

---

# 被遗弃的数据- 实现AI高精度预测多孔材料分离性能的有效拼图

作者：writer 来源：爱科学

本文原地址：<https://www.iikx.com/news/progress/19734.html>

*本文仅供学习交流之用，版权归原作者所有，请勿用于商业用途！*

被遗弃的数据-实现AI高精度预测多孔材料分离性能的有效拼图。

2022年8月22日，浙江大学邢华斌教授团队在Matter期刊上发表了一篇题为Machine-learning-assisted exploration of anion-pillared metal organic frameworks for gas separation的研究成果。

课题组提出通过结合废弃实验数据和计算化学参数实现对多孔材料吸附性能高精度预测的通用机器学习训练策略，并首次实现了对阴离子柱撑超微孔材料乙炔、乙烯、二氧化碳实验吸附性能的高精度预测，筛选并制备出两种兼具高吸附容量和高分离选择性的新型阴离子柱撑多孔材料。浙江大学邢华斌教授为论文通讯作者；胡建波为论文第一作者。

近年来，机器学习方法被引入到材料科学研究中，能显著提高材料研发效率。机器学习可以使计算机能够从训练数据集中挖掘材料结构与性质间的关联，并预测测试集中新材料的性能。机器学习效果的关键在于实现对材料性能的高精度预测，而其预测精度在很大程度上取决于训练数据的质量。在吸附分离领域，训练数据，包括材料结构特征描述符和材料性能数据，可以通过高通量计算模拟或实验表征测试获得。高通量模拟数据是目前吸附分离领域机器学习训练数据的主要来源，具有数据量大、一致性高、完整性佳等特点，但受限于力场参数，计算模拟方法获取的材料吸附性能数据准确性较低，特别是对于具有强作用位点和柔性特征的材料，限制了机器学习模型的预测精度。相较于计算模拟数据，基于实验表征测试方法可以获得更高精度的材料性能数据，但现有文献中报道的实验数据产生于不同时间和不同研究人员，且所报道的数据大多是性能优异的材料性能数据，数据一致性不足，完整性不佳，难以满足机器学习训练需求。基于机器学习实现对材料实际性能的高精度预测仍然具有很大的挑战。

在这项工作中，浙江大学邢华斌团队提出通过结合废弃实验数据和计算化学参数实现对材料性能高精度预测的通用机器学习训练策略，解决了当前吸附分离领域难以获取一致、完整、准确的数据导致机器学习模型预测精度低的难题。基于此通用机器学习策略，该研究首次实现了对阴离子柱撑超微孔材料乙炔、乙烯、二氧化碳实验吸附性能的高精度预测，筛选并制备出两种兼具高吸附容量和高分离选择性的新型离子杂化多孔材料。同时，定量描述了材料结构特征与吸附性能的关系，为新型吸附剂的设计提供更精确、直观的指导。

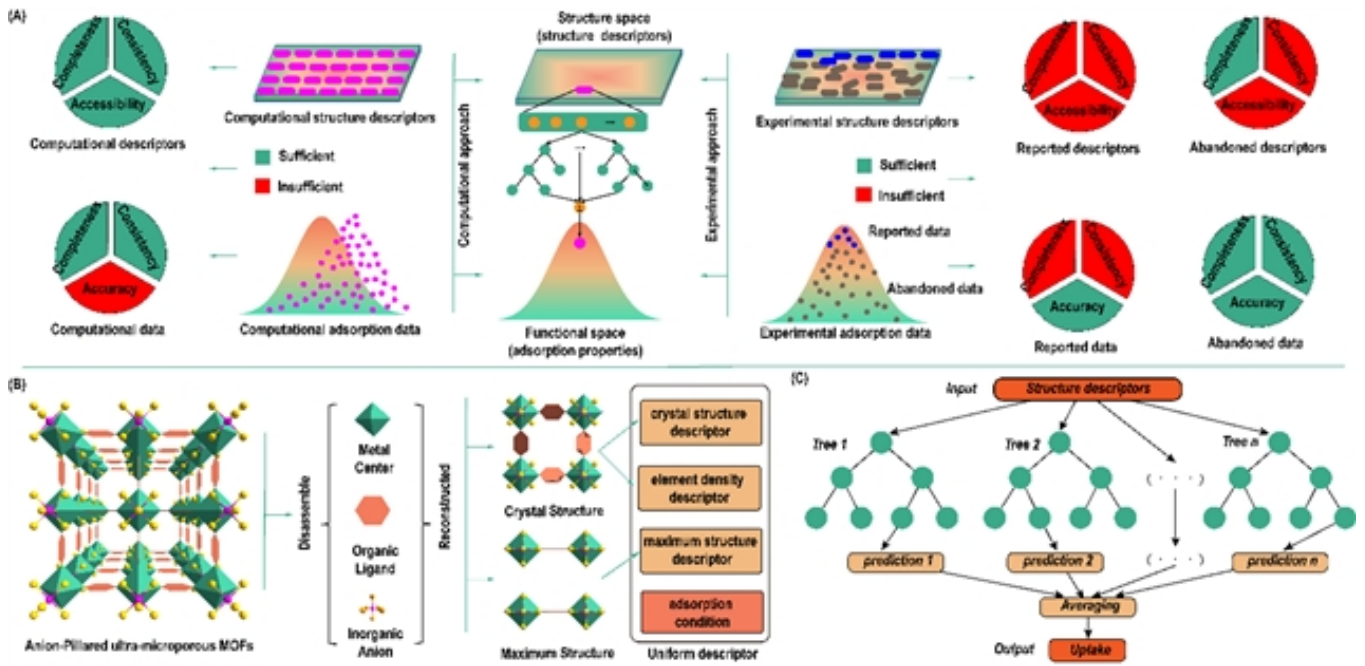


图1：机器学习 workflow 示意图

首先收集课题组过往研究过程中废弃的实验吸附等温线，并进一步收集已发表实验吸附等温线数据作为补充。进一步通过计算化学方法计算材料结构特征描述符。对于每一个金属有机骨架材料，计算其三种结构特征描述符，包括1) 晶体结构特征描述符，用于描述材料孔道结构特征；2) 化学元素密度，用于描述材料表面化学特征；3) 最大结构特征参数，用于描述材料柔性特征。采用随机森林算法构建机器学习预测模型，以识别材料结构特征描述符和材料吸附数据间的相关性。

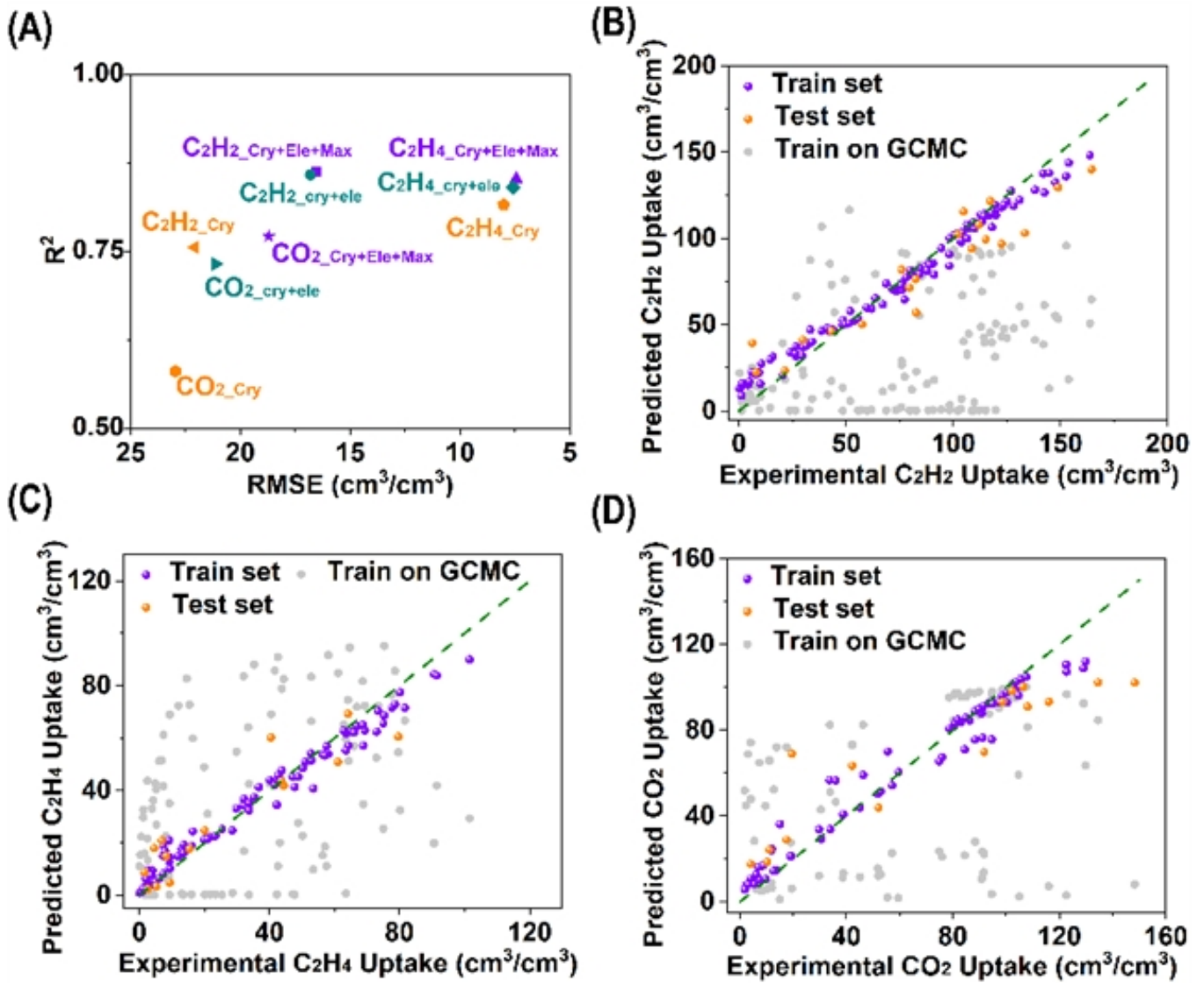


图2：机器学习模型对阴离子柱撑超微孔材料乙炔、乙烯、二氧化碳吸附性能预测效果

对训练集和测试集的预测结果表明最大结构特征参数的引入能显著提高机器学习模型的预测精度。同时训练集和测试集的预测结果都与实验数据吻合良好，这表明所构建的机器学习模型能够捕获特征描述符与吸附性能之间的相关性。不同训练数据集间的对比可以发现基于实验数据的预测精度要显著高于基于计算模拟数据的预测精度，且废弃数据的引入可以显著地提高机器学习预测精度。

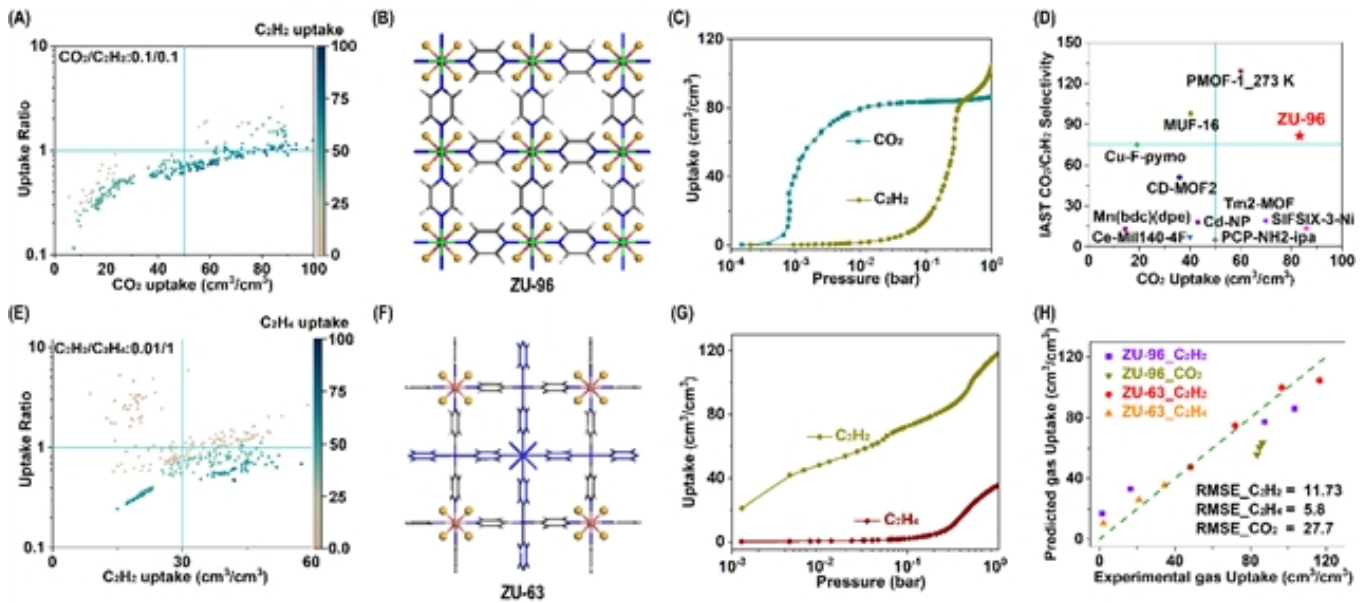


图3：机器学习辅助探索具有最佳吸附分离性能的阴离子柱撑超微孔材料

进一步结合材料基因组原理虚拟构建出300个阴离子柱撑超微孔材料，基于本文构建的高精度机器学习模型预测材料对乙炔、乙烯、二氧化碳的吸附性能，根据材料吸附性能数据和分离性能数据筛选并合成出两种兼具高选择性和高吸附容量的吸附分离材料ZU-96和ZU-63。其中，ZU-96材料具有标杆性的二氧化碳/乙炔分离性能，在298 K和0.1 bar条件下，其二氧化碳吸附量达到了83.2 cm<sup>3</sup>/cm<sup>3</sup>，二氧化碳/乙炔IAST选择性达到了81.5。

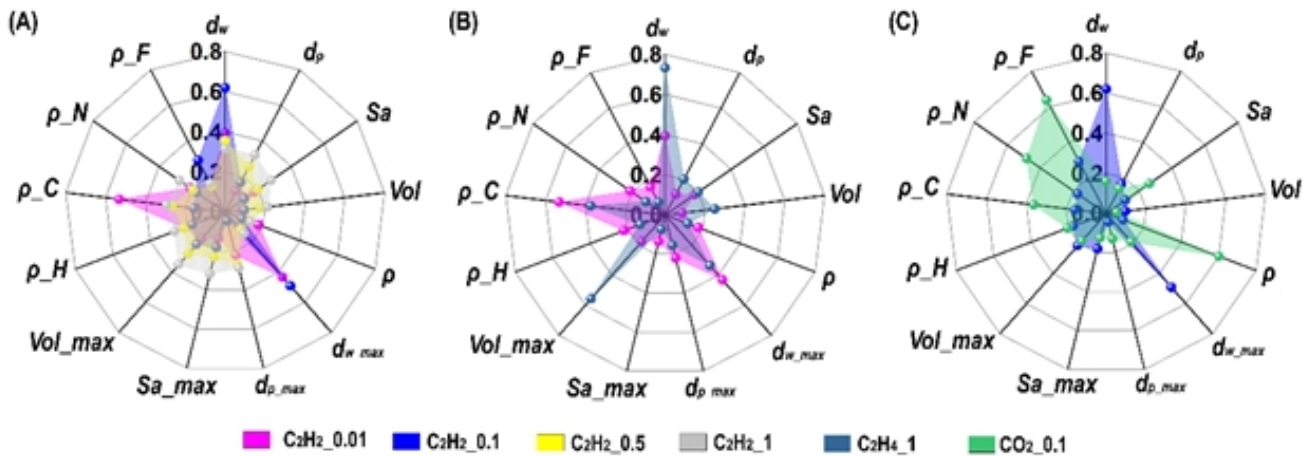


图4：阴离子柱撑多孔材料结构特征对材料气体吸附性能的相对重要性

进一步基于随机森林算法定量计算了阴离子柱支撑多孔材料结构特征对材料乙炔、二氧化碳、乙烯吸附性能的影响权重，为后续设计出具有理想吸附分离性能的吸附剂提供了更精确、直观的指导。

该研究工作表明废弃数据是实现AI精准预测材料实际性能的有效拼图。基于该研究发现，论文作者呼吁材料科学家们在未来发表的文献中提供更多的材料吸附数据，特别是那些在材料设计过程

---

中被遗弃的材料数据，这些数据将极大加快材料的发展。（来源：科学网）

相关论文信息：<https://doi.org/10.1016/j.matt.2022.07.029>

作者：邢华斌等 来源：《物质》

更多 科学进展 请访问 <https://www.iikx.com/news/progress/>

本文版权归原作者所有，请勿用于商业用途，[爱科学iikx.com](https://www.iikx.com)转发