

2026 年

“认证杯”数学中国数学建模网络挑战赛

第二阶段

A 题 水系电解液配方

电解液是电化学储能器件中的关键组成部分, 其配方会显著影响离子传导、酸碱环境、电化学稳定性以及后续器件的实际性能。一个电解液体系通常由若干溶质、溶剂及其不同比例组成, 不同组分之间往往存在非线性作用和协同效应, 因此仅凭经验进行配方筛选, 往往效率较低、成本较高。随着自动化实验与数据驱动方法的发展, 人们希望利用已有实验数据, 建立“配方—性能”之间的定量关系, 能够做到识别影响性能的关键因素, 理解不同组分之间的交互作用, 在有限实验预算下更高效地寻找优良配方, 为后续实验设计提供依据。

现有一份公开的水系电解液实验数据集, 包含 251 条实验记录。每条记录对应一种电解液配方, 并给出若干实验测量结果, 包括组成信息、导电率、pH 值以及电化学测试数据等。请你参考该数据集, 对水系电解液的配方与性能进行建模与分析。

第一阶段问题:

1. 在配方优化中, 首先需要明确“什么样的配方可以被认为是好的”。请你对数据进行分析, 建立一个合理的性能评价指标。需要在讨论中包含如下问题:
 - 单独使用电导率作为评价指标是否足够?
 - 请结合 pH 与电化学测试曲线, 讨论如何构造更合理的性能表征指标;

- 请设计一个或多个具有比较意义的稳定性相关指标, 并说明其物理或工程意义。
2. 在已知部分实验结果的基础上, 希望利用配方信息预测其性能。请你建立合理的数学模型, 由配方组成信息来预测以下的一个或多个目标量:
- 电导率;
 - pH 值;
 - 你在问题 1 中构造的性能和稳定性相关指标。

注: 如果你尝试了多种模型, 请比较不同模型在预测精度、稳定性和可解释性上的差异。分析模型在不同性能目标上(如在高性能区、低性能区或特殊样本上的表现)的适用性差异, 并讨论造成差异的原因。

3. 对配方的建模不仅要预测性能, 还应帮助研究者理解“为什么某些配方更优”。请基于数据和模型, 对以下问题进行尝试性的讨论:
- 哪些组分或特征对电导率影响最大?
 - 哪些组分或特征会更主要地影响你构造的稳定性相关指标?
 - 是否存在明显的组分交互作用或协同效应? 即某些组分单独作用不突出, 但与其他组分组合后性能显著变化。
 - 这些规律是否只在孤立的极个别样本中成立, 还是在较大范围内具有稳定性?

第二阶段问题:

1. 在第一阶段中已经建立了由配方组成预测性能的模型。请你基于第一阶段建立的性能指标和预测模型, 进一步分析模型的可信度与适用范围。需要在讨论中涉及如下问题:
- 随机划分训练集和测试集是否足以评价模型对电解液性能的预测能力?
 - 是否可以根据配方结构、样本聚类结果、主要组分或配方复杂度, 设计更合理的验证方式来验证模型的预测能力?

- 模型在不同配方区域中的预测误差是否一致?
 - 模型在哪些区域的预测结果相对可信, 哪些区域的预测可信度较低?
2. 在实际研究中, 实验资源通常是有限的。已有的实验记录能够帮助研究者判断哪些区域值得进一步探索, 但仍然不能覆盖所有可能的配方组合。请你基于已有数据、性能指标和预测模型, 设计一个下一轮实验的候选方案。包括如下问题:
- 如果只能新增少量实验, 例如 5 组或 10 组配方, 你会优先选择哪些候选配方或配方区域?
 - 你的方案更偏向于在已有高性能区域附近继续优化, 还是更偏向于探索当前数据覆盖不足的区域? 是否有某些原则来平衡二者?
 - 你将如何利用模型预测值、预测不确定性、样本分布密度或组分差异来指导下一步的实验配方? 请在至少同时考虑到电导率、pH 值和稳定性指标的条件下, 设计下一轮实验候选方案。
 - 你的方案与随机选点(或其他值得对比的简单实验设计方案)相比有哪些优势与风险?
3. 一个配方即使在已有数据中表现较好, 也未必适合进一步开发。例如若该配方对组分比例的小幅变化非常敏感, 这可以称为稳健性太差, 其实用价值通常有限。请你建立合理的数学模型来回答这些问题:
- 你推荐的候选配方是在相对稳定的高性能区域内, 还是只是少数孤立样本? 是否存在“性能较高但对扰动相当敏感”的候选配方?
 - 当组分比例发生小幅扰动时, 你的模型预测的准确度是否仍能保持较好水平?
 - 如何定义和预测候选配方的稳健性?(可以从附近样本表现、模型对扰动的敏感性、配方的复杂度等角度来入手讨论)