

华北电力大学

专业硕士学位论文

数据-机理双驱动的水氧化产过氧化氢电催化
剂设计

**Data and Mechanism Co-Guided Design of Water-
Oxidation Electrocatalysts Toward H₂O₂ Generation**

刘燕

2025年6月

国内图书分类号：TK01+8
国际图书分类号：620

学校代码：10079
密级：公开

专业硕士学位论文

数据-机理双驱动的水氧化产过氧化氢电催化 剂设计

硕士研究生：刘燕

导师：刘志坚教授

企业导师：段志洁高级工程师

申请学位：能源动力硕士

专业领域：动力工程

学习方式：全日制

所在学院：能源动力与机械工程学院

答辩日期：2025年5月

授予学位单位：华北电力大学

Classified Index: TK01+8

U.D.C: 620

Thesis for the Professional Master's Degree

Data and Mechanism Co-Guided Design of Water-Oxidation Electrocatalysts Toward H₂O₂ Generation

| | |
|---|---|
| Candidate: | Liu Yan |
| Supervisor: | Prof. Liu Zhijian |
| Enterprise mentor: | Duan Zhijie Senior Engineer |
| Professional Degree Applied for: | Master of Energy and Power |
| Speciality: | Power Engineering |
| Cultivation ways: | Full-time |
| School: | School of Energy Power & Mechanical Engineering |
| Date of Defence: | May, 2025 |
| Degree-Conferring-Institution: | North China Electric Power University |

华北电力大学学位论文原创性声明

本人郑重声明：此处所提交的学位论文《数据机理双驱动的水氧化产过氧化氢的催化剂筛选》，是本人在导师指导下，在华北电力大学攻读学位期间独立进行研究工作所取得的成果。论文中除已注明部分外不包含他人已发表或完成的研究成果。对本文的研究工作做出重要贡献的个人和集体，均已在文中以明确方式注明。论文由本人撰写，未使用人工智能代写。本声明的法律结果将完全由本人承担。

作者签名：

日期： 年 月 日

华北电力大学学位论文使用授权书

学位论文系研究生在华北电力大学攻读学位期间在导师指导下完成的成果，知识产权归属华北电力大学所有，学位论文的研究内容不得以其它单位的名义发表。研究生完全了解华北电力大学关于保存、使用学位论文的规定，同意学校保留并向有关部门送交论文的复印件和电子版本，同意学校将学位论文的全部或部分内容编入有关数据库进行检索，允许论文被查阅和借阅，学校可以为存在馆际合作关系的兄弟高校用户提供文献传递服务和交换服务。本人授权华北电力大学，可以采用影印、缩印或其他复制手段保存论文，可以公布论文的全部或部分内容。

保密论文在保密期内遵守有关保密规定，解密后适用于此使用权限规定。
本人知悉学位论文的使用权限，并将遵守有关规定。

本学位论文属于（请在相应方框内打“√”）

公开 内部 秘密 机密 绝密

作者签名：

日期： 年 月 日

导师签名：

日期： 年 月 日

摘 要

H_2O_2 作为绿色氧化剂与能源载体受到了广泛关注。在制备 H_2O_2 的过程中, 传统的蒽醌法面临着高能耗、重污染等突出问题。相比之下, 绿色且高效的电化学合成方法为实现可持续 H_2O_2 生产提供了保障。特别是二电子水氧化反应 (2e^- -WOR), 仅利用水作为原料, 无需外加能量供给, 是一种很有前景的路径。但是, 目前用于 2e^- -WOR 的催化剂数量较少, 且产率较低, 因此系统性高效筛选高性能 2e^- -WOR 催化剂至关重要。当前, 机器学习凭借其可解释性和高效率计算等优点, 能够快速筛选性能优越的催化剂。为此, 本文提出了一种数据与机理双驱动的设计策略, 通过构建新型通用描述符、机器学习模型以及微观动力学模型, 对 4274 种材料进行系统性筛选, 识别出多种高活性 2e^- -WOR 催化剂, 并通过实验制备验证了具有最优性能的新材料。此外, 本研究开发的自动化特征提取工具与机器学习模型, 为高效筛选 2e^- -WOR 催化剂提供了理论支撑与实践指导。

首先, 构建了一个涵盖金属间合金、金属氧化物、钙钛矿和单原子催化剂的 962 种材料的数据集。提出了加权原子中心对称函数 (wACSF) 描述符, 在传统 ACSF 基础上引入配位原子的电子属性作为权重因子, 以全面表征局域化学环境。借助开发的自动化特征提取工具包, 实现了 wACSF 的高效批量计算与原子本征属性提取, 形成了维度为 962×134 的特征矩阵。随后, 采用皮尔逊相关性分析和递归特征消除 (RFE) 方法, 筛选出一套最优特征集。基于 XGBoost 算法构建了最佳预测模型, 在测试集上, ΔG_{OH^*} 和 ΔG_{O^*} 的 R^2 分别为 0.84 和 0.91, RMSE 分别为 0.52 和 0.65 eV。该模型进一步应用于数据库外的 3312 种候选材料, 实现了吸附能的快速预测。

其次, 建立了一个针对 2e^- -WOR 反应的微观动力学模型, 揭示了 ΔG_{OH^*} 与理论交换电流密度之间的函数关系。该模型与已发表的实验数据高度吻合, 验证了火山图预测活性趋势的可靠性。依据 2e^- -WOR 反应的选择性条件 (即 $\Delta G_{\text{O}^*} \geq 3.52$ eV 和 1.41 eV $\leq \Delta G_{\text{OH}^*} < 2.4$ eV) 与微观动力火山图, 从庞大的数据集和预测集中分别筛选出了 49 种和 12 种高性能催化剂。

最后, 实验验证显示, 预测性能最佳的 LiScO_2 催化剂表现出优异的活性、选择性和稳定性。在 $2.2 \text{ V}_{\text{RHE}}$ 的电位下, 其 H_2O_2 的法拉第效率高达 90%, 并且在 168 小时的稳定性测试后, 其效率仍保持在 82% 至 86% 之间, 展现了出色的长期稳定性。此外, 实验测得的活性 ($\log(j)=1.56 \text{ mA cm}^{-2}$) 与理论预测 ($\log(j)=1.28 \text{ mA cm}^{-2}$) 之间的偏差小于 0.3。同时, LiScO_2 在 10 mA cm^{-2} 电流

密度下的过电位为 240 mV，较同类催化剂降低了 70%以上。

这一通用框架显著提高了催化剂的设计效率，并已证实其在预测和筛选催化剂方面的有效性，因此具有广阔的催化材料设计应用潜力。

关键词：局域环境描述符；机器学习；微观动力学模型；水氧化； H_2O_2

Abstract

H₂O₂ as a green oxidant and energy carrier, has received widespread attention. In the process of preparing H₂O₂, the traditional anthraquinone method faces prominent issues such as high energy consumption and severe pollution. In contrast, green and efficient electrochemical synthesis methods provide a guarantee for achieving sustainable H₂O₂ production. Especially the two-electron water oxidation reaction (2e⁻ WOR), which uses only water as a raw material and does not require external energy supply, is a very promising pathway. However, currently, the number of catalysts used for 2e⁻ WOR is limited, and their yield is low. Therefore, systematically and efficiently screening high-performance 2e⁻ WOR catalysts is crucial. Currently, machine learning, with its advantages of interpretability and high computational efficiency, can quickly screen for catalysts with superior performance. To this end, this paper proposes a dual-driven design strategy combining data and mechanisms. By constructing novel universal descriptors, machine learning models, and microscopic kinetic models, a systematic screening of 4274 materials was conducted, identifying several highly active 2e⁻ WOR catalysts. And through experimental preparation and verification, new materials with optimal performance were developed. In addition, the automated feature extraction tools and machine learning models developed in this study provide theoretical support and practical guidance for the efficient screening of 2e⁻ WOR catalysts.

First, a dataset comprising 962 materials, including intermetallic alloys, metal oxides, perovskites, and single-atom catalysts, was constructed. The weighted atomic-centered symmetry function (wACSF) descriptor was proposed, introducing the electronic properties of coordinating atoms as weight factors based on the traditional ACSF to comprehensively characterize the local chemical environment. With the help of the developed automated feature extraction toolkit, efficient batch computation of wACSF and extraction of atomic intrinsic properties were achieved, resulting in a feature matrix with dimensions of 962×134. Subsequently, Pearson correlation analysis and recursive feature elimination (RFE) methods were used to select an optimal feature set. Based on the XGBoost algorithm, the optimal prediction model was constructed. On the test set, the R² values for ΔG_{OH*} and ΔG_{O*} were 0.84 and 0.91, respectively, with RMSE values of 0.52 and 0.65 eV, respectively. The model was further applied to 3312 candidate materials outside the database, achieving rapid prediction of adsorption energy.

Secondly, a microkinetic model for the 2e⁻ WOR reaction was established, revealing the functional relationship between ΔG_{OH*} and the theoretical exchange

current density. The model is highly consistent with the published experimental data, validating the reliability of the volcano plot in predicting activity trends. Based on the selectivity conditions of the $2e^-$ WOR reaction (i.e., $\Delta G_{O^*} \geq 3.52$ eV and 1.41 eV $\leq \Delta G_{OH^*} < 2.4$ eV) and the microkinetic volcano plot, 49 and 12 high-performance catalysts were screened from the large dataset and prediction set, respectively.

Finally, experimental validation shows that the LiScO_2 catalyst, predicted to have the best performance, exhibits excellent activity, selectivity, and stability. At a potential of 2.2 V_{RHE}, its Faradaic efficiency for H_2O_2 reaches up to 90%, and after 168 hours of stability testing, its efficiency remains between 82% and 86%, demonstrating excellent long-term stability. In addition, the deviation between the experimentally measured activity ($\log(j)=1.56$ mA cm⁻²) and the theoretical prediction ($\log(j)=1.28$ mA cm⁻²) is less than 0.3. Furthermore, the overpotential of LiScO_2 at a current density of 10 mA cm⁻² is 240 mV, which is more than 70% lower than that of similar catalysts.

This universal framework significantly improves the efficiency of catalyst design and has been proven effective in predicting and screening catalysts, thus holding great potential for applications in catalyst material design.

Keywords: Local Environment Descriptor, Machine Learning, Microkinetic Model, Water Oxidation, H_2O_2

目 录

| | |
|--|----|
| 摘 要 | I |
| Abstract | I |
| 第 1 章 绪 论 | 1 |
| 1.1 研究背景 | 1 |
| 1.2 国内外研究现状 | 2 |
| 1.2.1 $2e^-$ WOR 的反应路径和机理 | 2 |
| 1.2.2 $2e^-$ WOR 产 H_2O_2 的研究现状与挑战 | 3 |
| 1.2.3 数据驱动的催化剂材料筛选 | 4 |
| 1.3 本文研究内容 | 5 |
| 第 2 章 研究方法 | 8 |
| 2.1 火山图的计算 | 8 |
| 2.1.1 $2e^-$ WOR 微观动力模型的计算 | 8 |
| 2.1.2 理论火山图的推导计算 | 9 |
| 2.2 机器学习 | 10 |
| 2.2.1 Scikit-learn 库介绍 | 10 |
| 2.2.2 极限梯度提升算法 | 11 |
| 2.2.3 机器学习模型的评估指标 | 12 |
| 2.2.4 特征之间的相关性系数 | 12 |
| 2.3 密度泛函理论 | 13 |
| 2.3.1 Hohenberg-Kohn 定理 | 13 |
| 2.3.2 Kohn-Sham 方程 | 14 |
| 2.3.3 计算软件介绍及参数设置 | 14 |
| 2.4 实验方法 | 15 |
| 2.4.1 实验材料 | 15 |
| 2.4.2 催化剂的合成 | 15 |
| 2.4.3 材料表征 | 16 |
| 2.4.4 电化学性能测试 | 16 |
| 2.5 本章小结 | 17 |
| 第 3 章 描述符和自动化提取策略 | 18 |
| 3.1 数据机理双驱动的催化剂设计框架 | 18 |
| 3.2 加权的对称函数描述符 | 19 |

| | |
|-------------------------------------|-----------|
| 3.2.1 径向和角向对称函数的计算 | 20 |
| 3.2.2 配位环境加权的对称函数介绍 | 21 |
| 3.3 结构与吸附能数据库的建立 | 22 |
| 3.4 自动化特征提取模型 | 22 |
| 3.4.1 特征提取策略 | 23 |
| 3.4.2 工作流程的实现 | 24 |
| 3.4.3 机器学习数据集 | 24 |
| 3.5 本章小结 | 25 |
| 第 4 章 基于机器学习的催化剂快速预测 | 26 |
| 4.1 机器学习的特征工程 | 26 |
| 4.1.1 特征选择的步骤 | 26 |
| 4.1.2 特征相关性分析 | 28 |
| 4.2 机器学习模型的构建与优化 | 29 |
| 4.2.1 多种机器学习算法的对比 | 29 |
| 4.2.2 模型的超参数优化 | 31 |
| 4.2.3 截断半径对模型性能的优化 | 31 |
| 4.3 模型验证与可解释性分析 | 33 |
| 4.3.1 机器学习模型展示 | 33 |
| 4.3.2 模型性能的对比及分析 | 33 |
| 4.3.3 特征重要性分析及物理化学解释 | 34 |
| 4.4 本章小结 | 36 |
| 第 5 章 催化剂的筛选与实验验证 | 38 |
| 5.1 催化剂筛选标准 | 38 |
| 5.1.1 微观动力火山图的验证分析 | 38 |
| 5.1.2 $2e^-$ WOR 催化剂的选择性条件 | 41 |
| 5.2 高活性和高选择性催化剂的筛选 | 42 |
| 5.2.1 数据集的筛选与结果分析 | 42 |
| 5.2.2 机器学习驱动预测集生成 | 43 |
| 5.2.3 预测集的筛选与结果分析 | 44 |
| 5.3 筛选结果的理论验证 | 45 |
| 5.3.1 预测集 12 种催化剂 DFT 验证 | 45 |
| 5.3.2 最佳催化剂的性能对比 | 46 |
| 5.4 实验验证预测的高性能催化剂 | 47 |
| 5.4.1 LiScO_2 的制备及表征 | 47 |

| | |
|---|-----------|
| 5.4.2 LiScO ₂ 的性能分析及验证 | 49 |
| 5.5 本章小结 | 52 |
| 第 6 章 结论与展望 | 54 |
| 6.1 总结 | 54 |
| 6.2 创新点 | 55 |
| 6.3 展望 | 55 |
| 参考文献 | 57 |
| 攻读硕士学位期间发表的论文及其它成果 | 67 |
| 致谢 | 68 |
| 华北电力大学学位论文答辩委员会情况 | 70 |

第 1 章 绪 论

1.1 研究背景

目前，全球能源结构以煤炭、石油为主，引发了一系列严峻的能源与环境问题。伴随着不可再生化石能源的逐年消耗，二氧化碳的排放量呈明显增长态势^[1]。部分观点认为像全球变暖引发的极端天气这类气候变化，并非当下最紧迫的挑战，但考虑到全球人口迅速增加与经济高速发展的需求，化石燃料的实际消耗率仍在持续增加。预计到本世纪末，化石能源逐步耗尽已成必然趋势^[2]。除上述因素外，过度依赖传统化石能源带来的环境污染问题，也将对经济、社会及环境的可持续发展造成重大制约。面对石油等不可再生资源的逐渐枯竭以及碳排放引发的严重环境问题，各国正加紧步伐，积极寻求更加绿色且高效的能源替代方案^[3]。

在过去近两个世纪中， H_2O_2 凭借其广泛的适用性和环保特性，赢得了极大的关注，被誉为全球最重要的 100 种化学物质之一^[4]。作为替代石油和氢气的理想能源载体， H_2O_2 的分解过程只产生氧气和水，不会生成其他有害的化学废物，其能量密度约为 96 kJ mol^{-1} ^[5, 6]。通过燃料电池， H_2O_2 能够实现电能的高效转换，70 wt% 的 H_2O_2 溶液能量密度约为 0.75 kWh L^{-1} ，接近压缩氢气的能量密度^[7, 8]。

在石油工业中， H_2O_2 经热分解后会产生大量高温蒸汽，被视为高耗能、污染重的传统蒸汽技术的替代方案，在稠油开采与油藏增产中具有重要应用前景。这一技术路径对推动能源生产的绿色低碳转型具有积极意义^[9]。截至 2023 年，全球 H_2O_2 市场消费量约为 700 万吨，年均复合增长率（CAGR）为 4.7%^[10]。由此可见，开发高效、便捷的 H_2O_2 生产方法，以满足不断扩大的市场需求，已变得愈发迫切。

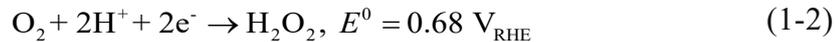
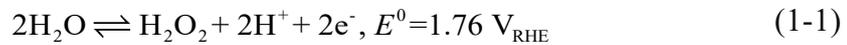
目前，全球大约 95% 的过氧化氢是通过蒽醌法生产的。这一工艺依赖于昂贵的 Pd 催化剂，并且会产生大量的有机废物。因此通常需要依赖大型化工设备进行集中生产。此外，为了避免 H_2O_2 在储运和运输过程中分解，还需添加缓冲剂，并在实际应用前再将其去除，这无疑增加了额外的成本和时间^[11]。

与传统的蒽醌工艺相比，电化学合成法在安全性、连续性及环保性等方面展现出明显优势。其反应过程无需引入任何有毒化学品，仅以水或氧气为反应物，就能在电极表面原位生成 H_2O_2 ，实现稳定而持续的产出。尤其是在质子交换膜反应器的辅助下，该方法能够高效地将电能转化为化学能。随着水能、风

能及太阳能光伏等可再生能源技术的快速发展，电化学方法为绿色能源的储存与利用开辟了一条新路径^[12-14]。总之，凭借其良好的环境友好性、低能耗特性以及与传统化工生产的兼容性，电化学方法在清洁能源开发和绿色化学合成领域中占据了重要地位。

1.2 国内外研究现状

在当前的电化学合成 H_2O_2 过程中，既可以利用水作为反应物进行二电子水氧化反应 ($2\text{e}^- \text{WOR}$, 式 1-1)，也可以利用氧气作为反应物进行二电子氧还原反应 ($2\text{e}^- \text{ORR}$, 式 1-2)^[15]。当利用太阳能和风能等经济高效的可再生能源作为供能条件时，这两种方法都能在常温常压下实现较高的反应效率，适合分散式生产及多种应用场景^[16-18]。

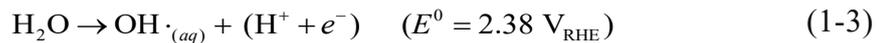


此中， $2\text{e}^- \text{ORR}$ 被充分研究。在固态电解质中，可通过它制备出浓度高达 20% 的纯 H_2O_2 水溶液，可是在此过程中需要供应稳定的氧气，并且阴极上的绿色能源 H_2 很难获得^[19-21]。相较而言， $2\text{e}^- \text{WOR}$ 反应能够直接利用水来制备 H_2O_2 ，不依赖气态氧。若将阳极的 $2\text{e}^- \text{WOR}$ 与阴极的 $2\text{e}^- \text{ORR}$ 串联于同一电解槽中，可显著提高电化学生产 H_2O_2 的整体效率^[22]。

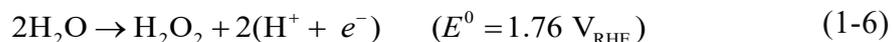
1.2.1 $2\text{e}^- \text{WOR}$ 的反应路径和机理

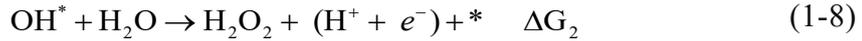
图 1-1 显示了催化表面电化学水氧化反应 (WORs) 包含的三种途径。目前研究最为深入的是水电解生成氧气的四电子水氧化反应 ($4\text{e}^- \text{WOR}$)，见反应式 (1-9) 至 (1-13)。相对而言，用于制备 H_2O_2 的两电子水氧化途径 ($2\text{e}^- \text{WOR}$) 的研究较少，仅有少数金属氧化物被深入探讨。此外，由反应式 (1-3) 至 (1-5) 可知，单电子水氧化反应会生成羟基自由基 ($\text{OH}\cdot$)。应当明确的是，三种途径的初始步骤均为羟基中间体 (OH^* ，其中*表示吸附态) 的吸附^[13]。

单电子水氧化反应 ($1\text{e}^- \text{WOR}$):

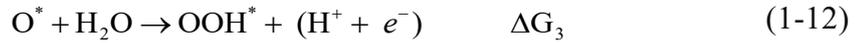
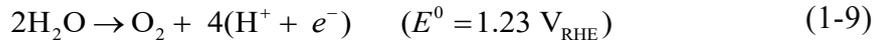


两电子水氧化反应 ($2\text{e}^- \text{WOR}$):





四电子水氧化反应 (4e⁻ WOR):



其中, E^0 为总反应的标准电极电势, ΔG_i 为每个分反应的吉布斯自由能变化。

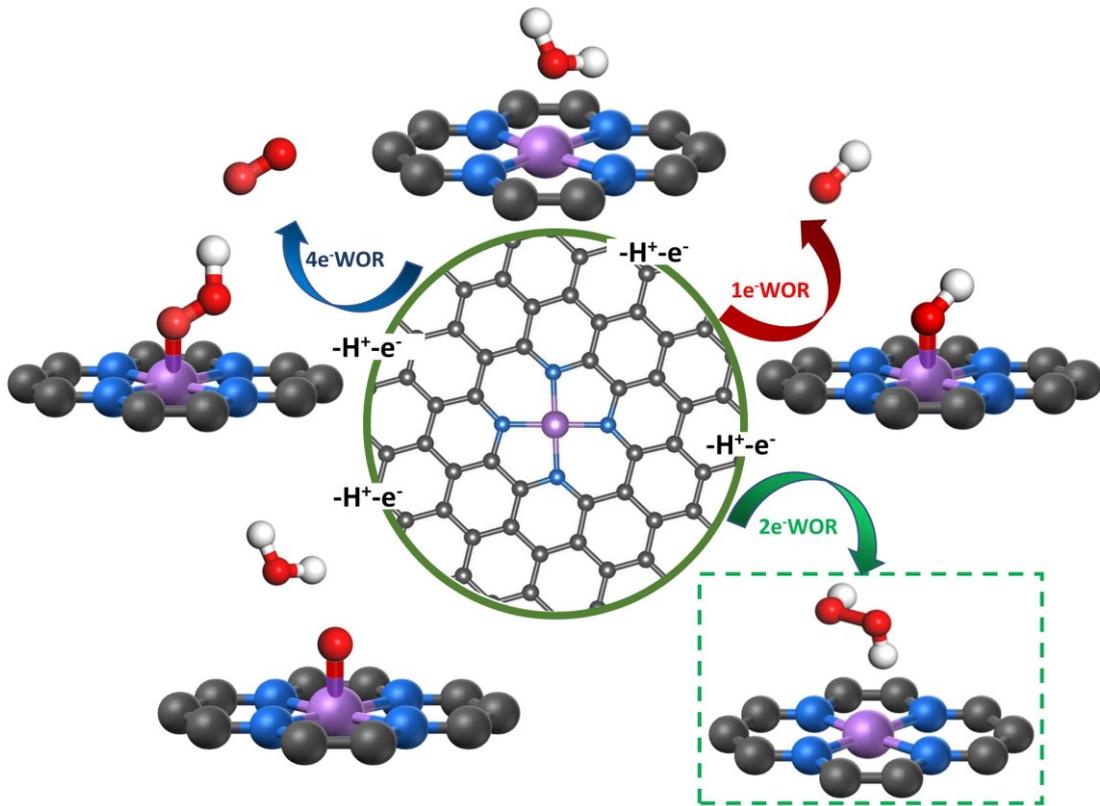


图 1-1 电化学水氧化反应的三种竞争反应示意图

1.2.2 2e⁻ WOR 产 H₂O₂ 的研究现状与挑战

到目前为止, 由于金属氧化物在氧化及碱性条件下展现出良好的稳定性, 它已成为 2e⁻ WOR 制备 H₂O₂ 的常用材料^[23]。研究人员已经报道了多种可有效产 H₂O₂ 的金属氧化物, 如二氧化钛 (TiO₂)^[24]、二氧化锡 (SnO₂)^[25]、三氧化二锑 (Sb₂O₃)^[26]、氧化锌 (ZnO)^[27]、钒酸铋 (BiVO₄)^[28]、钙锡矿 (CaSnO₃)^[29]、钨酸铜 (CuWO₄)^[30] 以及铝酸镧 (LaAlO₃)^[31] 等。但是, 这些催化剂大多

本征活性不高，导电性也极差。部分催化剂还需克服较大的过电位（通常超过 1.0 V），这导致其产率较低（一般低于 $5 \mu\text{mol}\cdot\text{min}^{-1}\cdot\text{cm}^{-2}$ ）^[32,33]。因此，对高性能 $2e^-$ WOR 催化剂的大规模筛选和开发的需求日益迫切。

先前研究已经证明，不同反应中间体（ O^* 、 OH^* 、 OOH^* ）的吸附能可以作为水氧化反应活性和选择性的有效描述符，其中 OH^* 的吸附能被认为是 $2e^-$ WOR 活性和选择性的关键指标^[13]。除此之外，催化剂对 O^* 的较弱吸附有助于抑制 $4e^-$ WOR 途径，防止过氧化氢被进一步氧化为氧气。目前，大多数研究主要采用基于反应中间体吸附能的理论极限电位模型，从热力学角度探讨理论过电位与羟基吸附自由能（ ΔG_{OH^*} ）之间的关系^[13,34]。不过，针对高电势下的电催化过程，仅依靠热力学模型会导致理论与实验结果之间存在较大差异。例如，早在 2007 年，Vassilev 与 Koper 的研究就指出，单纯依靠密度泛函理论（DFT）的热力学分析难以充分解释氧还原反应中的实验结果^[35]。此外，实验中过电位的量化指标较为模糊，这无疑增加了理论预测与实验数据之间直接对比的难度，进而难以有效验证模型的精确度。

为解决上述问题，微观动力学模型应运而生。该方法通过解析催化剂表面复杂的反应机制，将动力学与热力学相结合，不仅精确地阐明了反应路径，还能有效识别关键中间体，同时揭示了反应物吸附能与催化剂性能之间的关系。特别需要指出的是，微观动力学方法可量化包括电流密度在内的核心电化学参数（如转换频率、塔菲尔斜率等），这使得理论结果与实验数据能更好地进行对比验证^[36]。例如，Nørskov 团队借助氧还原反应动力学，创新性地开发了微观动力学火山模型。这一模型在目标电位下预测的过渡金属的电流密度与实验数据高度一致，突破了二十年前经典极限电位火山图的局限^[37,38]。值得一提的是，研究者针对 $4e^-$ WOR 反应（析氧反应，OER）也建立了相应的微观动力学火山模型，弥补了传统过电位模型在解释反应机理方面的不足^[39]。即便如此，目前对 $2e^-$ WOR 反应的理解仍以热力学极限电位模型为主，这表明迫切需要微观动力学模型提供更有效精确的理论支撑^[40,41]。

1.2.3 数据驱动的催化剂材料筛选

尽管微观动力学模型已经建立，但催化剂的筛选在很大程度上仍依赖密度泛函理论（DFT）以直接计算反应描述符（例如，关键反应中间体在催化剂表面的吸附自由能），这一过程既消耗大量计算资源又耗时。近年来，机器学习这一数据驱动的先进技术为催化剂材料的迅速筛选开辟了崭新的路径^[42]。尤其是基于原子中心对称函数（ACSF）描述符的机器学习模型，在能量和力的预测方面展现出了卓越的准确性，并且在许多情况下其预测误差极小^[43-46]。通过表征局

部原子环境的特征，如原子空间分布与元素类型，该类模型能够可靠地预测各种分子和晶体材料的原子能量和力。此外，这些模型在预测材料基本属性方面也表现出很高的准确性，使其在涉及多种材料体系的研究中具有广泛的适用性^[45, 46]。机器学习的核心优势在于能将现有材料信息映射到特定属性，从而在广阔的化学空间中迅速定位到目标性质。事实上，这种方法已在光伏材料^[47, 48]、催化剂^[49]和锂电池^[50]等多个领域成功应用并表现出不错的效果。

例如，Zhou 等人^[51]创建了一套基于贝叶斯优化的主动学习框架，在 3000 余种 Ni 基金属间化合物中，他们高效地筛选出了 15 种高性能低碳炔烃加氢催化剂。特别值得注意的是，NiIn 催化剂在实验中展现出 97% 的乙烯/丙烯选择性。在金属氧化物的研究中，Wang 等人^[52]利用四维结构描述符与梯度提升决策树模型，从 9095 种金属氧化物中筛选出 13 种低能垒催化剂。这一策略不仅显著提高了预测精度，还揭示了金红石型 IrO₂ 结构的独特优势。

为了更高效地进行催化剂筛选，Gong 等人^[53]开发了一种可解释的通用描述符模型 (ARSC)。他们通过物理特征工程实现了对氧还原 (ORR)、析氧 (OER)、二氧化碳还原 (CRR) 和氮还原 (NRR) 等电催化反应活性的同时预测。仅用 4500 次 DFT 计算，该模型就能筛选出高效双功能催化剂 Co-Co/Ir-Qv3，结果显示其 ORR 半波电位高达 0.941 V，而 OER 过电位低至 330 mV。在锂硫电池领域，Li 等人^[54]结合晶体图卷积神经网络 (CGCNN) 与 DFT 计算，对 800 余种单原子催化剂的吸附构型进行了分类，并预测了吸附能，最终成功筛选出了能有效抑制多硫化物穿梭效应的候选材料，其预测误差低至 0.14 eV。

综上所述，借助数据驱动策略，科研人员在催化剂的设计开发方面取得了显著进展。无论是在材料筛选还是性能优化层面，均展现出强大的优势，同时也为揭示反应机理提供了全新视角。尽管如此，即便将研究范围局限于单一催化剂体系，现有的机器学习方法仍需要复杂且耗时的特征构建工程，这极大地制约了其在大规模催化剂筛选任务中的适用性与效率。因此，迫切需要建立一个自动化且通用的设计框架，以提升催化剂筛选的规模、速度与精度。此类框架的构建有望从根本上突破传统基于密度泛函理论与机器学习方法所固有的局限，从而革新催化剂发现的范式。

1.3 本文研究内容

目前 2e⁻ WOR 催化剂体系依然面临着本征活性较低、过电位偏高及选择性调控机制不明等核心难题。传统的热力学模型在阐释高电位环境中的反应动力学特性时存在一定的局限性。与此同时，尽管依赖数据驱动的材料筛选方法在 ORR、OER 等反应中取得了突破，却尚未系统地应用于 2e⁻ WOR 催化剂，这在

一定程度上限制了高性能催化剂的快速筛选。

据此，本文提出了一种基于加权原子中心对称函数（wACSF）描述符的机器学习（ML）模型，并开发了自动化特征提取工具，实现了对金属合金、金属氧化物、钙钛矿及单原子催化剂（SACs）等材料体系中 ΔG_{OH^*} 和 ΔG_{O^*} 吸附自由能的精准预测，打破了传统单一材料的局限。此外，本文还建立了 $2e^-$ WOR 微观动力学模型和热力学选择性筛选标准，通过对数据集和预测集的快速筛选，识别出多种潜在的高活性材料。经实验验证，筛选出的最优催化剂在理论计算和实验结果上均显示出极高的一致性。本文整体研究框架如图 1-2 所示：

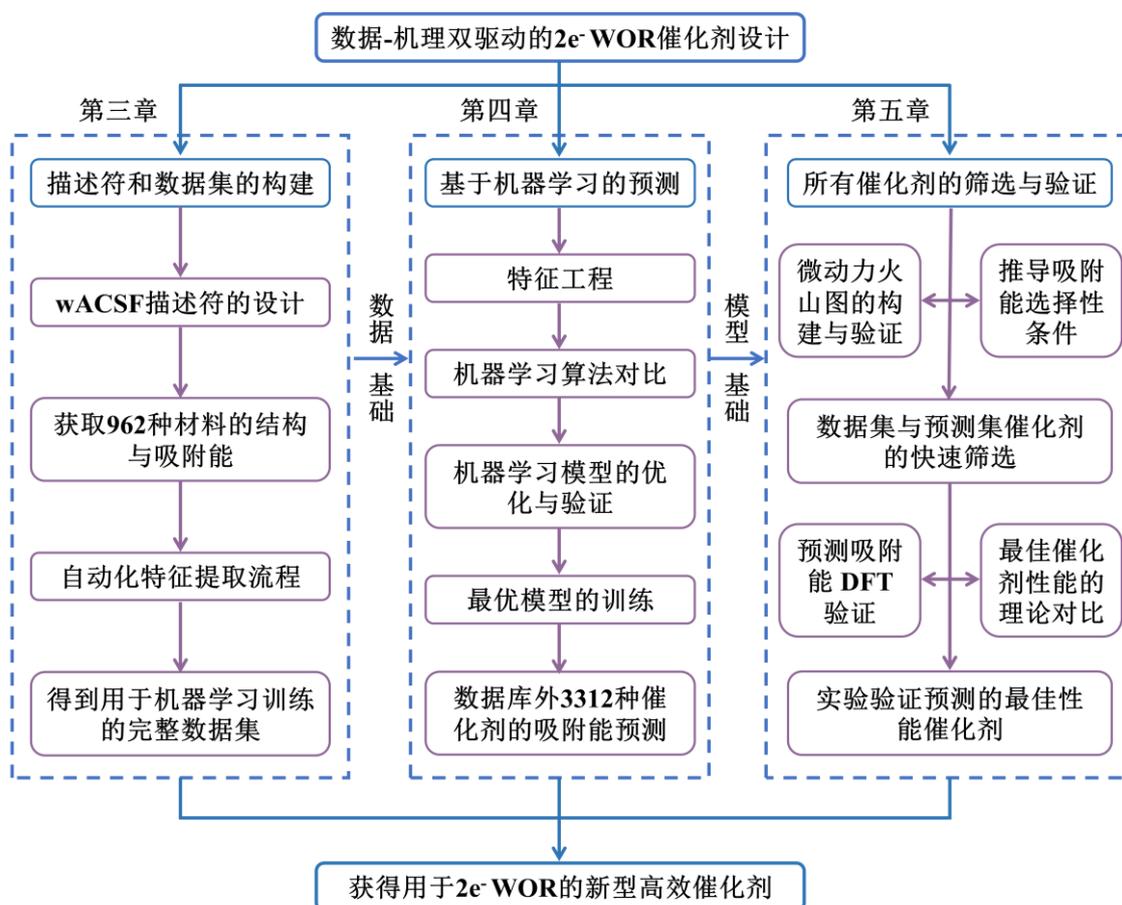


图 1-2 本文整体研究框架

具体研究内容如下：

(1) 在传统中心原子对称函数（ACSF）用于表征催化剂几何特征的基础上，本文进一步引入配位原子的电负性（Nm）和价电子数（V）作为权重因子，由此得到了两类加权描述符（wACSF），以实现对接域化学环境的多维解析。同时，与其他描述符的机器学习模型进行对比，验证了研究模型的精度和实用性。构建了 962 种催化材料的吸附自由能及吸附位点数据集，其中包括 426 种金属

间合金、160种金属氧化物（不含钙钛矿结构材料）、250种钙钛矿（ ABO_3 ）以及126种单原子催化剂。为了提高筛选效率，开发了一套自动化特征提取工具。该代码仅通过解析CIF文件中的晶体结构和元素信息，就可同时获取各类材料的全部描述符，进而得到了一个维度为 962×134 的特征矩阵。

（2）采用皮尔逊相关性分析与递归特征消除（RFE）方法，对134维描述符进行筛选，最终分别选出21个和19个核心描述符，用以构建 ΔG_{OH^*} 和 ΔG_{O^*} 预测模型。在与其他七种算法进行比较后，选择XGBoost回归模型进行训练。模型优化后，在测试集上对 OH^* 和 O^* 的吸附自由能预测表现优异， R^2 分别达到0.84和0.91，RMSE分别为0.52 eV和0.65 eV。另外，借助该机器学习模型，对数据集外新构建的3312种材料的吸附能进行快速预测。最后，通过SHAP值的分析，探究了各特征对预测值的影响强度和作用方向，并阐释了各描述符的物理化学特性。

（3）建立了 $2e^-$ WOR 微观动力学火山图，并与已报道的实验数据进行对比验证。通过微动力-热力双火山图的对比分析，证明了微观动力火山图在预测 $2e^-$ WOR 催化性能方面的准确性与适用性。基于此火山图及推导出的热力学选择性条件（ $\Delta G_{O^*} \geq 3.52$ eV，且 1.41 eV $\leq \Delta G_{OH^*} < 2.4$ eV），本文建立了 $2e^-$ WOR 催化剂的筛选标准。随后，对数据集和预测集中的催化剂进行了快速筛选，分别获得了49种和12种具有高选择性和高活性的材料。其中，经DFT计算验证的12种材料的预测值偏差均低于0.2 eV，且LiScO₂位于微观动力火山图的顶端。将前16种高性能催化剂的过电位与文献中报道的最佳 $2e^-$ WOR 催化剂相比较，以充分评估其催化性能的潜在优势。最后，通过固相烧结法制备了预测性能最佳的LiScO₂催化剂，并利用X射线衍射（XRD）、透射电子显微镜（TEM）、X射线光电子能谱（XPS）及电化学测试等手段，对其结构、活性及稳定性进行了全面表征和性能评估。实验结果显示，LiScO₂催化剂表现出优异的活性、选择性和稳定性。

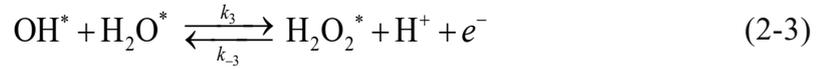
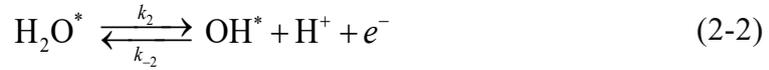
第 2 章 研究方法

2.1 火山图的计算

2.1.1 $2e^-$ WOR 微观动力模型的计算

微观动力学模型是一种对反应过程进行定量分析的方法，它以分子层面的反应机理为依据，用于描述催化剂表面发生的复杂反应过程。该方法结合了动力学和热力学参数，将催化反应过程分解为一系列基元反应步骤，以预测催化反应的速率和选择性^[55]。微观动力学方法还能模拟分析在真实反应条件下的催化反应具体状况，有助于理解催化剂的活性来源^[56]。这为开发具有优异性能的催化剂提供了坚实的理论基础，为其合理设计指明了方向。

建立微观动力学模型的首要步骤是确定初级反应的路径和相关速率常数。在此基础上，通过数值求解一组稳态常微分方程（ODEs），便可得到活性位点覆盖度、电子催化转化频率（TOF）等重要参数^[37]。本研究构建了一个 $2e^-$ WOR 微观动力学模型，其反应过程包含多个基元步骤，具体如下：



对于非电学步骤，正反应的速率常数 k_i 、平衡常数 K_i 和逆反应的速率常数 k_{-i} 由以下公式确定：

$$k_i = \nu_i \exp\left(-\frac{E_{\alpha}}{k_B T}\right) = \frac{k_B T}{h} \exp\left(-\frac{\Delta G_{\alpha,i}}{k_B T}\right) \quad (2-5)$$

$$K_i = \exp\left(-\frac{\Delta G_i}{k_B T}\right) \quad (2-6)$$

$$k_{-i} = \frac{k_i}{K_i} \quad (2-7)$$

对于电学步骤，它们由以下公式表示：

$$k_i = A_i \exp\left(-\frac{E_{\alpha,i}}{k_B T}\right) \exp\left(-\frac{e\beta_i(U - U_i)}{k_B T}\right) \quad (2-8)$$

$$K_i = \exp\left(-\frac{\Delta G_i + eU}{k_B T}\right) \quad (2-9)$$

$$U_i = -\frac{\Delta G_i}{e} \quad (2-10)$$

$$k_{-i} = \frac{k_i}{K_i} \quad (2-11)$$

式中， ΔG_i 、 E_a 和 A_i 分别代表自由能变化、活化能和前指数因子。由于溶剂重组，电化学反应的活化能被估算为 0.28 eV ^[57]。结合本研究的具体数据，对 A_i 的值作了适当修正。频率因子 ν_i 被估算为 $1 \times 10^8 \text{ s}^{-1}$ ^[58]。 β_i 是对称因子，取值为 0.5 ^[59]。

TOF_{e^-} 是电子的催化转化频率（TOF），通过数值求解在稳态条件下的速率微分方程来确定。速率微分方程的公式如下：

$$\frac{\partial \theta_*}{\partial t} = -k_1 \chi_{H_2O} \theta_* + k_{-1} \theta_{H_2O^*} + k_4 \theta_{H_2O_2^*} - k_{-4} \theta_* \chi_{H_2O_2} \quad (2-12)$$

$$\frac{\partial \theta_{H_2O^*}}{\partial t} = k_1 \chi_{H_2O} \theta_* - k_{-1} \theta_{H_2O^*} - k_2 \theta_{H_2O^*} + k_{-2} \theta_{OH^*} \quad (2-13)$$

$$\frac{\partial \theta_{OH^*}}{\partial t} = k_2 \theta_{H_2O^*} - k_{-2} \theta_{OH^*} - k_3 \theta_{OH^*} \chi_{H_2O} + k_{-3} \theta_{H_2O_2^*} \quad (2-14)$$

$$\frac{\partial \theta_{H_2O_2^*}}{\partial t} = k_3 \theta_{OH^*} \chi_{H_2O} - k_{-3} \theta_{H_2O_2^*} - k_4 \theta_{H_2O_2^*} + k_{-4} \theta_* \chi_{H_2O_2} \quad (2-15)$$

式中， χ 为摩尔分数， t 为时间。 χ_{H_2O} 和 $\chi_{H_2O_2}$ 分别取 1 和 1×10^{-5} 。 θ 是物种的覆盖度，假设每个物种的表面浓度不变^[59]。这些中间体的覆盖度需要满足以下条件：

$$\theta_* + \theta_{H_2O^*} + \theta_{OH^*} + \theta_{H_2O_2^*} = 1 \quad (2-16)$$

最后，电流密度的计算公式为：

$$j = e \rho TOF_{e^-} \quad (2-17)$$

式中， e 为基本电荷， ρ 为催化剂表面活性位点的密度。假设它与 Pt(111) 的表面密度相同，那么 $e\rho$ 的值估算为 $80.3 \mu\text{C}/\text{cm}^2$ ^[60, 61]。

2.1.2 理论火山图的推导计算

理想的高活性 $2e^-$ WOR 催化剂应具有低过电位。反应的过电位（ η_{WOR} ）是由理论极限电位（ U_L ）计算而来的。所谓极限电位，是指在所有基元反应步骤中，自由能降幅最小的电位值。计算得到的 U_L 或 η_{WOR} 与 ΔG_{OH^*} 之间的函数关系，能够绘制出 $2e^-$ WOR 反应的活性火山图。

根据 $2e^-$ WOR 反应式 (1-6) 至 (1-8) 可得到 U_L 的计算公式:

$$U_L = \max(\Delta G_1, \Delta G_2) / e \quad (2-18)$$

极限电位与平衡电位之间的差值称为过电位:

$$\eta_{WOR} = [\max(\Delta G_1, \Delta G_2) / e] - 1.76 \quad (2-19)$$

对于每个反应步骤, 自由能变化量 ΔG 为初始态与终态的自由能之差, 其计算表达式如下^[38]:

$$\Delta G = \Delta E + \Delta ZPE - T\Delta S + \Delta G_U + \Delta G_{pH} \quad (2-20)$$

式中, ΔE 为反应物和产物之间的总能量差, ΔZPE 为零点能校正值, ΔS 为有限温度 T 下的振动熵变。 $\Delta G_U = -eU$, 其中 e 为基本电荷, U 为电极电位。 ΔG_{pH} 为 H^+ 的自由能校正值。

2.2 机器学习

作为人工智能的重要组成部分, 机器学习依托多种算法和统计模型, 使计算机系统能够借助数据进行自我训练, 从而不断优化任务完成的效果^[62]。其核心在于利用数据训练模型, 使其能够对新的、未见过的数据做出精准的预估或判断。

本研究利用 Python 3 脚本和 Scikit-learn 库构建了机器学习模型^[63]。在完整的特征集上对评估器进行了训练, 随后根据特征的重要性评分, 逐步剔除贡献较低的特征。这个剪枝过程以递归的方式不断重复, 直至剩余特征的数量达到预设标准。在模型训练前, 对所有描述符进行了标准化处理, 以减小由于特征尺度不一致所引发的误差。每轮训练中, 为提升模型的表现, 采用了网格搜索算法优化超参数, 并调用 GridSearchCV 函数实现 10 折交叉验证, 对模型的稳定性与准确性进行了综合评估。

2.2.1 Scikit-learn 库介绍

Scikit-learn 是一个用 Python 编写的开源机器学习库, 广泛应用于数据挖掘、特征工程、模型训练与评估等众多领域^[64]。其主要优势在于提供了简洁高效的 API 接口、种类繁多的算法库 (如回归、分类、聚类、降维等) 以及完善的数据预处理工具 (如标准化、特征选择、交叉验证)。

在催化剂设计与性能预测的研究中, Scikit-learn 常被用来构建监督学习模型 (如回归模型)。通过输入催化剂的结构特征 (如原子环境描述符、热力学参数) 来预测其吸附自由能、反应活性或选择性等关键指标^[65]。本研究采用 Scikit-learn 库完成了数据预处理、特征工程以及模型的训练与评估。依靠 Scikit-learn

的灵活性和高效性，能够迅速构建出具备可解释性和高精度的模型。

2.2.2 极限梯度提升算法

极限梯度提升算法 (XGBoost) 作为一种先进的集成学习算法，通过组合多个学习器而非依赖单一模型来实现更高的预测精度^[66]。随机森林 (RF) 和 XGBoost 虽然都属于集成学习算法，但两者的核心逻辑截然不同。RF 是通过“并行”生成多棵独立的决策树 (类似“集体投票”)。而 XGBoost 采用的是提升策略，其子树是按顺序构建并通过加权方式合并输出的。在复杂数据场景下，后者往往能展现出更高的预测准确性。另外，尽管梯度提升决策树 (GBDT) 也属于提升方法，但 XGBoost 在优化过程中引入了二阶泰勒展开，而 GBDT 仅依赖于二阶导数^[67]。这一区别极大地提高了预测实际值的精确度，并提升了运算效率。最终，XGBoost 通过聚合多个决策树的输出获得最终预测结果，其公式如下：

$$\hat{y}_j = \sum_1^L f_1(x_j), f_1 \in F \quad (2-21)$$

在该公式中， \hat{y}_j 为模型最终输出的预测值， L 表示组合中决策树的总数 (即实际应用中需要调优的超参数 `n_estimators`)。 f_1 对应第 1 棵决策树， x_j 指的是第 j 个输入样本。 F 代表所有树模型的集合。

模型所采用的目标函数和正则化项在公式 (2-22) 和 (2-23) 中给出：

$$Obj^{(t)} = \sum_{j=1}^n loss(y_j, \hat{y}^{(t-1)} + f_t(x_j)) + \Omega(f_t) + c \quad (2-22)$$

$$\Omega(f_t) = \gamma T + \frac{1}{2} \lambda \sum_{o=1}^T W_o^2 \quad (2-23)$$

式中， $Obj^{(t)}$ 表示构建第 t 棵树时所采用的目标函数， $loss$ 表示损失函数，通常取均方误差。 $\hat{y}^{(t-1)}$ 是由前 $t-1$ 棵树计算出的预测值， c 为常数项。 $\Omega(f_t)$ 表示第 t 棵树的正则化项 (决定了待调整树的深度 `max_depth` 超参数)。 γ 和 λ 是正则化系数， T 表示一棵树中叶子节点的总数，而 W_o 则代表该树中第 i 个叶子节点的权重。

$Obj^{(t)}$ 的泰勒公式的展开式如下所示：

$$Obj^{(t)} \approx \sum_{j=1}^n [loss(y_j, \hat{y}^{(t-1)}) + g_j f_t(x_j) + \frac{1}{2} h_j f_t^2(x_j)] + \Omega(f_t) + c \quad (2-24)$$

在这里，所有 h_j 值的累加值表示待调整叶子节点的最小样本权重和，即超参数 `min_child_weight`。

2.2.3 机器学习模型的评估指标

在机器学习模型的构建和训练中，评估指标的选择尤为关键。合适的评估参数能全面考察模型的表现，识别出过拟合或欠拟合偏差，并在模型精度和泛化能力之间权衡，从而选择出最合适的方案。本研究采用均方根误差（RMSE）与决定系数（ R^2 ）作为核心评估指标^[68]，计算方法见式（2-25）和式（2-26）。

（1）RMSE 反映真实值与预测结果之间的平均偏差程度，其数值越大表明预测精度越低，且对异常值尤为敏感，常被用作损失函数来评估模型的误差。

$$RMSE = \sqrt{\frac{1}{m} \sum_{i=1}^m (y_i - \hat{y}_i)^2} \quad (2-25)$$

式中： y_i 表示真实值， \hat{y}_i 表示预测值。

（2） R^2 反映了模型的拟合程度。当模型的预测性能提升时， R^2 值趋近于 1。若测试集的 R^2 值显著低于训练集，说明模型可能存在过拟合现象；若测试集和训练集的 R^2 值均较低，则模型可能为欠拟合。一般来说， R^2 值越大，表示模型效果越好。

$$R^2 = 1 - \frac{\sum_i (y_i - \hat{y}_i)^2}{\sum_i (\bar{y}_i - y_i)^2} \quad (2-26)$$

式中： \bar{y}_i 表示实际值的平均值。该表达式的分子部分是实际值与预测值的平方差之和，分母部分为实际值与平均值的平方差之和。

2.2.4 特征之间的相关性系数

皮尔逊（Pearson）相关系数是一种用于量化两个连续变量间线性关系的统计量，其数学定义见式（2-27）。它既能反映线性关联的强度，又能显示其方向（即正相关或负相关）^[69]。当相关系数达到 1 时，表明两变量间存在极强的正向线性关系；若其值为 -1 时，则说明存在极强的负向线性关系；如果相关系数接近 0，可视为两变量之间几乎不存在显著的线性相关性。

$$cor(x, y) = \frac{1}{n-1} \cdot \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sigma_x \sigma_y} \quad (2-27)$$

其中， x_i 表示特征 x 的第 i 个样本值， y_i 表示特征 y 的第 i 个样本值， \bar{x} 和 \bar{y} 分别为 x 和 y 的均值， σ_x 与 σ_y 则分别代表 x 和 y 的标准差， n 为样本总量。

2.3 密度泛函理论

密度泛函理论 (Density Functional Theory, DFT) 是一种基于量子力学的电子结构计算方法, 已被广泛应用于从头计算原子、分子、晶体、表面结构及相互作用^[70]。通过将多体问题转化为单电子在有效势场中运动的问题, DFT 显著降低了计算复杂度。其基本理念是用电子密度替代波函数作为描述体系基态性质的主要变量, 从而解决了传统多体波函数方法在高维度问题上的难题^[71]。DFT 的关键在于交换-关联泛函 (XC functional), 用于描述电子间的交换关联效应。DFT 通过 Hohenberg-Kohn 定理证明电子密度包含了系统的所有信息^[72]。Kohn-Sham 方程则通过引入虚拟的单电子系统, 将原本复杂的问题转化为求解单电子方程的问题, 同时考虑了交换-关联效应^[73]。DFT 的计算效率高, 特别适合处理较大系统, 但其结果的准确性依赖于所采用的交换-关联泛函。

2.3.1 Hohenberg-Kohn 定理

Hohenberg-Kohn 定理作为 DFT 的基石, 由 Walter Kohn 和 Pierre Hohenberg 于 1964 年提出。这一定理为多电子体系的电子结构研究提供了全新的视角, 即通过电子密度而非复杂的多电子波函数来描述体系的性质^[74]。这一创新性理论巧妙地降低了量子力学多体问题的求解难度, 从而使 DFT 成为计算化学和材料科学等领域的关键工具。该定理主要适用于非均匀电子气模型, 其核心内容由两个子定理共同构成:

(1) 在外部势场作用下的多电子体系中, 其基态能量可唯一地表示为电子密度的泛函。这意味着, 只要给定一个电子密度, 就能以唯一的方式确定相应的外部势和基态能量。

(2) 对于特定的外部势场, 体系的基态能量作为电子密度的泛函, 当电子密度取基态密度时, 能量达到最低值。换句话说, 基态能量是能量泛函的最小值。

根据 Hohenberg-Kohn 定理, 体系的能量泛函的数学表达式如下:

$$E(\rho, V) = \int V(\mathbf{r})\rho(\mathbf{r})d\mathbf{r} + T[\rho(\mathbf{r})] + \frac{e^2}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}[\rho(\mathbf{r})] \quad (2-28)$$

其中, 等式右侧四项依次是外部势场中的电子势能、动能、电子间的库伦作用和交换-关联势能。需要强调的是, 该定理并未给出具体的能量泛函形式, 因此其本身无法直接用于实际计算。

2.3.2 Kohn-Sham 方程

Kohn-Sham 方程由 W. Kohn 和 L. J. Sham 于 1965 年共同提出。这一方程将原本复杂的多体电子系统等效为单电子在有效势场中的运动模型，从而大大降低了计算难度。其核心思想在于构建了一个无相互作用的虚拟粒子模型，将真实体系的交换关联作用纳入有效势场中。由此，通过求解单电子方程，便可获得与实际体系相匹配的电子密度，并进而推导出基态能量^[73]。

在理论推导中，若将 ρ 的变分用 $\Phi_i(r)$ 的变分来代替，拉格朗日乘子用 E_i 代替，则单电子方程可表示为：

$$\begin{aligned} V_{KS}[\rho(r)] &= V(r) + V_{Coul}[\rho(r)] + V_{XC}[\rho(r)] \\ &= V(r) + \int dr' \frac{\rho(r')}{|r-r'|} + \frac{\delta E_{XC}[\rho(r)]}{\delta \rho(r)} \end{aligned} \quad (2-29)$$

Kohn-Sham 方程为除交换关联泛函以外的各个部分提供了明确的数学表达式，同时将复杂的相互作用项也整合在内。如此一来，对实际材料的电子结构及性质的计算成为可能。随着问题焦点的转变，研究主要集中在如何合理准确地表达交换关联泛函，而交换关联势的近似表达则直接决定了密度泛函理论整体的计算精度。

2.3.3 计算软件介绍及参数设置

本研究采用维也纳从头计算软件包（Vienna ab-initio simulation package, VASP）进行结构弛豫与电子性质的密度泛函理论计算。VASP 是由维也纳大学 Hafner 团队开发的一款基于第一性原理的电子结构计算和量子力学-分子动力学模拟软件^[75-77]。它采用平面波基组和赝势方法，在材料科学、化学和物理等领域均有广泛应用，是材料模拟和计算物质科学研究中最流行的商用软件之一。

计算前，预先准备包括 INCAR、POSCAR、POTCAR 和 KPOINTS 在内的多种输入文件^[78]，通过命令行提交任务。其中，POTCAR 文件包含了体系中各元素的赝势信息，用于描述电子与原子核之间的相互作用。POSCAR 文件呈现了晶格参数、原子坐标以及系统的空间结构。KPOINTS 文件用于指定在布里渊区积分时采用的 k 点网格，直接影响计算的精度与所需时间。INCAR 文件则用于设定计算类型、算法参数及收敛标准。

计算完成后，利用 VESTA（Visualization for Electronic and Structural Analysis）、VASPKIT 和 Wolfram Mathematica 等软件对生成的输出文件（如 OUTCAR、CONTCAR）进行结果处理及结构优化后的可视化分析。VESTA 能够读取、编辑及处理晶体结构数据，并以多种方式展示原子排列和晶格信息^[79]。

VASPKIT 提供了便捷的后处理功能,包括生成输入文件、进行能带结构及态密度分析和电荷密度计算^[80]。Wolfram Mathematica 则主要服务于数据分析与函数可视化。

本研究采用 DFT 对最终的吸附能预测结果进行了验证。计算中选用改进的广义梯度近似 (GGA) 泛函 Revised Perdew-Burke-Ernzerhof (RPBE),并结合 Hubbard U 修正方法来评估催化剂的自由能,同时考虑自旋极化^[81]。计算采用投影缀加波 (PAW) 赝势,平面波能量截断值设置为 500 eV,高斯展宽参数设为 0.05 eV。在体相金属优化时,利用共轭梯度算法,采用 11×11×11 的 Monkhorst-Pack k 点网格。动能截断值设定为 700 eV,并应用能量宽度为 0.02 eV 的高斯展宽方法^[76]。为更好地体现每种催化剂的简并态,选用尺寸为 19 Å×20 Å×21 Å 的非对称超胞,在 Γ 点进行能量优化。同时,对金属氧化物超胞采用 7×7×1 的 k 点网格进行采样。最后,通过几何弛豫,直至体系的力收敛至 0.01 eV/Å。

2.4 实验方法

2.4.1 实验材料

在本实验中,所采用的材料包括:碳酸锂 (Li_2CO_3 , 纯度 99.99%)、三氧化二钪 (Sc_2O_3 , 纯度 99.995%)、碳酸氢钾 (KHCO_3 , 纯度 $\geq 99.95\%$)、盐酸 (HCl, 浓度 37%)、异丙醇 (IPA, 纯度 99%) 以及无水 N,N-二甲基甲酰胺 (DMF, 纯度 99.8%)。这些试剂均购自 Sigma-Aldrich 公司,并在使用前未经过任何特殊处理。此外,镍 (II) 酞菁 (NiPc, 纯度 >95%) 购自 PorphyChem 公司,并在使用前经过三次升华处理以提高纯度。多壁碳纳米管 (CNT), 直径约为 10 nm, 由 CheapTubes Inc 提供;使用前对 CNT 进行了如下处理:先将样品在 400°C 下于空气中煅烧 30 分钟;随后用 3 M HCl 溶液进行酸洗 (超声浴 1 小时和在 80°C 下搅拌 12 小时);最后在流量为 200 sccm 的 5% H_2/Ar 混合气中,于 1200°C 下热退火 2 小时。氩气 (Ar, 纯度 5.0 级) 及 5% H_2/Ar 混合气均由 BOC Australia 公司提供。

2.4.2 催化剂的合成

本研究中, LiScO_2 采用固态反应法合成^[82]。具体而言,按化学计量比称取 Li_2CO_3 (过量 20%) 与 Sc_2O_3 , 先在丙酮中充分混合均匀,然后在 500 MPa 压力下压制成圆片。接着将该圆片以 5 K/min 的升温速率加热至 1473 K, 在空气中保温 12 h 后缓慢冷却至室温。所得产物进一步研磨成细粉,用于后续的表征和测试。

NiPc/CNT 的制备则是将 NiPc 负载于经过纯化处理的多壁碳纳米管(CNT)载体上,然后将两者按照 1:5 的质量比 (NiPc:CNT) 混合溶解于 DMF 中。经过 30 分钟的超声处理后,将混合物持续搅拌一整夜。之后通过过滤回收固体,并用 DMF 和乙醇进行洗涤。最后,在真空环境下将固体产物干燥一夜,即得 NiPc/CNT 催化剂。

2.4.3 材料表征

本研究采用 Perkin Elmer Avio 500 型电感耦合等离子体原子发射光谱(ICP-AES)仪器对催化剂的元素组成进行测定。LiScO₂ 的 X 射线衍射 (XRD) 图谱利用 Stoe Stadi P 型衍射仪 (Cu-K α 辐射源, $\lambda=1.5406 \text{ \AA}$) 采集。X 射线光电子能谱 (XPS) 则由 Thermo Fisher Scientific K-Alpha+型光谱仪 (Al-K α 激发源, 1486.3 eV) 获得。在 Cs 校正的 Hitachi HF5000 型透射电子显微镜 (加速电压 200 kV) 上采集了透射电子显微镜 (TEM) 图像和 X 射线能量色散谱 (EDX) 元素面分布图。所有样品均预先放置在碳膜支撑的铜网上,经过等离子体和紫外线照射处理后,再装载到低温样品架上,以保护样品在测试过程中免受电子束辐射损伤。

2.4.4 电化学性能测试

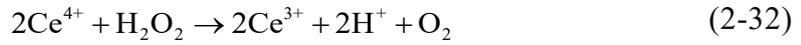
所有电化学测试均在 Autolab PGSTAT302N 电化学工作站上进行,同时采用旋转环盘电极 (RRDE, 型号为 E6R1) 进行测试。使用的玻碳盘电极直径为 5 mm,铂环电极的外径与内径分别为 7.5 mm 和 6.5 mm,收集效率设定为 $N=0.25$ 。对电极为铂网,参比电极为 Ag/AgCl (以 3M KCl 溶液填充)。所有电位均以可逆氢电极 (V_{RHE}) 为基准进行校准。催化剂以 5 mg/mL 的浓度分散于含有 0.05% Nafion 117 离子聚合物的水/异丙醇混合液 (体积比 1:9) 中,经超声浴处理后滴涂在盘电极上,质量负载量为 0.2 mg/cm^2 。水氧化性能测试在氩气饱和的 0.5 M KHCO₃ 电解液中进行,电极转速为 1600 rpm。测试前,催化剂在 1.0~1.6 V_{RHE} 范围内以 50 mV/s 的扫描速率循环 50 次以达到稳定状态,随后在新鲜电解液中以 5 mV/s 扫描速率采集线性扫描伏安 (LSV) 曲线。环电极偏压设为 0.4 V_{RHE} ,用以还原盘电极上产生的 O₂ 或 H₂O₂。

H₂O₂ 的选择性 ($X_{H_2O_2}$) 和法拉第效率 ($FE_{H_2O_2}$) 的计算公式如下:

$$X_{H_2O_2} (\%) = 200 \times \frac{\frac{i_{ring}}{N}}{i_{disk} + \frac{i_{ring}}{N}} \quad (2-30)$$

$$FE_{H_2O_2} = \frac{i_{ring}}{i_{disk}} \times 100\% \quad (2-31)$$

在流动电解池三电极体系中，对 $LiScO_2$ 的稳定性进行了评估。具体操作为：将 $LiScO_2$ 以 2 mg/cm^2 的负载量沉积于面积为 4 cm^2 的碳纸上作为工作电极，并采用预活化的 Nafion 212 膜与铂网对电极隔开。 0.5 M KHCO_3 电解液通过蠕动泵以 1 mL/min 的流速持续供给阳极室。通过 $Ce(SO_4)_2$ 滴定法定期检测回收电解液中 H_2O_2 的浓度，其中黄色的 Ce^{4+} 被 H_2O_2 还原为无色的 Ce^{3+} 。随后，在 316 nm 处利用紫外光谱测量确定 H_2O_2 的浓度 ($C_{H_2O_2}$) 以及其法拉第效率，其中 $\Delta C_{Ce^{4+}}$ 表示 Ce^{4+} 浓度的变化量，相关计算公式如下：



$$C_{H_2O_2} = \frac{1}{2} \cdot \Delta C_{Ce^{4+}} \quad (2-33)$$

2.5 本章小结

本章一开始阐述了微观动力模型和理论火山图的计算过程，接着介绍了机器学习的基本概念，并重点说明了所用的 Scikit-learn 库、XGBoost 算法以及皮尔逊相关系数分析方法。随后，讨论了密度泛函理论的基本概念和核心公式，包括 Hohenberg Kohn 定理和 Kohn-Sham 方程。接下来介绍了本研究中所采用的理论计算软件及其后处理工具，并详细说明了各计算软件中相关参数的具体设置。最后，对实验部分涉及的材料和合成方法进行了介绍，包括具体操作步骤和参数设置。同时，在材料的表征与测试部分，介绍了各类仪器的应用以及测试过程中各环节的具体步骤。

第 3 章 描述符和自动化提取策略

目前，二电子水氧化 ($2e^-$ WOR) 产 H_2O_2 的核心瓶颈在于所用催化剂普遍表现出过电位较高 ($>1.0 V$) 且产率偏低 ($<5 \mu mol \cdot min^{-1} \cdot cm^{-2}$)。尽管机器学习方法可加速催化剂筛选，但现有描述符 (如原子固有描述符) 尚未能实现多类型非贵金属催化剂的普适性评价。因此，迫切需要开发一种自动化且通用的设计框架，以便有效筛选各类催化剂。这一框架有望革新基于密度泛函理论和机器学习的筛选方法，从根本上提升筛选的规模、速度和准确性，从而有效克服传统方法的局限。

本章首先介绍了催化剂筛选的通用设计框架，并开发了一种基于加权原子中心对称函数 (wACSF) 的局域原子环境描述符。该描述符结合了配位原子的电负性和价电子数，构造出适用于多类催化剂的通用描述符。通过 Python 程序自动从 962 种材料的 CIF 文件中提取特征，最终构建了一个 962×134 的特征矩阵。这一方法为跨材料体系的高通量筛选提供了高效且可解释的解决方案。

3.1 数据机理双驱动的催化剂设计框架

本研究提出了如图 3-1 所示的通用催化剂设计框架，通过融合数据驱动技术与机理建模策略，实现了跨材料体系的高效筛选。整个流程的核心步骤如下：

(1) 构建四类典型材料 (包括金属间合金、金属氧化物、钙钛矿及单原子催化剂) 的数据库，解析其活性位点的晶体结构信息 (图 3-1 步骤 1)。之后在不同对称函数参数组合条件下，计算各催化剂活性位点的径向和角向对称函数值 (G_i)^[83]。接着结合催化剂中心原子与配位原子的电负性加权值 ($\sum N m_i$) 和价电子数加权值 ($\sum V_i$)，构建出描述化学环境的特征。将计算得到的 G_i 与这些加权值相乘，生成加权原子中心对称函数 (wACSF) 描述符 (图 3-1 步骤 2)。

(2) 将这两类加权对称函数值 (A_i 和 B_i) 与催化剂固有特征值合并，作为自变量，最终整合形成包含 134 维特征的矩阵 (图 3-1 步骤 3)。以羟基 (ΔG_{OH^*}) 和氧 (ΔG_{O^*}) 的吸附自由能作为目标变量，建立起催化位点吸附特性与其数值特征之间的对应关系。整个过程由本研究编写的代码实现自动化处理。

(3) 采用皮尔逊相关性分析与递归特征消除 (RFE) 方法剔除冗余特征 (图 3-1 步骤 4)。并针对 ΔG_{OH^*} 与 ΔG_{O^*} 的预测任务分别筛选出 21 个和 19 个关键描述符。接着，利用 XGBoost 算法构建机器学习模型，通过网格搜索优化超参数，从而实现了对 3312 种新催化剂吸附自由能的高精度预测 (如图 3-1 步骤 5 所示)。

(4) 基于 ΔG_{OH^*} 与交换电流密度 (j) 之间的函数关系，构建了微观动力学

火山图(图 3-1 步骤 6)。通过与文献报道的多种催化剂实验数据进行对比,验证了微观动力学模型在描述活性趋势及反应路径方面的预测精度和可靠性。

(5) 最后,遵循 ΔG_{OH^*} (1.41~2.4 eV) 与 ΔG_{O^*} (≥ 3.52 eV) 的选择性条件和微观动力学活性模型,对所有材料进行筛选,快速识别出 61 种潜在候选催化剂(图 3-1 步骤 7)。实验验证了预测性能最优的催化剂,结果显示理论与实验性能一致,从而证实了筛选结果的可靠性。

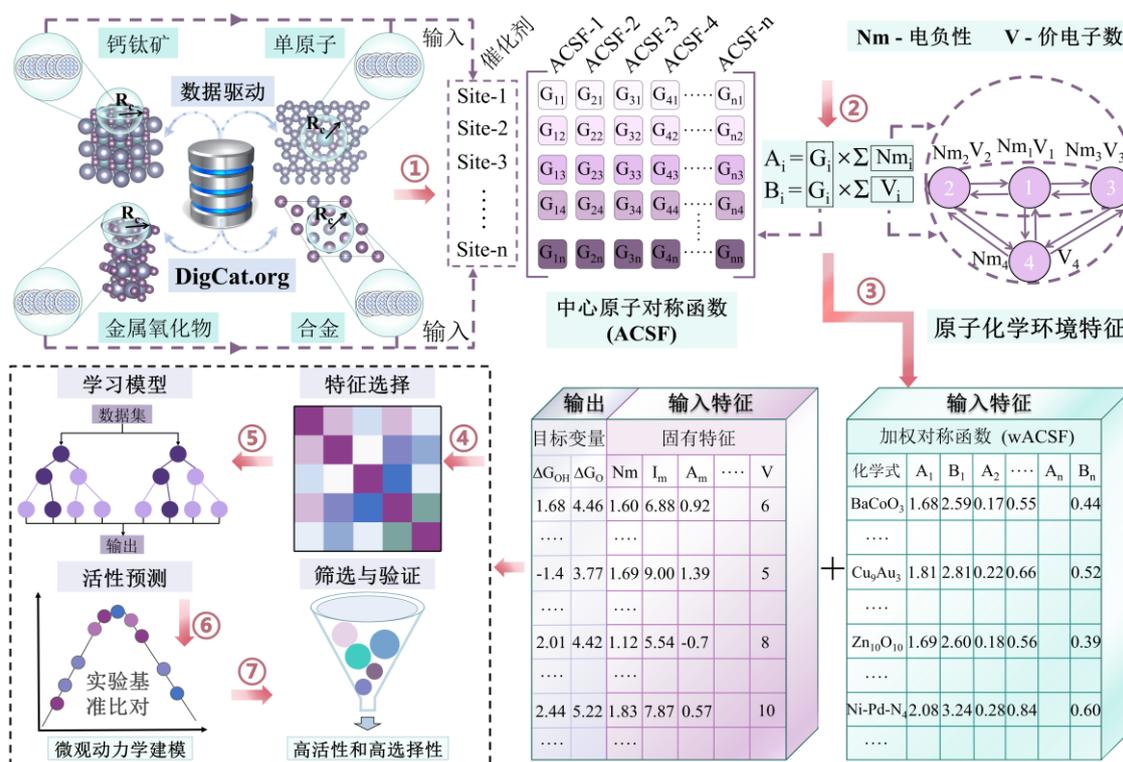


图 3-1 数据与机理双驱动的催化剂设计筛选框架示意图

3.2 加权的对称函数描述符

为突破传统单一材料标度关系模型的固有局限性,并提升描述符在多种材料体系中的适用性,本研究引入了加权对称函数和本征原子特性描述符。在原子中心对称函数(Atomic-Centered Symmetry Function, ACSF)的基础上,进一步将中心原子与其配位环境的电负性或价电子数之和作为加权因子,从而得到了配位环境加权的对称函数(Weighted Atomic-Centered Symmetry Function, wACSF)。此外,采用了六种仅取决于主体金属元素且可直接从元素周期表获取的次级特征,包括电负性、原子序数、原子半径、第一电离能、电子亲和能与价电子数。这些本征属性参数与局域环境无关,对于构建适用于多类材料吸附位点的普适性预测模型具有重要作用。

3.2.1 径向和角向对称函数的计算

原子中心对称函数（ACSF）通过构建原子核周围的径向和角向分布函数，模拟了原子的几何环境，实现了对活性中心原子周围不同化学元素空间排布特征的精确表征^[84, 85]。其中，径向对称函数的数学表达式为：

$$G_i^{rad} = \sum_{j \neq i}^N e^{-\eta(r_{ij}-R_s)^2} f_c(r_{ij}) \quad (3-1)$$

公式中， N 代表原子总数， r_{ij} 表示原子 i 与 j 之间的距离。参数 η (eta) 和 R_s (rshift) 分别用于调节高斯函数的宽度和位置。截断函数 $f_c(r_{ij})$ (cutoff function) 确保仅考虑中心原子邻近区域对体系能量的贡献，其典型定义见式 (3-2)。式中的截断半径 R_c (cutoff radius) 界定了中心原子作用域的空间范围。

$$f_c(r_{ij}) = \begin{cases} \frac{1}{2} \left[\cos\left(\frac{\pi r_{ij}}{R_c}\right) + 1 \right], & r_{ij} \leq R_c \\ 0, & r_{ij} > R_c \end{cases} \quad (3-2)$$

描述角向环境的对称函数形式更为复杂，其公式为：

$$G_i^{ang} = 2^{1-\zeta} \sum_{\substack{j,k \neq i \\ j < k}}^N (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta((r_{ij}-R_s)^2 + (r_{ik}-R_s)^2)} f_c(r_{ij}) f_c(r_{ik}) \quad (3-3)$$

式中 θ_{ijk} 表示原子 i 、 j 、 k 所构成的键角。参数 λ (lambda) 调控函数的对称特性，通过取值 $\lambda = \pm 1$ 实现函数极值在 0° 与 180° 之间的转换。参数 η 和 R_s 分别调节对称函数的宽度和中心位置，参数 ζ (zeta) 用于控制角向分布的锐度。 $f_c(r_{ij})$ 和 $f_c(r_{ik})$ 为相应的截断函数。

关键参数 η 和 ζ 分别用于控制径向和角向“指纹”特征的宽度特性。例如，较大的 η (或 ζ) 值会导致径向 (或角向) 对称函数分布更窄。为准确描述每个原子中心的化学环境，需构建包含多种 η 和 ζ 参数组合的对称函数集合。表 3-1 列出了 ACSF 关键参数 (η 、 R_s 、 R_c 、 λ 、 ζ) 的取值。采用 Behler 方法^[43]时，仅对 R_c 进行优化筛选，其余参数则通过合理组合以构建完整的对称函数集合。具体的参数组合见 <https://github.com/Weijie-Yang/FG-LAED> 中的配置文件。

表 3-1 原子中心对称函数参数列表

| 径向对称函数参数 | | | 角向对称函数参数 | | | | |
|----------|-------|-------|----------|-----------|---------|-------|-------|
| η | R_s | R_c | η' | λ | ζ | R_c | R_s |
| 0.007 | 0 | 4 | 0.007 | -1 | 1 | 5 | 4 |
| 0.012 | 0 | 5 | 0.043 | 1 | 6 | 7 | 5 |
| 0.022 | 0 | 6 | 0.135 | — | — | — | 6 |

表 3-1 (续表)

| η | R_s | R_c | η' | λ | ζ | R_c | R_s |
|--------|-------|-------|---------|-----------|---------|-------|-------|
| 0.040 | 0 | 7 | 0.236 | — | — | — | 7 |
| 0.072 | 0 | 8 | — | — | — | — | 8 |
| 0.131 | 0 | 9 | — | — | — | — | 9 |
| 0.239 | 0 | ∴ | — | — | — | — | ∴ |
| 0.428 | 0 | 16 | — | — | — | — | 16 |

3.2.2 配位环境加权的对称函数介绍

配位环境加权的对称函数 (wACSF) 是一种新型描述符, 它融合了几何结构与化学特性, 以增强对多类催化剂活性位点的普适性描述。其基本原理是在 ACSF 表征几何特征的基础上, 引入配位原子的固有电子参数 (即电负性和价电子数) 作为权重因子。通过对 ACSF 进行加权, 能够将配位原子的电子特性纳入局部化学环境中, 从而更全面地捕捉活性中心周围的几何与电子结构特征, 实现局域化学环境的多维度解析。构建的两类加权对称函数描述符分别见公式 (3-4) 和 (3-5)。

(1) 电负性加权的对称函数表征了配位原子电负性对中心原子电子分布的调控效应。它能够反映配位原子对局域化学环境的影响, 电负性较高的配位原子通过加权项增强了描述符对电子吸引作用的敏感度。

$$A_i = G_i^{rad/ang} \times \sum_{i=1}^n Nm_i \quad (3-4)$$

(2) 价电子数加权的对称函数反映了配位原子价电子数对局域电荷转移的影响。用于表征配位原子在局部电子密度中的贡献, 价电子数较大的配位原子会增强中心原子与吸附质间的电荷转移。加权项量化了配位原子的供电子能力。

$$B_i = G_i^{rad/ang} \times \sum_{i=1}^n V_i \quad (3-5)$$

以上公式中, n 为配位数, A_i 与 B_i 分别对应电负性加权对称函数与价电子数加权对称函数的计算结果。 $\sum_{i=1}^n Nm_i$ 表示中心原子及其配位原子的电负性之和,

$\sum_{i=1}^n V_i$ 为中心原子及其配位原子的价电子数之和。

通过上述加权, wACSF 不仅能描述活性位点的几何构型 (如配位数、键长和键角), 还能量化配位原子的电子效应 (如电荷分布与电子转移趋势)。加权处理使得相同的几何构型在不同电子环境下也能产生差异化描述, 显著提升了对不同活性位点的辨识能力。图 4-7 展示了机器学习特征的重要性排名, 其中

电负性或价电子数加权的对称函数(如 A_{27} 、 A_{16} 、 B_6 、 B_{21})排名靠前,且对 ΔG_{OH^*} 和 ΔG_{O^*} 的预测贡献度很高,进一步验证了电子加权描述符的有效性。

3.3 结构与吸附能数据库的建立

数据库的建立主要由两个环节构成:第一个是结构数据的收集,包括获取晶体结构信息(CIF 文件)、催化活性位点(Adsorption Site)及吸附自由能(ΔG_{OH^*} 和 ΔG_{O^*})数据;第二部分是利用自动化特征提取算法,对上述材料的结构和位点信息进行加权对称函数计算和固有特征的提取。完成数据收集和特征计算后,将所有数据汇总整理成表格,以便后续分析和使用。

对于材料结构与吸附能的获取,本研究中的催化剂数据库涵盖了四类材料。这些材料的 CIF 结构是从文献中获取的,其结构构建过程如下:采用 ABO_3 型立方钙钛矿结构,通过 A、B 位元素的不同组合(如 $LaAlO_3$ 、 $CaSnO_3$)实现结构多样性^[86]。其中, A 和 B 代表不同的金属元素。单原子催化剂(SACs)是通过将单个金属原子锚定在以石墨烯为基底的载体表面而形成的^[87]。在选择单原子催化剂时,以过渡金属作为活性中心。在单缺位 SACs 中,主要采用 N 原子作为配位非金属原子进行掺杂。金属间合金由 37 种选定金属与过渡金属组合而成,形成具有 A_3B 和 AB 化学计量比的面心立方结构。其中结构类型包括桥位(bridge)、面心立方 A 位(FCC-A)、六方最密堆积 A 位(HCP-A)、顶位 A(top-A)以及顶位 B(top-B)等吸附位点^[88]。除双金属合金外,数据库还包括了 37 种纯金属催化剂。对于金属氧化物,则是通过金属与氧结合得到的。根据不同金属种类、金属与氧的化学计量比以及晶体结构,可构建出多种金属氧化物^[89]。本研究选择了简单的二元金属氧化物(如 MgO 和 ZnO)以及通式为 ABO_2 的复杂金属氧化物(如 $LiScO_2$ 和 $AgMoO_2$)。

最终,构建了包含 962 种催化材料的结构、吸附自由能及吸附位点信息的数据库。该数据集涵盖了四大类材料:426 种金属间合金、160 种金属氧化物、250 种钙钛矿(ABO_3)以及 126 种单原子催化剂。对各催化剂的吸附位点信息和吸附自由能数据进行了统计整理,最终汇总成了 Excel 表格文件,命名为“原始数据集”,访问链接为: <https://github.com/Weijie-Yang/H2O2>。

3.4 自动化特征提取模型

为了快速获取 962 种催化剂的 wACSF 值和固有特征,编写了 Python 代码。该代码能够自动解析不同材料活性位点的配位元素结构信息,并为每种材料建立独特的数据渠道。程序会批量计算各种材料的 wACSF 值,同时自动导出对应的固有特征,最后将所有信息整合成表格,形成完整的特征集。通过有效获

取目标结构的众多描述符,该方法不仅能够量化活性位点及其局域环境的特征,还为不同材料体系的一致性表征提供了独特路径,显著提升了特征提取效率,为后续研究奠定了坚实基础。

3.4.1 特征提取策略

在 962 种催化剂的结构文件及吸附位点信息的基础上,通过自动化特征提取策略构建了催化材料数据集,其流程如图 3-2 所示,主要分为以下两部分:

(1) 对称函数值的计算:将结构文件(CIF)输入,程序读取位置信息并生成 962 个独立的输入文件(input.data),同时获取催化剂的元素信息。随后批量修改所有催化剂配置文件(config.data)中对应的元素信息。将这两个文件提交计算后,会输出各种材料活性位点的对称函数值。

(2) 元素固有特征提取:程序解析催化剂名称中的元素信息,从特征库(见:<https://github.com/Weijie-Yang/H2O2>,其中包含 88 种元素的六个基本特征)中提取相应的电负性和价电子数,然后根据配位信息对对称函数值进行加权。同时引入六种原子中心的次要特征(电负性、原子序数、原子半径、第一电离能、电子亲和能和价电子数),这些属性仅与金属本身有关,可直接从元素周期表中获取。这些与环境无关的属性对于构建涉及不同金属吸附位点的模型至关重要。

最终,该流程将两部分模块的输出结果整合为一张特征表。与传统手动方法相比,这种自动化策略仅需极少输入参数即可生成一套全面的特征集。通过使用特征库和标准化配置文件,用户只需提供 CIF 文件和活性位点信息,系统便能自动生成加权对称函数值和元素固有属性的完整特征集。由于所有结构的初始对称函数值均采用同一组参数计算,该方法确保了不同材料体系特征维度的一致性,且通过加权优化原始特征并不改变其维度,这对于机器学习模型的稳定性和适用性至关重要。

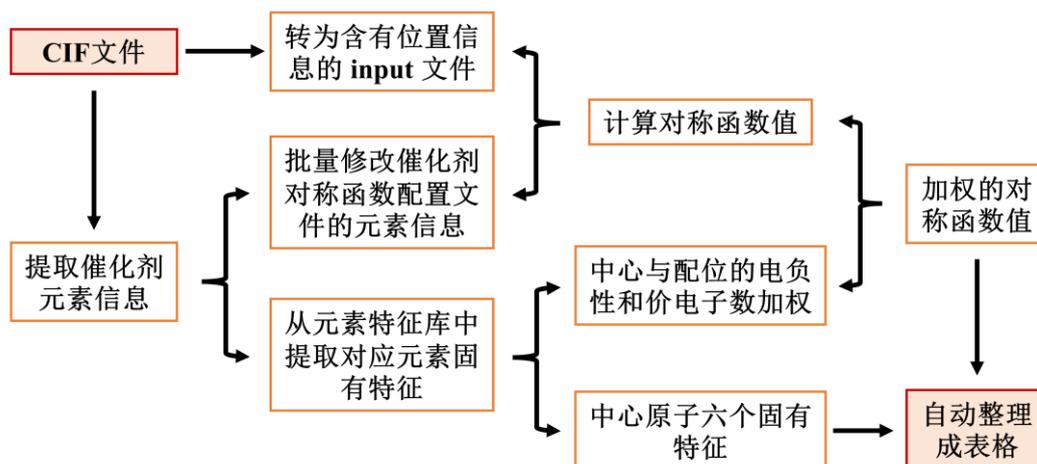


图 3-2 自动化特征提取流程图

3.4.2 工作流程的实现

通过标准化特征生成流程，为机器学习预测提供了高精度且可扩展的数据基础，这一方法同样适用于其他催化反应体系的特征工程。该流程通过局部原子环境描述符的特征生成器（FG-LAED）实现，这是一个用 Python 开发的工具包，它利用外部库 ACSF.so、cutoff.so 和 cif2input.so 来提取晶体学信息文件（CIF）中的结构信息并计算对称函数。在使用前，用户需将 CIF 文件放置于 input 目录中，并在 FG-LAED 提供的 config.data 文件中设置相关参数。此外，吸附位点信息应录入到 input.xlsx 电子表格中。此工具包集成了工作流程所需的所有文件和自动化特征提取代码，具体内容可在 GitHub 上查看，链接为：<https://github.com/Weijie-Yang/FG-LAED>。所有材料的 CIF 文件已上传至该链接。FG-LAED 能够为不同类型的催化剂定制专属特征，同时有效降低了成本。

3.4.3 机器学习数据集

最终，每种材料都生成了 134 个描述符，其符号及定义详见表 3-2。随后，将吸附自由能数据整合到特征集中，形成一个维度为 962×134 并关联两个目标向量（ ΔG_{OH^*} 和 ΔG_{O^*} ）的特征矩阵。该特征矩阵将用于后续的特征工程。本特征设计策略的关键优势在于这些描述符能适应由不同元素构成的多种材料。

表 3-2 描述符的单位及物理化学意义

| 符号 | 单位 | 描述符名称 | 物理化学意义 |
|--------------------|----|-----------------------|--------------------------|
| Nm | — | 电负性 | 反映中心原子吸引电子的能力 |
| V | — | 价电子数 | 原子最外层电子数 |
| Im | eV | 第一电离能 | 中性原子移除最外层电子所需的能量 |
| Am | eV | 电子亲和能 | 中性原子获得一个电子形成负离子的能量变化 |
| N | — | 原子序数 | 原子核中质子数，决定元素在周期表中的位置 |
| R | Å | 原子半径 | 从原子核到周围电子云边界的平均距离 |
| A ₁₋₁₆ | — | 不同参数组合下，电负性加权的径向对称函数 | 配位原子的电负性对中心原子径向环境电荷转移的影响 |
| A ₁₇₋₆₄ | — | 不同参数组合下，电负性加权的角向对称函数 | 配位原子的电负性对中心原子角向环境电荷转移的影响 |
| B ₁₋₁₆ | — | 不同参数组合下，价电子数加权的径向对称函数 | 配位原子价电子数对中心原子径向环境电子分布的影响 |
| B ₁₇₋₆₄ | — | 不同参数组合下，价电子数加权的角向对称函数 | 配位原子价电子数对中心原子角向环境电子分布的影响 |

3.5 本章小结

本章展示了催化剂筛选的通用设计框架，包括加权对称函数描述符（wACSF）的构建过程。首先计算出径向和角向对称函数值，并进一步引入配位原子的电负性（Nm）和价电子数（V）作为权重因子，以此构建两类加权对称函数描述符。wACSF 不仅能描述活性位点的几何构型，还能量化配位原子的电子效应，因此在多种催化体系中具有广泛的适用性，可推广至更多类型的催化剂研究领域。同时，构建了包括 962 种催化材料的吸附自由能及吸附位点的数据库，这些材料分为四大类：426 种金属间合金、160 种金属氧化物、250 种钙钛矿（ ABO_3 ）以及 126 种单原子催化剂。相关数据集可在以下链接获取：<https://github.com/Weijie-Yang/H2O2>。

在此基础上，开发了一个名为 FG-LAED 的自动化特征提取工具包（链接为：<https://github.com/Weijie-Yang/FG-LAED>）。该工具包可解析 CIF 文件中的催化剂位置和元素信息，计算不同类型材料的 wACSF 值，并提取六种原子中心的固有特征（电负性、原子序数、原子半径、第一电离能、电子亲和能和价电子数）。最终，这些特征和吸附自由能被自动整合，生成一个维度为 962×134 并包含两个目标向量（ ΔG_{OH^*} 和 ΔG_{O^*} ）的完整数据集。

本章的研究为不同催化体系的描述符构建提供了理论工具与方法支持，弥补了纯数据驱动模型在物理可解释性方面的不足，为催化剂的高效设计与筛选奠定了重要基础。

第 4 章 基于机器学习的催化剂快速预测

在第 3 章中，借助加权对称函数描述符以及自动化特征提取流程，构建了一个适用于机器学习建模的特征数据集。这一特征集为后续吸附能预测模型的训练提供了坚实的数据基础。然而，原始特征数量多达 134 维，存在部分信息冗余与相关性较强的问题，若直接用于建模可能会影响模型的泛化性能与预测效率。因此，有必要对特征进行有效筛选，并探索更加高效准确的建模策略，以实现催化剂吸附能的快速预测。

因此，本章围绕吸附能的高效预测问题展开深入研究。在第 3 章构建的特征数据集基础上，采用特征工程方法对原始特征进行降维与优化筛选，并用 XGBoost 算法进行建模。经过超参数调优与性能评估，建立了一个无需依赖 DFT 计算即可快速预测各类催化剂吸附能的机器学习模型。该模型显著提升了筛选效率，为实现大规模、数据驱动的 $2e^-$ WOR 催化剂筛选提供了坚实的技术支撑。

4.1 机器学习的特征工程

特征工程旨在通过优化输入特征集的数量与质量，从数百种材料属性中获取包含充足信息而又精简的特征子集，从而降低特征维度并防止模型过拟合。本研究的目标是从 134 个原始特征中筛选出信息密度高且冗余度低的特征子集，这一步骤在构建机器学习模型时至关重要，因为输入特征的选择决定了机器学习模型所能达到的最高精度。为此，设计了两个步骤的特征选择策略：第一步利用皮尔逊相关系数分析来剔除高度相关的冗余特征；第二步采用递归特征消除（RFE）算法从剩余特征中选取对吸附能预测贡献大的关键描述符，最终得到能准确表征材料特性的合适子集。

4.1.1 特征选择的步骤

数据预处理完成后，经过数据清洗和标准化，整合得到了 134 维特征的数据集（涵盖 6 个原子固有属性、64 个电负性加权对称函数以及 64 个价电子数加权对称函数）。高度相关的特征会降低模型的准确性和效率，为实现最佳预测效果，构建一个相关性较低的特征集十分关键。相关性被定义为两个可量化特征之间关系的强度和方向，而皮尔逊相关系数是衡量这一关系的常用指标^[90]。

最初，应用皮尔逊相关系数对原始描述符进行筛选，以排除高度相关的描述符。通过对阈值的分析调整，从 134 个特征中选取了相关性较低子集。具体来说，对于 ΔG_{OH^*} 预测模型，设定皮尔逊相关系数阈值为 0.05 后，最终保留

了 111 个特征；对于 ΔG_{O^*} 预测模型，设定 0.1 的阈值后选出了 114 个特征。一些近乎完全相关的描述符被移除，但特征数量仍然相对较多。

尽管有众多因素影响催化剂性能，但需合理控制特征的数量。RFE 算法以训练数据集为基础构建模型，依据权重逐步淘汰贡献较低的描述符，以生成新的子集。之后针对该子集再次重复上述操作，直至生成最优的特征集。这种逆向特征消除策略能够有效降低冗余数据对交叉验证评分的负面影响^[91]。

在完成相关性筛选后，采用 XGBoost 算法进行特征选择。通过 RFE 剔除冗余信息，从而避免过拟合并提升模型性能。该方法包括以下四个步骤：

(1) 利用相关性筛选后保留下来的 111 个（针对 ΔG_{OH^*} ）和 114 个（针对 ΔG_{O^*} ）特征来训练预测模型；

(2) 去除最不重要的特征，并用剩余特征重新训练模型；

(3) 重复步骤 (2) 直到特征数量达到理想目标。图 4-1 展示了 XGBoost 模型在不同特征数量下的预测性能变化，呈现先上升后下降的趋势。当特征数量分别达到 21（针对 ΔG_{OH^*} ）和 19（针对 ΔG_{O^*} ）个时，出现明显的转折点，表明此时的特征维度能有效确保 RFE 优化后模型的预测精度和效率。在此期间，共构造了 91 个（针对 ΔG_{OH^*} ）和 96 个（针对 ΔG_{O^*} ）不同特征组合的模型，特征数量范围分别为 21 至 111 个和 19 至 114 个；

(4) 最终选出在测试集上均方根误差（RMSE）最小且决定系数（ R^2 ）最高的模型，对应的特征即为最优描述符。

通过以上特征选择，依次为 ΔG_{OH^*} 和 ΔG_{O^*} 模型选出了 21 个（5 个原子固有属性和 16 个 wACSF 特征）和 19 个（3 个原子固有属性和 16 个 wACSF 特征）最优描述符。图 4-1 显示，进一步增加描述符数量并未带来模型性能的提升。相关特征信息可查阅特征集文件，链接为：<https://github.com/Weijie-Yang/H2O2>。

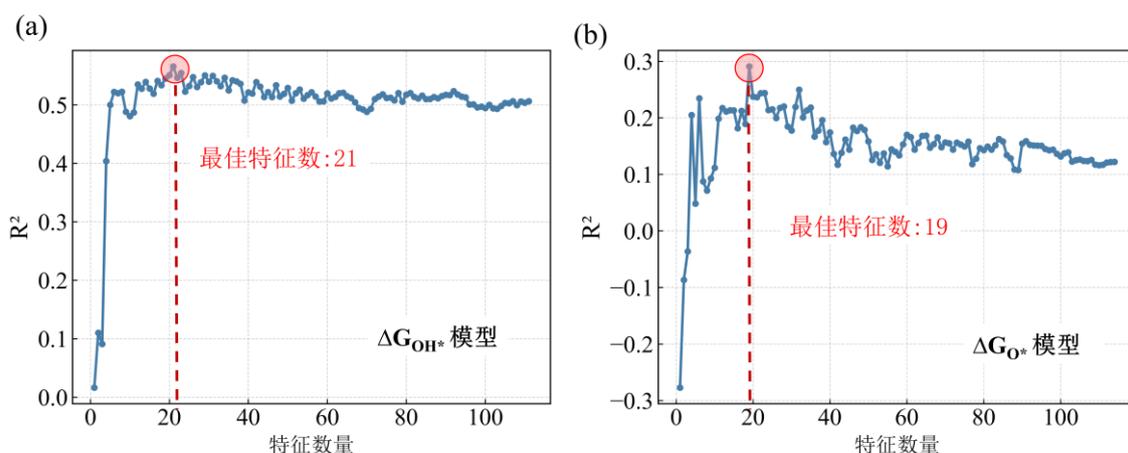


图 4-1 RFE 处理后不同特征数量下 RFR 模型的 R^2 值；a) ΔG_{OH^*} 模型中特征数量和预测性能关系图；b) ΔG_{O^*} 模型中特征数量和预测性能关系图

4.1.2 特征相关性分析

筛选出的描述符在机器学习训练中展现出很强的预测能力(测试集 $R^2 > 0.8$, 见表 4-1)。如图 4-2 所示, 皮尔逊相关系数热力图展示了特征间的相关性, 其程度与颜色深度成正比, 最深颜色对应的相关系数为 1, 表示完全相关(特征名称见表 3-2)。从图中可以发现部分 wACSF 特征组(如 A_1 与 A_{20})之间的相关性绝对值超过 0.8。这种较高的相关性可能源于两个主要因素: 首先, 对称函数的数学定义(公式 3-1 和 3-3)决定了其特征值受多参数的协同影响; 其次, 不同参数的内在联系会导致特征空间的聚类现象。例如 $A_{1\sim 16}$ 描述符表示在特定参数(表 3-1)组合下的 16 种电负性加权径向对称函数值, 其中的加权过程进一步加强了特征间的固有联系, 因此衍生出的特征之间出现了较强的相关性。

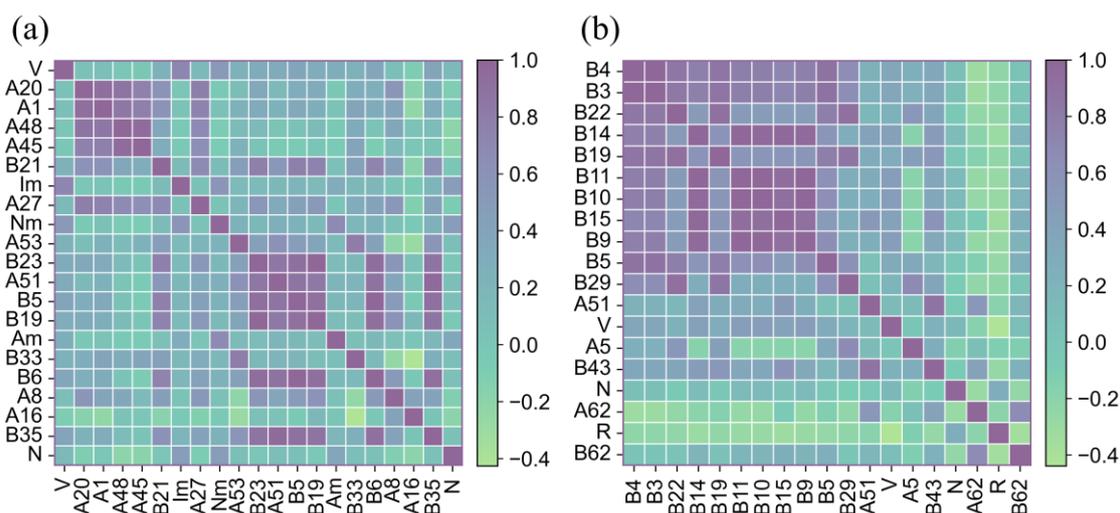


图 4-2 经特征工程筛选后的描述符的皮尔逊相关系数热力图; a) 预测 ΔG_{OH^*} 的描述符之间的皮尔逊相关系数热力图; b) 预测 ΔG_{O^*} 的描述符之间的皮尔逊相关系数热力图

考虑到高度相关的特征可能会增加模型复杂度并引发过拟合风险, 通过设定相关系数阈值, 剔除了皮尔逊相关系数大于 0.8 的特征, 并重新训练了模型。图 4-3 显示, 精简后的特征间相关性绝对值均低于 0.8。表 4-1 对比了特征删除前后的模型性能, 结果表明去除高度相关特征后, 模型的预测准确性有所下降 (R^2 降低约 0.11, RMSE 增加约 0.1)。这一现象表明, 尽管部分描述符之间具有较高的相关性, 但它们仍然能提供独特的信息, 在模型中具有不可替代的作用。对于 XGBoost 这样的非线性模型^[66], 其正则化项能有效处理多重共线性问题, 树状结构的特征划分方式能够充分挖掘相关特征间的有效信息。因此, 在电催化剂设计等对模型精度要求较高的应用场景中, 适当保留部分高相关的特征有助于提升模型预测能力, 避免过度简化特征空间。

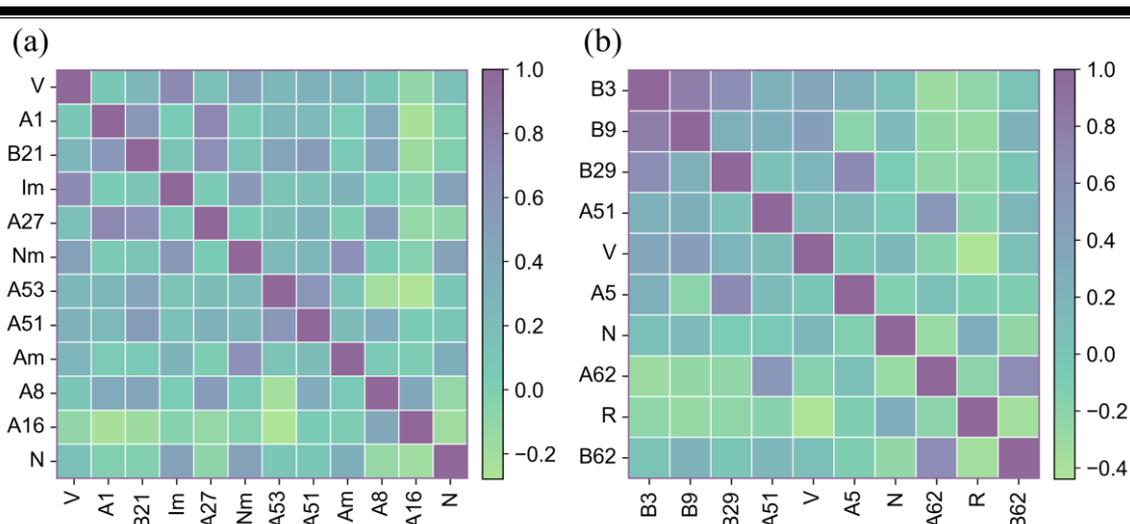


图 4-3 去掉高度相关特征后的皮尔逊相关系数热力图；a) 预测 ΔG_{OH^*} 的描述符之间的皮尔逊相关系数热力图；b) 预测 ΔG_{O^*} 的描述符之间的皮尔逊相关系数热力图

表 4-1 高度相关特征去除前后模型性能对比

| 类型 | $R^2 (\Delta G_{OH^*})$ | RMSE (eV) (ΔG_{OH^*}) | $R^2 (\Delta G_{O^*})$ | RMSE (eV) (ΔG_{O^*}) |
|-----|-------------------------|---------------------------------|------------------------|--------------------------------|
| 去除前 | 0.84 | 0.52 | 0.91 | 0.65 |
| 去除后 | 0.73 | 0.62 | 0.82 | 0.73 |

注： R^2 和RMSE为测试集的性能指标。

4.2 机器学习模型的构建与优化

采用特征选择后的 21 个 ΔG_{OH^*} 关键描述符和 19 个 ΔG_{O^*} 关键特征作为输入变量，将数据库中所有催化剂的 ΔG_{OH^*} 和 ΔG_{O^*} 数据分别作为目标变量，构建预测模型。数据集通过随机划分方式分为两部分，其中 80% 作为训练集用于优化模型参数，20% 作为测试集用于评估模型在未见过数据上的表现，以此验证模型的泛化能力。为防止数据分布不均导致性能偏差，在划分过程中设定随机种子数，并采用 K 折交叉验证，确保各类材料在训练集与测试集中的分布合理。

4.2.1 多种机器学习算法的对比

为选择最优预测模型，对八种机器学习算法进行了对比分析，并评估了这些算法训练模型的 R^2 和 RMSE。这些模型如图 4-4 所示，包括随机森林回归 (RFR)^[92]、核岭回归 (KRR)^[93]、极端梯度提升回归 (XGBoost 记作 XGBR)^[67]、岭回归 (RR)^[94]、k-近邻 (KNN)^[95]、梯度提升回归 (GBR)^[96]、支持向量回归 (SVR)^[97] 和极端随机树回归 (ETR)^[98]。所有模型均经过网格搜索超参数优化和 10 折交叉验证，以确保在模型比较时的准确性和可靠性。

图 4-4 显示，ETR、SVR、KNN 和 RR 训练的模型在预测性能上明显低于

RFR、XGBR、GBR 和 KRR。其中，XGBR 模型在吸附能预测中表现最佳。具体而言，该模型在训练集和测试集中具有最高的 R^2 分数 (ΔG_{OH^*} 为 0.93 和 0.84， ΔG_{O^*} 为 0.97 和 0.91) 和最低的 RMSE (ΔG_{OH^*} 为 0.32 和 0.52 eV， ΔG_{O^*} 为 0.36 和 0.65 eV)，这说明 XGBR 对数据集有良好的解释和预测能力。其次是 GBR 模型，在 ΔG_{OH^*} 模型中训练集和测试集的 R^2 分别为 0.90 和 0.80，RMSE 分别为 0.39 和 0.59 eV；在 ΔG_{O^*} 模型中训练集和测试集的 R^2 评分分别为 0.95 和 0.89，RMSE 为 0.47 和 0.73 eV。相较之下，RFR 和 KRR 模型的性能中等，其 R^2 分数 (0.72~0.90) 较低和 RMSE 相对较高 (0.55~0.94 eV)。值得注意的是，SVR 的表现最差，其 R^2 最低 (ΔG_{OH^*} 为 0.44 和 0.47， ΔG_{O^*} 为 0.79 和 0.81)，RMSE 最高 (ΔG_{OH^*} 为 0.94 和 0.82 eV， ΔG_{O^*} 为 0.90 和 0.98 eV)。

综合各模型的性能评估，优选 XGBR 为吸附自由能预测模型。尽管 GBR 等基于树的算法在部分任务中表现出相近的性能，但 XGBR 在测试集上的优势尤为明显 (如 ΔG_{O^*} 预测的 R^2 高达 0.91)。此外，选择依据不仅基于性能指标，还考虑模型复杂度、可解释性及计算效率等因素。相较于 GBR 等一阶梯度优化方法，XGBR 通过二阶梯度优化目标函数，具备更快的收敛速度，能更精准地捕捉数据中复杂的非线性关系。本研究的重点是构建不同类型催化剂的高精度预测模型，XGBR 因具备卓越的泛化能力和计算效率，被确定为最佳算法^[99,100]。

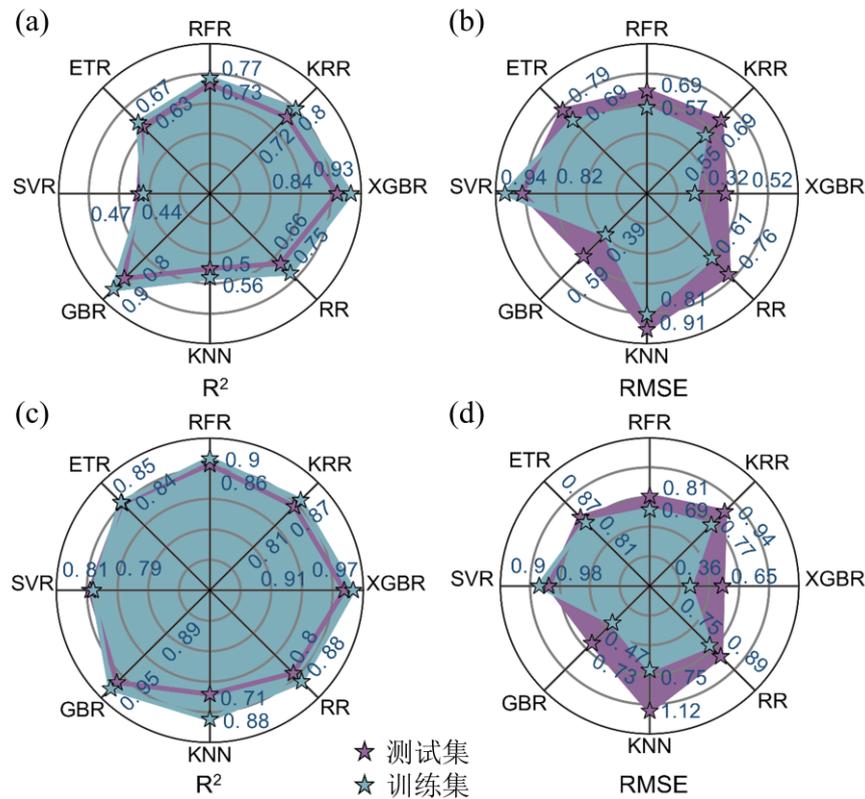


图 4-4 不同机器学习模型的吸附自由能预测性能比较； a) 预测 ΔG_{OH^*} 时的 R^2 得分； b) 预测 ΔG_{OH^*} 时的 RMSE； c) 预测 ΔG_{O^*} 时的 R^2 得分； d) 预测 ΔG_{O^*} 时的 RMSE

4.2.2 模型的超参数优化

超参数优化被认为是提升模型预测性能的一个重要环节。本研究采用网格搜索算法(Grid Search Algorithm)^[101]对 ΔG_{OH^*} 和 ΔG_{O^*} 预测模型进行超参数调优。通过全面搜索所有可能的参数组合, 寻找出最优参数配置, 从而提高模型的预测精度与稳定性。

在参数调优过程中, 使用了 Scikit-learn 库中的 GridSearchCV 函数进行 10 折交叉验证, 确保模型在不同数据集上的泛化能力, 降低过拟合风险。调参过程以最小化 RMSE 和最大化 R^2 为优化目标, 针对 XGBoost 回归器中的 10 项关键参数进行优化, 包括: 弱评估器数量(`n_estimators`)、复杂度的惩罚项(`gamma`)、树的最大深度 (`max_depth`)、叶子结点最小样本权重 (`min_child_weight`)、L2 正则化系数 (`reg_lambda`)、L1 正则化系数 (`reg_alpha`)、每次生成树时随机抽样特征的比例 (`colsample_bytree`)、每次生成树的一层时随机抽样特征比例 (`colsample_bylevel`)、样本有放回采样率(`subsample`)和学习率(`learning_rate`)。最终确定的最佳超参数组合如表 4-2 所示, 各超参数均经过精细调整, 在挖掘复杂数据特征和控制过拟合风险之间达到了最佳平衡, 从而为催化活性预测提供了可靠的模型基础。

表 4-2 XGBoost 模型的最终超参数表

| 参数名称 | ΔG_{OH^*} 模型的参数值 | ΔG_{O^*} 模型的参数值 |
|--------------------------------|--------------------------|-------------------------|
| <code>n_estimators</code> | 134 | 229 |
| <code>gamma</code> | 0.02 | 0.26 |
| <code>max_depth</code> | 6 | 8 |
| <code>min_child_weight</code> | 19 | 13 |
| <code>colsample_bytree</code> | 0.8 | 1 |
| <code>colsample_bylevel</code> | 0.9 | 1 |
| <code>subsample</code> | 0.7 | 0.8 |
| <code>reg_lambda</code> | 0.5 | 0.5 |
| <code>reg_alpha</code> | 0.03 | 0.01 |
| <code>learning_rate</code> | 0.13 | 0.05 |

4.2.3 截断半径对模型性能的优化

除了超参数调优外, 对称函数的截断半径 (R_c) (见公式 3-2) 的选择也会对机器学习模型的预测性能产生显著影响。 R_c 定义了中心原子邻近区域的边界范围, 并决定了局域化学环境特征提取的有效性。因此, 优化 R_c 参数是提升模

型性能的必要步骤^[84, 85]。在此，系统测试了不同的 R_c 值（取值范围为 4 Å 至 16 Å）对性能指标（ R^2 和 RMSE）的影响。

图 4-5 的结果表明，当 R_c 取 7 Å 时， ΔG_{OH^*} 模型在测试集上展现出最优性能（ R^2 为 0.84，RMSE 为 0.52 eV）；而 ΔG_{O^*} 模型的最佳 R_c 值为 5 Å（ R^2 为 0.91，RMSE 为 0.65 eV）。值得注意的是，当 R_c 超过临界值（ ΔG_{OH^*} 为 7 Å， ΔG_{O^*} 为 5 Å）时，模型性能趋于稳定，并未进一步提升。过大的 R_c 会引入冗余信息，从而增加计算复杂度及机器学习训练的难度。反之，当 R_c 过小时，重要的局部相互作用信息可能会丢失，例如当 R_c 低于 4 Å 时，对称函数计算结果大多为 0，导致输入特征缺乏有效信息。这一现象表明，选择合适的 R_c 值不仅能够节省计算资源和训练时间，还能在预测中保持较高准确性。

该结果进一步验证了吸附自由能预测精度受局部环境主导的观点，这与文献[63]的研究结论一致。该文献指出，采用局部描述符替代全局描述符时可显著提升模型预测的准确性。这源于吸附强度与局部原子配位结构、电子密度分布等微观特征紧密相关，而全局描述符往往难以精准捕捉这些关键细节。所以根据局部环境特征来构建描述符是提高吸附能预测模型精度的一个重要策略。此外， ΔG_{OH^*} 比 ΔG_{O^*} 需要的 R_c 更大，这可能是因为羟基（OH）作为吸附质具有较高的电子亲和性，其键合强度不仅受局域环境相互作用的影响，还受到长程电子相互作用的影响^[102]。

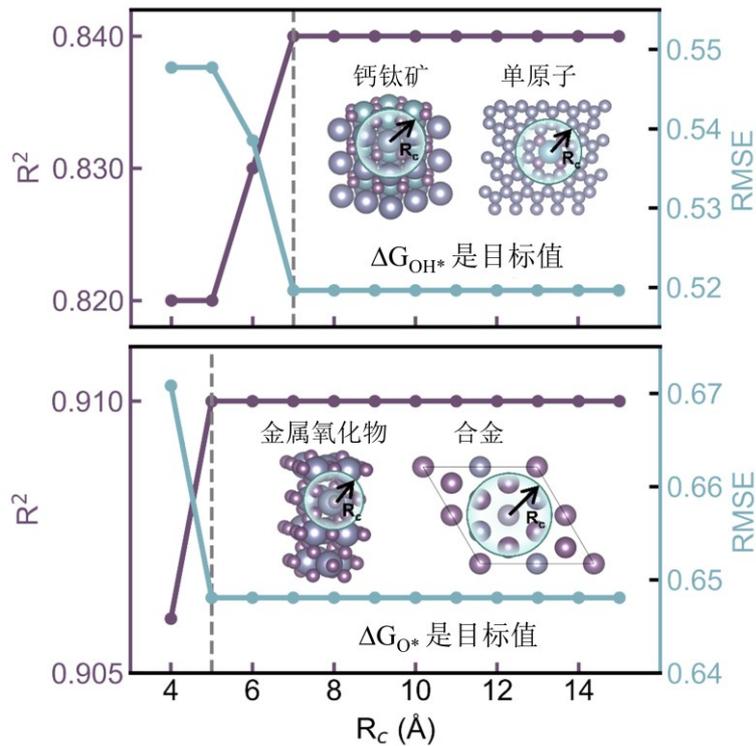


图 4-5 截断半径 R_c 对 ΔG_{OH^*} 与 ΔG_{O^*} 模型吸附自由能预测性能的影响

4.3 模型验证与可解释性分析

4.3.1 机器学习模型展示

通过调整模型超参数并优化截断半径，得到了性能最优的 XGBoost 模型。图 4-6 展示了最优 XGBoost 模型的吸附自由能预测值与 DFT 计算结果之间的拟合效果。图中蓝色和紫色数据点分别代表模型在训练集和测试集上的表现。在预测 ΔG_{OH^*} 时，该模型在训练集中的 R^2 为 0.93，RMSE 为 0.32 eV；在测试集中 R^2 为 0.84，RMSE 为 0.52 eV。预测 ΔG_{O^*} 时，模型在训练集上的 R^2 为 0.97，RMSE 为 0.36 eV；在测试集中 R^2 为 0.91，RMSE 为 0.65 eV。这些结果表明机器学习模型具备较强的泛化能力与预测性能，能够较为准确地预测吸附自由能。

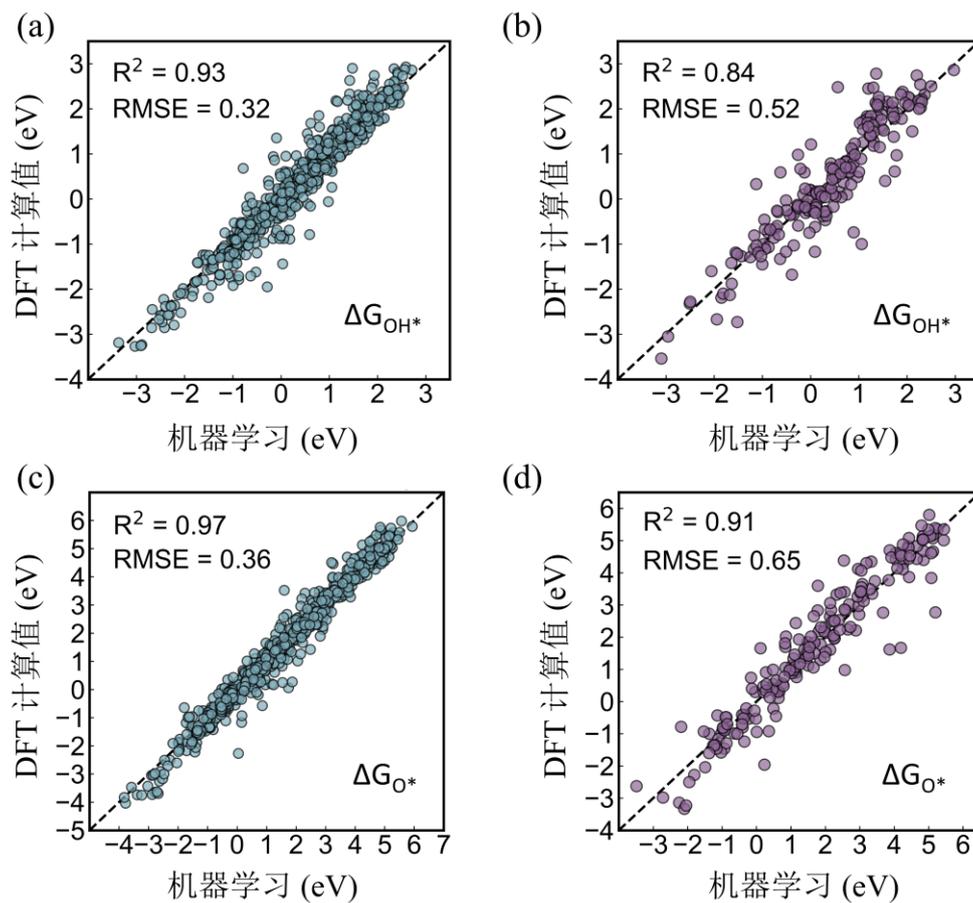


图 4-6 数据集催化剂吸附能预测模型表现：a) 数据集催化剂的 ΔG_{OH^*} 预测模型在训练集上的表现；b) 数据集催化剂的 ΔG_{OH^*} 预测模型在测试集上的表现；c) 数据集催化剂的 ΔG_{O^*} 预测模型在训练集上的表现；d) 数据集催化剂的 ΔG_{O^*} 预测模型在测试集上的表现

4.3.2 模型性能的对比较及分析

如表 4-3 所示，本研究构建的吸附自由能预测模型在多个方面展现出显著

优势。前人采用的表面中心-环境描述符机器学习模型中^[86]， ΔG_{OH^*} 和 ΔG_{O^*} 的 RMSE 分别为 0.54 和 0.93 eV， R^2 分别为 0.79 和 0.80；而本模型使用更简单的描述符却展现出更低的误差和更高的 R^2 ： ΔG_{OH^*} 和 ΔG_{O^*} 的 RMSE 分别为 0.52 和 0.65 eV， R^2 依次为 0.84 和 0.91。尽管与采用固有属性及电子/几何描述符构建的最佳模型（其中 RMSE 介于 0.10~0.34 eV）相比^[63, 87, 99, 103]，本模型在误差指标上略显不足，但两者 R^2 值差距均在 0.12 以内。在适用范围方面，与以上仅适用于特定材料体系的模型相比，本模型在保证可靠预测能力的同时，创建了适用于各种类型材料的通用描述符。此外，还通过简化的自动化流程，为多类型材料特征的构建与模型训练提供了有效的工具。

此外，对比发现，模型预测的准确性与采用三维坐标特征（RMSE 为 0.63 eV）和腔体特征（RMSE 为 0.56 eV）的方法相当^[104, 105]，但两者均需要大量密度泛函理论计算与复杂特征提取过程。相较之下，加权对称函数描述符更易获取且实用，为相关研究提供了高性价比的替代方案。

综上，所提出的基于通用描述符的机器学习模型能有效预测各类材料的吸附自由能，在大规模、多样化材料筛选任务中具有显著的应用优势。

表 4-3 不同研究预测性能的比较

| 描述符 | 目标 | 催化剂类型 | 模型 | R^2 | RMSE (eV) | Ref. |
|---------|---------------------------|----------------|-------|-------|-----------|-------|
| wACSF | ΔG_{OH^*} | 四类催化剂 | XGBR | 0.84 | 0.52 | 本研究 |
| wACSF | ΔG_{O^*} | 四类催化剂 | XGBR | 0.91 | 0.65 | 本研究 |
| 表面中心-环境 | ΔG_{OH^*} | ABO_3 | RFR | 0.79 | 0.54 | [86] |
| 表面中心-环境 | ΔG_{O^*} | ABO_3 | RFR | 0.80 | 0.93 | [86] |
| 金属及合金特征 | ΔG_{O^*} | 二元合金 | NN | — | 0.34 | [63] |
| 固有特征 | $E_{\text{ads,OH}}$ | 二维材料 | XGBR | 0.85 | 0.18 | [99] |
| 固有特征 | $E_{\text{ads,H}}$ | 二维材料 | XGBR | 0.89 | 0.1 | [99] |
| 三维坐标特征 | ΔG_{OOH^*} | SACs | MLP | — | 0.63 | [104] |
| 腔体特征 | E_{ads} | SACs | RFR | 0.93 | 0.56 | [105] |
| 固有特征 | ΔG_{OH^*} | SACs | XGBR | 0.96 | 0.29 | [87] |
| 固有特征 | ΔG_{H^*} | SACs | XGBR | 0.91 | 0.27 | [87] |
| 电子和几何特征 | 吸附焓 | 金属氧化物 | SISSO | — | 0.18 | [103] |

注： R^2 和 RMSE 为测试集的性能指标。

4.3.3 特征重要性分析及物理解释

为进一步理解这些描述符在预测过程中的作用机制，计算了 SHAP (SHapley Additive exPlanations) 值，从特征对输出结果的影响强度（即特征的

重要性排序)和特征对预测值的作用方向(正面或负面)两个方面来量化特征对目标值的影响^[106]。如图 4-7 a 和 c 所示,左侧的条形图展示了各特征的重要性排名,用于衡量它们在模型中的总体贡献。在图 4-7 b 和 d 中,横轴代表 SHAP 值,正值表示该特征对吸附自由能的正向贡献。每个实心圆点对应各特征的 SHAP 值,颜色梯度(从青色到紫色)表示特征值的相对大小,其中青色代表高值,紫色代表低值。

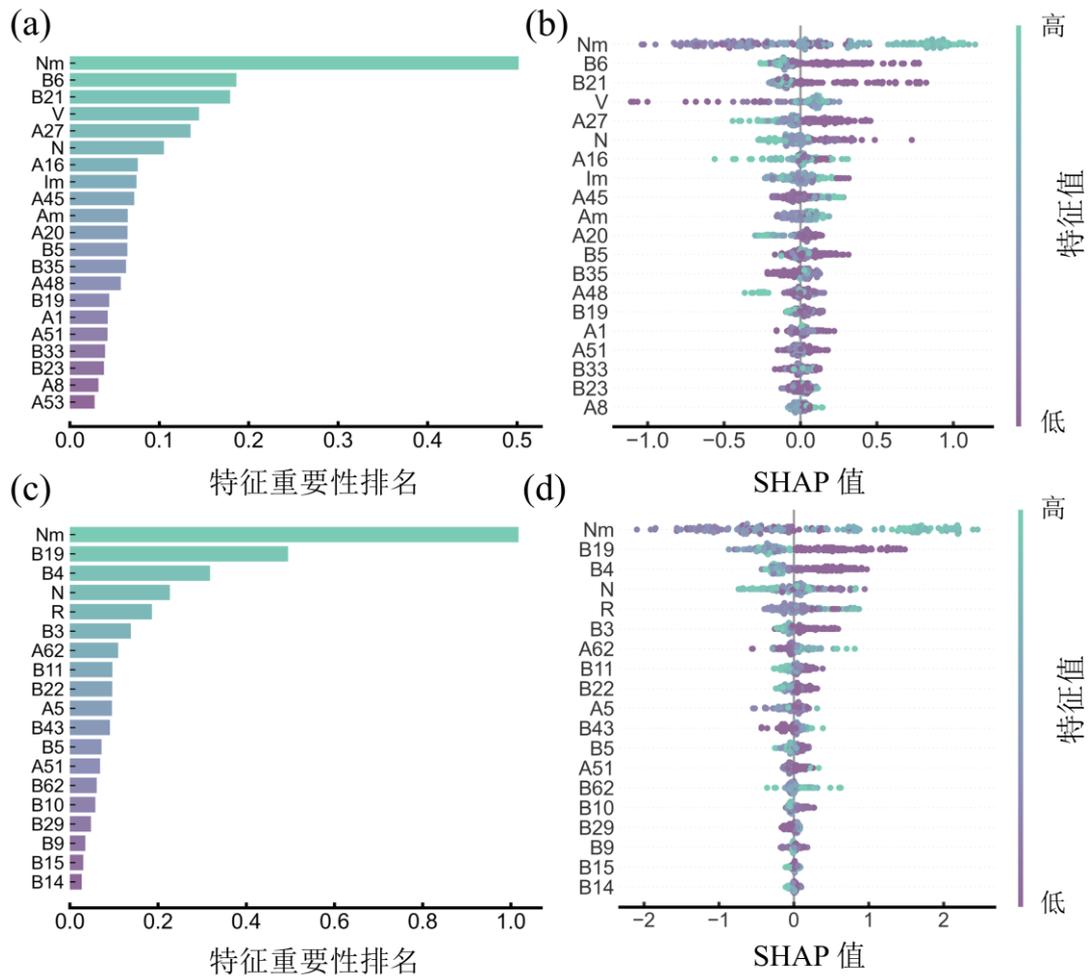


图 4-7 基于 XGBoost 模型的特征重要性和 SHAP 分析; a) ΔG_{OH^*} 模型基于平均 SHAP 值的特征重要性排序; b) ΔG_{OH^*} 模型的单个特征 SHAP 敏感性分析; c) ΔG_{O^*} 模型基于平均 SHAP 值的特征重要性排序; d) ΔG_{O^*} 模型的单个特征 SHAP 敏感性分析

SHAP 值分布越广,表明对应特征对目标吸附自由能越敏感。对于 ΔG_{OH^*} , SHAP 值分布最广的三个描述符依次为:中心原子电负性(Nm,重要性得分:0.50)、配位价电子数加权的径向对称函数(B₆, 0.20)以及配位价电子数加权的角对称函数(B₂₁, 0.18);对于 ΔG_{O^*} , SHAP 值分布最广的三个描述符依次为:中心原子电负性(Nm, 1.02)、配位价电子数加权的角对称函数(B₁₉, 0.50)以及配位价电子数加权的径向对称函数(B₄, 0.32)。这些特征的 SHAP 分布显著

宽于其他特征，说明它们对模型预测起主导作用。

N_m 的 SHAP 值分布最广且重要性得分最高，所以它对输出结果有显著影响。电负性反映了原子得失电子的能力，高电负性原子倾向于获得电子，而低电负性原子易失去电子。图 4-7 b 和 d 的颜色分布显示，较低的 N_m 值（紫色）对 ΔG_{OH^*} 和 ΔG_{O^*} 有负面影响，而较高的 N_m 值（青色）则有正面影响，这意味着电负性与吸附自由能呈正相关。即电负性越强，原子越倾向于吸附反应中间体，反映出电子得失能力对吸附质间的电子转移和相互作用具有显著的调控作用^[107]，表明了 ΔG_{OH^*} 和 ΔG_{O^*} 与材料原子的供/受电子特性紧密相关。这与 Zhu 等人^[108]的研究结论一致，即元素电负性差异会引发电荷重新分布，进而影响吸附能大小。

此外， B_6 、 B_4 、 B_{21} 和 B_{19} 等加权对称函数特征也表现出显著影响。其中， B_6 和 B_4 为配位价电子数加权的径向对称函数描述符， B_{21} 和 B_{19} 为配位价电子数加权的角对称函数描述符。图 4-7 b 和 d 的颜色分布表明，较低的 B_6 、 B_4 、 B_{21} 和 B_{19} 值（紫色）对 ΔG_{OH^*} 和 ΔG_{O^*} 有正向影响。这是因为较小的 B_6 、 B_4 、 B_{21} 和 B_{19} 值意味着中心原子周围的电子环境更均匀，与配位原子间的相互作用较弱，有利于反应物的吸附并使反应中间体更稳定，从而有更大的吸附能。这一现象的本质在于配位价电子数加权的径/角向对称函数描述符考虑了配位原子的电子贡献，反映了中心原子与配位原子在径/角向的相互作用强度和电子分布。这为描述吸附过程中的复杂相互作用提供了重要信息，是模拟原子局部化学环境的关键^[109]。

综上所述，所构建的模型不仅在性能上具备竞争力，其核心特征展现出良好的物理化学可解释性，验证了新型加权对称函数描述符在预测吸附行为及揭示构效关系中的实际价值。

4.4 本章小结

本章在前期构建的数据集特征的基础上，通过数据预处理、特征选择、模型构建和优化等步骤，建立了 XGBoost 的高效预测模型。首先采用皮尔逊相关系数分析结合 RFE 算法对 134 维特征进行选择，最终选出了 21 个和 19 个核心描述符分别用于预测 ΔG_{OH^*} 和 ΔG_{O^*} 模型。然后对八种机器学习算法进行对比，确定 XGBoost 为最优算法。经过超参数优化、截断半径优化和性能评估后，模型展现出优异的预测性能。最后通过 SHAP 分析深入理解了特征的重要性及其对吸附自由能的影响，突显了中心原子与配位原子之间的局部化学环境和电子分布的重要性。主要结论如下：

(1) XGBoost 模型能够准确预测 ΔG_{OH^*} 和 ΔG_{O^*} ，在测试集上 R^2 分别为

0.84 和 0.91, RMSE 分别为 0.52 eV 和 0.65 eV。相较于其他算法, 如 RFR、KRR 等, XGBoost 展现出显著优势, 验证了其在处理非线性数据与多重共线性方面的出色能力。

(2) 截断半径的选择对模型性能有显著影响, ΔG_{OH^*} 和 ΔG_{O^*} 模型分别在 R_c 为 7 Å 和 5 Å 时达到最佳预测效果, 表明局部环境特征在吸附自由能预测中起主导作用。

(3) SHAP 分析表明, N_m 与吸附能呈正相关, 反映了电子转移能力对吸附强度的调控作用。配位价电子数加权的径向/角对称函数 (如 B_6 、 B_{19}) 进一步揭示了局部化学环境与电子分布对吸附过程的贡献, 为设计高效催化剂提供了理论依据。

(4) 与现有研究相比, 本模型通过简化的通用描述符实现了对多种催化剂吸附能的准确预测, 其性能与依赖复杂特征提取的模型相当, 但计算成本显著降低, 为大规模催化剂筛选提供了高效工具。

第 5 章 催化剂的筛选与实验验证

第 4 章通过特征工程与超参数优化, 创建了高精度、高泛化性的吸附自由能预测模型, 为催化材料的筛选提供了可靠的理论支持。此外, 水氧化反应包含单电子、双电子及四电子三种竞争反应路径, 且所有反应的初始步骤均涉及 OH^* 的吸附过程。因此, 可根据数据集中的 ΔG_{OH^*} 数据对催化剂进行选择性筛选, 以找出能产 H_2O_2 的 2e^- WOR 催化剂。前人的研究表明, 在筛选高活性催化剂时, 微观动力学火山模型展现出显著优势, 但尚未实现在 2e^- WOR 反应中的应用。

本章对第 2 章方法部分构建的 2e^- WOR 微观动力火山图进行了验证, 结合选择性条件, 建立了催化剂筛选标准。利用数据集中的催化剂信息构建预测集, 借助机器学习模型快速预测了数据集外新构建的 3312 种合金、金属氧化物、钙钛矿和单原子催化剂的 ΔG_{OH^*} 和 ΔG_{O^*} 。结合筛选标准, 对数据集和预测集中的催化剂进行了快速筛选。筛选出的 12 种催化剂的预测值与 DFT 计算的结果偏差均小于 0.2 eV, 验证了机器学习模型的准确性。最后, 通过实验制备了预测性能最佳的 LiScO_2 催化剂, 结合表征及电化学测试, 该催化剂展现出高达 90% 的 H_2O_2 法拉第效率 ($\text{FE}_{\text{H}_2\text{O}_2}$), 并在长时间运行中保持良好的稳定性, 实验结果与理论预测高度吻合。

5.1 催化剂筛选标准

5.1.1 微观动力火山图的验证分析

为准确预测 2e^- WOR 催化剂的活性并筛选材料, 本研究开发了针对 2e^- WOR 反应的微观动力学模型。该模型通过电极电位与羟基吸附自由能 (ΔG_{OH^*}) 的函数关系来预测反应交换电流密度 (j)。 ΔG_{OH^*} 与 j 之间的关系可绘制成如图 5-1 虚线所示的催化活性火山曲线 (图中的电流密度 j 记为 J)。该火山图在评估 2e^- WOR 电催化剂的催化活性方面起着至关重要的作用。计算微观动力学火山图相关的反应步骤及方程详见公式 (2-1) 至 (2-17) (第二章的方法部分), 相关代码已集成至 CatMath 云端平台 <http://81.71.102.117:8080/>。

该火山图与已发表的 2e^- WOR 实验数据进行了基准比对, 这些实验数据以散点形式展示在图 5-1 中, 其中紫色圆圈代表金属氧化物、蓝色五角星代表钙钛矿、青色六边形代表金属掺杂的氧化物、蓝绿色菱形代表无金属有机网络。实验电流密度及计算极限电位详见表 5-1。结果显示, 模拟的微观动力学火山模

型与实验数据点高度吻合。其中, $\text{CaSnO}_3\text{@CF}$ ^[110]、 CuWO_4/Sn ^[111]、 LaAlO_3 ^[16]、 ZnO (1010)^[27]及 Gd 掺杂 BiVO_4 ($\text{Gd}:\text{BiVO}_4$)^[28]等催化剂在微观动力学火山模型(图 5-1)的峰值附近。因此, 这些催化剂被认为是性能最佳的 $2e^-$ WOR 电催化剂。除传统金属氧化物外, $\text{Pr}_{1.0}\text{Sr}_{1.0}\text{Fe}_{0.75}\text{Zn}_{0.25}\text{O}_{4-\delta}$ (D-PSFZ) 钙钛矿^[112]、共价三嗪骨架六氮杂苯无金属有机网络 (HAT-CTF)^[113]以及掺杂或改性的金属氧化物也表现出与文献报道一致的良好性能。以上结果展示了模型与各种材料实验结果的高度一致性, 进一步验证了微观动力学模型的准确性和适用性。

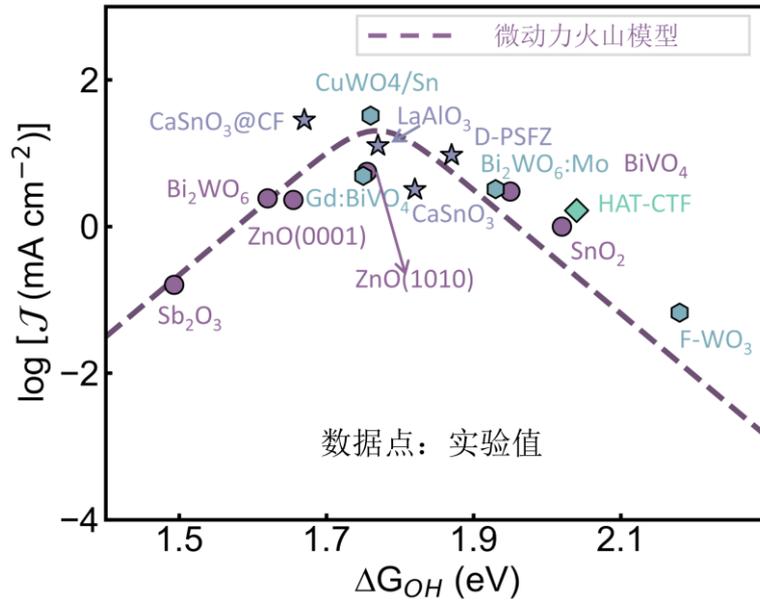


图 5-1 在电位为 $2.4 V_{\text{RHE}}$ 下 $2e^-$ WOR 微观动力学火山模型

表 5-1 文献报道的用于 $2e^-$ WOR 材料的实验电流密度和计算极限电位

| 化学式 | ΔG_{OH^*} (eV) | j (mA cm^{-2} @ V_{RHE}) | U_{LWOR} (V) | Ref. |
|---|-------------------------------|--|-----------------------|-------|
| ZnO (0001) | 1.66 | 2.3@2.4 | 1.87 | [114] |
| ZnO (1010) | 1.76 | 5.57@2.4 | 1.77 | [27] |
| BiVO_4 | 1.95 | 3@2.4 | 1.95 | [28] |
| SnO_2 | 2.02 | 1@2.4 | 2.02 | [25] |
| $\text{CaSnO}_3\text{@CF}$ | 1.67 | 22@2.4 | 1.85 | [110] |
| LaAlO_3 | 1.77 | 12.69@2.4 | 1.77 | [31] |
| CaSnO_3 | 1.82 | 3.2@2.4 | 1.82 | [29] |
| $\text{Pr}_{1.0}\text{Sr}_{1.0}\text{Fe}_{0.75}\text{Zn}_{0.25}\text{O}_{4-\delta}$ (D-PSFZ) | 1.87 | 9.4@2.4 | 1.87 | [112] |
| 6% Gd: BiVO_4 | 1.75 | 4.9@2.4 | 1.77 | [28] |
| $\text{Bi}_2\text{WO}_6:5\% \text{ Mo}$ | 1.93 | 3.23@2.4 | 1.93 | [115] |
| $\text{F}_{0.4}\text{-WO}_3$ | 2.18 | 0.0667@2.4 | 2.18 | [116] |

表 5-1 (续表)

| 化学式 | ΔG_{OH^*} (eV) | j (mA cm ⁻² @V _{RHE}) | U_{LWOR} (V) | Ref. |
|---------------------------------|------------------------|--|----------------|-------|
| CC/CuWO ₄ /Sn | 1.76 | 32.5@2.4 | 1.76 | [111] |
| HAT-CTF | 2.04 | 1.66@2.4 | 2.04 | [113] |
| Sb ₂ O ₃ | 1.49 | 0.16@2.4 | 2.03 | [26] |
| Bi ₂ WO ₆ | 1.62 | 2.42@2.4 | 1.9 | [115] |

图 5-2 展示了微观动力学-热力学双火山模型。图中的青色实线表示由极限电位确定的热力学活性火山，计算公式见式 (2-18) 至 (2-20) (第二章的方法部分) [13]。散点是根据催化剂的 ΔG_{OH^*} 通过微观动力学计算公式得到的，其中圆圈代表计算的电流密度，菱形代表热力学极限电位 (见表 5-1)。研究发现，两类火山图模型都呈现出相似的活性趋势。在 2.4 V_{RHE}下，热力学火山峰值位于 $\Delta G_{OH^*}=1.76$ eV，微观动力学火山峰值为 1.77 eV。此外，热力学火山左侧的决速步骤是OH^{*}氧化生成H₂O₂，右侧为H₂O活化生成OH^{*}；而微动力学火山的决速步骤相反，左侧为H₂O活化生成OH^{*}，右侧为OH^{*}氧化成H₂O₂。

以上表明，简化的极限电位模型虽能展现 2e⁻ WOR催化活性的整体趋势，但微观动力学模型通过量化反应路径中各步骤的动力学参数(如电子转移速率、活化能等)，能更精确地描述各类催化剂在不同条件下的速率控制步骤，进而更准确地预测催化剂的活性。因此，在复杂催化体系中，微观动力学模型为催化剂设计和预测提供了可靠的理论工具，而单一的热力学模型只能用于趋势性评估。

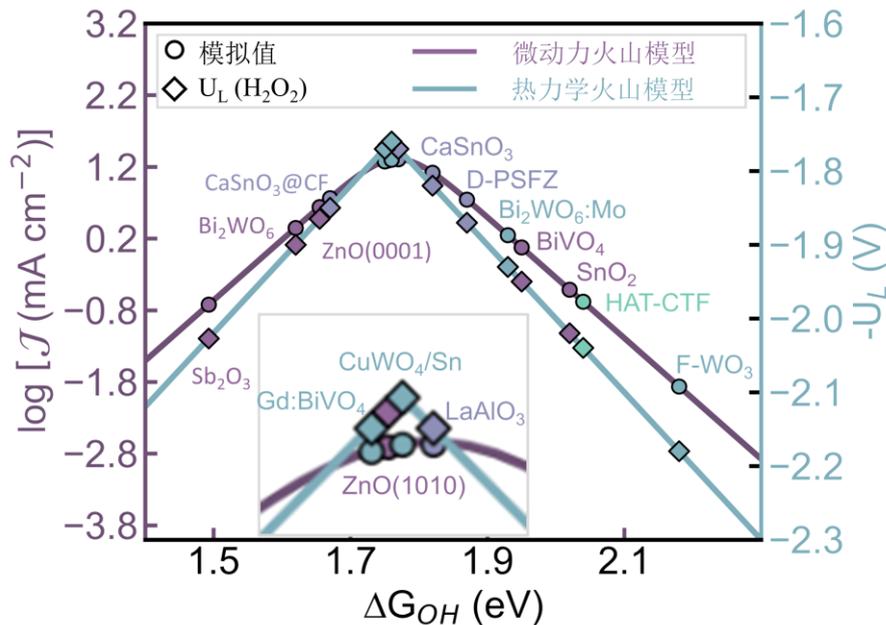


图 5-2 微观动力学和热力学火山模型

5.1.2 2e⁻ WOR 催化剂的选择性条件

在 2e⁻ WOR 催化剂筛选中,除活性外,催化剂的选择性也是核心指标之一。已有研究表明,各种反应中间体(如 O*和 OH*)的吸附自由能是评估 2e⁻ WOR 活性和选择性的有效描述符^[13]。因此,理想的 2e⁻ WOR 产 H₂O₂ 催化剂的中间体吸附能应处于合适范围。根据反应热力学原理推导出吸附能选择性筛选范围。

首先,为选择性生成 H₂O₂,催化剂表面的 OH*自由能需低于 OH·的自由能(即 $\Delta G_{OH^*} < 2.4 \text{ eV}$),以避免 OH·自由基的形成,这也为 ΔG_{OH^*} 设定了上限。换句话说,为避免单电子氧化并实现双电子氧化, ΔG_{OH^*} 必须小于 2.4 eV。此外,由于 H₂O₂ 的生成自由能($\Delta G_{H_2O_2}$)为 3.52 eV,对 O*吸附强度较弱($\Delta G_{O^*} \geq 3.52 \text{ eV}$)的催化剂有利于 H₂O₂ 的形成,即 2e⁻ WOR;而吸附能力过强($\Delta G_{O^*} < 3.52 \text{ eV}$)时,O*中间体难以脱附,则更倾向于 OER 反应路径,即 4e⁻ WOR。

图 5-3 展示了数据集中 962 种材料的 O*与 OH*吸附自由能的比例关系,即 $\Delta G_{O^*} = 2\Delta G_{OH^*} + 0.71$ 。其中比例线斜率约为 2,与现有研究结果一致^[117, 118]。根据这一比例关系,进一步推导出 ΔG_{OH^*} 的下限为 1.41 eV。具体推导过程如下:由 $\Delta G_{O^*} \geq 3.52 \text{ eV}$,结合比例关系式,可得 $2\Delta G_{OH^*} + 0.71 \geq 3.52 \text{ eV}$,进而得出 $\Delta G_{OH^*} \geq (3.52-0.71)/2=1.41 \text{ eV}$ 。

综上所述,2e⁻ WOR 反应高选择性催化剂需同时满足以下条件: $\Delta G_{O^*} \geq 3.52 \text{ eV}$,且 $1.41 \text{ eV} \leq \Delta G_{OH^*} < 2.4 \text{ eV}$ 。随后,结合微观动力学火山图,将这些条件应用于本文的数据集和预测集,进行初步筛选。

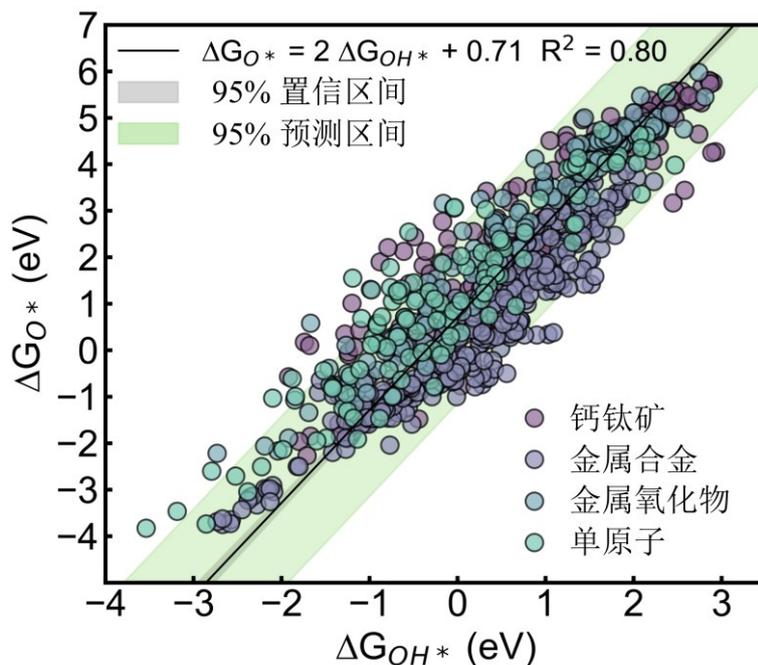


图 5-3 962 种催化剂表面上 ΔG_{OH^*} 和 ΔG_{O^*} 之间的线性关系

5.2 高活性和高选择性催化剂的筛选

5.2.1 数据集的筛选与结果分析

根据吸附能筛选范围和微观动力活性火山图,数据集筛选流程如图 5-4 a 所示,图中显示出每步筛选后的催化剂数量。最初,从 962 种催化剂中,依据 ΔG_{OH^*} (1.41~2.4 eV) 与 ΔG_{O^*} (≥ 3.52 eV) 的热力学条件进行初步筛选,得到 145 种符合 $2e^-$ WOR 路径的候选催化剂。这些催化剂的 OH^* 吸附能处于 1.41 至 2.4 eV 的最佳范围内,满足高选择性产 H_2O_2 的条件。接着,在图 5-5 的微观动力火山图中,对这 145 种催化剂进行更严格的筛选,即要求 $\log(J) \geq 0 \text{ mA cm}^{-2}$,选取图中虚线框内位于火山顶峰附近的催化剂,最终筛选出 49 种。以上筛选出的催化剂的数据详见 <https://github.com/Weijie-Yang/H2O2> 中名为“预测集和筛选出的催化剂”的文件。

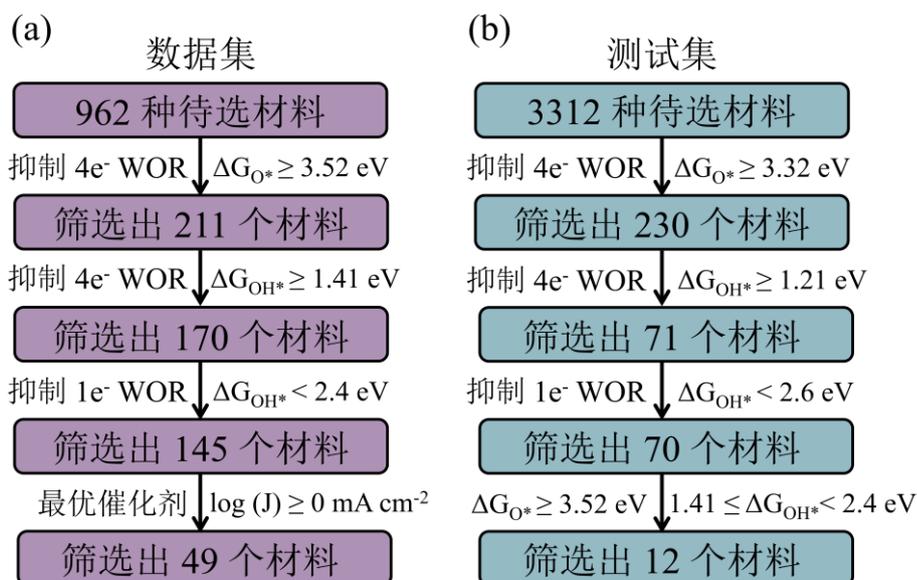
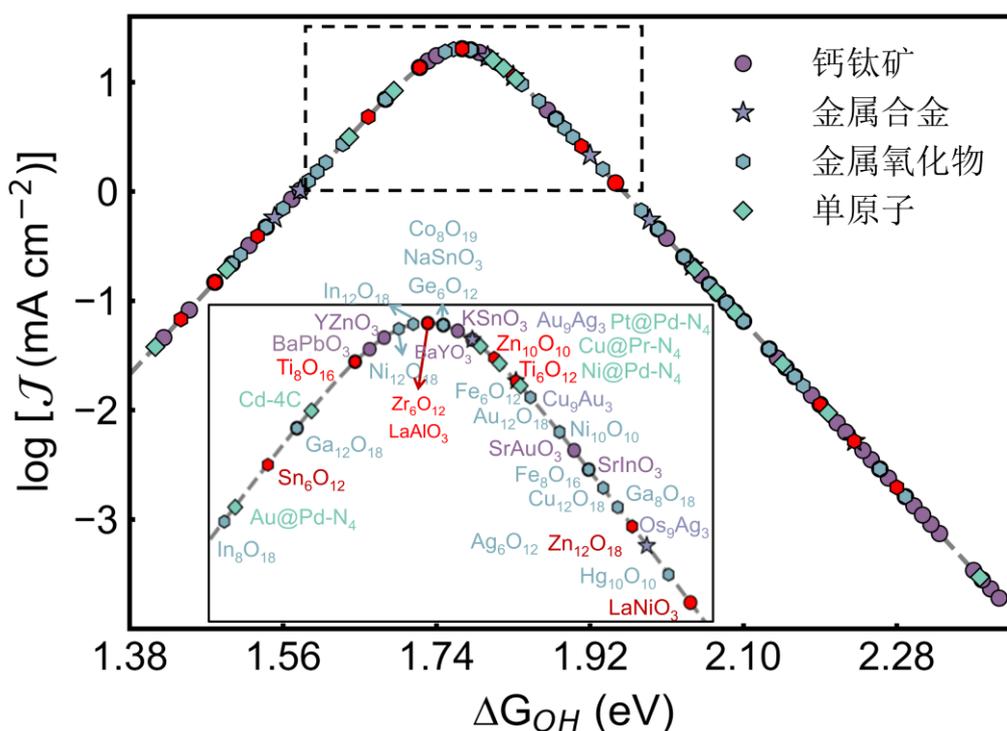


图 5-4 催化剂的筛选流程图: a) 数据集筛选流程图; b) 预测集的筛选流程图

经过进一步的文献调研,图 5-5 中在火山顶部附近标红的催化剂,如 LaAlO_3 ^[31]、 $\text{Zn}_{10}\text{O}_{10}$ ^[27]和 Ti_6O_{12} ^[23]等,已在实验中合成并表现出优异的 $2e^-$ WOR 活性。这表明,基于吸附能和微观动力学火山图的筛选策略具有较高的可靠性和预测能力,能有效筛选出具备实际催化活性的候选材料,为高效 $2e^-$ WOR 催化剂的开发提供了有力的理论支持。

图 5-5 149 种催化剂在微观动力学火山图中的 $2e^-$ WOR 活性分布

5.2.2 机器学习驱动的预测集生成

利用第 4 章建立的机器学习模型,可快速预测催化剂的吸附自由能 (ΔG_{OH^*} 和 ΔG_{O^*})。结合吸附能的筛选条件和微观动力火山模型,就能筛选出高活性催化剂材料。因此,借助构建的机器学习模型预测数据集之外催化剂的吸附能,可实现潜在高活性催化剂的快速筛选。

本文利用数据集的四类催化剂构建预测集。在金属间合金中,不同金属种类及比例会改变电子分布,进而影响反应物的吸附;对于金属氧化物,掺杂不同金属可调整其表面化学性质,增强对反应中间体的吸附能力;在钙钛矿结构中,替换 A 位和 B 位离子可改变晶体的稳定性和电子传导性;单原子催化剂中,中心和配位原子的种类会影响催化位点的电子结构。综上,通过调控这些因素,可改变催化剂与反应中间体的相互作用,从而影响催化活性。

结合以上讨论,通过调整金属间合金中的元素比例,以及改变金属氧化物、钙钛矿和单原子催化剂中的中心原子和配位原子的种类和组合方式,本文整合了 3312 种新结构预测集,其中包括:2381 种金属间合金、311 种金属氧化物、570 种钙钛矿和 50 种单原子催化剂。这些新结构丰富了数据库,对筛选高活性 $2e^-$ WOR 材料具有重要意义。最后借助自动化特征提取工具构建特征后,利用机器学习模型高效预测了 3312 种催化剂的 ΔG_{OH^*} 和 ΔG_{O^*} 吸附自由能。

5.2.3 预测集的筛选与结果分析

依据选择性和活性条件，对 3312 种催化剂进行筛选。为提高筛选的准确性，针对热力学筛选条件，考虑了 0.20 eV 的预测误差范围^[34]，以降低模型预测的不确定性。因此，预测集的选择性条件为： $1.21 \text{ eV} \leq \Delta G_{\text{OH}}^* < 2.6 \text{ eV}$ 且 $\Delta G_{\text{O}}^* \geq 3.32 \text{ eV}$ 。据此筛选出 70 种符合条件的催化剂。随后，采用更严格的筛选标准（ $1.41 \text{ eV} \leq \Delta G_{\text{OH}}^* < 2.4 \text{ eV}$ 且 $\Delta G_{\text{O}}^* \geq 3.52 \text{ eV}$ ），其中有 12 种催化剂满足该标准。筛选流程如图 5-4 b 所示。这些候选材料具有通过 2e^- WOR 产 H_2O_2 的潜力。其中每步筛选出的催化剂的具体数据可在 <https://github.com/Weijie-Yang/H2O2> 中查看。

此外，考虑预测误差后，选择性条件的筛选范围有所扩大，如图 5-6 所示，青色虚线选择的区域包括金属氧化物、钙钛矿和合金。这使得能够识别出更多类型和数量的具有实际潜力的候选材料。采用更严格的筛选标准后，符合条件的催化剂只有金属氧化物和钙钛矿。图中红色圆点表示的 Zn_8O_{18} ^[89] 和 Al_2O_3 ^[114] 已通过实验证实具有 2e^- WOR 催化活性。这表明筛选标准能有效地筛选出具有实际催化活性的材料。

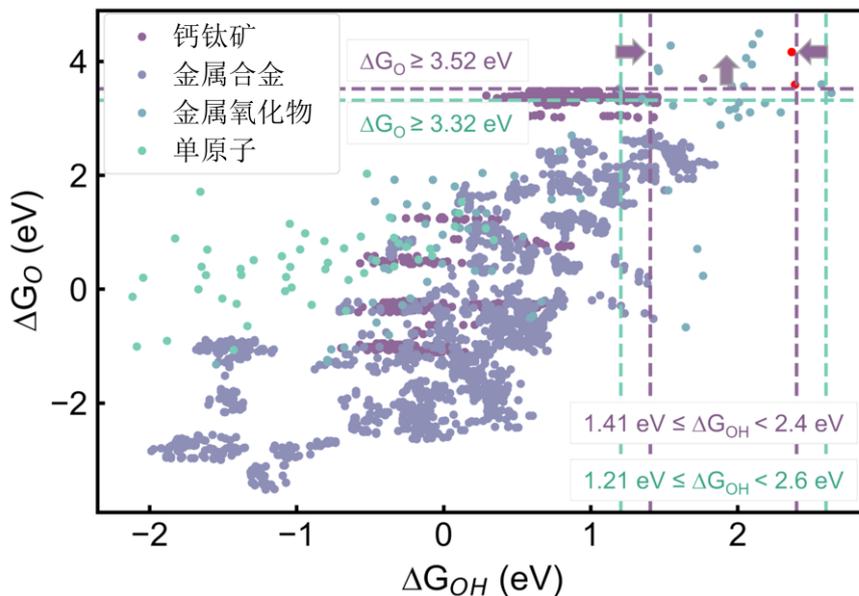


图 5-6 3312 种催化剂的选择性分布图

最后，利用微观动力学模型对筛选出的 12 种候选材料进行活性评估，以确定最优的高活性催化剂。如图 5-7 所示， LiScO_2 表现出最高的活性，在火山图中位于顶部区域，表明其具有卓越的催化活性和选择性。此外，图中标识的如 Y_6O_{12} 、 AgPtO_2 、 TbAlO_3 等催化剂，也展现出良好的 2e^- WOR 催化活性，为进一步的应用研究提供了有价值的参考。

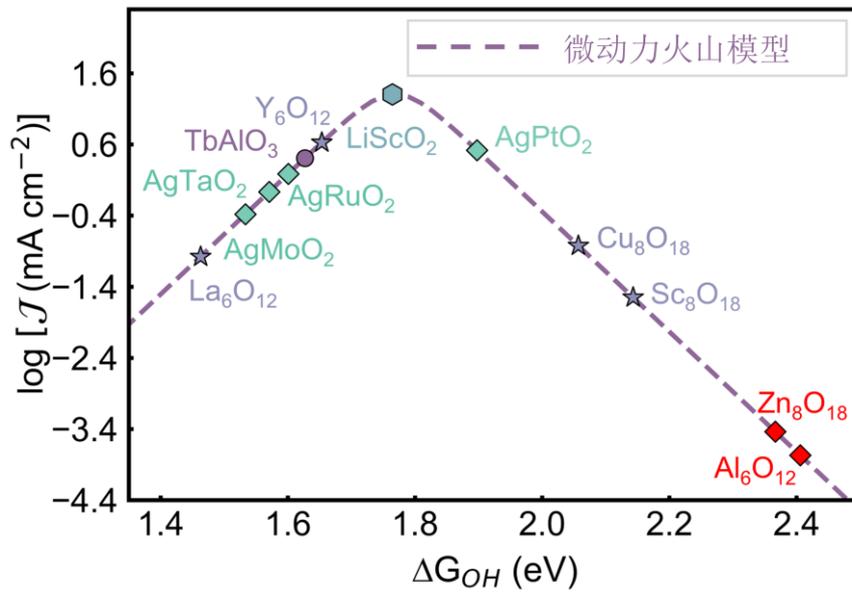


图 5-7 12 种催化剂在微观动力学火山图中的 $2e^-$ WOR 活性分布

5.3 筛选结果的理论验证

5.3.1 预测集 12 种催化剂 DFT 验证

采用 DFT 计算了预测集筛选出的 12 种催化剂的吸附自由能，并与预测结果进行对比分析，以验证预测的准确性。图 5-8 展示了预测的 ΔG_{OH^*} 和 ΔG_{O^*} 值与 DFT 计算结果之间的差值分布。结果显示，这些偏差均在 0.2 eV 以内，最大偏差约为 0.17 eV，标准差为 0.1 eV 左右，表明预测值与 DFT 计算值具有高度一致性。这种一致性证明了机器学习模型能够合理且准确地预测不同体系的吸附自由能，说明筛选结果具备高度可靠性。

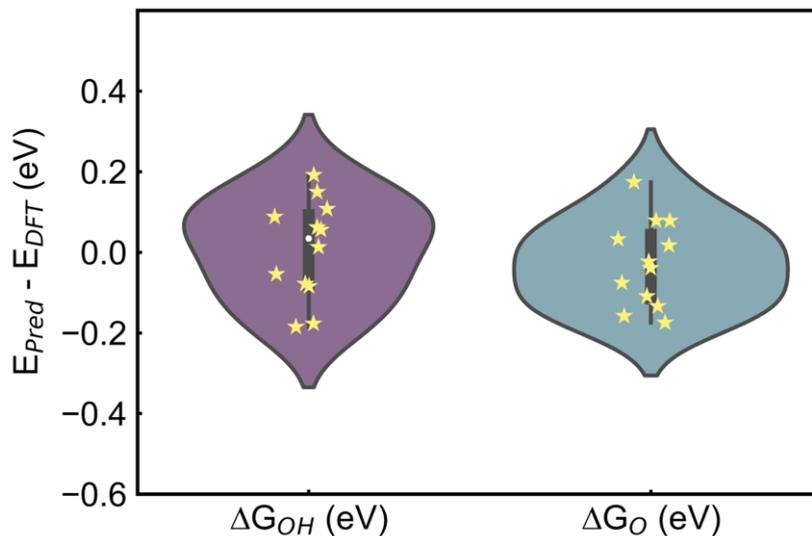


图 5-8 12 种催化剂吸附自由能预测值与 DFT 计算值的差值分布小提琴图

5.3.2 最佳催化剂的性能对比

筛选完成后，将数据集和预测集中前 16 种高性能催化剂的过电位与文献中的最佳值进行了比较。如表 5-2 所示，除合金外，这些催化剂均为非贵金属催化剂。其中，值得关注的催化剂有 LiScO_2 、 $\text{In}_{12}\text{O}_{18}$ 、 $\text{Ni}_{12}\text{O}_{18}$ 、 Zr_6O_{12} 、 NaSnO_3 和 YZnO_3 等，它们的过电位 (η_{WOR}) 均接近 0 V，催化活性和选择性与文献中报道的 $\text{CuWO}_4/\text{Sn}^{[111]}$ 、 $\text{ZnO}^{[27]}$ 、 $\text{LaAlO}_3^{[31]}$ 和 $\text{Gd:BiVO}_4^{[28]}$ 等最优催化剂相当。尽管文献中报道的材料展现出了出色的 $2e^-$ WOR 活性，但其过电位略高于本研究筛选出的部分有潜力材料。这些结果充分验证了本研究筛选方法的有效性，能够筛选出在催化活性和选择性方面表现优异的新型催化剂，为实际应用中的催化剂开发提供了有力的理论支持。

此外，如表 5-2 所示，四种材料均列其中，显示出卓越的催化活性和选择性。该结果不仅展示了金属氧化物与钙钛矿体系在 $2e^-$ WOR 反应中已被广泛认可的优势，还强调了 SACs 和合金作为高效 $2e^-$ WOR 催化剂的潜力，为材料探索和实验研究提供了新的视角和参考。同时，多轮筛选验证结果表明了该筛选方法的有效性和预测模型的准确性，具有重要的实用价值。

表 5-2 本研究最佳催化剂的理论过电位与其他研究结果的对比

| 化学式 | ΔG_{OH^*} (eV) | η_{WOR} (V) | 化学式 | ΔG_{OH^*} (eV) | η_{WOR} (V) | Ref. |
|-------------------------------|-------------------------------|-------------------------|------------------------------------|-------------------------------|-------------------------|-------|
| $\text{In}_{12}\text{O}_{18}$ | 1.76 | 0 | CuWO_4/Sn | 1.76 | 0 | [111] |
| LiScO_2 | 1.765 | 0.005 | ZnO (1010) | 1.755 | 0.005 | [27] |
| $\text{Ni}_{12}\text{O}_{18}$ | 1.75 | 0.01 | Gd:BiVO_4 | 1.75 | 0.01 | [28] |
| Zr_6O_{12} | 1.77 | 0.01 | LaAlO_3 (100)(110) | 1.77 | 0.01 | [31] |
| NaSnO_3 | 1.78 | 0.02 | CaSnO_3 (001) | 1.82 | 0.06 | [29] |
| YZnO_3 | 1.78 | 0.02 | $\text{CaSnO}_3@CF$ | 1.67 | 0.09 | [110] |
| Co_8O_{19} | 1.78 | 0.02 | WO_3 (100) | 1.66 | 0.1 | [116] |
| Ge_6O_{12} | 1.78 | 0.02 | ZnO (0001) | 1.655 | 0.105 | [114] |
| BaPbO_3 | 1.73 | 0.03 | D-PSFZ | 1.87 | 0.11 | [112] |
| KSnO_3 | 1.79 | 0.03 | Bi_2WO_6 | 1.62 | 0.14 | [115] |
| Cu@Pr-N_4 | 1.72 | 0.04 | $\text{Bi}_2\text{WO}_6:\text{Mo}$ | 1.93 | 0.17 | [115] |
| Au_9Ag_3 | 1.8 | 0.04 | BiVO_4 (111) | 1.95 | 0.19 | [28] |

表 5-2 (续表)

| 化学式 | ΔG_{OH^*} (eV) | η_{WOR} (V) | 化学式 | ΔG_{OH^*} (eV) | η_{WOR} (V) | Ref. |
|---------------------------------|------------------------|------------------|--|------------------------|------------------|-------|
| Cd-4C | 1.69 | 0.07 | LaGaO3(100) | 1.99 | 0.23 | [31] |
| Cu ₉ Au ₃ | 1.83 | 0.07 | SnO ₂ (110) | 2.02 | 0.26 | [25] |
| Fe ₆ O ₁₂ | 1.83 | 0.07 | Sb ₂ O ₃ (Cubic) | 1.493 | 0.267 | [26] |
| Ni@Pd-N ₄ | 1.85 | 0.09 | HAT-CTF | 2.04 | 0.28 | [113] |

注：表格左侧的数据来源于本研究，右侧的数据则来源于已发表的文献。

5.4 实验验证预测的高性能催化剂

5.4.1 LiScO₂ 的制备及表征

为验证理论预测，选取预测集中位于火山图顶端、活性最高的 LiScO₂ 作为目标化合物。该材料采用固相烧结法制备（详见第二章的实验方法）^[82]。经电感耦合等离子体原子发射光谱（ICP-AES）测定，证实产物的化学计量比符合理论值。图 5-9 的 X 射线衍射（XRD）图谱分析表明，样品呈四方晶系，空间群为 I41/amd，与标准卡片（PDF#76-1701）匹配度良好。

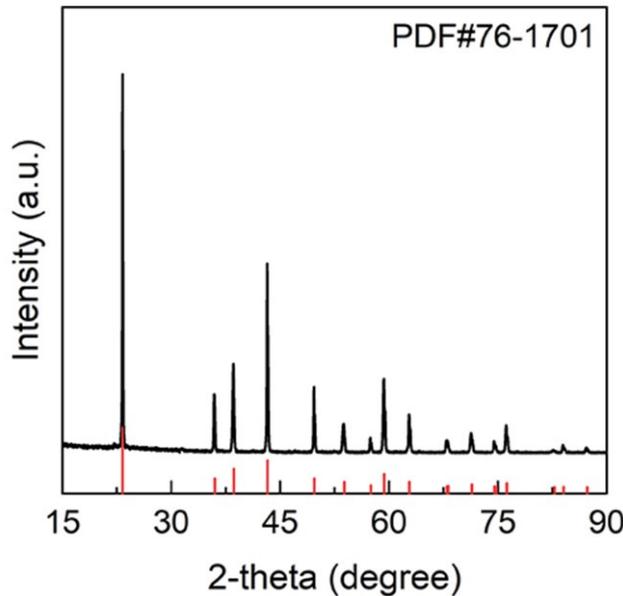


图 5-9 X 射线衍射（XRD）图谱

在高分辨率图像中可以清晰地观察到晶格条纹，这些条纹延伸至晶体内部，如图 5-10 a 所示。低倍透射电子显微镜（TEM）表征如图 5-10 b 所示，材料由纳米颗粒团聚形成。图 5-10 c 为图 5-10 a 中红色矩形区域的进一步放大图，其中晶格条纹对应于 LiScO₂ 的（101）晶面，晶面间距为 0.382 nm。此外，还可

观察到一个厚度约 1 nm 的低结晶度表面层（红色虚线标记），该层可归因于表面钝化。后续通过表面敏感的 X 射线光电子能谱（XPS）测量得到了进一步证实。图 5-10 d 的能量色散 X 射线（EDX）元素映射结果显示，样品中 Sc 和 O 元素分布均匀；由于 Li 难以被 EDX 准确检测，其分布未显示。

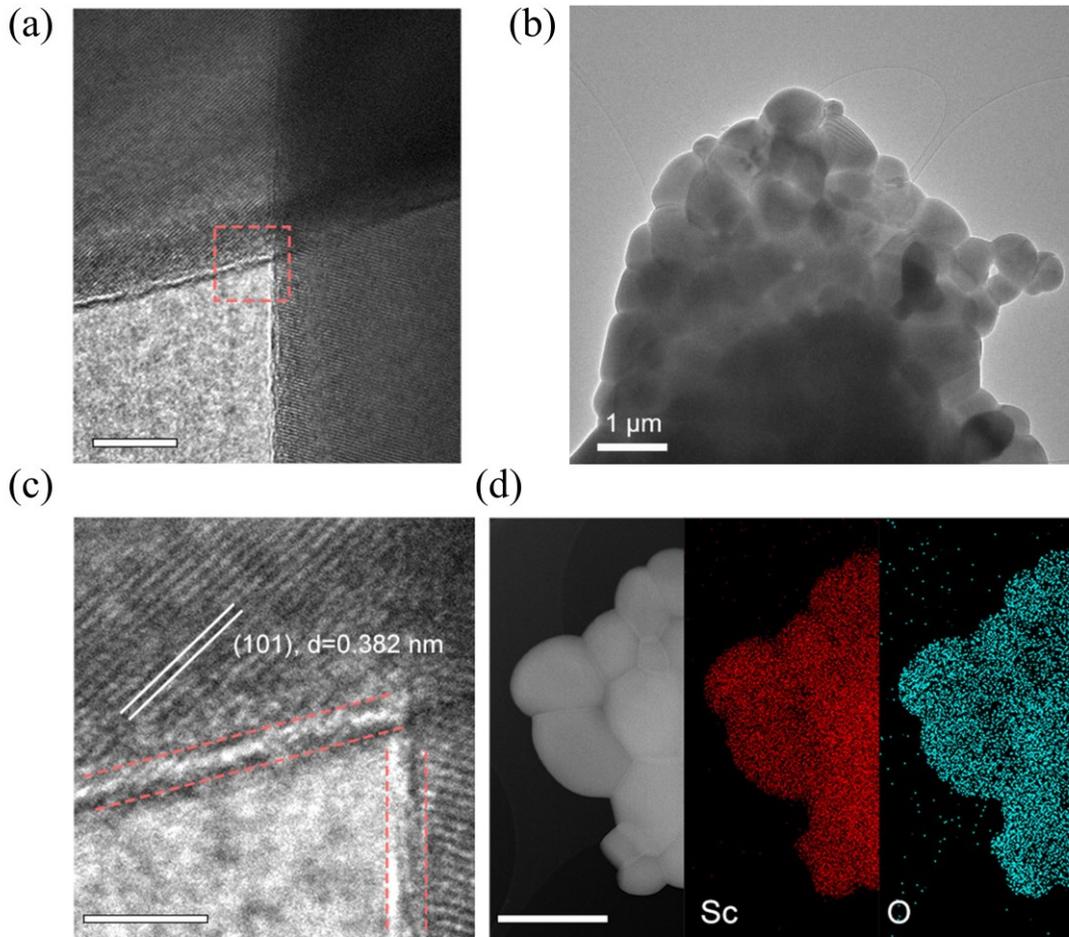


图 5-10 LiScO₂ 催化剂的结构表征；a) 催化剂表面附近的高分辨率透射电子显微镜（TEM）图像，比例尺为 10 nm；b) 合成态的 LiScO₂ 的低倍透射电子显微镜（TEM）图像；c) 图 a) 中矩形区域的放大图，比例尺为 4 nm；d) 扫描透射电镜（STEM）图像及其对应的能量色散 X 射线光谱（EDX）元素面分布结果，比例尺为 1 nm

通过 X 射线光电子能谱（XPS）全谱扫描进一步证实了样品中存在 Li、Sc 和 O 元素，如图 5-11 a 所示。对 Li 1s、Sc 2p 和 O 1s 区域采集的高分辨率光谱图进行分析，结果如图 5-11 b 所示。Li 1s 谱显示出位于约 53.4 eV 处的单峰，对应于 Li⁺。对 Sc 2p 光谱进行解卷积后，在 Sc 2p_{3/2} 和 Sc 2p_{1/2} 区域均出现一对峰，峰面积比为 1.67。位于 401.0 和 405.2 eV 的峰归因于氧化态 Sc，而结合能较高（402.4 和 406.8 eV）的峰则归因于样品表面水合 Sc-OOH 相的形成^[119]。这一观察结果与 O 1s 光谱相符，该光谱经解卷积后分为三个峰，即氧化物中的

Sc-O (约 529.8 eV)、氢氧化物和水合物 (531.8 eV) 以及少量表面吸附水 (533.8 eV)。这种水合表面与 TEM 观察到的非晶层相对应。

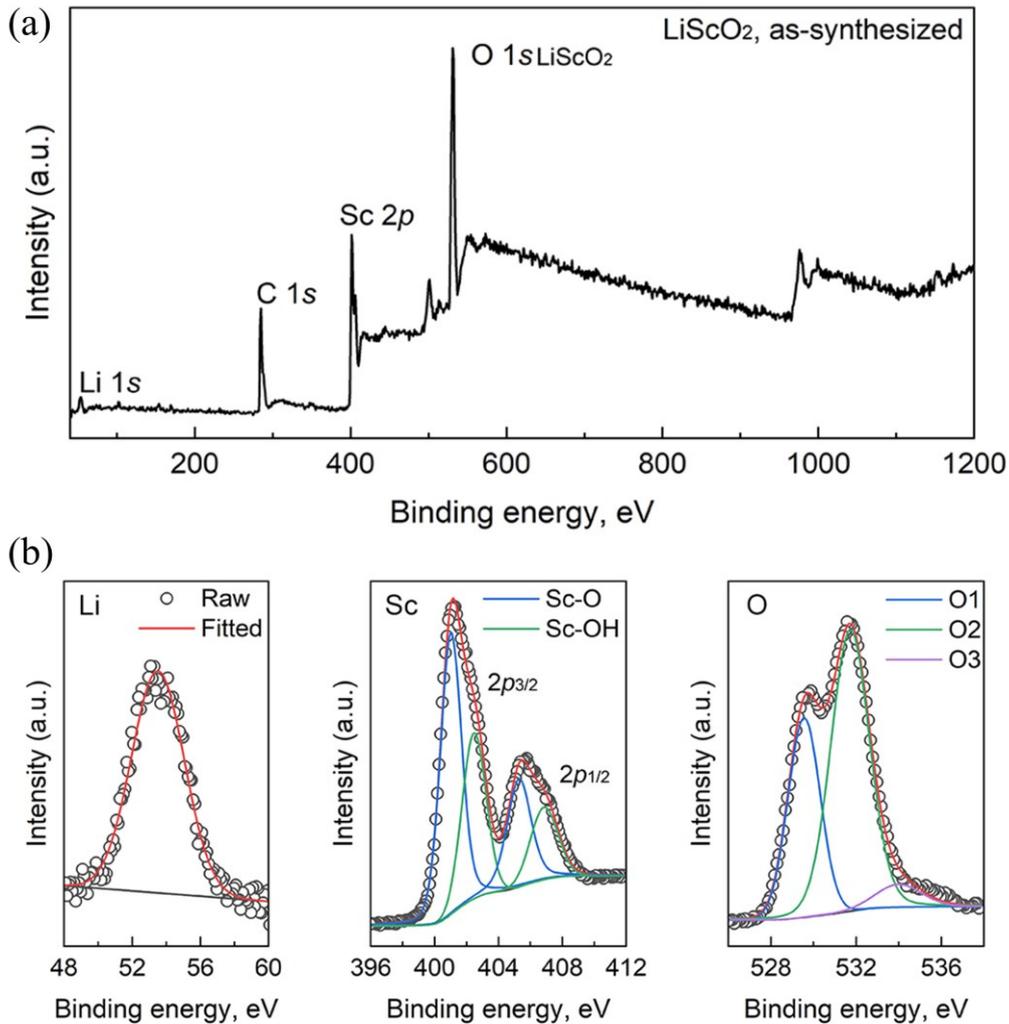


图 5-11 LiScO₂ 的 XPS 分析；a) X 射线光电子能谱 (XPS) 全谱扫描图；b) Li 1s、Sc 2p 和 O 1s 特征的高分辨光谱

5.4.2 LiScO₂ 的性能分析及验证

采用旋转环盘电极 (RRDE) 方法, 在氩气 (Ar) 饱和的 0.5 M KHCO₃ 电解液中评估了催化剂的 WOR 性能 (详见第二章的方法部分)。同时, 将镍 (II) 酞菁分子负载于纯化的碳纳米管上, 合成了一种异相分子催化剂 (NiPc/CNT) 作为对比参照^[120]。该催化剂优先遵循 4e⁻ WOR 路径。图 5-12 a 的上半部分展示了 LiScO₂ 和 NiPc/CNT 催化剂的线性扫描伏安法 (LSV) 曲线。在弱碱性电解液 (pH=9.0) 中, 两种催化剂在盘电极上的起始电位相近, 而 LiScO₂ 的电流密度能迅速超过 NiPc/CNT。

图 5-12 a 下半部分显示了 H₂O₂ 法拉第效率 (FE_{H₂O₂}) 的计算结果。在约 2.2

V_{RHE} 时, $LiScO_2$ 的 $FE_{H_2O_2}$ 可达约 90% 的最大值。尽管此后 $FE_{H_2O_2}$ 略有下降, 但在高达 2.4 V_{RHE} 时, 它仍能保持在 80% 以上。相比而言, $NiPc/CNT$ 倾向于 $4e^-$ WOR, 其 $FE_{H_2O_2}$ 低于 2~4%。此外进一步将催化剂的实验活性与理论计算进行了比较。如图 5-12 b 所示, $LiScO_2$ 的 ΔG_{OH^*} 为 1.75 eV, 在 2.4 V_{RHE} 下, 其实验 ($\log(j)=1.56 \text{ mA cm}^{-2}$) 和计算 ($\log(j)=1.28 \text{ mA cm}^{-2}$) 得到的 $2e^-$ WOR 电流密度几乎一致。在 $NiPc/CNT$ 参照催化剂上也有类似发现, 其实验值为 $\log(j)=0.53 \text{ mA cm}^{-2}$, 计算值为 $\log(j)=0.60 \text{ mA cm}^{-2}$ 。这进一步验证了我们的理论模型。

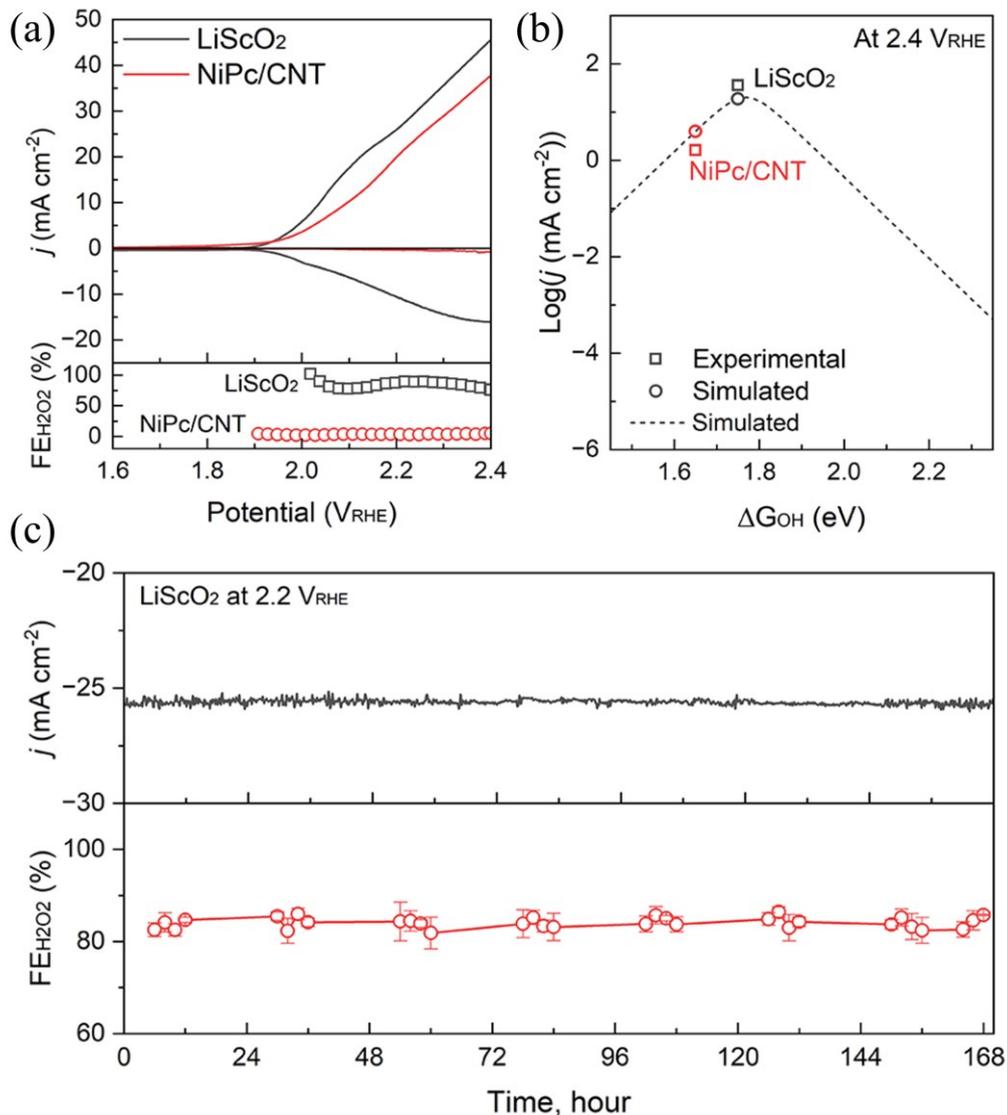


图 5-12 $LiScO_2$ 催化剂性能分析; a) $LiScO_2$ 和 $NiPc/CNT$ 的线性扫描伏安 (LSV) 曲线及 H_2O_2 法拉第效率; b) 实验结果与理论计算的性能对比; c) 流动电解池的稳定性测试

在流动池电解池中, 使用 0.5 M $KHCO_3$ 电解液评估了 $LiScO_2$ 的稳定性。将该催化剂负载在碳纸上, 并施加 2.2 V_{RHE} 的偏压。在 168 小时的测试期间,

电流保持稳定，如图 5-12 c 上半部分所示。定期采集电解液后，通过 Ce^{2+}/Ce^{4+} 滴定法测定 H_2O_2 浓度，进而计算 $FE_{H_2O_2}$ ，结果如图 5-12 c 下半部分所示。该催化剂的 $FE_{H_2O_2}$ 能稳定维持在 82%~86% 的范围内。对测试后的催化剂进行表征，XPS 分析如图 5-13 所示，其谱图与原始催化剂一致；TEM 观察也表明催化剂表面形貌无明显变化（图 5-14），这进一步证明了该催化剂的良好稳定性。

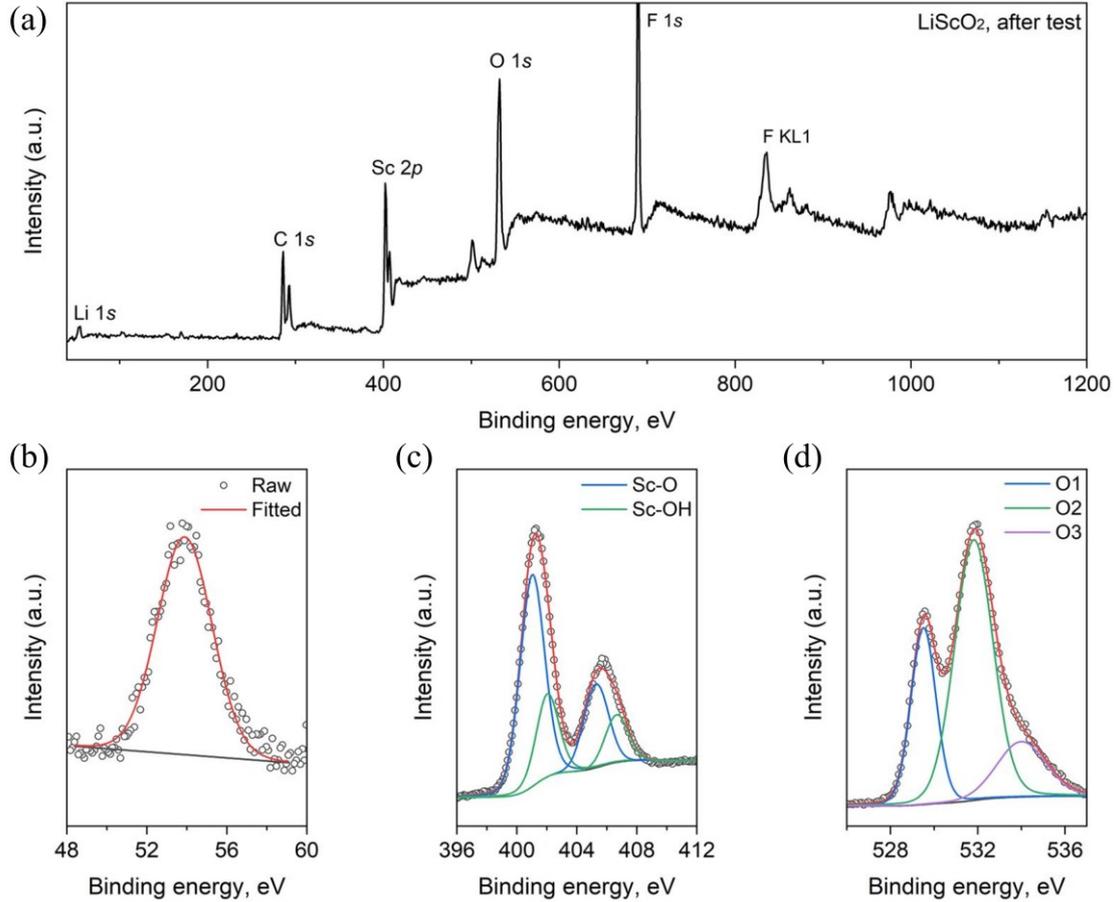


图 5-13 $LiScO_2$ 稳定性测试后的 XPS 分析；a) 测试后样品的 XPS 全谱扫描图；b) $Li\ 1s$ 的高分辨 XPS 光谱；c) $Sc\ 2p$ 的高分辨 XPS 光谱；d) $O\ 1s$ 的高分辨 XPS 光谱

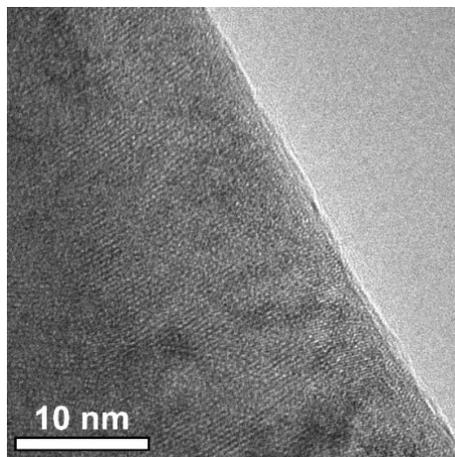


图 5-14 测试后 $LiScO_2$ 表面的透射电子显微镜 (TEM) 图像

为进一步评估本研究合成的 LiScO_2 在 $2e^-$ WOR 反应中的催化性能，将 LiScO_2 与文献中多种性能优异的 $2e^-$ WOR 催化剂的实验性能进行了对比。如表 5-3 所示，在 $2.4 V_{\text{RHE}}$ 的电位下， LiScO_2 的 $\text{FE}_{\text{H}_2\text{O}_2}$ 高达 82%，明显优于钙钛矿 ($\text{Pr}_{1.0}\text{Sr}_{1.0}\text{Fe}_{0.75}\text{Zn}_{0.25}\text{O}_{4.8}$, 72.8%) 和改良后的金属氧化物 ($\text{Gd}:\text{BiVO}_4$, 74% 和 $\text{Sn}:\text{CuWO}_4$, 72%) 等代表性材料。此外，在 10 mA cm^{-2} 的电流密度下， LiScO_2 的过电位仅为 240 mV，远低于其他高效催化剂，如 $\text{Gd}:\text{BiVO}_4$ (760 mV) 和 $\text{ZnO}@\text{CF}$ (710 mV)，表明其具有优异的催化活性。

尽管 $\text{ZnO}@\text{CF}$ 的 $\text{FE}_{\text{H}_2\text{O}_2}$ 可达 78%，但较高的过电位限制了其实际应用。此外， BiFeO_3 和 Bi_2MoO_6 等同类的金属氧化物分别表现出 6.05% 和 17.6% 的 $\text{FE}_{\text{H}_2\text{O}_2}$ ，催化选择性极差，几乎不具备应用价值。相比之下碳基材料 S-CFP 虽具有较为适中的过电位 (340 mV)，但其 $\text{FE}_{\text{H}_2\text{O}_2}$ 仅为 62.1%，在选择性上仍难以与 LiScO_2 匹敌。

综上所述， LiScO_2 是目前已知催化剂中少数能在保证高 $\text{FE}_{\text{H}_2\text{O}_2}$ (> 80%) 的同时实现最低过电位的材料之一。其层状结构中 $\text{Li}^+/\text{Sc}^{3+}$ 的协同效应为设计高效 $2e^-$ WOR 电催化体系提供了新方向。上述实验数据充分验证了该催化剂的结构稳定性和高催化活性。

表 5-3 $2e^-$ WOR 催化剂实验性能对比

| 化学式 | $\text{FE}_{\text{H}_2\text{O}_2}$ (%@ $2.4 V_{\text{RHE}}$) | 过电位 (mV@ 10 mA cm^{-2}) | Ref. |
|--|---|------------------------------------|-------|
| LiScO_2 | 82 | 240 | 本研究 |
| $\text{Sn}:\text{CuWO}_4$ | 72 | 340 | [111] |
| $\text{Pr}_{1.0}\text{Sr}_{1.0}\text{Fe}_{0.75}\text{Zn}_{0.25}\text{O}_{4.8}$ (D-PSFZ) | 72.8 | 410 | [112] |
| $\text{Gd}:\text{BiVO}_4$ | 74 | 760 | [28] |
| BiFeO_3 | 6.05 | 640 | [115] |
| Bi_2MoO_6 | 17.6 | 1060 | [115] |
| $\text{ZnO}@\text{CF}$ | 78 | 710 | [121] |
| S-CFP | 62.1 | 340 | [122] |

5.5 本章小结

本章建立了 $2e^-$ WOR 催化剂的筛选标准，并对数据集和预测集中的催化剂进行了快速筛选。在数据集中发现了 49 种具有高选择性和高活性的候选材料；在预测集中获得 12 种候选催化剂（筛选出的催化剂信息可在 <https://github.com/Weijie-Yang/H2O2> 中查看），其预测值与 DFT 计算结果的偏差均低于 0.2 eV。其中，发现 LiScO_2 是具有高活性和高选择性的新型 $2e^-$ WOR

催化剂。将前 16 种高性能催化剂的过电位与已报道的各类最佳 $2e^-$ WOR 催化剂进行了对比，以充分评估其催化性能的潜在优势。最后，通过固相烧结法制备了预测性能最佳的 LiScO_2 催化剂，并通过 XRD、TEM、XPS 以及电化学测试等手段对其结构、活性及稳定性进行了全面表征和性能分析，同时将其性能与已报道的优异 $2e^-$ WOR 催化剂的实验性能进行了对比。本章主要结论如下：

(1) 微观动力火山图与已报道的实验数据高度吻合，特别是 $\text{CaSnO}_3@CF$ 、 CuWO_4/Sn 、 LaAlO_3 、 ZnO (1010) 及 Gd 掺杂 BiVO_4 等已知高性能催化剂，均位于火山图峰值附近。这验证了模型的准确性和适用性。同时，与热力学火山模型相比，微观动力学模型通过量化电子转移速率与活化能，能够精准地描述不同条件下的速率控制步骤，从而更准确地预测催化剂的活性。

(2) $2e^-$ WOR 的选择性条件为： $\Delta G_{O^*} \geq 3.52 \text{ eV}$ ，且 $1.41 \text{ eV} \leq \Delta G_{OH^*} < 2.4 \text{ eV}$ 。结合微观动力火山图的严格条件： $\log(j) \geq 0 \text{ mA cm}^{-2}$ ，从 962 种材料中逐层筛选出 49 种高性能催化剂。其中 LaAlO_3 、 $\text{Zn}_{10}\text{O}_{10}$ 等已被实验证实具有优异 $2e^-$ WOR 性能的材料，均位于火山顶点，验证了筛选策略的可靠性。

(3) 针对合金、金属氧化物、钙钛矿及单原子催化剂设计了 3312 种新结构，并利用机器学习模型预测了它们的 ΔG_{OH^*} 与 ΔG_{O^*} 。结合 0.2 eV 的误差范围优化筛选条件后，最终获得了 12 种高潜力材料。经 DFT 验证，这些材料的预测值与计算值偏差均小于 0.2 eV 。其中 LiScO_2 位于火山峰值，具有高活性和高选择性。

(4) 实验合成的 LiScO_2 在电化学测试中表现出色。在 0.5 M KHCO_3 溶液中， $\text{FE}_{\text{H}_2\text{O}_2}$ 在 $2.2 \text{ V}_{\text{RHE}}$ 时达到 90% 的最大值，在 $2.4 \text{ V}_{\text{RHE}}$ 时仍保持在 80% 以上。经过 168 小时的稳定性测试，电流与 $\text{FE}_{\text{H}_2\text{O}_2}$ (82%~86%) 均保持稳定，且表面结构无明显变化，表明该催化剂稳定性良好。在 $2.4 \text{ V}_{\text{RHE}}$ 下， LiScO_2 实验 ($\log(j)=1.56 \text{ mA cm}^{-2}$) 和计算 ($\log(j)=1.28 \text{ mA cm}^{-2}$) 得到的 $2e^-$ WOR 活性几乎一致，这验证了理论预测的准确性。

(5) LiScO_2 在 $2e^-$ WOR 反应中展现出卓越活性和选择性。在 $2.4 \text{ V}_{\text{RHE}}$ 电位下实现 82% 的 $\text{FE}_{\text{H}_2\text{O}_2}$ ，显著优于典型钙钛矿和掺杂的金属氧化物 (72-74%)。同时在 10 mA cm^{-2} 电流密度下，仅需 240 mV 过电位，较同类催化剂降低了 70% 以上。其层状氧化物中 $\text{Li}^+/\text{Sc}^{3+}$ 的协同效应为高效 H_2O_2 电催化材料的设计提供了新方向。

第 6 章 结论与展望

6.1 总结

本文主要围绕数据驱动策略下 $2e^-$ WOR 的各类催化剂筛选开展研究。其一，建立了一个包含金属间合金、金属氧化物、钙钛矿以及单原子催化剂的 962 种材料数据库 (<https://github.com/Weijie-Yang/H2O2>)。其二，依据中心与配位原子的几何和电子特性，设计了加权对称函数描述符 (wACSF)，并利用自动化特征提取工具包 (<https://github.com/Weijie-Yang/FG-LAED>)，批量计算了催化剂的 wACSF 值，输出了维度为 962×134 的特征矩阵。其三，对特征数据集进行特征工程，建立 XGBoost 高效预测模型，快速预测了数据库外 3312 种催化剂的 ΔG_{OH^*} 与 ΔG_{O^*} 吸附能。最关键的是，利用微观动力学模型和热力学选择性条件对数据集和预测集进行筛选，分别选出 49 种和 12 种高选择性和高活性的候选材料 (<https://github.com/Weijie-Yang/H2O2>)，从理论和实验两个方面验证了筛选策略的可靠性和预测结果的准确性。本文建立的 wACSF 描述符、自动特征提取工具包与机器学习模型，为数据驱动策略下各类催化剂的快速筛选奠定了理论基础；筛选出的若干种高活性 $2e^-$ WOR 催化剂，为电化学产 H_2O_2 提供了新的见解，推动了这一清洁能源载体在可持续能源领域的开发与应用。本文主要结论如下：

(1) 针对 134 维原始特征，采用皮尔逊相关系数和 RFE 进行特征选择，确定 ΔG_{OH^*} 和 ΔG_{O^*} 的最优特征数量分别为 21 维和 19 维。RFE 测试结果表明，增加额外描述符并不能提升模型性能。对筛选出的特征，进一步剔除相关系数大于 0.8 的特征后，模型的 R^2 和 RMSE 与原模型相比有所下降。说明在保证精度的前提下，适当保留相关特征有助于提升预测性能。

(2) 采用 XGBoost 算法构建模型，在测试集上展现出优异性能 (ΔG_{OH^*} : $R^2=0.84$, RMSE=0.52 eV; ΔG_{O^*} : $R^2=0.91$, RMSE=0.65 eV)，显著优于传统表面中心-环境描述符模型 (ΔG_{OH^*} : $R^2=0.79$, RMSE=0.54 eV; ΔG_{O^*} : $R^2=0.80$, RMSE=0.93 eV)，并且与采用三维坐标 (RMSE=0.63 eV) 和腔体特征 (RMSE=0.56 eV) 的模型预测精度相当。

(3) 截断半径对模型性能影响显著， ΔG_{OH^*} 模型在 $R_c=7 \text{ \AA}$ 时预测效果最佳， ΔG_{O^*} 模型在 $R_c=5 \text{ \AA}$ 时表现最优，表明局部环境特征在吸附自由能预测中起主导作用。通过 SHAP 分析，Nm 和配位价电子数加权的径/角向对称函数 (B_6 、 B_4 、 B_{21} 和 B_{19}) 是影响吸附自由能的关键因素。Nm 与输出呈正相关，反映了

出电子转移对吸附强度的调控作用；较低的 B_6 、 B_4 、 B_{21} 和 B_{19} 值对输出有正向影响，说明局部化学环境和电子分布在吸附过程中发挥着重要作用。

(4) $2e^-$ WOR 的选择性条件为： $\Delta G_{O^*} \geq 3.52 \text{ eV}$ 且 $1.41 \text{ eV} \leq \Delta G_{OH^*} < 2.4 \text{ eV}$ 。结合微观动力火山图的严格条件，从数据集中筛选出 49 种高性能催化剂，其中包含 LaAlO_3 、 $\text{Zn}_{10}\text{O}_{10}$ 等实验已验证的催化剂，说明了筛选策略的可靠性。从预测集中筛选出 12 种候选材料，其中 LiScO_2 位于火山图顶端，表现出卓越的活性和选择性。实验合成的 LiScO_2 在 $2.2 \text{ V}_{\text{RHE}}$ 下 $\text{FE}_{\text{H}_2\text{O}_2}$ 最高达 90%，在 $2.4 \text{ V}_{\text{RHE}}$ 时仍超过 80%。在 168 小时流动池测试中，电流和 $\text{FE}_{\text{H}_2\text{O}_2}$ 均保持稳定。 LiScO_2 的 $2e^-$ WOR 实验活性 ($\log(j)=1.56 \text{ mA cm}^{-2}$) 与理论计算 ($\log(j)=1.28 \text{ mA cm}^{-2}$) 偏差小于 0.3，这验证了理论预测的准确性。同时在 10 mA cm^{-2} 电流密度下， LiScO_2 的过电位仅为 240 mV，较同类催化剂降低了 70%以上。其层状氧化物中 $\text{Li}^+/\text{Sc}^{3+}$ 的协同效应为高效 H_2O_2 电催化材料的设计提供了新方向。

6.2 创新点

本研究围绕 $2e^-$ WOR 产 H_2O_2 的催化剂设计，构建了数据与机理双驱动的通用研究框架，突破了传统催化剂筛选的效率瓶颈与理论局限，主要创新性工作如下：

(1) 加权原子中心对称函数 (wACSF) 描述符：在传统 ACSF 基础上引入配位原子的电负性与价电子数权重，设计通用且可解释的局域环境描述符；还开发 FG-LAED 自动化工具包实现批量特征提取。

(2) 高效机器学习预测模型：采用 XGBoost 算法构建吸附能预测模型，通过皮尔逊相关性分析与递归特征消除 (RFE) 筛选核心特征，最终确定 21 维 (ΔG_{OH^*}) 和 19 维 (ΔG_{O^*}) 最优特征集。测试集 R^2 分别达 0.84 和 0.91、RMSE 0.52 和 0.65 eV，并对数据库外 3312 种新材料完成快速预测。

(3) 实验验证高性能催化剂：合成并测试 LiScO_2 ，实测 $2.2 \text{ V}_{\text{RHE}}$ 下 H_2O_2 法拉第效率高达 90%，168 h 稳定保持 82~86%，活性与计算值偏差 < 0.3 ，过电位仅 240 mV，大幅优于同类催化剂。

本文涵盖了从“数据与描述符设计”到“模型构建”、再到“机理筛选”与“实验验证”全流程，为高效电催化材料设计提供了新范式。

6.3 展望

本文研究了数据驱动策略下 $2e^-$ WOR 的高活性材料筛选，本研究仍有进一步完善与提升之处，在本研究基础上还可以进行如下几个方面的研究：

(1) 本文构建的数据库主要包括四类材料（金属间合金、金属氧化物、钙

钛矿以及单原子催化剂)的 OH^* 和 O^* 吸附能信息。后续可进一步扩展 OOH^* 、 H^* 等关键中间体吸附能数据,构建多反应描述符矩阵,实现氧还原(ORR)与析氧(OER)等反应的协同筛选。还可以引入金属有机框架(MOFs)、双原子催化剂(DACs)等新型材料,建立跨尺度的普适性预测模型。

(2)未来的研究可进一步探索更复杂的深度学习模型,如卷积神经网络、图神经网络等,这类模型能够更好地捕捉催化界面上微小的结构变化和电子分布特征信息,从而实现线对线及面对面的吸附能精准预测。

(3)在处理更大规模的数据集及应对复杂多任务学习时,提升计算效率至关重要。未来的研究可探索更高效且运算负担更轻的机器学习算法,并结合更为灵活的计算资源分配优化策略。现代计算框架(如GPU加速和多节点分布式计算平台)有望实现对大规模数据集的高效处理,加速材料筛选过程。

(4)本研究在开发与验证 wACSF 描述符方面取得了积极成果。后续工作可尝试设计更具多样性和维度的原子环境特征,以全面反映催化剂的电子和几何结构,进而更精准地捕捉活性位点信息。催化反应过程涉及物理、化学和材料科学等多个领域,未来可推进跨学科合作,从多角度深入解析催化剂的作用机制与构效关系,为高性能催化剂的开发提供更为丰富的理论支持和技术手段。

参考文献

- [1] 朱欢, 郑洁, 赵秋运, 等. 经济增长、能源结构转型与二氧化碳排放——基于面板数据的经验分析[J]. 经济与管理研究, 2020, 41(11): 19-34
- [2] Arutyunov V S, Lisichkin G V. Energy resources of the 21st century: problems and forecasts. Can renewable energy sources replace fossil fuels[J]. Russian Chemical Reviews, 2017, 86(8): 777
- [3] Liu L, Cheng S Y, Li J B, et al. Mitigating environmental pollution and impacts from fossil fuels: The role of alternative fuels[J]. Energy Sources, Part A, 2007, 29(12): 1069-1080
- [4] Sato K, Aoki M, Noyori R. A "green" route to adipic acid: direct oxidation of cyclohexenes with 30 percent hydrogen peroxide[J]. Science, 1998, 281(5383): 1646-1647
- [5] Yamada Y, Yoneda M, Fukuzumi S. High and robust performance of H₂O₂ fuel cells in the presence of scandium ion[J]. Energy & Environmental Science, 2015, 8(6): 1698-1701
- [6] An L, Zhao T S, Zeng L, et al. Performance of an alkaline direct ethanol fuel cell with hydrogen peroxide as oxidant[J]. international journal of hydrogen energy, 2014, 39(5): 2320-2324
- [7] Tang J, Zhao T, Solanki D, et al. Selective hydrogen peroxide conversion tailored by surface, interface, and device engineering[J]. Joule, 2021, 5(6): 1432-1461
- [8] Hu S. Membrane-less photoelectrochemical devices for H₂O₂ production: efficiency limit and operational constraint[J]. Sustainable Energy & Fuels, 2019, 3(1): 101-114
- [9] 海云, 褚宝增, 王连习. 过氧化氢在石油工业上的应用[J]. 断块油气田, 2000, (06): 63-65
- [10] 佚名. 过氧化氢市场现状和技术发展趋势[R]. 北京: 化工网, 2023
- [11] Chen Q. Development of an anthraquinone process for the production of hydrogen peroxide in a trickle bed reactor—From bench scale to industrial scale[J]. Chemical Engineering and Processing: Process Intensification, 2008, 47(5): 787-792
- [12] Shi X, Siahrostami S, Li G L, et al. Understanding activity trends in electrochemical water oxidation to form hydrogen peroxide[J]. Nature communications, 2017, 8(1): 701
- [13] Shi X, Back S, Gill T M, et al. Electrochemical synthesis of H₂O₂ by two-electron water oxidation reaction[J]. Chem, 2021, 7(1): 38-63

- [14] Siahrostami S, Verdaguer-Casadevall A, Karamad M, et al. Enabling direct H₂O₂ production through rational electrocatalyst design[J]. *Nature materials*, 2013, 12(12): 1137-1143
- [15] Sun X, Yang J, Zeng X, et al. Pairing Oxygen Reduction and Water Oxidation for Dual-Pathway H₂O₂ Production[J]. *Angewandte Chemie International Edition*, 2024, 63(52): e202414417
- [16] Mavrikis S, Perry S C, Leung P K, et al. Recent advances in electrochemical water oxidation to produce hydrogen peroxide: a mechanistic perspective[J]. *ACS Sustainable Chemistry & Engineering*, 2020, 9(1): 76-91
- [17] Ponce de León C. In situ anodic generation of hydrogen peroxide[J]. *Nature Catalysis*, 2020, 3(2): 96-97
- [18] Perry S C, Pangotra D, Vieira L, et al. Electrochemical synthesis of hydrogen peroxide from water and oxygen[J]. *Nature Reviews Chemistry*, 2019, 3(7): 442-458
- [19] Lu Z, Chen G, Siahrostami S, et al. High-efficiency oxygen reduction to hydrogen peroxide catalysed by oxidized carbon materials[J]. *Nature Catalysis*, 2018, 1(2): 156-162
- [20] Kim H W, Ross M B, Kornienko N, et al. Efficient hydrogen peroxide generation using reduced graphene oxide-based oxygen reduction electrocatalysts[J]. *Nature Catalysis*, 2018, 1(4): 282-290
- [21] Xia C, Xia Y, Zhu P, et al. Direct electrosynthesis of pure aqueous H₂O₂ solutions up to 20% by weight using a solid electrolyte[J]. *Science*, 2019, 366(6462): 226-231
- [22] Fan L, Bai X, Xia C, et al. CO₂/carbonate-mediated electrochemical water oxidation to hydrogen peroxide[J]. *Nature Communications*, 2022, 13(1): 2668
- [23] Fuku K, Miyase Y, Miseki Y, et al. Enhanced oxidative hydrogen peroxide production on conducting glass anodes modified with metal oxides[J]. *ChemistrySelect*, 2016, 1(18): 5721-5726
- [24] Li J, Solanki D, Zhu Q, et al. Microstructural origin of selective water oxidation to hydrogen peroxide at low overpotentials: a study on Mn-alloyed TiO₂[J]. *Journal of Materials Chemistry A*, 2021, 9(34): 18498-18505
- [25] Guo W, Xie Y, Tang S, et al. H₂O₂ formation mechanisms on the (112) and (310) facets of SnO₂ via water oxidation reaction with the participation of bicarbonate: DFT and experimental investigations[J]. *Applied Surface Science*, 2022, 596: 153634
- [26] Guo W, Xie Y, Liu Y, et al. Effects of Sb₂O₃ polymorphism on the performances for electrocatalytic H₂O₂ production via the two-electron water oxidation reaction[J]. *Applied Surface Science*, 2022, 606: 155006

-
- [27] Kelly S R, Shi X, Back S, et al. ZnO as an active and selective catalyst for electrochemical water oxidation to hydrogen peroxide[J]. ACS Catalysis, 2019, 9(5): 4593-4599
- [28] Baek J H, Gill T M, Abroshan H, et al. Selective and efficient Gd-doped BiVO₄ photoanode for two-electron water oxidation to H₂O₂[J]. ACS Energy Letters, 2019, 4(3): 720-728
- [29] Park S Y, Abroshan H, Shi X, et al. CaSnO₃: an electrocatalyst for two-electron water oxidation reaction to form H₂O₂[J]. ACS Energy Letters, 2018, 4(1): 352-357
- [30] Li L, Xiao K, Wong P K, et al. Hydrogen peroxide production from water oxidation on a CuWO₄ anode in oxygen-deficient conditions for water decontamination[J]. ACS Applied Materials & Interfaces, 2022, 14(6): 7878-7887
- [31] Baek J, Jin Q, Johnson N S, et al. Discovery of LaAlO₃ as an efficient catalyst for two-electron water electrolysis towards hydrogen peroxide[J]. nature communications, 2022, 13(1): 7256
- [32] Dong K, Lei Y, Zhao H, et al. Noble-metal-free electrocatalysts toward H₂O₂ production[J]. Journal of Materials Chemistry A, 2020, 8(44): 23123-23141
- [33] Wang Y, Lian X, Zhou Y, et al. Synthesis and characterization of Sb₂O₃: a stable electrocatalyst for efficient H₂O₂ production and accumulation and effective degradation of dyes[J]. New Journal of Chemistry, 2021, 45(20): 8958-8964
- [34] Chaudhary P, Evazzade I, Belosludov R, et al. Computational Discovery of Active and Selective Metal-Nitrogen-Graphene Catalysts for Electrooxidation of Water to H₂O₂[J]. ChemCatChem, 2023, 15(10): e202300055
- [35] Vassilev P, Koper M T M. Electrochemical reduction of oxygen on gold surfaces: a density functional theory study of intermediates and reaction paths[J]. The Journal of Physical Chemistry C, 2007, 111(6): 2607-2613
- [36] Zhang D, Hirai Y, Nakamura K, et al. Benchmarking pH-field coupled microkinetic modeling against oxygen reduction in large-scale Fe-azaphthalocyanine catalysts[J]. Chemical Science, 2024, 15(14): 5123-5132
- [37] Hansen H A, Viswanathan V, Nørskov J K. Unifying kinetic and thermodynamic analysis of 2e⁻ and 4e⁻ reduction of oxygen on metal surfaces[J]. The Journal of Physical Chemistry C, 2014, 118(13): 6706-6718
- [38] Nørskov J K, Rossmeisl J, Logadottir A, et al. Origin of the overpotential for oxygen reduction at a fuel-cell cathode[J]. The Journal of Physical Chemistry B, 2004, 108(46): 17886-17892
- [39] Man I C, Su H Y, Calle-Vallejo F, et al. Universality in oxygen evolution

- electrocatalysis on oxide surfaces[J]. *ChemCatChem*, 2011, 3(7): 1159-1165
- [40] Hu J M, Zhang J Q, Cao C N. Oxygen evolution reaction on IrO₂-based DSA® type electrodes: kinetics analysis of Tafel lines and EIS[J]. *International Journal of Hydrogen Energy*, 2004, 29(8): 791-797
- [41] Zhang J J, Hu J M, Zhang J Q, et al. IrO₂-SiO₂ binary oxide films: Geometric or kinetic interpretation of the improved electrocatalytic activity for the oxygen evolution reaction[J]. *international journal of hydrogen energy*, 2011, 36(9): 5218-5226
- [42] Schlexer Lamoureux P, Winther K T, Garrido Torres J A, et al. Machine learning for computational heterogeneous catalysis[J]. *ChemCatChem*, 2019, 11(16): 3581-3601
- [43] Behler J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials[J]. *The Journal of chemical physics*, 2011, 134(7): 074106
- [44] Behler J, Parrinello M. Generalized neural-network representation of high-dimensional potential-energy surfaces[J]. *Physical review letters*, 2007, 98(14): 146401
- [45] Hu Z, Guo Y, Liu Z, et al. Aisnet: A universal interatomic potential neural network with encoded local environment features[J]. *Journal of Chemical Information and Modeling*, 2023, 63(6): 1756-1765
- [46] Zhang K, Yin L, Liu G. Physically inspired atom-centered symmetry functions for the construction of high dimensional neural network potential energy surfaces[J]. *Computational Materials Science*, 2021, 186: 110071
- [47] Sahu H, Rao W, Troisi A, et al. Toward predicting efficiency of organic solar cells via machine learning and improved descriptors[J]. *Advanced Energy Materials*, 2018, 8(24): 1801032
- [48] Li Z, Xu Q, Sun Q, et al. Thermodynamic stability landscape of halide double perovskites via high-throughput computing and machine learning[J]. *Advanced Functional Materials*, 2019, 29(9): 1807280
- [49] Gu G H, Choi C, Lee Y, et al. Progress in computational and machine-learning methods for heterogeneous small-molecule activation[J]. *Advanced materials*, 2020, 32(35): 1907865
- [50] Severson K A, Attia P M, Jin N, et al. Data-driven prediction of battery cycle life before capacity degradation[J]. *Nature Energy*, 2019, 4(5): 383-391
- [51] Ge X, Yin J, Ren Z, et al. Atomic design of alkyne semihydrogenation catalysts via active learning[J]. *Journal of the American Chemical Society*, 2024, 146(7): 4993-5004
- [52] Zhou C, Chen C, Hu P, et al. Topology-determined structural genes enable data-driven discovery and intelligent design of potential metal oxides for inert

- C–H bond activation[J]. *Journal of the American Chemical Society*, 2023, 145(40): 21897-21903
- [53] Lin X, Du X, Wu S, et al. Machine learning-assisted dual-atom sites design with interpretable descriptors unifying electrocatalytic reactions[J]. *Nature Communications*, 2024, 15(1): 8169
- [54] Lian Z, Yang M, Jan F, et al. Machine learning derived blueprint for rational design of the effective single-atom cathode catalyst of the lithium–sulfur battery[J]. *The Journal of Physical Chemistry Letters*, 2021, 12(29): 7053-7059
- [55] Bustillo-Lecompte C. *Advanced oxidation processes: Applications, trends, and prospects*[M]. IntechOpen, 2020
- [56] Motagamwala A H, Dumesic J A. Microkinetic modeling: a tool for rational catalyst design[J]. *Chemical Reviews*, 2020, 121(2): 1049-1076
- [57] Limmer D T, Willard A P, Madden P, et al. Hydration of metal surfaces can be dynamically heterogeneous and hydrophobic[J]. *Proceedings of the National Academy of Sciences*, 2013, 110(11): 4200-4205
- [58] Gong S, Wang C, Jiang P, et al. Designing highly efficient dual-metal single-atom electrocatalysts for the oxygen reduction reaction inspired by biological enzyme systems[J]. *Journal of Materials Chemistry A*, 2018, 6(27): 13254-13262
- [59] Wang Y, Tang Y J, Zhou K. Self-adjusting activity induced by intrinsic reaction intermediate in Fe–N–C single-atom catalysts[J]. *Journal of the American Chemical Society*, 2019, 141(36): 14115-14119
- [60] Marković N M, Ross Jr P N. Surface science studies of model fuel cell electrocatalysts[J]. *Surface Science Reports*, 2002, 45(4-6): 117-229
- [61] Wang J X, Markovic N M, Adzic R R. Kinetic analysis of oxygen reduction on Pt (111) in acid solutions: intrinsic kinetic parameters and anion adsorption effects[J]. *The Journal of Physical Chemistry B*, 2004, 108(13): 4127-4133
- [62] 夏杰桢, 曹蓉, 吴琪. 机器学习结合密度泛函理论计算在材料科学中的研究进展[J]. *化学通报*, 2022, 85(10): 1224-1232
- [63] Wang T, Li J, Shu W, et al. Machine-learning adsorption on binary alloy surfaces for catalyst screening[J]. *Chinese Journal of Chemical Physics*, 2020, 33(6): 703-711
- [64] Pedregosa F, Varoquaux G, Gramfort A, et al. Scikit-learn: Machine learning in Python[J]. *the Journal of machine Learning research*, 2011, 12: 2825-2830
- [65] Abraham A, Pedregosa F, Eickenberg M, et al. Machine learning for neuroimaging with scikit-learn[J]. *Frontiers in neuroinformatics*, 2014, 8: 14
- [66] Friedman J H. Stochastic gradient boosting[J]. *Computational statistics &*

- data analysis, 2002, 38(4): 367-378
- [67] Friedman J H. Greedy function approximation: a gradient boosting machine[J]. *Annals of statistics*, 2001: 1189-1232
- [68] Belyadi H, Haghghat A. *Machine learning guide for oil and gas using Python: A step-by-step breakdown with data, algorithms, codes, and applications*[M]. Gulf Professional Publishing, 2021
- [69] Schober P, Boer C, Schwarte L A. Correlation coefficients: appropriate use and interpretation[J]. *Anesthesia & analgesia*, 2018, 126(5): 1763-1768
- [70] Argaman N, Makov G. Density functional theory: An introduction[J]. *American Journal of Physics*, 2000, 68(1): 69-79
- [71] 李震宇, 贺伟, 杨金龙. 密度泛函理论及其数值方法新进展[J]. *化学进展*, 2005, 17(02): 192-202
- [72] Hohenberg P, Kohn W. Inhomogeneous electron gas[J]. *Physical review*, 1964, 136(3B): B864
- [73] Kohn W, Sham L J. Self-consistent equations including exchange and correlation effects[J]. *Physical review*, 1965, 140(4A): A1133
- [74] Gilbert T L. Hohenberg-Kohn theorem for nonlocal external potentials[J]. *Physical Review B*, 1975, 12(6): 2111
- [75] Kresse G, Furthmüller J. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set[J]. *Physical review B*, 1996, 54(16): 11169
- [76] Kresse G, Furthmüller J. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set[J]. *Computational materials science*, 1996, 6(1): 15-50
- [77] Krause M, Dent E W, Bear J E, et al. Ena/VASP proteins: regulators of the actin cytoskeleton and cell migration[J]. *Annual review of cell and developmental biology*, 2003, 19(1): 541-564
- [78] Lodeiro L, Rauch T. DensityTool: A post-processing tool for space-and spin-resolved density of states from VASP[J]. *Computer Physics Communications*, 2022, 277: 108384
- [79] Momma K, Izumi F. VESTA 3 for three-dimensional visualization of crystal, volumetric and morphology data[J]. *Applied Crystallography*, 2011, 44(6): 1272-1276
- [80] Wang V, Xu N, Liu J C, et al. VASPKIT: A user-friendly interface facilitating high-throughput computing and analysis using VASP code[J]. *Computer Physics Communications*, 2021, 267: 108033
- [81] Hensley A J R, Ghale K, Rieg C, et al. DFT-based method for more accurate adsorption energies: an adaptive sum of energies from RPBE and vdW density

- functionals[J]. *The Journal of Physical Chemistry C*, 2017, 121(9): 4937-4945
- [82] Zhao G, Muhammad I, Suzuki K, et al. Synthesis, crystal structure, and the ionic conductivity of new lithium ion conductors, M-doped LiScO₂ (M= Zr, Nb, Ta)[J]. *Materials Transactions*, 2016, 57(8): 1370-1373
- [83] Li L, Li H, Seymour I D, et al. Pair-distribution-function guided optimization of fingerprints for atom-centered neural network potentials[J]. *The Journal of Chemical Physics*, 2020, 152(22): 224102
- [84] Gallegos M, Isamura B K, Popelier P L A, et al. An Unsupervised Machine Learning Approach for the Automatic Construction of Local Chemical Descriptors[J]. *Journal of Chemical Information and Modeling*, 2024, 64(8): 3059-3079
- [85] Jäger M O J, Morooka E V, Federici Canova F, et al. Machine learning hydrogen adsorption on nanoclusters through structural descriptors[J]. *npj Computational Materials*, 2018, 4(1): 37
- [86] Wang X, Xiao B, Li Y, et al. First-principles based machine learning study of oxygen evolution reactions of perovskite oxides using a surface center-environment feature model[J]. *Applied Surface Science*, 2020, 531: 147323
- [87] Xu H, Cheng D, Cao D, et al. Revisiting the universal principle for the rational design of single-atom electrocatalysts[J]. *Nature Catalysis*, 2024, 7(2): 207-218
- [88] Mamun O, Winther K T, Boes J R, et al. High-throughput calculations of catalytic properties of bimetallic alloy surfaces[J]. *Scientific data*, 2019, 6(1): 76
- [89] Comer B M, Bothra N, Lunger J R, et al. Prediction of O and OH adsorption on transition metal oxide surfaces from bulk descriptors[J]. *ACS Catalysis*, 2024, 14(7): 5286-5296
- [90] Zhou H, Deng Z, Xia Y, et al. A new sampling method in particle filter based on Pearson correlation coefficient[J]. *Neurocomputing*, 2016, 216: 208-215
- [91] Su R, Liu X, Wei L. MinE-RFE: determine the optimal subset from RFE by minimizing the subset-accuracy-defined energy[J]. *Briefings in Bioinformatics*, 2020, 21(2): 687-698
- [92] Breiman L. Random forests[J]. *Machine learning*, 2001, 45: 5-32
- [93] Robert C. Machine learning, a probabilistic perspective[J]. *CHANCE*, 2014, 27: 62-63
- [94] Saleh A K M E, Arashi M, Kibria B M G. Theory of ridge regression estimation with applications[M]. John Wiley & Sons, 2019
- [95] Zhang Z. Introduction to machine learning: k-nearest neighbors[J]. *Annals of translational medicine*, 2016, 4(11): 218
- [96] Keprate A, Ratnayake R M C. Using gradient boosting regressor to predict

- stress intensity factor of a crack propagating in small bore piping[A]. 2017 IEEE International Conference on Industrial Engineering and Engineering Management (IEEM)[C]. IEEE, 2017: 1331-1336
- [97] Smola A J, Schölkopf B. A tutorial on support vector regression[J]. *Statistics and computing*, 2004, 14: 199-222
- [98] Geurts P, Ernst D, Wehenkel L. Extremely randomized trees[J]. *Machine learning*, 2006, 63: 3-42
- [99] Gao X J, Yan J, Zheng J J, et al. Clear-box machine learning for virtual screening of 2D nanozymes to target tumor hydrogen peroxide[J]. *Advanced Healthcare Materials*, 2023, 12(10): 2202925
- [100] Zong X, Vlachos D G. Exploring structure-sensitive relations for small species adsorption using machine learning[J]. *Journal of Chemical Information and Modeling*, 2022, 62(18): 4361-4368
- [101] Tarwidi D, Pudjaprasetya S R, Adytia D, et al. An optimized XGBoost-based machine learning method for predicting wave run-up on a sloping beach[J]. *MethodsX*, 2023, 10: 102119
- [102] Li H, Nørskov J K. Effects of a conductive support on the bonding of oxygen containing molecules to transition metal oxide surfaces[J]. *Physical Chemistry Chemical Physics*, 2020, 22(45): 26216-26222
- [103] Xu W, Andersen M, Reuter K. Data-driven descriptor engineering and refined scaling relations for predicting transition metal oxide reactivity[J]. *ACS Catalysis*, 2020, 11(2): 734-742
- [104] Deng B, Chen P, Xie P, et al. Iterative machine learning method for screening high-performance catalysts for H₂O₂ production[J]. *Chemical Engineering Science*, 2023, 267: 118368
- [105] Minotaki M G, Geiger J, Ruiz-Ferrando A, et al. A generalized model for estimating adsorption energies of single atoms on doped carbon materials[J]. *Journal of Materials Chemistry A*, 2024, 12(18): 11049-11061
- [106] Lundberg S M, Lee S I. A unified approach to interpreting model predictions[J]. *Advances in neural information processing systems*, 2017, 30
- [107] An Y, Ouyang M, Kong S, et al. Electronegativity principle for hydrogen evolution activity using first-principles calculations[J]. *Physical Chemistry Chemical Physics*, 2023, 25(19): 13289-13296
- [108] Hao J, Zhuang Z, Cao K, et al. Unraveling the electronegativity-dominated intermediate adsorption on high-entropy alloy electrocatalysts[J]. *Nature Communications*, 2022, 13(1): 2662
- [109] Gastegger M, Schwiedrzik L, Bittermann M, et al. wACSF—Weighted atom-centered symmetry functions as descriptors in machine learning potentials[J]. *The Journal of chemical physics*, 2018, 148(24): 241709

- [110] Zhang C, Lu R, Liu C, et al. High yield electrosynthesis of hydrogen peroxide from water using electrospun CaSnO_3 @carbon fiber membrane catalysts with abundant oxygen vacancy[J]. *Advanced Functional Materials*, 2021, 31(26): 2100099
- [111] Li L, Xu L, Chan A W M, et al. Direct hydrogen peroxide synthesis on a Sn-doped CuWO_4/Sn anode and an air-breathing cathode[J]. *Chemistry of Materials*, 2021, 34(1): 63-71
- [112] Kim C, Park S O, Kwak S K, et al. Concurrent oxygen reduction and water oxidation at high ionic strength for scalable electrosynthesis of hydrogen peroxide[J]. *Nature Communications*, 2023, 14(1): 5822
- [113] Khan R, Chakraborty J, Singh Rawat K, et al. Super-Oxidizing Covalent Triazine Framework Electrocatalyst for Two-Electron Water Oxidation to H_2O_2 [J]. *Angewandte Chemie*, 2023, 135(47): e202313836
- [114] Xu J, Wang Y, Dong C, et al. Possibility of hydrogen peroxide production by $2e^-$ WOR with non-oxygen evolution catalysts and its application for organic dye decomposition via a self-cycled Fenton System[J]. *Journal of Environmental Chemical Engineering*, 2024, 12(3): 112779
- [115] Li L, Hu Z, Yu J C. On-demand synthesis of H_2O_2 by water oxidation for sustainable resource production and organic pollutant degradation[J]. *Angewandte Chemie*, 2020, 132(46): 20719-20725
- [116] Chen Z, Geng S, Wang Y, et al. Boosting $2e^-$ water oxidation reaction on WO_3 by F modification and revealing the mechanism by probing interfacial water structure[J]. *Applied Catalysis B: Environmental*, 2022, 317: 121756
- [117] Li H, Kelly S, Guevarra D, et al. Analysis of the limitations in the oxygen reduction activity of transition metal oxide surfaces[J]. *Nature Catalysis*, 2021, 4(6): 463-468
- [118] Liu H, Zhang D, Holmes S M, et al. Origin of the superior oxygen reduction activity of zirconium nitride in alkaline media[J]. *Chemical Science*, 2023, 14(34): 9000-9009
- [119] Biesinger M C, Lau L W M, Gerson A R, et al. Resolving surface chemical states in XPS analysis of first row transition metals, oxides and hydroxides: Sc, Ti, V, Cu and Zn[J]. *Applied surface science*, 2010, 257(3): 887-898
- [120] Zhang D, She F, Chen J, et al. Why Do Weak-Binding M–N–C Single-Atom Catalysts Possess Anomalously High Oxygen Reduction Activity?[J]. *Journal of the American Chemical Society*, 2025, 147(7): 6076 - 6086
- [121] Zhang C, Lu J, Liu C, et al. ZnO nanoparticles embedded in hollow carbon fiber membrane for electrochemical H_2O_2 production by two-electron water oxidation reaction[J]. *Environmental Research*, 2022, 206: 112290
- [122] Ling C, Liang A, Li C, et al. Coupling functional anodes with natural air-

diffused cathodes enables highly efficient hydrogen peroxide electrosynthesis[J]. Journal of Zhejiang University-SCIENCE A, 2023, 24(4): 377-386

攻读硕士学位期间发表的论文及其它成果

发表的学术论文

- [1] A Ni single-atom catalyst for advanced environmental disinfection based on electrochemical production of hydrogen peroxide[J]. *Journal of Materials Chemistry A*, 2025. (WOS: 001445844900001, 影响因子: 10.6, 除导师外第二作者)

致谢

当笔锋游至此处，砚池微澜泛起三载寒暑的光影。值此论文付梓之际，谨以寸笺尺素，向漫漫长路上洒落星辉的引路人致意。回首这段求学之路，自踏入科研之门，至今顺利完成学业，这一路走来，既有探索的喜悦，也有困顿的迷茫，而正是这些经历，塑造了今日的我。在此，谨以至诚之心，向给予我指导、陪伴与支持的师长、亲友与同窗，献上最真挚的谢意。

承蒙刘志坚教授不弃，以高山仰止的学术襟怀为我启迪。先生治学如松立危崖，既有穿云破雾的凌厉笔锋，亦怀润物无声的谆谆教诲。您让我领略了科研的深邃与严谨。您的博学睿智、严谨治学精神，引领我在学术的海洋中乘风破浪，教会我如何以科学的态度面对问题、探索未知。

杨维结副教授则似江南细雨，于我如春风化雨，温润中自有千钧之力。您是我学术道路上的灯塔，以渊博的学识指引我在科研的海洋中锚定方向，从论文选题到思路设计，从数据分析到文章撰写，您的悉心指导与耐心点拨，使我在科研的道路上少走了许多弯路。在您的引领下，我学会了如何思考，如何求索，如何以持之以恒的精神追求真理。这份恩情，铭刻于心，必将伴随我未来的人生旅程。

椿萱之恩，重若岱岳。父母半生操劳，将青丝熬成月色，把掌纹刻满风霜，以田舍翁媪之身，托举我触摸学术星斗。你们的无私付出与默默支持，是我前行的最大动力。从幼时的启蒙到今日的学有所成，你们不仅赋予我生命，更以言传身教让我明白了责任、坚韧与担当。故乡的炊烟虽远，血脉中流淌的温情始终萦绕着你们的牵挂与鼓励，在异乡的檐角化作深切的叮咛，治愈我的每一次的忧伤与疲倦。

手足情深如窗前常青藤，细碎光阴里缠绕着无声守候。弟弟妹妹用纯粹的笑容为我驱散疲惫，用无条件的支持为我筑起最温暖的港湾。那些深夜的畅谈、困境中的鼓励，甚至是偶尔的拌嘴，都成了我生命中最鲜活的光亮。你们的天真烂漫让我学会用简单的心看待世界，你们的勇敢坚韧也时刻鞭策我成为更好的自己。未来的路，愿我们依然能彼此守护，分享喜悦，共渡风雨。谢谢你们，让我的人生因“手足”二字而充满底气与幸福。

同窗诸君，恰似星辉共映。感谢我的同门与舍友。在这段求学岁月里，我们一起奋斗，一起探讨，一起分享科研的喜悦与困惑。正是有你们的陪伴，我的研究生生活才充满了欢声笑语，科研的路途也因此少了些孤独，多了些温暖。那些在工位的日日夜夜，那些讨论交流的思维碰撞，那些课余闲谈的片刻欢愉，

已成为我心底最珍贵的记忆。

寒门学子逐梦路上，这册拙作或许只是学术瀚海中的蜉蝣芥子，却承载着太多人的星河璀璨。毕业不仅意味着一段旅程的结束，更象征着新的起点。未来的道路或许仍然充满挑战，但我相信，今日所获的知识与经历，将成为我继续前行的指引。在此，再次感谢所有给予我帮助和支持的人，愿我们在各自的征途上乘风破浪，勇往直前！

华北电力大学学位论文答辩委员会情况

| | 姓名 | 职称 | 工作单位 | 学科专长 |
|--------|-----|-----|------------------------|--------------------|
| 答辩主席 | 纪文静 | 副教授 | 北京科技大学 建筑环境与能源应用工程系 | 供热、供燃气、 通风及空调工程 |
| 答辩委员 1 | 高月芬 | 副教授 | 华北电力大学 动力工程系 | 动力工程及工程 热物理 |
| 答辩委员 2 | 杜亚荣 | 副教授 | 华北电力大学 动力工程系 | 动力工程及工程 热物理 |
| 答辩委员 3 | 韩旭 | 副教授 | 华北电力大学 动力工程系 | 动力工程及工程 热物理 |
| 答辩委员 4 | 李晓静 | 讲师 | 华北电力大学 动力工程系 | 动力工程及工程 热物理 |