

大语言模型在科学研究中应用指导手册

吴思远 李泓[†]

(中国科学院物理研究所, 北京 100190)

S1. 引言

本文是《大语言模型在电池科研全流程应用的测评与无机固态电解质综合数据库构建》的补充材料, 用于复现文章中所有的内容, 也希望作为初入语言模型电池研究人员的入门手册; 同时也可以作为物质科学其他领域使用语言模型协助科研过程的入门手册。

本文的语言模型均以中国科学院凝聚态物质科学数据中心开发的 MatElab 平台为介绍 (以下简称 MatElab 平台), 主要以 DeepSeek R1 蒸馏的 70B/QwQ 32B (以下简称为模型) 介绍。

本手册所涉及到的内容与正文实验部分完全一致, 主要从语言模型的知识储备、实时学习、指令遵从和结构化输出能力四个角度阐明如何利用语言模型辅助科研, 同时以文本分类、信息抽取、文档问答和推理四个实例来展示过程。

本手册第二节为模型的知识储备、实时学习、指令遵从和结构化输出能力的介绍; 第三节为文本分类、信息抽取、文档问答和推理四个实例; 第四节为 MatElab 集成的有关电池的其他功能。对于其他领域, 只需要将电池部分换成对应领域即可。

S2. 概念解释

S2.1 知识储备

知识储备是测试模型在预训练和微调过程中训练语料的质量, 主要考察模型对知识的学习程度。如模型是否知道固态电池的定义, 能否判断该体系是否属于固态电池。

S2.2 实时学习

实时学习是指针对模型没有学习过的概念, 通过上下文的提示或者例子来让模型快速学习的能力, 主要考察模型泛化能力。例如对于固态电池的定义在业界和学界尚未统一, 通过上下文的实时学习能力来提升模型的知识。

S2.3 指令遵从

指令遵从是指通过提示的方法让模型按照用户的思路解决问题或者生成指定格式的内容, 特别是需要逻辑思考和推理的问题上, 一个具备指令遵从能力的模型可以通过提示的方

法优化思考内容，提升推理能力。

S2.4 结构化输出

结构化输出和指令遵从的内容相似，唯一的区别是结构化输出增加了输出结构的要求，特别在信息抽取、函数调用和 MCP(模型上下文协议)拥有重要应用。

S3. 实例

S3.1 体验地址

本指导手册中的体验地址在 MatElab 平台均可以使用，中国科学院物理研究所内师生可以直接在 <https://in.iphy.ac.cn/eln/#/> 登录，其他科研和教育用户均可以在 <https://matelab.iphy.ac.cn/signup.php> 注册后登录。

入口：在右上角”chat”位置选择后进入（图 S1）



图 S1 操作入口

S3.2 知识储备能力测试——以文本分类为例

为了考核模型的知识储备能力，我们只需要将内容和分类标签按照一下格式输入即可。如图 S2 所示，我们需要定义出一些可能的标签，并让模型在该组类别中选择。



图 S2 模型知识储备能力测试和对话模型文本分类方法介绍

S3.3 实时学习能力测试——以文本分类为例

在 S3.2 中，我们让模型在无机固态电解质、聚合物电解质、复合电解质、其他四个类别中判断，但是对于复合电解质、其他模型可能会有歧义，特别是“其他”类别，可能是不属于以上三类别的固态电解质，也可能是电池其他部分。因此我们需要通过提示的方法来修正该内容，让模型实习学习“其他”类别的范围，既可以让“其他”定义为不属于以上三类别的固态电解质，也可以将“其他”定义为电池其他部分，详细参考图 S3。



图 S3 模型实时学习能力测试

S3.4 结构化输出能力测试——以提取固态电解质离子电导率为例

对于一些信息抽取的内容，需要让模型以固定格式输出，常见的是 json 格式，同时结构化输出也有利于模型调用外部函数或者支持 MCP 协议的服务。在图 S4 中，我们让模型提取出数据中的离子电导率数据。



图 S4 模型结构化输出能力测试和信息提取实例

S3.5 指令遵从能力测试——以计算一元二次方程为例

除了 S3.4 属于指令遵从外，本节和 S3.6 节会介绍其他形式指令遵从。

考虑到大部分模型对求解一元二次方程已经十分容易（对于推理模型该功能仅为演示，见图 S5），本小节仅作介绍，更详细指令遵从参考 S3.6。以下为求解一元二次方程示例。

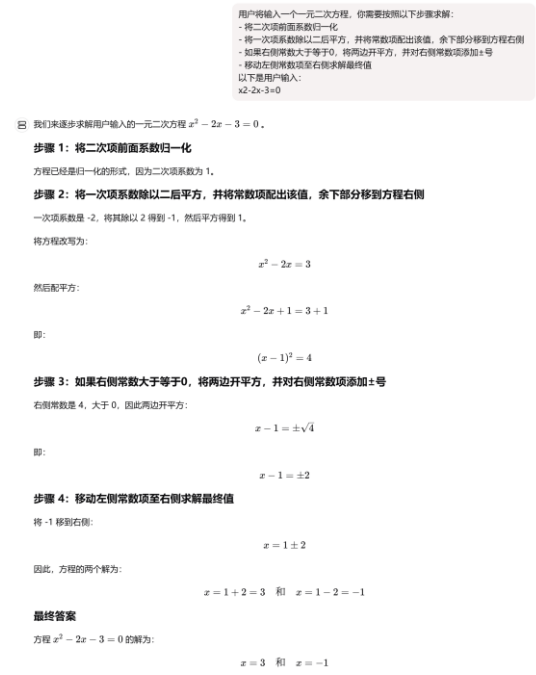


图 S5 模型指令遵从能力测试

S3.6 指令遵从能力测试——以总结为例

除了让模型按照我们的思路求解问题外，也可以让模型按照指定模板生成固定格式，以总结文章为例，我们可以让模型按照指定模板输出（见图 S6）。

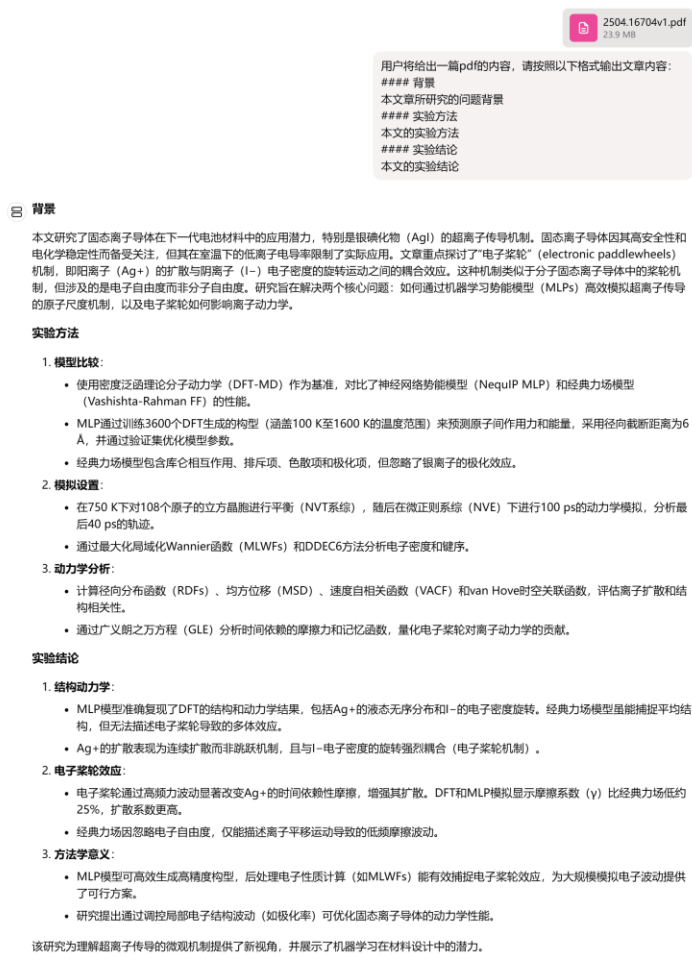


图 S6 模型指令遵从能力测试，以总结文章为例

S4. MatElab 其他与电池相关的内容

S4.1 文本分类

在“文献推荐”栏目里选择“我的订阅”并选择“电池”即可找到模型分类出电池相关的文献（图 S7）。

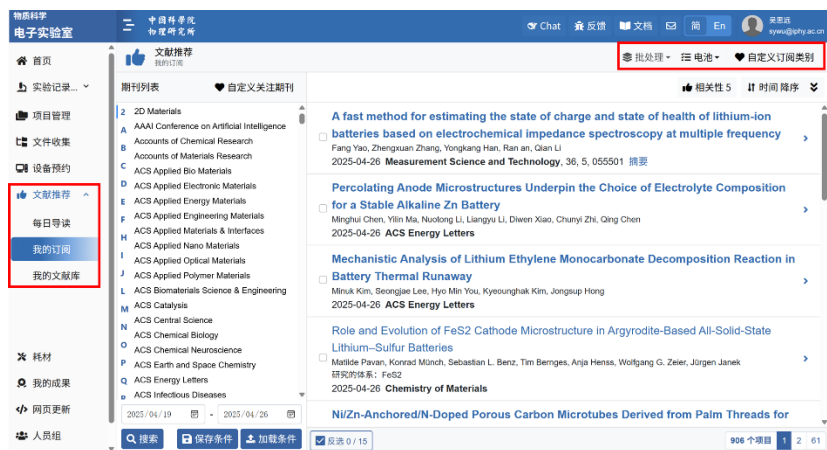


图 S7 电池最新文献入口，左侧选择“我的订阅”后右上角选择“电池”

S4.2 每日导读（公共版）

在“文献推荐”栏目里选择“每日导读”并选择“电池”即可找到生成的电池文献总结。

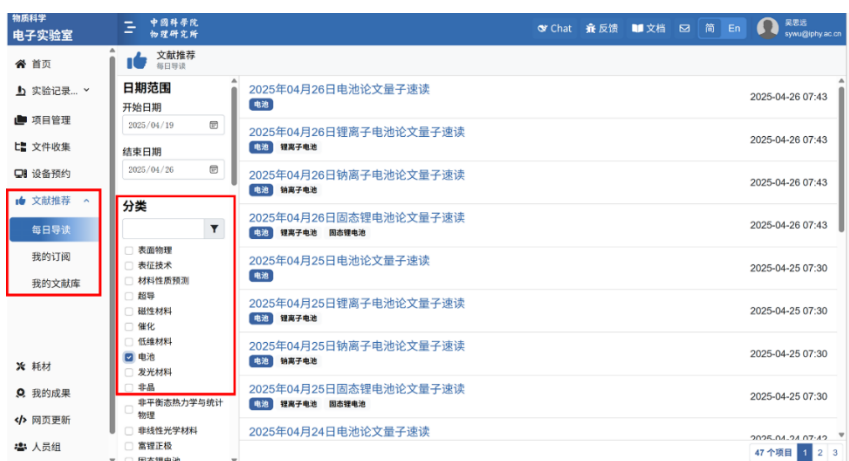


图 S8 电池每日导读入口，左侧选择“每日导读”后选择“电池”

S4.3 每日导读（个人版）

打开 chat，点击“通用回答”切换至“智能助手”后选择“每日导读个人版”，将“个人简历”、近两年发表论文或者收藏的前 10 篇摘要内容复制到对话框输入后等待 1-2 分钟即可生成（图 S9）。

同时，每日导读个人版支持调用记忆功能生成文献日报（见图 S10）。我们在设置中添加我们的记录并开启，选择“每日导读个人版”后只需要输入“推荐我的个性化文献”即可。



图 S9 每日导读 (个人版) 入口, 点击切换到“智能助手”后输入“个人简历”、近两年发表论文或者收藏的前 10 篇摘要内容后等待 1-2 分钟即可

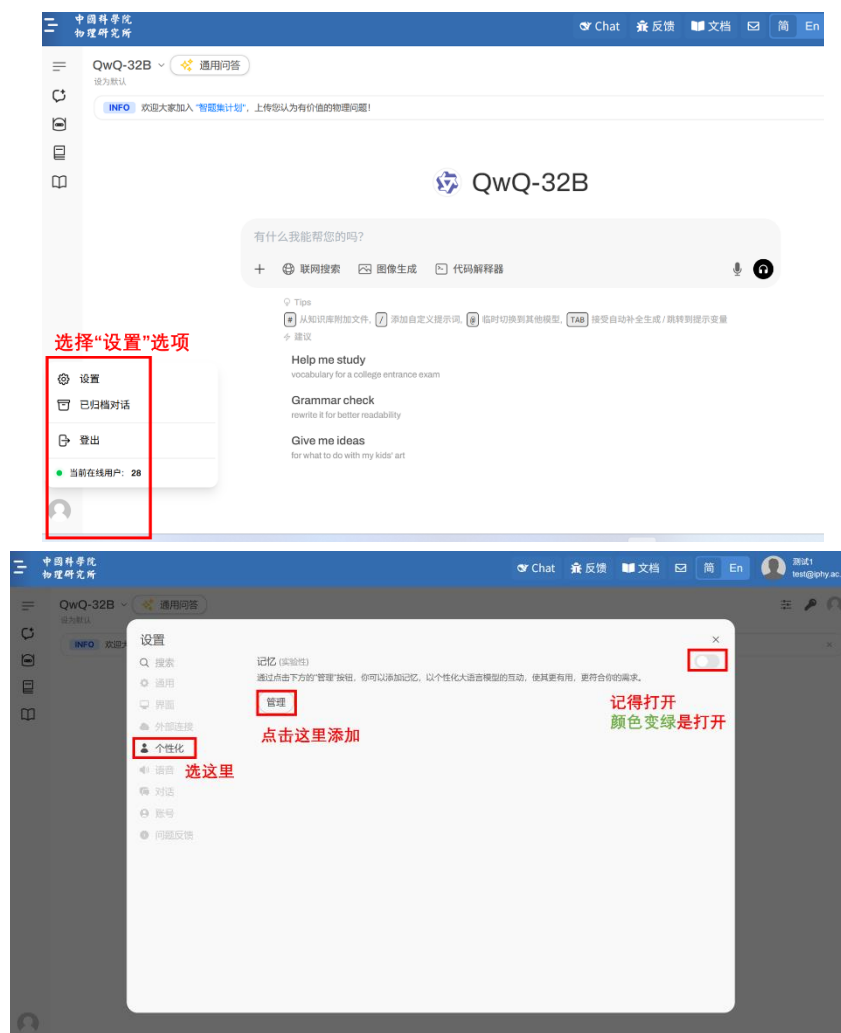




图 S10 通过“记忆”功能配置每日导读个人版

S4.4 电池知识库和 Deep Research 功能

打开 chat, 点击“通用回答”切换至“智能助手”后选择“电池专家模型”或者“DeepBattery”, 将电池相关问题输入后等待回答 (图 S11)。

电池专家模型(R1) 智能助手

LiCoO2正极改性策略

电池专家模型(R1)

思考用时 17 秒

LiCoO₂作为锂离子电池的重要正极材料，在高压和高循环次数下往往面临容量衰减和结构不稳定的问题。为了改善这些缺陷，研究者提出了多种正极改性策略，主要包括以下几种方法：

1. 表面涂层
表面涂层是通过在LiCoO₂表面涂覆其他材料以改善其电化学性能的常用方法。常见的涂层材料包括氧化铝（Al₂O₃）、二氧化硅（SiO₂）、磷酸锂（Li₃PO₄）、钛酸锂（Li₄TiO₄）等。这些涂层材料能够有效减少正极与电解液之间的副反应，抑制钴的溶解，提高电池的循环稳定性和倍率性能。
例如：

- **Al₂O₃涂层**：通过溶胶-凝胶法在LiCoO₂表面形成Al₂O₃涂层，显著提高了其循环性能和界面稳定性（DOI: 10.1149/2.107311jes）。
- **SiO₂涂层**：在LiCoO₂中添加SiO₂形成LiCoO₂/SiO₂复合材料，优化了其结构稳定性和电化学性能（DOI: 10.1016/j.electacta.2005.09.021）。

DeepBattery 智能助手

LiCoO2正极改性策略

DeepBattery

思考用时 4 分钟

LiCoO₂正极改性策略的全面分析
LiCoO₂作为锂离子电池的传统正极材料，因其高理论容量、良好的可逆容量和相对成熟的制备工艺，广泛应用于商业锂离子电池中。然而，其在高压下存在结构不稳定性、循环性能差、热不稳定性等问题，限制了其在高性能电池中的应用。为了克服这些缺陷，研究者提出了多种改性策略，包括表面改性、掺杂、纳米结构设计、界面工程等。本文将从改性策略的分类、改性后材料的电化学性能、实际应用中的挑战及未来的研究方向进行分析。

一、LiCoO₂正极的改性策略
1. 表面改性
表面改性是通过在LiCoO₂表面涂覆或修饰一种或多种材料，抑制正极与电解液之间的副反应，改善其循环稳定性和热稳定性。常见的表面改性材料包括：

- **氧化锆（Li₂ZrO₃）**：Li₂ZrO₃涂层因其高化学稳定性、锂离子导电性，能够有效抑制LiCoO₂与电解液的反应，降低表面阻抗，提升电池性能（DOI:10.1016/j.electacta.2016.05.066）。

图 S11 “电池专家模型”或者“DeepBattery”，点击切换到“智能助手”后输入电池相关问题输入后等待回答