



第六单元

AI在材料研究中的应用

31.机器学习、神经网络、大语言模型、智能体简介；

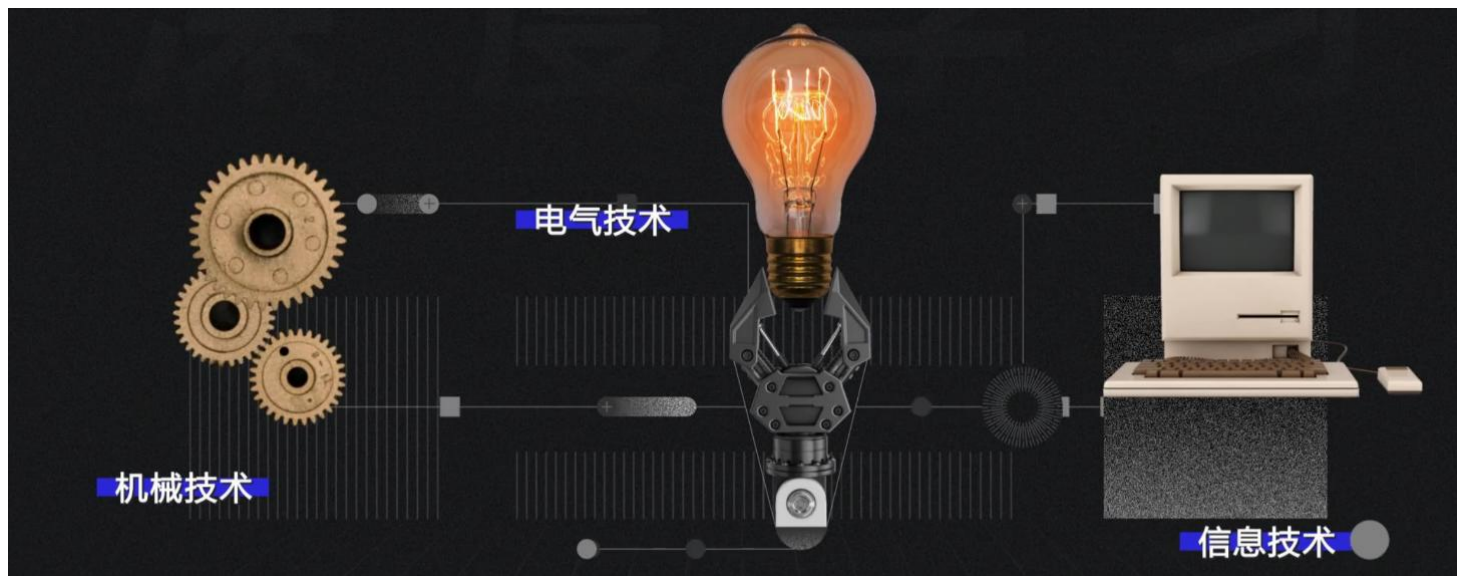
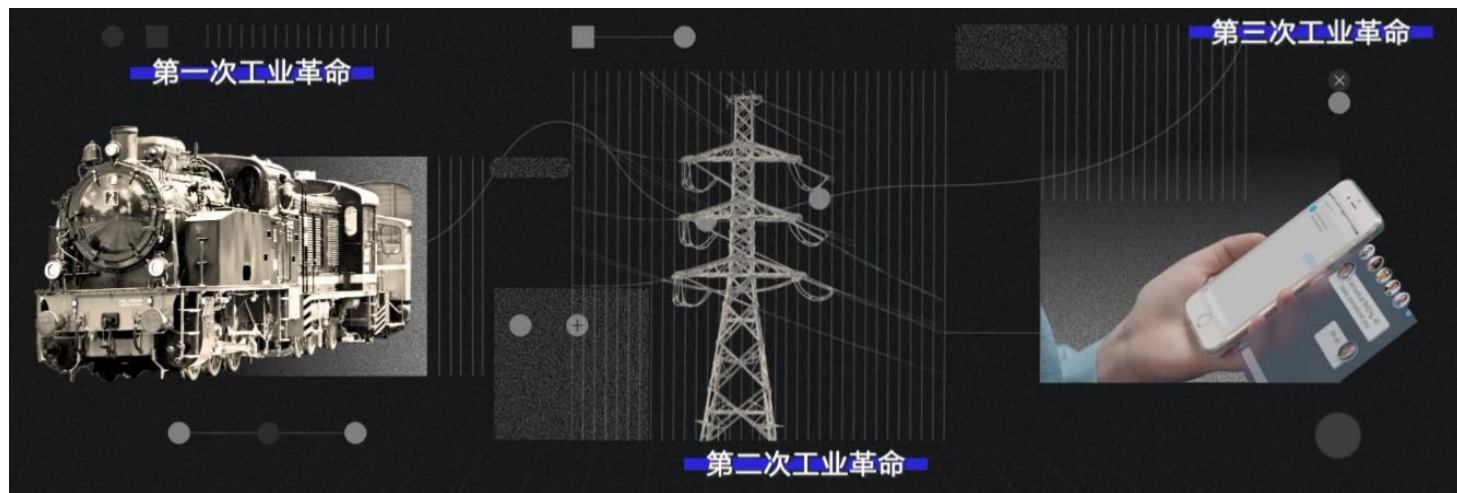
32.AI专家系统：IntelliMat；

33.机器学习建模：材料物性预测；

34.VibeCoding、Claw

35. A-Lab

31. 机器学习、神经网络、大语言模型、智能体



AI在材料研究中的应用——机器学习

机器学习 (Machine Learning, ML)

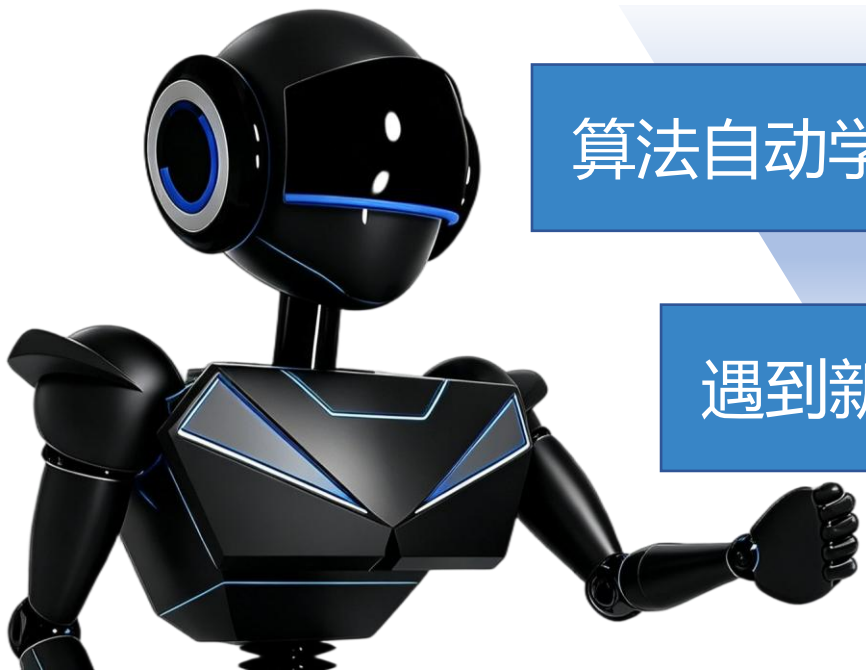
人工智能的重要分支，可从数据与经验中自动挖掘规律、完成预测决策，无需人工逐条编写执行规则。

机器学习流程

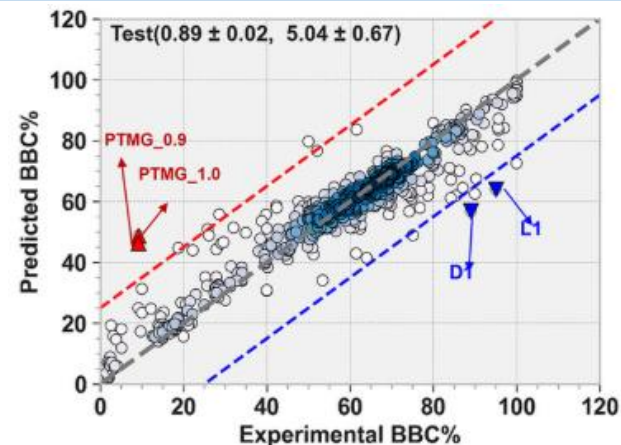
输入原始数据 + 对应结果

算法自动学习数据背后的规律

遇到新数据自动输出预测结果

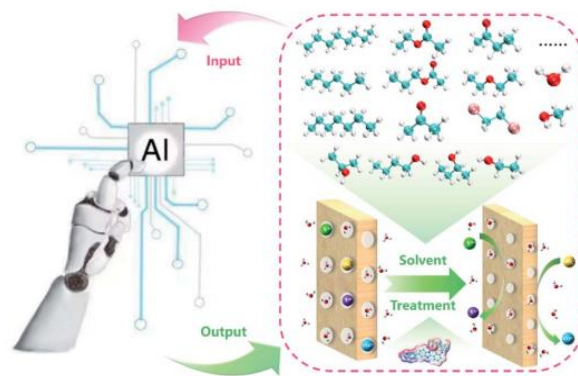


预测材料物化性质¹



1. *Macromol. Rapid Commun.* 2025, 2500054

筛选高性能材料²

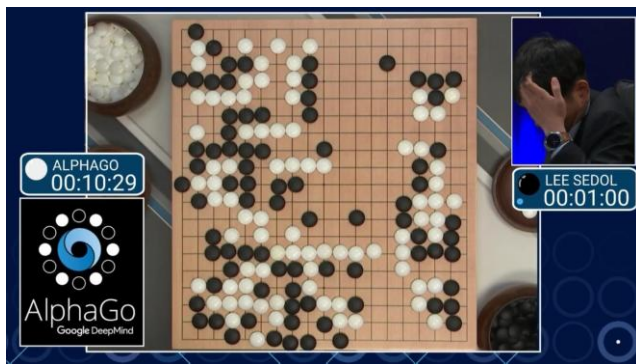


2. *J. Mater. Chem. A.* 2021, 9, 14545-14552

AI在材料研究中的应用——机器学习

AI三要素：**数据**、算法、算力

量变 → 质变

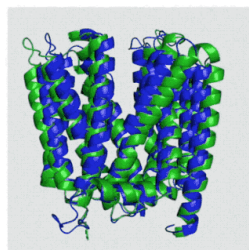


AlphaGo战胜围棋冠军

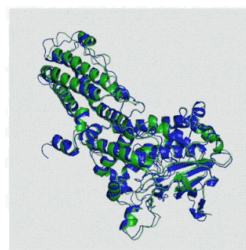
2016.3

小模型时代

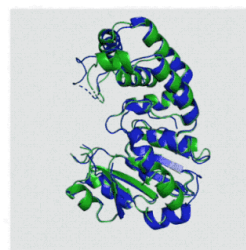
- 人工提取特征
- 模型参数量较少
- 适用范围窄



T1024/6t1z
TM-score 0.923
multidrug transporter LmrP



T1053/7m7a
TM-score 0.975
Phosphoinositide 3-kinase MavQ



T1057/7m6b
TM-score 0.933
N4-Cytosine Methyltransferase

AlphaFold 2预测蛋白质结构

2020.11

深度学习、强化学习大模型时代

- 自动提取特征
- 模型参数量大
- 特定领域能力超越人类

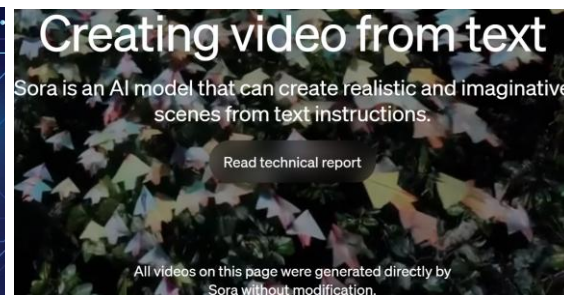


ChatGPT、Sora生成式大语言模型

2022.11

生成式、决策式大模型时代

- 覆盖领域更加广泛
- 可生成数据样例
- 在动态系统中采取行动决策



2024.2



AI在材料研究中的应用——机器学习



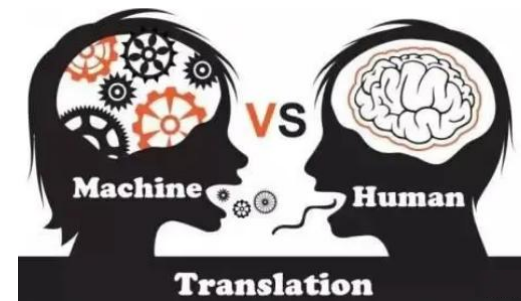
语音识别



人脸识别



智能客服



机器翻译



智能家居



智能驾驶



智慧城市



智能工业



神经网络 (Neural Network, NN)

是一种模仿生物大脑神经元连接方式的ML模型，通过多层非线性变换处理输入信息，能够从海量数据中自动学习复杂的特征和模式，是深度学习的核心基础。

不同模型架构

ANN: 基础网络，全连接。应用：简单分类 / 回归、表格数据分析。

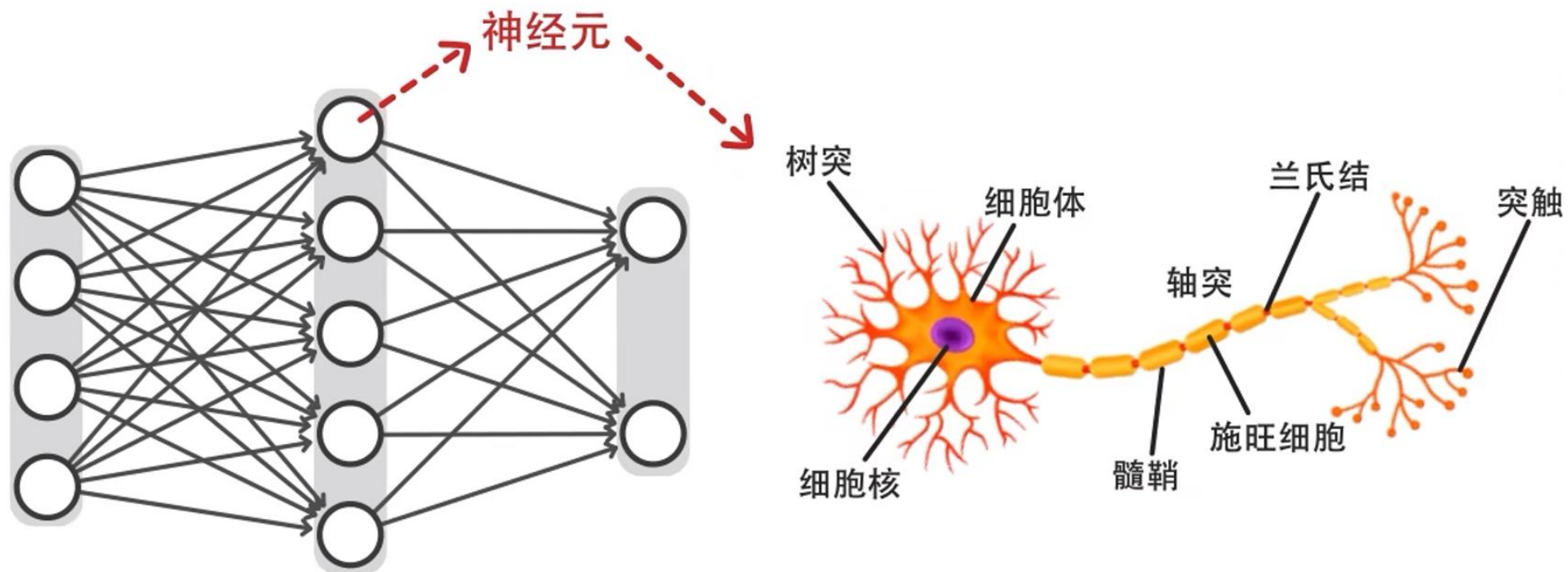
CNN: 处理图像，卷积提取特征。应用：图像识别、自动驾驶、医疗影像。

RNN/LSTM/GRU: 处理序列，带记忆。应用：时间序列预测、语音识别、机器翻译。

Transformer: 处理序列，注意力机制。应用：大语言模型 (GPT)。

GNN: 处理图数据。应用：社交网络分析、推荐系统、药物发现。

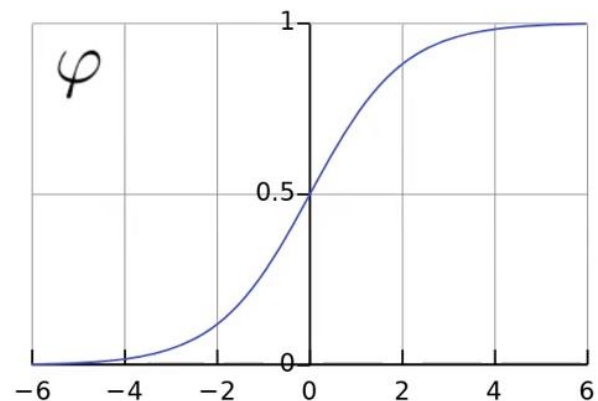
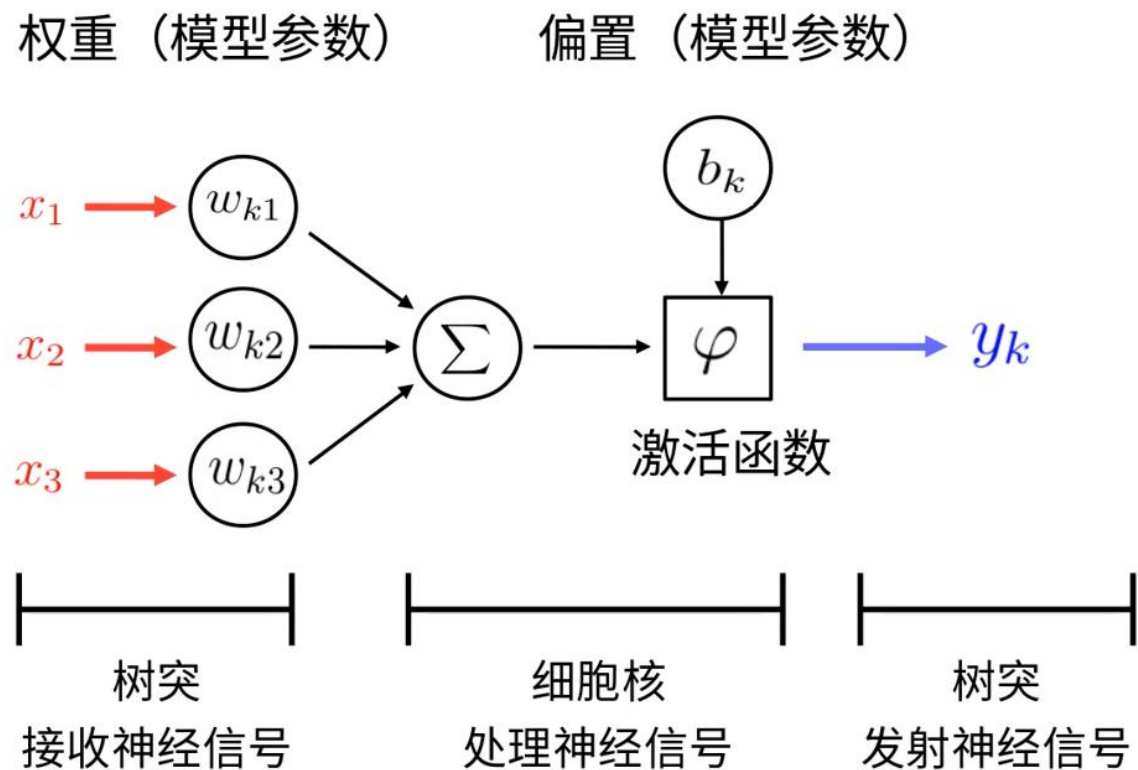
GAN/VAE: 生成新数据。应用：AI 绘画、数据增强、超分辨率。



- 神经网络由大量节点与节点间的连接组成
- **多层**结构，相邻层间呈**全连接**



AI在材料研究中的应用——神经网络



$$y_k = \varphi \left(\sum_i w_{ki} x_i + b_k \right)$$

上一层第 i 个神经元输入 x_i

本层第 k 个神经元输出 y_k

通用近似定理

- 一个具有足够多隐藏层神经元的多层前馈神经网络，可以以任意精度逼近任意连续函数



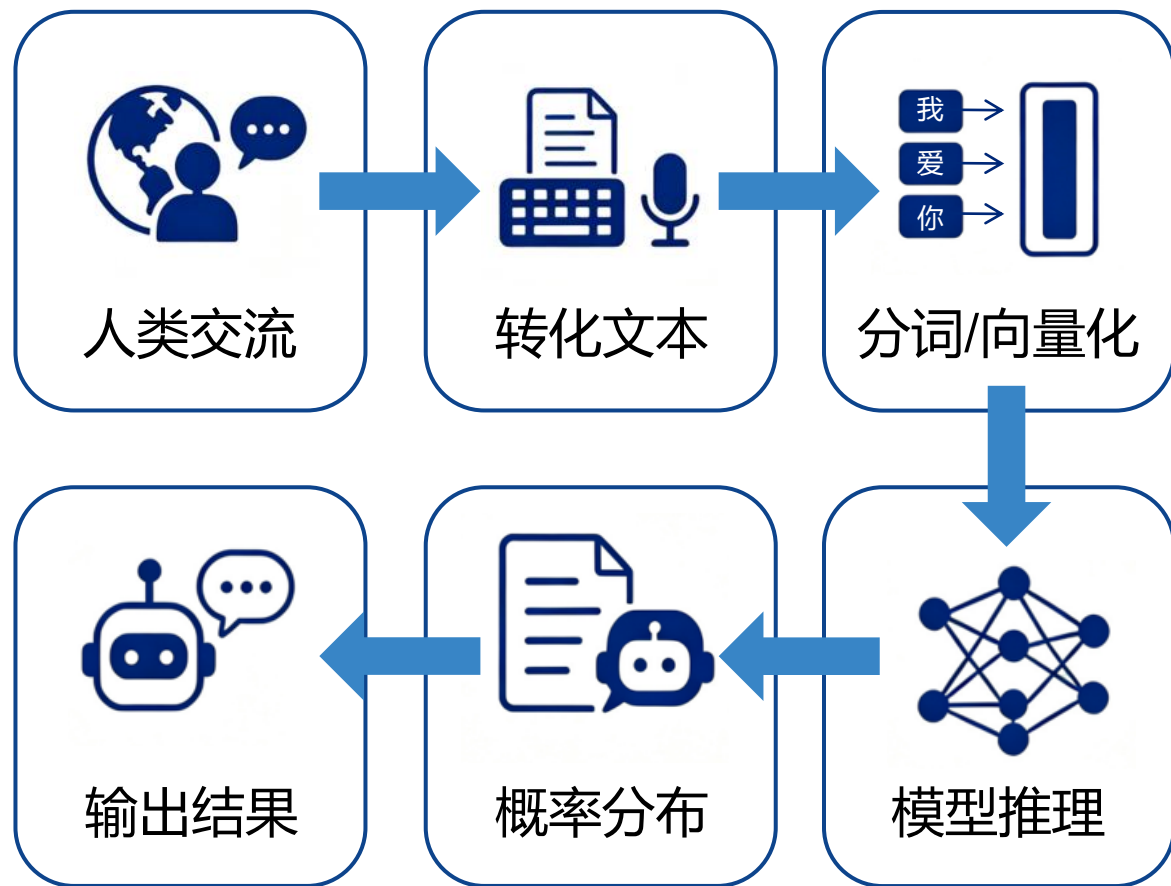
大语言模型（Large Language Model, LLM）



基于海量文本训练的大型神经网络，可理解复杂语义、生成连贯自然的人类语言。

基石：自然语言处理
(Natural Language Processing, NLP)

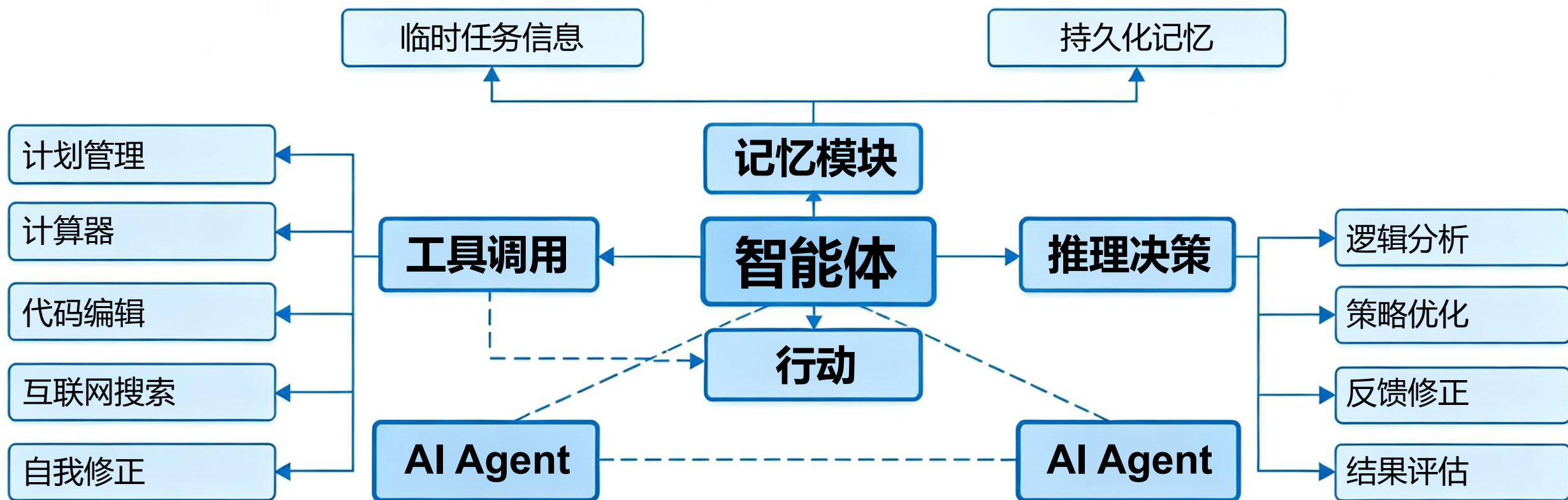
让计算机理解、生成人类语言的技术总称；LLM 是 NLP 的高级阶段，突破了传统模型的语义瓶颈



语言处理核心流程示意

智能体 (Agent)

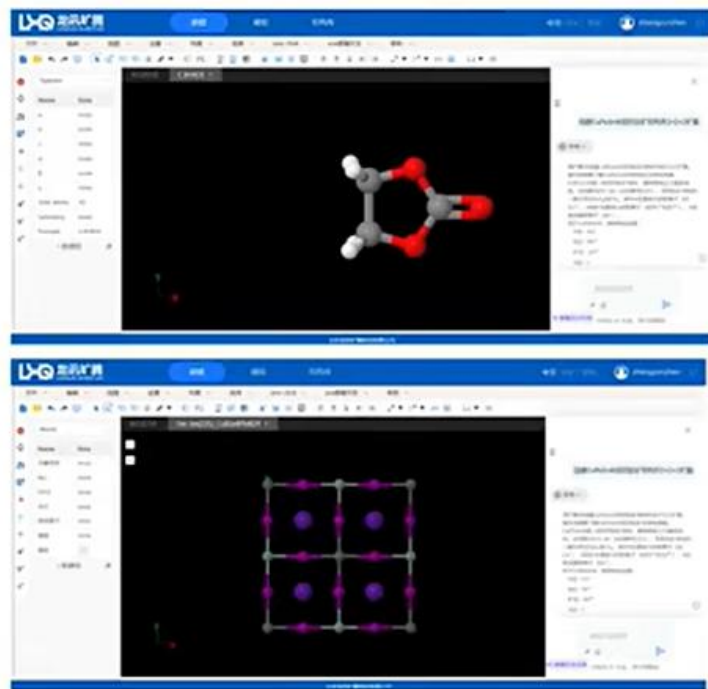
一个能够根据目标，自主理解任务、制定步骤、调用工具并完成行动的 AI 系统



31. 机器学习、神经网络、大语言模型、智能体

PWmat智能计算架构

前端可视化层



窗口模拟
实时同步



结构文件
渲染

文件管理层

模型工作目录

Al12O18.pwmat

C3H4O3.mol

Cs8Sn4Pb4I24.pwmat

.....

建模、计算任务目录

DFT calculations

Model generations

.....

读取



更新

SCINet 超算互联网
Supercomputing
Network

智能体层

用户: 构建Cs2PbSnI6立方相
晶体

AI: -> 好的, 我需要调用
crystal_from_spacegroup工具,
参数为...

工具实时检测->匹配->调用

调用

建模/计算工具函数

molecule_mixture

common_surface

DFT_optimization

.....

核心
计算
基座



Q-flow



Q-Studio

PWmat

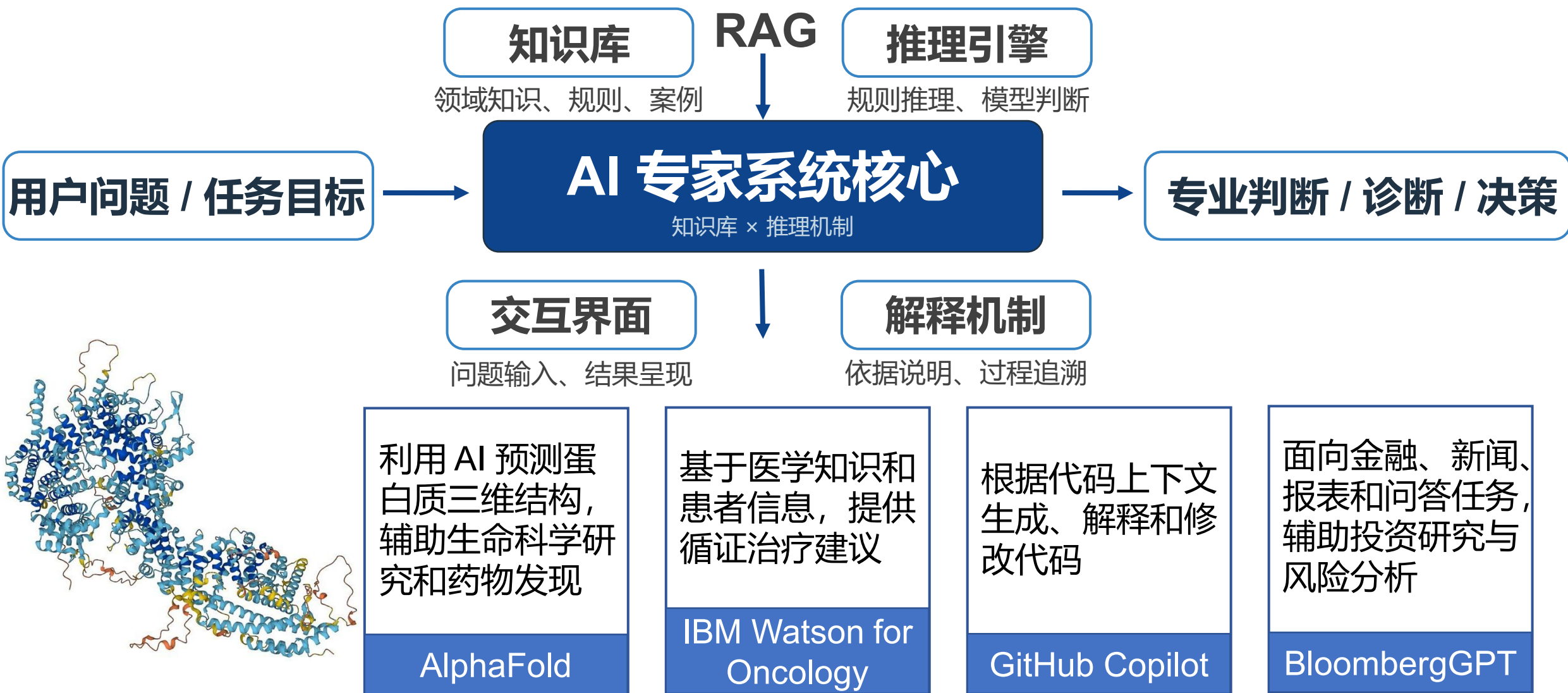
基座大
语言模
型

deepseek



32: AI专家系统Intellimat

基于专业知识库和推理机制，模拟领域专家进行判断、诊断和决策的人工智能系统。



32: AI专家系统Intellimat

Intellimat 是一款面向材料科学领域打造的智能化 AI 助手，聚焦科研与研发场景，融合大模型能力、领域知识检索能力与智能体定制能力，为材料科学研究人员、研发工程师及企业创新团队提供高效、专业、可扩展的智能支持。

依托材料科学领域海量文献资源、专业工具能力与智能分析技术，Intellimat 可围绕材料研发全过程提供多维度辅助，包括材料科学问题问答、文献综述撰写、机理分析、知识整理与研发决策支持等，帮助用户显著提升信息获取效率、知识沉淀能力与研发协同水平。

The screenshot displays the Intellimat AI interface, which is designed for material science research. The top navigation bar includes options for '知识库' (Knowledge Base), '搜索' (Search), 'Agent 代理人' (Agent), and '文件管理' (File Management). The main content area is divided into several sections:

- 搜索记录 (Search History):** A list of recent searches, including '帮我找10篇钛合金领域文献...' and '多主元合金 (Pd-30Ag-14Cu-1...'.
- 更懂材料的 AI 助手 (AI Assistant):** A section titled '智能聚合行业最新文献, AI助您快速获取' (Intelligent aggregation of the latest industry literature, AI assistant helps you quickly obtain it). It features a search bar and a '智能搜索' (Smart Search) button.
- 搜索结果 (Search Results):** A list of search results, including 'Cr₁₀Fe₃₀Co₃₃Ni₃₈Mo₁₀高熵合金...' and 'Cu₄₁2Ni₂₉7Si₁₅6Al₁₃5 高熵...'.
- 推荐搜索提示 (Recommended Search Prompts):** A section with prompts like '对比一下PtRh10和PtIr10在高温...' and '航空发动机、燃气轮机不用铂铑...'.
- 详细分析 (Detailed Analysis):** A section titled '方案三: 全新明星组合 (New All-Star Team)' (Scheme 3: New All-Star Team). It discusses the design concept of combining 'star' materials to improve the stability of the SEI layer. The analysis includes a list of references and a detailed discussion of the NaDFOB and FEC system.
- 参考文献 (References):** A list of references, including '1. NaDFOB and FEC as Electrolyte Additives Enabling Improved Cyclability of Sodium Metal Batteries and Sodium Ion Batteries' and '2. Black Phosphorus as a High-Capacity, High-Capability Negative Electrode for Sodium-Ion Batteries: Investigation of th...'.

33.机器学习建模：Intellimat 能做什么？

IntelliMat 通过机器学习模型实现预测、筛选和参数优化

建立材料性能预测模型

利用ML模型学习材料特征与目标性能之间的关系

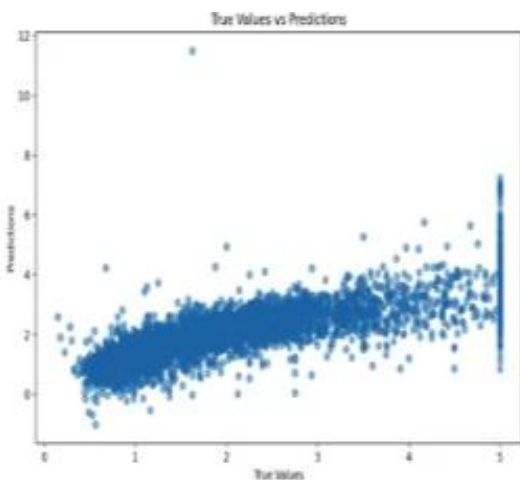
利用模型筛选材料

利用已经训练好的预测模型，对大量候选材料进行快速评估，并筛选出最有潜力的材料

利用模型优化工艺参数

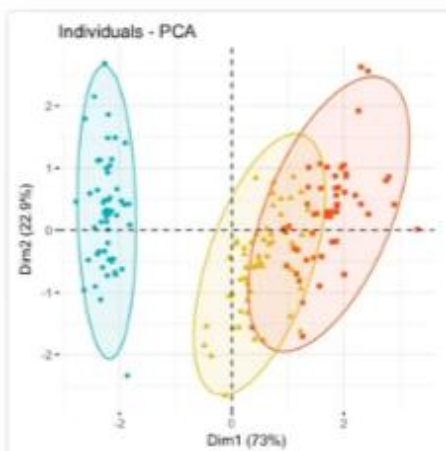
利用预测模型寻找更优的合成或加工条件，减少传统试错实验，提高材料性能

33.机器学习建模：典型机器学习模型



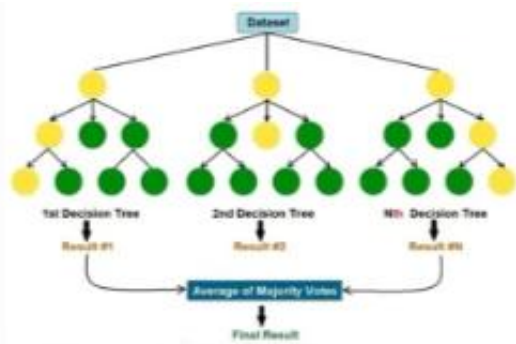
回归预测模型

Linear Regression,
SVR, Random Forest,
XGBoost, ANN



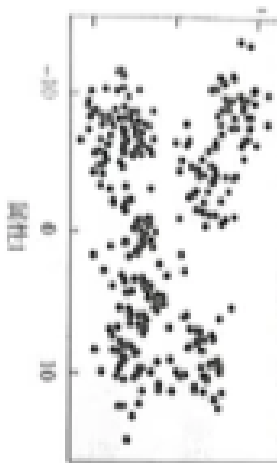
分类模型

Decision Tree, KNN,
SVM, Naive Bayes,
Random Forest



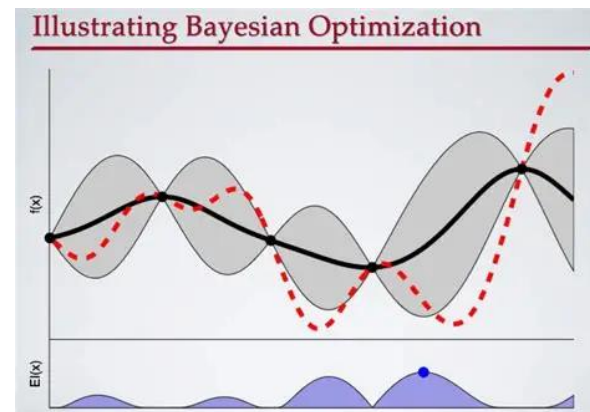
集成学习模型

Random Forest,
XGBoost



无监督学习

K-means, PCA



工艺优化方法

Bayesian optimization

33.机器学习建模：高通量计算生成计算数据

从结构出发，经过自动化计算与结果整理，形成可用于机器学习的标准化数据。

01 候选材料结构库

成分、晶体结构、掺杂位点、表面构型



02 高通量计算平台

批量DFT/MD计算，自动提交与结果解析



03 计算结果数据库

能量、稳定性、带隙、吸附能、扩散能垒等



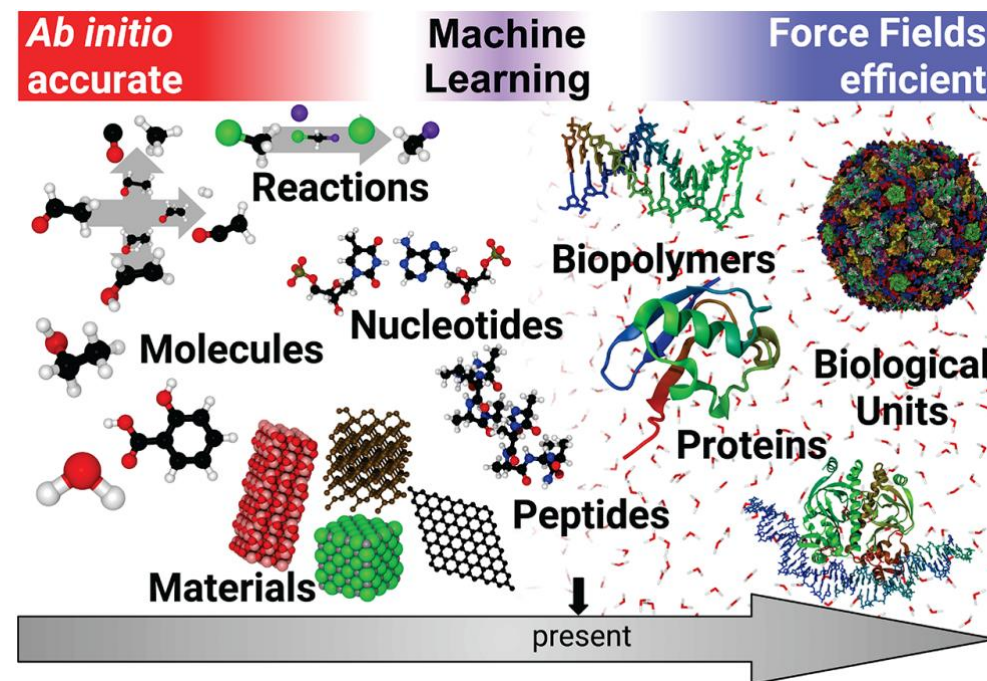
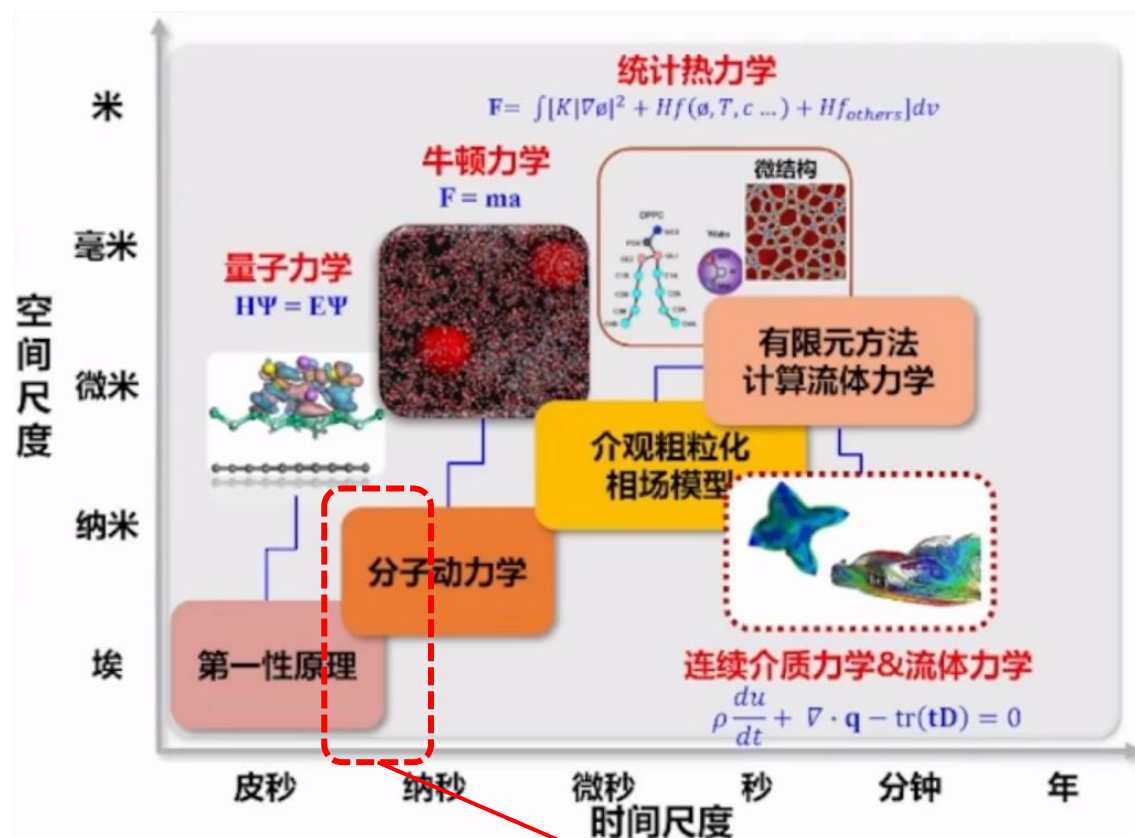
04 ML模型训练与预测

训练性能预测模型，支撑材料筛选与优化

高通量计算为机器学习提供大规模、标准化、可复用的材料数据基础

33.机器学习建模：材料物性预测

Some examples



Chem. Rev. **121**, 10142 (2021)

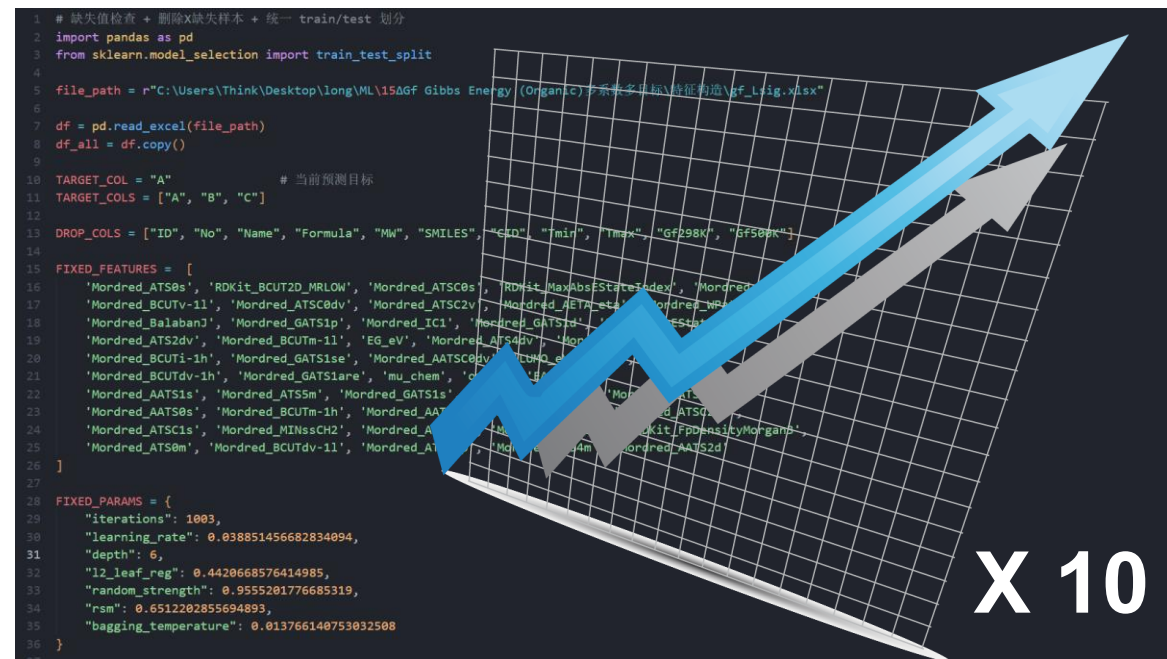
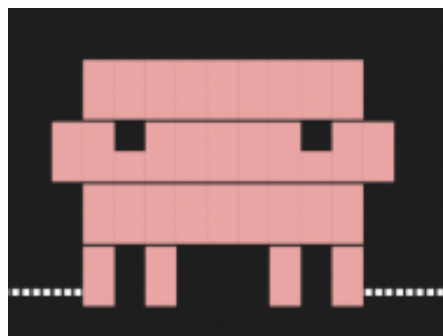
机器学习力场 (MLFF) 建立AIMD与MD之间的桥梁



34. AI 辅助编程与代码智能体

Vibe Coding

通过自然语言描述目标，让 AI 生成、修改和调试代码。用户不需要逐行手写代码，而是通过对话驱动程序开发。



Claude Code

Claude Code 是一种运行在终端中的 AI 编程工具，可以理解代码库，并通过自然语言命令完成代码生成、解释、调试、重构和 Git 工作流等任务。

35. A-Lab

A-Lab 是 劳伦斯伯克利国家实验室 (Lawrence Berkeley National Laboratory, LBNL) 提出的自主材料合成平台, 用计算、文献历史数据、机器学习、主动学习和机器人实验来规划并解释实验结果

文献与数据库辅助的初始决策

文献 / 数据库 / 经验公式

高通量实验平台

自动合成 + 自动表征

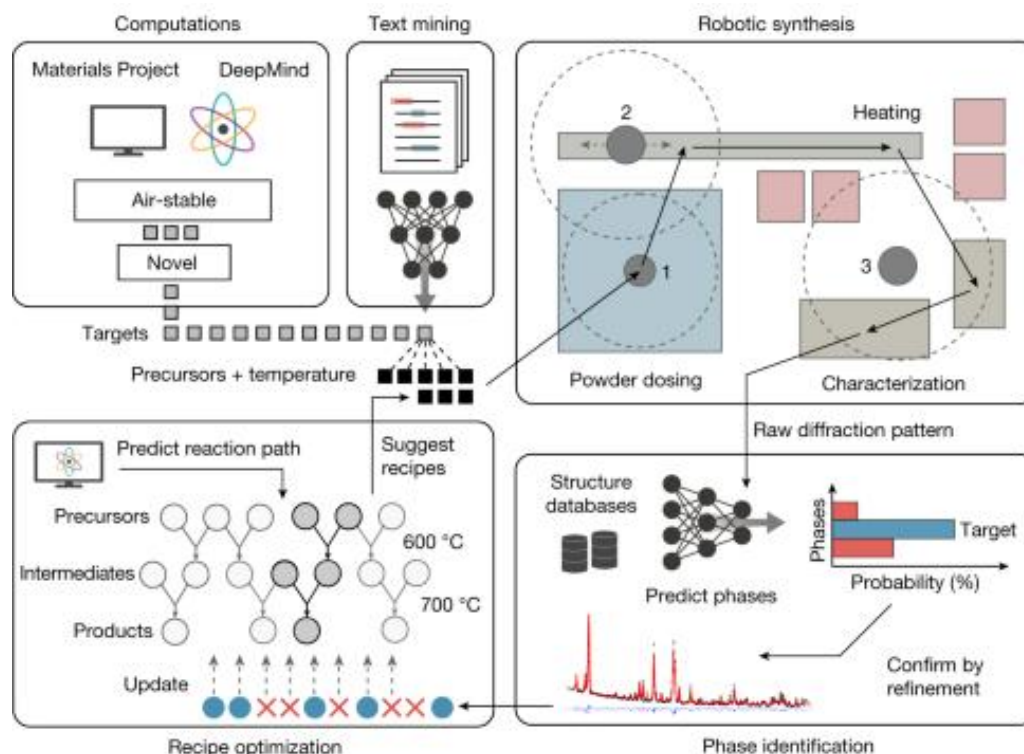
实验结果数据库

结构 / 相组成 / 性能

ML模型反馈

贝叶斯优化 / 主动学习

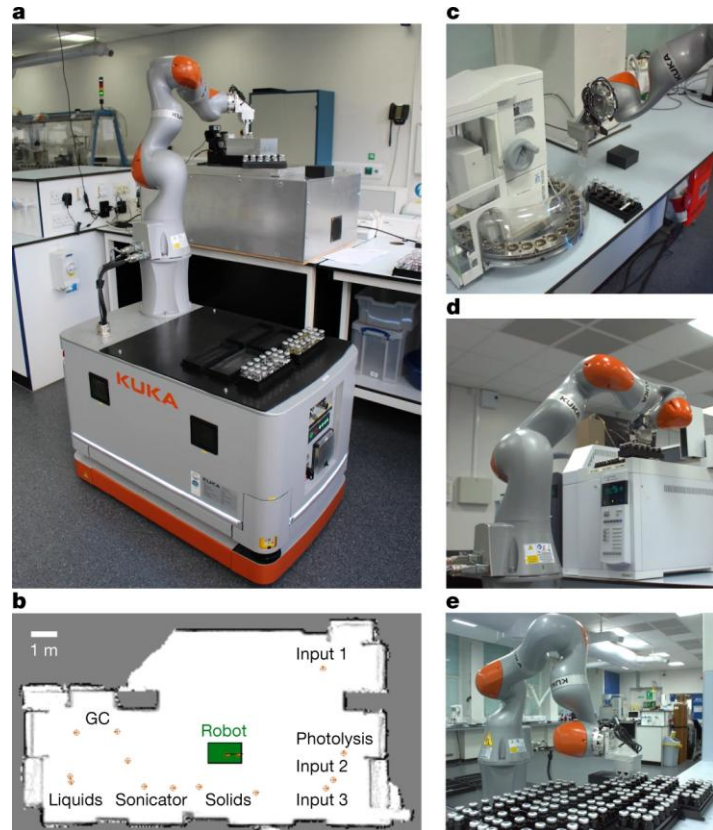
下一轮实验方案



A-Lab 在 17 天连续运行中, 从 57 个目标材料中实现了 36 个化合物¹

35. A-Lab应用

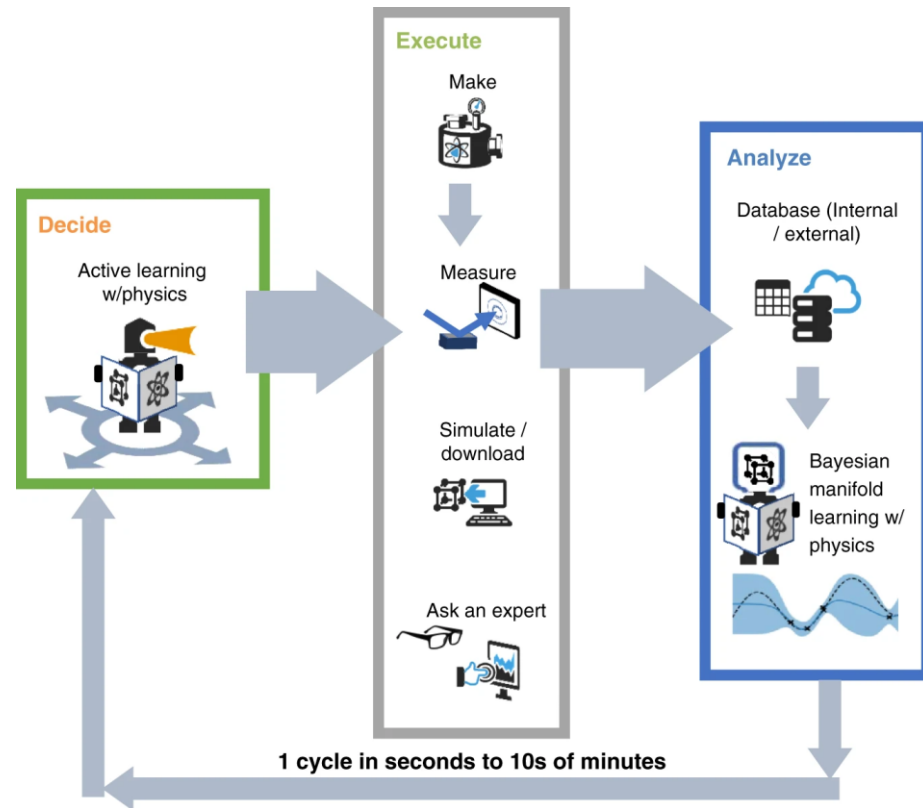
Mobile Robotic Chemist



1. *Nature*. 2020, 583, 237–241

移动机器人在实验室中自主移动和操作仪器，利用贝叶斯搜索优化光催化产氢体系；8 天完成 688 次实验，找到活性提高约 6 倍的光催化剂配方

CAMEO



2. *Nat. Commun.* 2020, 11, 5966

将高通量实验、同步辐射表征和贝叶斯主动学习结合，实现实时闭环材料探索与性能优化，并发现新的外延纳米复合相变存储材料