
新方法训练AI模型更具“人类智慧”

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/38601.html>

本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！

新方法训练AI模型更具“人类智慧”

。科技日报北京3月3日电（记者刘霞）加拿大滑铁卢大学科学家领导的国际科研团队，提出了一套训练人工智能（AI）大语言模型使其更具“人类智慧”的新方法，探索了支持明智推理的新型架构，同时提出了衡量AI智慧的基准。相关论文发表于新一期《认知科学趋势》杂志。



图片来源：物理学家组织网

新方法试图将人类智慧融入AI“血脉”，以打造更强大、更透明、协作性更强且更安全的智能系统。

滑铁卢大学心理学教授、论文共同作者萨姆·约翰逊博士解释道，AI日益聪慧，却始终缺失人类的一项核心素养——智慧。智慧不仅关乎知识积累或智力提升，更关乎应对生活挑战、作出艰难抉择、驾驭不可预测社交场域的能力。

尽管当前的AI系统在处理边界清晰的任务时游刃有余，但一旦遭遇混沌不明的难题，它们很可能变得束手无策。原因在于它们尚未习得人类应对不确定性的全方位策略。新方法聚焦于培养AI的“元认知能力”或思维方式，即认识到其知识的局限性，适应不同的环境，权衡多种观点，并对情况的发展保持灵活性。

研究团队认为，智慧听起来太过哲学化、人性化，似乎难以在机器中形式化。但他们可以把智慧拆解为具体策略，如多角度思考、情境适应等，为AI构建智慧蓝图便有了清晰路径。

团队展望，这种智慧AI可处理从未见过的新问题，适应陌生环境，在追逐共同目标时精诚协作，其决策过程对用户更具可解释性，且能更好地与人类价值观对齐，从而提升安全性。

作者：刘霞 来源：科技日报

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发