要程度正在逐年增加。

部分关注度和参与度较高的竞赛列举如下。

- Causal Insights for Learning Paths in Education
- The MineRL BASALT Competition on Fine-tuning from Human Feedback
- IGLU: Interactive Grounded Language Understanding in a Collaborative Environment
- Real Robot Challenge III -Learning Dexterous Manipulation from Offline Data in the Real World
- Second AmericasNLP Competition: Speech-to-Text Translation for Indigenous Languages of the Americas
- Multimodal Single-Cell Integration Across Time, Individuals, and Batches
- Weakly Supervised Cell Segmentation in Multimodality High-Resolution Microscopy Images

六、 总结展望

本届 NeurIPS 2022 大会中围绕时间序列预测、分类、异常检测、表征学习以及在医疗、生物、交通、金

融等方向的应用文章数量大大增加，竞争热度逐渐提升。3D 视觉、图像与视频的生成、Transformer 等领域依旧保持之前的高热度。相比于 2021 年强化学习、深度学习和表征学习等领域的火热，今年神经网络和时序预测的文章占比有明显提升，未来可能成为更加火热的研究方向。更值得关注的是，本次会议的最佳论文中有 3 篇属于随机梯度方向，相比于模型上的改进和优化，对于底层数学算法的优化更容易获得 NeurIPS 评选委员会的青睐。与会专家认为，未来人工智能技术的发展趋势包括更强大的模型和更高效的计算机架构；在模型方面，深度学习将继续成为研究和应用的主流，同时也会出现新的架构和方法来解决现有模型的限制和缺点；在计算机架构方面，人们将继续寻求更快、更节能、更可靠的硬件来支持更大、更复杂的模型；但同时也需要关注人工智能的伦理问题和安全问题，加强监管和规范，确保人工智能的安全可控和可持续发展。笔者认为回答好上述问题将会机器学习的新机遇，从而更好地迈向更高层级更具有创造力的智能化机器学习时代。

责任编辑 魏秀参

参考文献

[1] <https://neurips.cc/Conferences/2022>



许铮铎

河北工业大学共建电工装备可靠性与智能化国家重点实验室教授、博士生导师，河北省海外高层次人才计划入选者，河北省“优青”获得者，英国牛津大学计算机系博士、博士后、客座研究员、外聘博导。主要研究方向为：深度学习、强化学习、医学影像智能计算等。

Email: zhenghua.xu@hebut.edu.cn