

材料领域知识嵌入的 属性分而治之机器学习 平台

V1.0

用户操作手册

文档编号：2025-R0311

最后修订日期：2025 年 6 月 18 日

版权所有 2025-。保留所有权利。

本文档中所提到的其他所有公司和产品名，均是其所有人的商标或者注册商标。

本文档内容如有变动，恕不另行通知。

目 录

关于本文档 1

第 1 章 软件概述	2
1.1 模块介绍	2
1.2 用户层级、角色与相应权限	3
1.3 登录注册界面	3
1.4 主界面和导航栏	4
第 2 章 领域核特征识别	6
2.1 数据上传	6
2.2 选择领域核特征	7
第 3 章 数据核特征识别	9
3.1 构建特征决策表	9
3.2 特征决策表聚类	9
3.3 最简决策表构建	10
3.4 决策表求核	11
3.5 正数据核特征识别	11
第 4 章 特征图构建	12
4.1 核特征组合与训练	12
4.2 最优特征团筛选	12
4.3 特征团关系可视化	13
第 5 章 多方面构效关系分析	14
5.1 特征团分析	14
5.2 构效关系提取	14
5.3 AI 构效解释	15

关于本文档

主题

欢迎使用材料领域知识嵌入的属性分而治之机器学习平台的用户操作手册，该手册包含了需要了解和使用材料领域知识嵌入的属性分而治之机器学习平台的相关知识。

本文档包含了如下章节：

- **第 1 章 软件概述**
- **第 2 章 领域核特征识别**
- **第 3 章 数据核特征识别**
- **第 4 章 特征图构建**
- **第 5 章 多方面构效关系分析**

序言中介绍了使用该文档的一些帮助信息。

读者

使用材料领域知识嵌入的属性分而治之机器学习平台的最终用户和系统管理员。

第 1 章 软件概述

本章概述 本章介绍了材料领域知识嵌入的属性分而治之机器学习平台的功能和操作界面特点。

内容	主题	页码
	模块介绍	2
	用户层级、角色与相应权限	3
	登录注册界面	3
	主界面和导航栏	4

1.1 模块介绍

构效关系是揭示材料微观结构与宏观性能之间因果联系的核心桥梁，准确解析构效关系对于推动新材料的设计与机制理解具有重要意义。然而，仅依赖单一描述符组合构建机器学习模型难以准确有效挖掘多方面材料构效关系，尤其是在面对多尺度结构信息与属性耦合特征时，限制了构效关系的深层次挖掘与解释。为此，我们研发了融合材料领域知识的属性分而治之机器学习平台。该平台将领域知识与粗糙集理论、特征选择算法以及多类型机器学习模型深度集成，旨在从多种结构维度与描述符组合中，系统挖掘影响材料性能的关键结构特征与决策规则，构建可解释的多元构效映射。该平台致力于从复杂材料属性数据中深入挖掘材料结构、工艺参数与性能表现之间的本质关联机制。平台通过嵌入材料专业知识的特征选择方法，联合粗糙集理论中的求核算法，有效筛选出影响模型性能的核心描述符（特征核）与协同特征群体（特征团），提升了建模过程的可解释性与稳定性。进一步地，平台利用多种主流机器学习模型的建模原理，从高精度预测结果中提取可转化为材料设计逻辑的多元决策规则，结合与材料构效机制相关的知识图谱，实现了材料性能影响因素的多角度、多层次分析。该平台采用 B/S 架构，前端基于 Vue.js 与 Element Plus 框架开发，后端采用 Spring Boot 技术架构，并集成 MySQL 数据库，实现从领域核特征识别、数据核特征识别、特征团构建、多方面构效关系分析的全流程处理；平台内置多种领域知识嵌入方法，支持数据核与领域核特征的联合建模，结合 GPR、SVR、KNN、MLR 等模型进行分析与模型优化选择。同时在构效关系分析模块提供特征贡献分析、模型性能可视化、多视角结构关系挖掘、AI 大模型分析等模块，助力研究人员深入理解材料性能影响机制。

具体来说，材料领域知识嵌入的属性分而治之机器学习平台设计与实现了如图 1-1 所示的领域核特征识别、数据核特征识别、特征团构建与多方面构效关系分析这四大模块。

- （1）领域核特征识别模块
该模块结合材料化学与结构先验知识，对晶体结构、元素属性等信息进行解析与筛选，提取关键物理化学特征，形成领域特征库，为后续数据建模提供有针对性的知识支撑。
- （2）数据核特征识别模块
该模块通过对已知属性的实验数据与计算数据进行统计分析，自动挖掘显著性高、相关性强的数据驱动特征，补充领域核信息，增强模型输入的多样性与代表性。
- （3）特征团构建模块

该模块综合领域核与数据核特征，采用自研的分而治之策略，形成多维度、不同层次的特征组合（特征团），以捕捉复杂的非线性特征交互效应，提升预测模型的泛化能力。

（4）多方面构效关系分析模块

该模块对建模结果进行可视化与多角度解读，通过贡献度分析、关系网络挖掘与聚类分析等手段，揭示各特征对 BVSE 能量势垒值的综合影响机制，助力性能调控与新材料筛选。

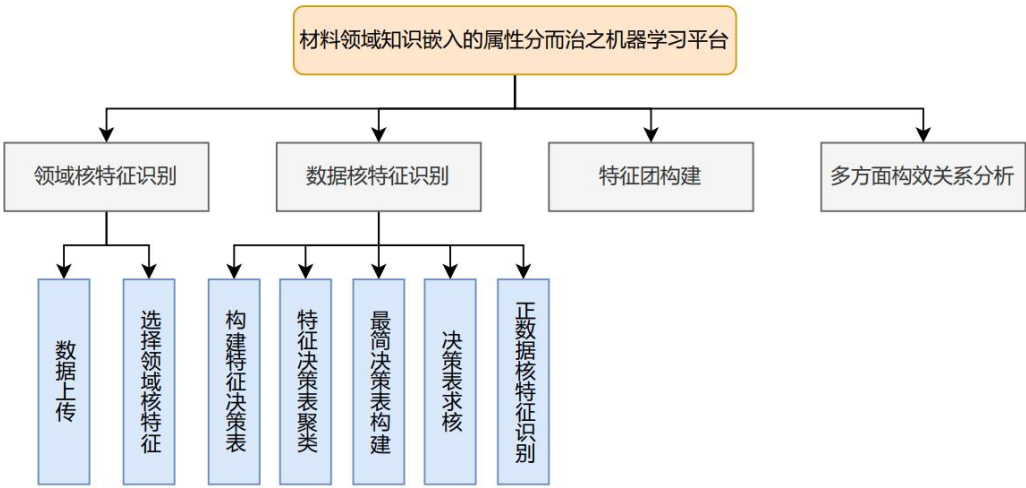


图 1-1 系统软件架构

1.2 用户层级、角色与相应权限

为保证平台使用安全与管理方便，本系统设定了单一用户层级，默认角色为开发者。

表 1-1 用户层级

用户层级	用户名	密码
开发者	123	111

1.3 登录注册界面

如图 1-2 所示是材料领域知识嵌入的属性分而治之机器学习平台的初始登录页面。
对于用户，需选择登录方式为统一认证号，输入账号密码（表 1-1），以登录进入主界面。

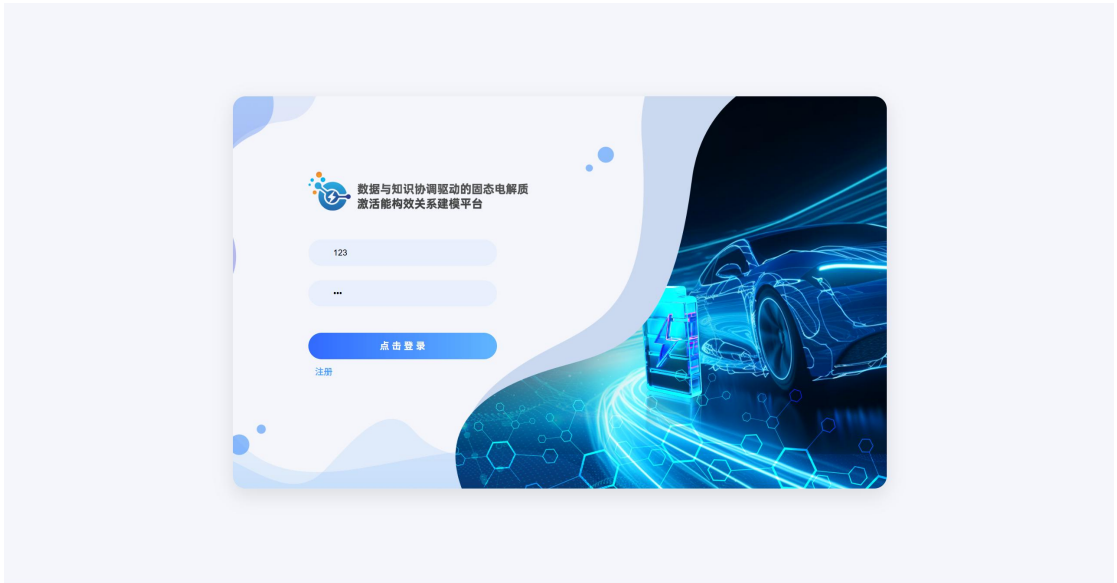


图 1-2 登录界面

如图 1-3 所示是材料领域知识嵌入的属性分而治之机器学习平台的注册页面。注册页面提供了用户名邮箱等用户信息，用户按提示输入后即可注册账号。

注册

* 用户名

* 邮箱(登录账号)
请输入合法的邮箱

* 登录密码

* 密码确认

用户信息(简介)

手机号

职业

研究领域

研究方向

* 研究机构/工作单位

* 国家

图 1-3 注册界面

1.4 主界面和导航栏

平台的主界面整体设计科技感十足，背景融合了新能源电动汽车和电解质材料元素，体现了研究的应用场景与前沿方向。首页中央清晰展示了平台的名称及入口按钮，方便用户一键进入核心功能模块。

用户在首页点击进入后，将跳转到平台的主要功能操作界面（如图 1-5 所示）。该页面以模块化形式展示了“领域核特征识别”、“数据核特征识别”、“特征团构建”和“多方面构效关系分析”等主要功能，配以示意图和文字说明，帮助用户快速理解每个模块的作用与执行流程。

同时，页面顶部的导航栏包含“首页”、“退出”、“关于我们”等选项，用户可以随时返回首页、退出登录或查看平台的相关信息介绍，界面布局合理、操作便捷，显著提升了用户体验。

验。



图 1-4 主界面

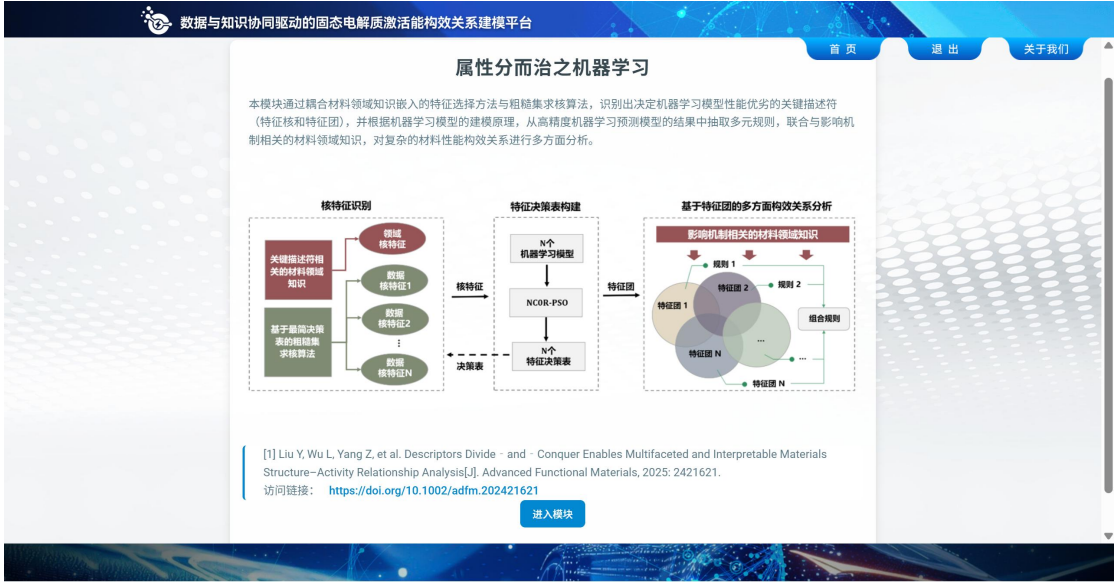


图 1-5 导航栏

第 2 章 领域核特征识别

本章概述

本章主要介绍了领域核特征识别模块的使用流程，包括数据上传、属性信息浏览、领域特征选择及结果查看，帮助用户快速筛选出对预测最有价值的领域核特征。

本章内容

主题	页码
数据上传	6
数据用户信息	7

2.1 数据上传

本模块支持用户通过上传本地数据文件的方式，将实验属性的各项条件属性及目标属性导入平台。系统目前仅支持 .xlsx 格式的文件，用户可点击“选择文件”按钮进行文件选取，并通过“上传”按钮将数据上传至服务器，平台会自动解析文件内容并在页面中以表格形式进行可视化展示，方便用户核对数据的完整性与准确性。同时，系统提供数据模板下载，用户可参考模板格式准备待上传的数据文件。数据内容信息上传完成后点击“下一步”。



图 2-1 数据上传与数据集展示界面



图 2-2 属性统计信息查看界面

2.2 选择领域核特征

上传数据后，用户可在“选择领域核特征”区域根据具体需求勾选与研究目标相关的领域核特征。平台预置了多种可选的特征，包括占据数、离子半径、电负性、价电子数等多种材料学属性，用户可单选或多选组合以用于后续的特征团构建与模型训练。选择完成后，点击“提交”按钮即可确认选择，系统将根据所选特征生成对应的候选核特征组合，并在后续步骤中用于训练和筛选最优模型。



图 2-3 领域核特征选择界面

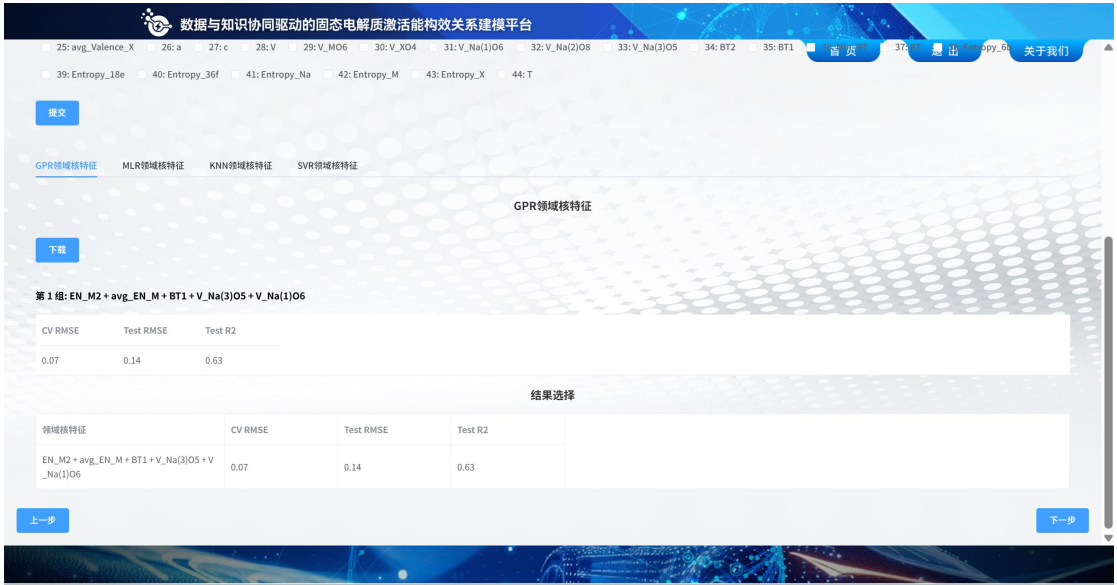


图 2-4 领域核特征组合与模型性能结果界面

第 3 章 数据核特征识别

本章概述 本章围绕属性分而治之平台的特征决策表构建与核特征识别展开介绍，包含特征决策表的生成、聚类、简化、求核以及正数据核特征识别等功能。通过本章内容，用户可以高效筛选材料性能相关的核心特征，为后续构效关系分析和模型预测提供支撑。

主题	页码
构建特征决策表	9
特征决策表聚类	9
最简决策表构建	10
决策表求核	11
正数据核特征识别	11

3.1 构建特征决策表

本步骤是特征处理流程的起点。在完成领域核特征和数据核特征识别后，系统将两类特征进行组合，形成包含候选特征子集与对应决策属性的完整特征决策表。用户可通过界面选择预训练数据或直接调用训练，选择不同模型（如高斯过程回归、支持向量机等），系统会根据所选模型输出每个候选特征组合的打分值、误差（RMSE）、决定系数（R²）及计算耗时等信息，帮助用户初步筛选较优特征组合，并可导出结果用于后续处理。



图 3-1 构建特征决策表界面

3.2 特征决策表聚类

在完成候选特征组合生成后，本步骤对特征决策表中的各个特征子集进行聚类分析，旨

在发现相似或冗余的特征组合。通过聚类结果，系统可有效识别信息高度相似的特征子集，将其归为一类，减少后续建模过程中的重复计算，提高特征选择的效率与结果的可靠性。界面中会显示每一类聚类后的代表性特征组合及对应的性能指标，用户可根据聚类结果进行筛选和下载。



图 3-2 特征决策表聚类结果界面

3.3 最简决策表构建

在完成聚类并筛除冗余后，本步骤自动对每个聚类类别内部进行特征优选，通过启发式算法或基于规则的最简化策略，提取对目标决策属性影响最大的核心特征，构建最简决策表。最简决策表保留了必要且充分描述决策信息的最小特征子集，有效避免了因高维特征带来的过拟合及模型复杂度增加问题。界面展示以 0/1 矩阵形式直观显示最简特征是否被选中，并支持结果下载。



图 3-3 最简决策表构建结果界面

3.4 决策表求核

在获取最简决策表后，本步骤进一步对其进行核求解操作。决策表求核是粗糙集理论中的重要步骤，用于挖掘出在不同条件属性下，始终保持一致的最小核特征子集。通过求核，系统可识别出不受其他属性干扰的稳定特征，对特征团构建及后续的可解释性分析至关重要。操作界面中可显示求核后的结果，以数组或特征索引的形式呈现，方便用户查看和应用。



图 3-4 决策表求核界面

3.5 正数据核特征识别

在完成核求解后，本步骤结合正数据（例如已知性能较好的属性）对核特征进行验证与筛选，从而最终识别出对材料性能提升最具指导价值的正数据核特征。这一过程保证了提取出的核特征不仅在理论上具备稳定性，还能在实验或实际数据中得到验证，提升模型的预测准确性和泛化能力。该步骤最终输出的正数据核特征可直接用于后续特征团构建、多方面构效关系分析及最终模型训练。



图 3-5 正数据核特征识别界面

第 4 章 特征图构建

本章概述

本章主要介绍平台在数据核特征识别的基础上，如何通过聚类、最简化及可视化等操作构建特征团。通过多轮组合与筛选，提取出与目标性能高度相关的核心特征群，并以交集分析等形式对结果进行可视化，为后续建模提供最优的特征组合支持。

本章内容

主题	页码
核特征组合与训练	12
最优特征团筛选	13
特征团关系可视化	13

4.1 核特征组合与训练

如图 4-1 所示，平台在确定了数据核特征、关键领域核特征后，进行多组组合训练，自动生成组合核特征，并计算每组的拟合分数、RMSE、CV_RMSE、R² 等指标，方便用户选择表现最优的特征组合。该步骤保证了多维信息的有效整合，提高了模型的泛化能力。



图 4-1 核特征组合与训练界面

4.2 最优特征团筛选

如图 4-2 所示，系统基于多个组合核特征的训练结果，按照拟合分数及误差指标自动筛选出表现最优的特征组合，形成最优特征团。用户可以清晰查看各组合的具体特征 ID、数量及对应评价指标，并支持一键选用最优组合进行后续分析。

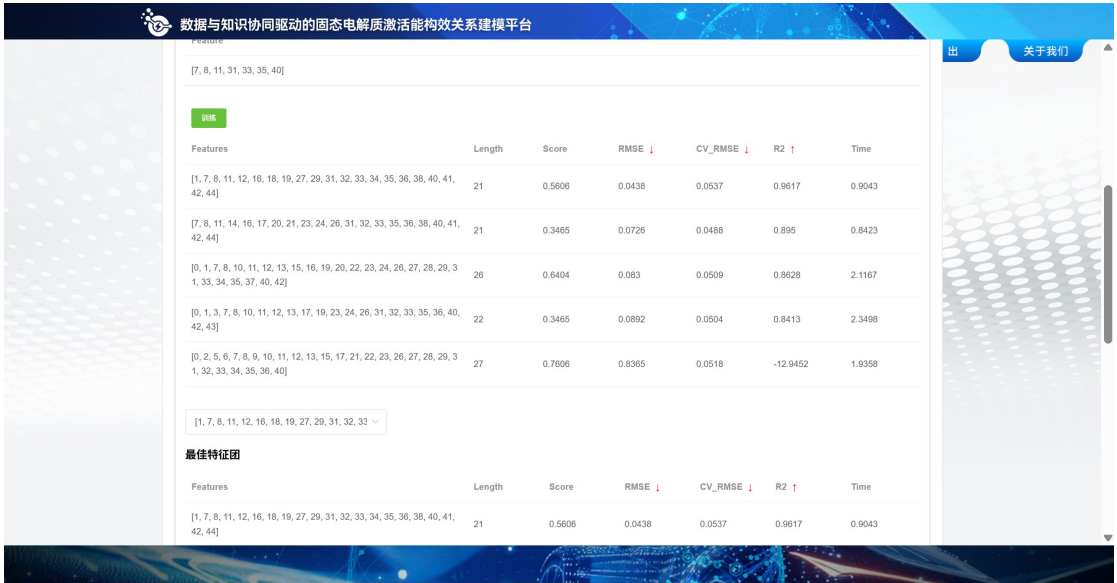


图 4-2 最优特征团筛选与展示界面

4.3 特征团关系可视化

如图 4-3 所示，平台将选出的最优特征团与所有核特征的关系用 Venn 图直观展示，帮助用户理解各特征之间的交集、差异及贡献度。同时，配套表格列出了各特征 ID 与含义，为后续解释性分析和领域知识验证提供依据。

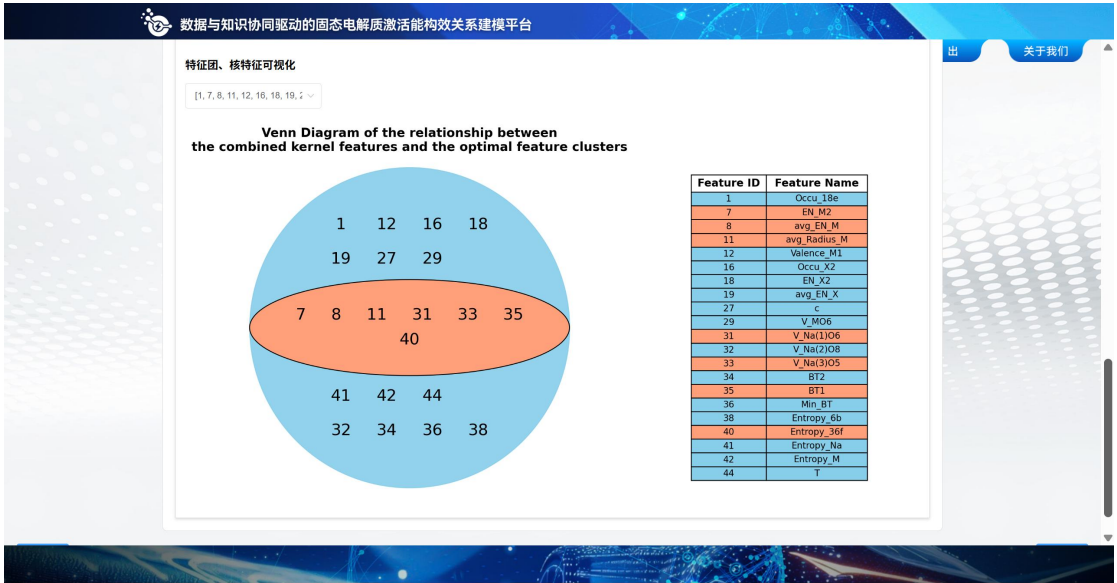


图 4-3 特征团与核特征关系可视化界面

第 5 章 多方面构效关系分析

本章概述

本章主要介绍平台在属性和属性分而治之分析完成后，如何对已识别出的特征、核特征及其组合进行多维度、多方法的构效关系分析。通过多模型回归、构效关系提取及 AI 解释等功能，全面探索材料结构、能带、电子态等属性与激活能之间的相互作用及其贡献，辅助研究者从多角度理解材料性能的关键影响因子，为后续材料设计与性能优化提供参考依据。

本章内容

主题	页码
特征团分析	14
构效关系提取	15
AI 构效解释	15

5.1 特征团分析

如图所示，平台采用多元线性回归（MLR）、KNN、SVR、GPR 等多种模型对各特征及其组合的重要性进行定量分析。可视化分析图展示了不同特征对目标值（如激活能）的正负贡献及贡献度大小，便于用户快速把握哪些属性是提升材料性能的核心驱动因子。



图 5-1 Multiple Linear Regression 分析图

5.2 构效关系提取

在完成回归分析后，平台基于内置的构效关系规则库，自动提取属性中符合规则的典型构效关系。用户可在列表中查看每条关系对应的关键特征、影响趋势（正向/负向）及关联解释，支持一键导出和进一步验证。



图 5-2 关键构效关系总览

5.3 AI 构效解释

为了提高构效关系的可解释性与可追溯性，平台集成了 AI 分析助手，自动对提取出的典型构效关系生成中文文字解释，阐述属性变化对材料性能（如离子迁移、电子云分布、结构稳定性等）的物理化学含义及作用机理，帮助研究者深入理解背后的科学机制。



图 5-3 AI 分析结果