机器学习助力发现高效 CO₂ 电催化还原 Cu-Al 合金催化剂

李晔飞*, 刘智攀*

复旦大学化学系, 上海市分子催化和功能材料重点实验室, 物质计算科学教育部重点实验室, 上海 200433

* 联系人, E-mail: yefeil@fudan.edu.cn; zpliu@fudan.edu.cn

Machine learning leads to the discovery of Cu-Al alloy for efficient CO₂ reduction

Ye-Fei Li* & Zhi-Pan Liu*

Key Laboratory of Computational Physical Science (Ministry of Education), Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Department of Chemistry, Fudan University, Shanghai 200433, China

* Corresponding authors, E-mail: yefeil@fudan.edu.cn; zpliu@fudan.edu.cn

doi: 10.1360/TB-2020-0632

CO₂大量排放造成的环境变化已逐渐成为制约现代社会发展的难题. 将CO₂还原为高附加值化学品(如乙烯)是解决这一问题的有效途径, 具有重要的科学意义和应用前景. 然而, 由于CO₂是一个非常惰性的分子, 高效还原难度高. 实验发现金属Cu可以作为将CO₂还原为C₂物种的催化剂, 但是催化选择性和产率都很低, 不能满足大规模工业化的要求^[1]. 因此尝试对Cu改性, 如形成合金, 是常见的研究思路, 但是Cu合金种类非常多, 元素组成和比例均可调, 如何理性高效寻找催化剂是一个重要科学挑战. 随着高通量实验技术的引入, 机器学习理论方法的兴起, 科学家开始尝试这些新技术、新方法辅助筛选催化剂.

人工智能方法应用于化学领域, 目前主要有两种思路: 其一, 通过构建基于人工智能技术的反应势函数, 替代繁琐费时的求解薛定谔方程的量子力学计算过程, 从而达到高效的大规模原子模拟的目的; 其二, 基于实验大数据或简单的理论模型构建数据集, 通过机器学习方法, 直接预测反应. 前者主要用于以模拟反应过程为目的, 得到反应能量、能垒的精确解. 在催化研究领域, 刘智攀课题组^[2,3]发展的SSW-NN全局势能面模拟方法, 就是通过构建全局机器学习势函数从而大规模快速筛选所有可能的热力学稳定结构, 形成从材料的体相相图出发, 到严格计算反应条件下表面结构, 再到反应动力学的“自下而上”的研究框架. 而后者是一种“自上而下”的实用研究策略, 在技术上能够集合多种先进的机器学习方法, 主要目的是直接

预测反应的可能性, 达到诸如反应路径选择(如逆合成), 预测反应产物分布等较为定性的结果. 比如, 依靠有机实验的数据, Waller等人^[4]采用类似Google发展的AlphaGo的策略, 发展了逆合成的人工智能预测方法.

近期, 多伦多大学的Sargent教授团队和卡耐基梅隆大学Ulissi教授团队^[5]合作, 在CO₂还原电催化剂设计方面取得进一步突破. 在Cu-Al合金上实现了迄今为止最高的电催化CO₂还原生成乙烯的活性和选择性, 达到了80%乙烯转化法拉第效率, 并且电流密度最高能达到600 mA/cm². 而前期最好的结果同样来自Sargent教授课题组^[6]所开发的碱性介质下含突变界面(abrupt interface)的铜催化剂, 其乙烯转化法拉第效率在275 mA/cm²电流密度下达到70%. 值得注意的是, 在最新的研究中他们正是采用了目前流行的后一种人工智能方法, 即通过简单理论模型构建数据集的机器学习, 替代高通量实验筛选, 达到预测催化性能的目的. 根据前期的研究, 建立了以CO吸附能为判据的简单火山型曲线(volcano curve)模型, 确定当CO的最优吸附能在-0.67 eV附近时, 能够达到最高的催化活性. 在简单模型的基础上, 机器学习方法被用来快速搜索最优的合金组分, 主要分为三步: (1) 构建机器学习模型. 他们选用了金属原子序数、泡林电负性、CO吸附配位数以及训练集中CO在特定金属位点上吸附能的中位数作为描述子, 结合random forest和boosted tree机器学习, 构建出CO吸附能和催化剂局部环境关联的模型. (2) 构建数据集. 用机器学习模型枚举可能结构, 采用第一性原理计算这些不同结构

图 1 (网络版彩色)Cu-Al合金表面结构示意图. 不同颜色代表不同元素, 其中橙色是Cu, 浅灰色是Al, 深灰色是C, 红色是O, 白色是H

Figure 1 (Color online) Scheme of the surface of Cu-Al alloy. Orange balls: Cu; light-gray balls: Al; dark-gray balls: C; red balls: O; and white balls: H

上的CO吸附能, 主动学习, 扩大数据集. (3) 迭代训练, 直到机器学习模型预测与第一性原理计算结果自洽收敛. 理论最终预测出Al含量在12%时具有最佳的CO吸附能. 实验最终合成了17个表面Al含量在10%左右的合金样品, 优化了最佳反应活性和选择性的条件. 催化剂稳定性良好, 在反应过程中只有不到1%的Cu/Al金属被剥离.

可见“自上而下”的人工智能研究策略, 既有实用快速的优点, 也有模型简单、基本原理不清等内禀的问题. 在Sargent组的研究中, 对于原子尺度反应机理的探索, 基本还是靠实验原位表征等方式, 通过原位同步辐射近边X射线吸收精细结构(XANES)等手段, 发现了反应时表面Cu-O物种的存在. 进一步通过理论计算表明, 当合金表面/次表层存在Cu-O物种时, 可有效降低C-C偶联的能垒, 解释了Cu-Al合金能够显著提高乙烯选择性的原因. 但是文中理论计算仅基于有限几种猜测的O原子覆盖度和构型, 因此其解释还不能被完全确定. 尤其是, 实验上观察到的Cu-O物种, 说明表面可能存在Cu₂O或CuO等氧化物, 而源头的机器学习模型的材料基础是金属状态的表面, 并由此训练得出的最优CO吸附合金. 显然, 机器学习最终对于实验的观测和解释还有明显的逻辑鸿沟.

今天的科学迎来了大数据时代, 建立在大数据基础上的机器学习方法已经逐渐证明了其在化学研究中作用. 如何结合大数据分析, 实验表征和合成、量子力学计算和机器学习手段来加速化学创新, 既需要多学科的融合, 也需要进一步基础理论上的突破, 才能让化学真正在电子和原子的尺度上展现其神奇的魅力. 我们相信, 两种人工智能方法, “自下而上”的大规模原子模拟和“自上而下”的实用筛选策略都会在未来展现不同的成功应用, 从晶体结构预测到有机合成路线规划, 从实验图谱识别到智能合成, 走向化学的各个研究领域.

参考文献

- 1 Nitopi S, Bertheussen E, Scott S B, et al. Progress and perspectives of electrochemical CO₂ reduction on copper in aqueous electrolyte. *Chem Rev*, 2019, 119: 7610–7672
- 2 Ma S, Huang S D, Liu Z P. Dynamic coordination of cations and catalytic selectivity on zinc-chromium oxide alloys during syngas conversion. *Nat Catal*, 2019, 2: 671–677
- 3 Ma S, Shang C, Liu Z P. Heterogeneous catalysis from structure to activity via SSW-NN method. *J Chem Phys*, 2019, 151: 050901
- 4 Segler M H S, Preuss M, Waller M P. Planning chemical syntheses with deep neural networks and symbolic AI. *Nature*, 2018, 555: 604–610
- 5 Zhong M, Tran K, Min Y, et al. Accelerated discovery of CO₂ electrocatalysts using active machine learning. *Nature*, 2020, 581: 178–183
- 6 Dinh C T, Burdyny T, Kibria M G, et al. CO₂ electroreduction to ethylene via hydroxide-mediated copper catalysis at an abrupt interface. *Science*, 2018, 360: 783