优化引文幻觉的AI来了

作者：writer 来源：科学网

本文原地址：https://www.iikx.com/news/progress/38228.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！

优化引文幻觉的AI来了。一个新推出的开源语言模型OpenScholar，在进行准确文献综述方面可超越商用大语言模型（LLM）。比如，实验中，GPT4o会在78%-90%的情况下出现引文幻觉，而新的人工智能（AI）工具的引文准确率却与人类专家相仿。虽然仍需进一步优化，但该工具有望帮助科学家处理复杂且日益繁重的科学文献综述任务。相关研究成果2月5日发表于《自然》。

科学文献综述对于支持循证决策、微调科学过程和引导新发现都很重要。然而，文献发表数量的增长使研究人员很难掌握全部资讯。LLM可以提供协助，但却很容易出错，如归因能力有限和引文幻觉。

为了生成准确、全面、透明的科学文献综述，美国西雅图华盛顿大学的Akari Asai、Hannaneh Hajishirzi和同事推出了OpenScholar。该模型是专为科研任务设计的检索增强语言模型。其他系统也用过这个框架，但作者将其与一个包含4500万篇最新开放获取科研论文的专业数据库以及一个自我评估机制相结合，从而优化其输出。作者还创建了名为ScholarQABench的基准工具来评估文献综述的自动化。研究显示，OpenScholar的准确率比GPT4o和PaperQA2（一个文献综述工具）这类现有系统分别高出6.1%和5.5%。此外，OpenScholar生成的答案在50%到70%的情况下比专家注释器的答案更有用。作者总结道，以上结果和引文幻觉大幅下降证明了OpenScholar有望支持和推动科学家的研究工作。

但作者指出该系统仍有局限性，并强调基于语言模型的系统无法使科学文献综述完全自动化。他们向学界同时开放了ScholarQABench和OpenScholar，以鼓励进一步研究和优化。（来源：中国科学报 赵熙熙）

相关论文信息：https://doi.org/10.1038/s41586-025-10072-4

作者：Akari Asai 来源：《自然》

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转发