
成功合成35种新化合物！化学家有了新AI“助手”

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/38125.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！

成功合成35种新化合物！化学家有了新AI“助手”。近日，一项发表于《自然》的研究报道了可以极大简化并加快化学合成过程的人工智能（AI）系统——MOSAIC。该系统由美国耶鲁大学研究团队与跨国制药公司勃林格殷格翰的研究人员合作开发。研究人员利用该系统推荐的条件下，无需再进行任何筛选或调整，就成功合成了35种具有成为药物、农用化学品或化妆品等产品潜力的化合物。

小分子的合成是药物研发以及其他许多重要领域中耗时较长的一个环节。论文作者、耶鲁大学化学家Timothy Newhouse说，MOSAIC能够打破这一瓶颈，起草完整、详细到足以让化学家遵循的实验室操作指南，帮助他们创造新分子。

对于化学家来说，寻找新药和新材料是一项艰巨的任务。为了合成这些有前景的化合物，他们必须在数以百万计的已知化学反应中进行筛选，同时每年有数十万种新的化学反应被添加进来，还要测试这些化合物是否能够被合成。

预测化学反应条件一直是AI在化学领域应用的重点，其中最突出的工具是美国IBM公司开发的基于AI的在线化学合成预测平台RXN for Chemistry。它利用简化分子线性输入规范（SMILES）系统，将化学三维结构转化为字母、数字和标点等更适合语言识别系统的符号。相比之下，瑞士洛桑联邦理工学院联合美国罗切斯特大学开发的大语言模型（LLM）ChemCrow则通过自然语言数据来训练完成化学任务。

SMILES方法使得处理诸如起始材料和溶剂之类的化学信息变得更加容易。我们的目标是建立一个通用模型，通过倾听实验步骤，能够像化学家书写化学式那样读取信息，将其迅速转化为实用建议。Newhouse说，将MOSAIC生成的分步指令整合到自动化系统中将是下一步。

Newhouse等研究人员利用此前研发的AI系统，将从专利中提取的约100万条反应记录分为2285个子集。利用这些子集，团队训练了Meta公司部分开源的Llama LLM，创建了2498个独立的专家模型。每个模型专门对应从一种分子开始的化学转化的一种组合。这可以在本地计算机上运行，因为使用的参数比目前主要的LLM要少。

美国北卡罗来纳州立大学的材料科学家Martin Seifrid表示，MOSAIC避免了用大的模型解决问题，反而选择专注于一个精心设计的由许多更小的专家模型组成的系统。

研究人员用MOSAIC提出的化学条件合成52种新物质。在实验室测试中，他们成功合成了其中的35种。此外，MOSAIC准确预测了这些化合物的颜色和形态。MOSAIC还提出了一种合成氮杂稠

噪分子的新方法，并且成功通过测试。

勃林格殷格翰已经在使用MOSAIC。他们对设计新的合成途径很感兴趣。论文作者、耶鲁大学的理论与计算化学家Victor Batista说，如果能减少合成步骤，就能节省大量资金。MOSAIC作为开源代码可供其他团队使用。

相关专家表示，MOSAIC是AI辅助化学领域的一个重要概念性进步，它将AI从预测推进到行动，通过将反应条件的选择定义为决策问题，而非单一输出的预测任务，使其更符合化学家在实验中的实际推理方式。

英国利物浦大学的计算机科学家Xenofon Evangelopoulos认为这种方法具有更广泛的潜力，除了作为可靠的化学合成AI工具外，MOSAIC还确立了一种利用全球化学知识的模块化与功能专化的可扩展范例。（来源：中国科学报 许悦）

相关论文信息：<https://doi.org/10.1038/s41586-026-10131-4>

作者：Timothy Newhouse 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](https://www.iikx.com)转发