
新型数据驱动反应优化策略

作者：writer 来源：科学网

本文原地址：<https://www.iikx.com/news/progress/26920.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！

新型数据驱动反应优化策略。 2024年4月19日，德国哥廷根大学Lutz Ackermann教授和浙江大学化学系洪鑫教授在Chem期刊上合作发表了Machine Learning-Guided Yield Optimization for Palladaelectro- Catalyzed Annulation Reaction的研究工作。该成果报道了一种新型的数据驱动反应优化策略，将基于物理有机描述符的产率预测和正交实验设计相结合，同时平衡了采样多样性和目标产率提升，实现了海量条件组合的高效探索，成功预测了电化学钯催化环化反应的优势条件，充分展示了人工智能建模技术在合成化学反应优化领域的潜力。

随着化学大数据和人工智能技术的跨越式发展，数据驱动的反应优化策略近年来引起了广泛关注。由于合成反应构效关系高维且复杂的特性，催化剂、反应物、甚至溶剂细微的结构变化都可能引起反应产率的显著改变，为合成反应的条件优化提出了重要挑战。

在该文中，作者采用主动学习框架对电化学Pd催化环化反应的条件（图1a）进行优化。该智能建模策略的核心在于，将传统的贪心优化策略中融入了正交实验设计作为约束，从而提升条件空间取样的多样性，并根据优化进度逐级放宽正交约束，从而平衡对于采样多样性和采样目标提升的双重需求。在该策略中，以少量符合正交约束的实验作为初始数据集，结合机器学习建模的产率预测和基于化学维度的正交性约束，通过主动学习指导下一轮实验设计，并将实验结果反馈于机器学习模型，不断更新迭代，直至产率符合优化目标为止（图1b）。

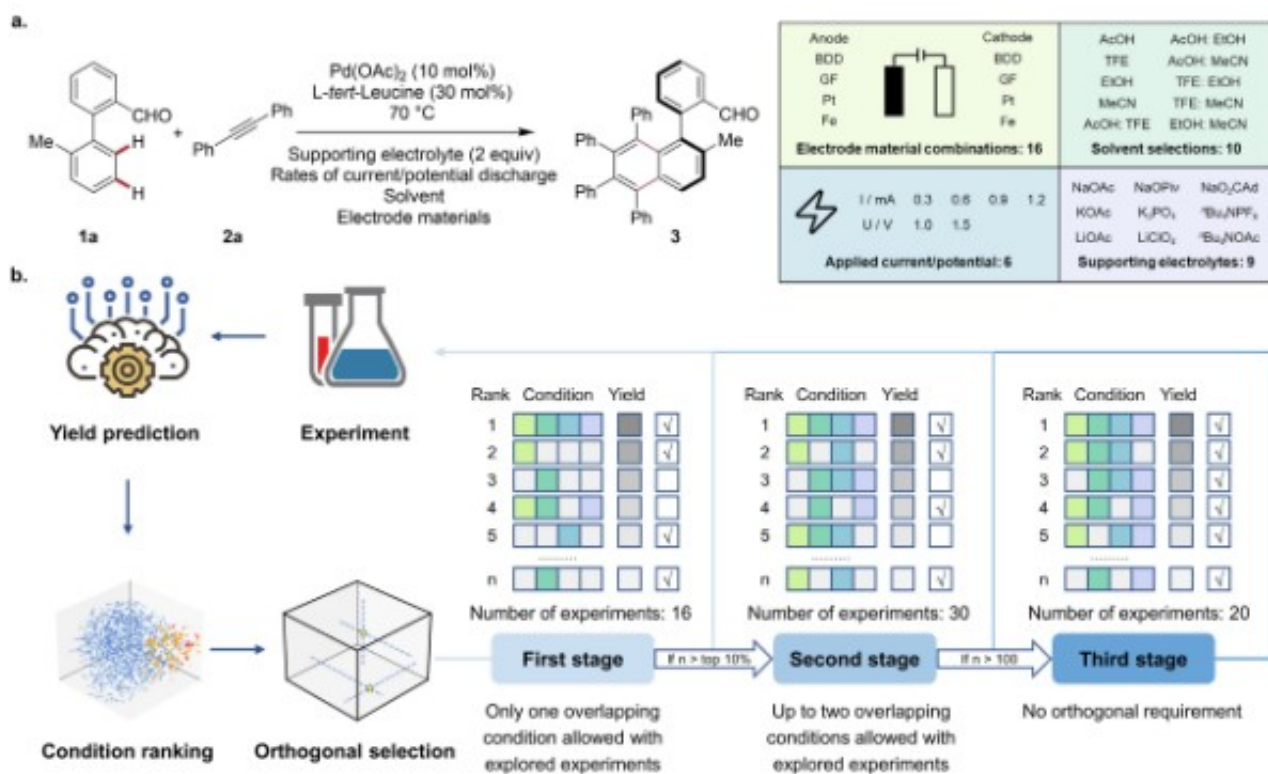


图1：数据驱动的反应产率优化策略。

在ML建模中，作者使用了一系列物理有机描述符（图2a）对反应体系进行编码，包括起始电位、电化学表面积、Tafel斜率、电导率等。每轮优化的结果如图2b所示，在进行了12轮共68个实验之后，模型就成功预测出了最优反应条件。此外，作者通过UMAP降维和特征重要性分析，揭示了产率优化的路径以及影响产率的关键特征。

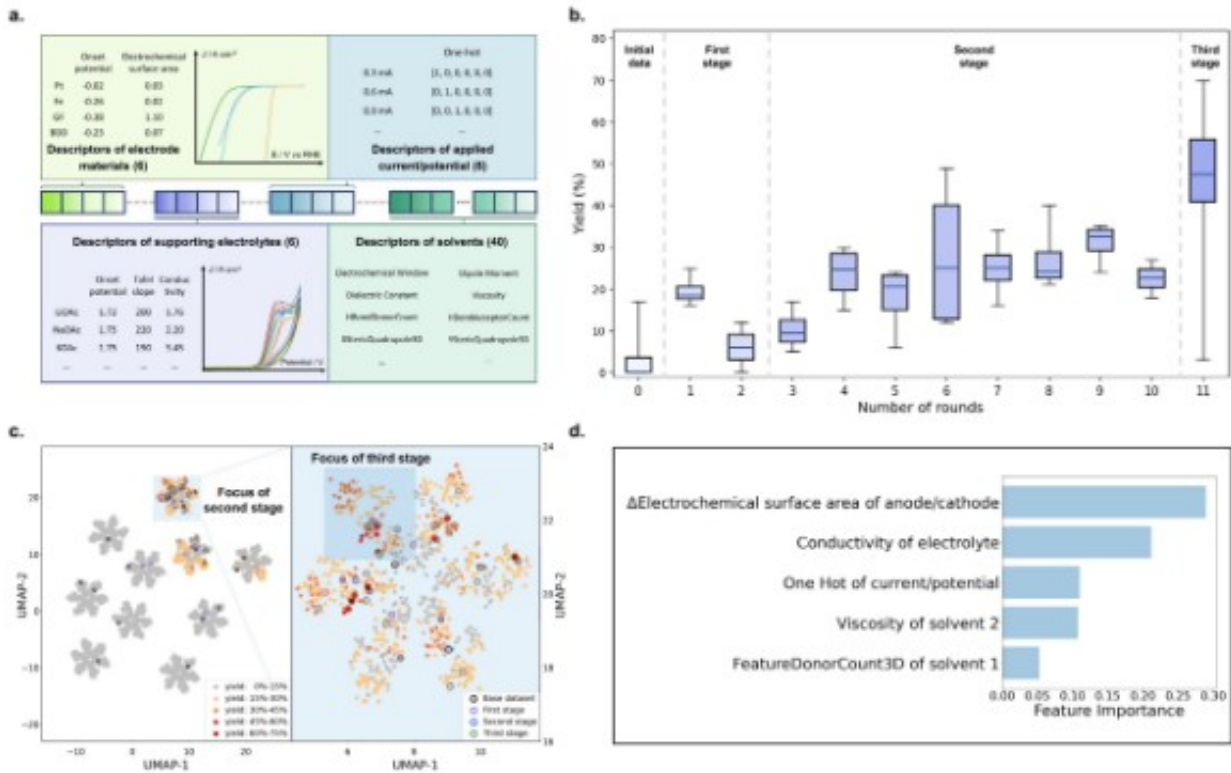


图2：机器学习指导的产率优化结果。

为进一步评估反应优化策略的普适性，作者进一步增加了电极材料和溶剂的可能组合，将化学空间从8640种可能扩大到了39690种组合（图 3a）。模型利用已有产率数据对扩大空间内的反应条件进行产率预测，预测的前四名如图3b所示，所有这些推荐的反应条件均具有较高的产率，最高组合的产率达到了83%，更加凸显了该策略在反应优化上的有效性。

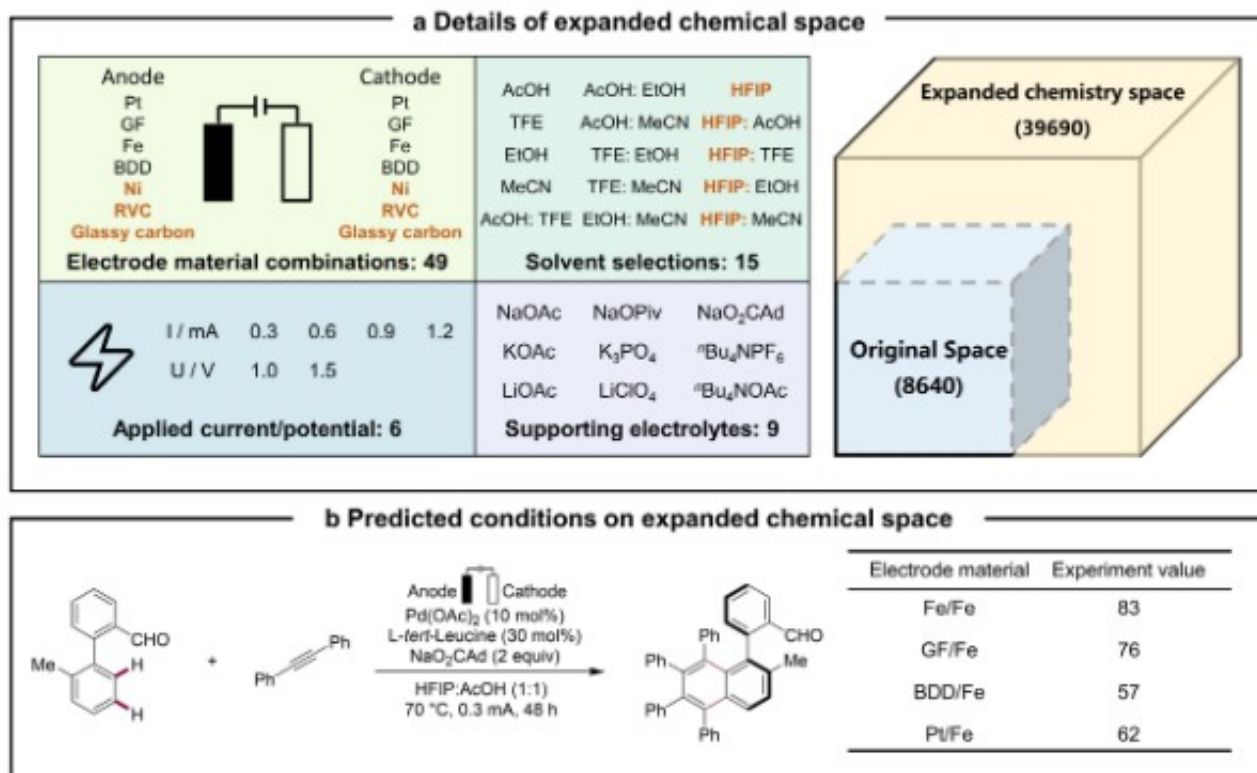


图3：基于机器学习的化学空间探索。

（来源：科学网）

相关论文信息：<https://doi.org/10.1016/j.chempr.2024.03.027>

作者：洪鑫等 来源：《化学》

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