DOI:10.20176/j.cnki.nxdz.20260903

机器学习在化工过程优化中的应用

刘馨羽[1], 唐旻萌[2], 李宏平[1], 吴沛文[1*], 朱文帅[2]
(1.江苏大学 化学化工学院, 江苏 镇江 212013;
2.中国石油大学(北京) 化学工程与环境学院, 北京 102249)

摘 要:化工过程与材料研发涉及复杂的多相作用和动态变化,传统机理模型在处理未知机理或非理想状态时存在局限。系统梳理机器学习在化学及其相关领域的应用现状与前沿进展,聚焦流程优化与自动控制、新材料逆向筛选与设计、智能化闭环实验平台及大尺度分子动力学模拟等核心方向。结果表明,数据驱动模型为复杂化工过程的故障诊断与非线性控制难题的解决提供了有效方法;生成式算法与自动化机器人深度结合,显著加快了物理实验的闭环迭代速度;等变图神经网络与机器学习力场突破了传统量子化学计算的时空尺度限制;机器学习推动传统试错研发模式转向全自动、可自校正的主动设计范式。该研究结果可为复杂化工系统的智能化转型与新一代化学精准设计提供理论参考。

关键词:机器学习;数据驱动;化工过程;过程控制;材料设计
分类号:(中图)TU43;O344 **文献标志码:**A

化工过程涉及多相反应及其复杂动力学,受到严格安全标准约束。传统数学模型依赖完备的反应机理知识,在非理想状态或未知机理下预测结果不够准确,这限制了其在过程优化和实时控制中的应用[1]。近年来,随着测量技术不断发展、计算能力大幅度提升,在化工生产和研究领域积累了大量数据。数据驱动建模逐渐成为一种有效方法,能够弥补机理模型中存在的不足[2-3]。机器学习是当前研究热点,主要通过深入分析数据特征,建立预测模型;其核心在于转化科学数据,将实验或模拟得到的物理性质和工作条件转化为计算机可识别的特征[4]。目前,很多经典机器学习算法已经被广泛用于化工系统建模[5-6],其中包括支持向量机、随机森林和各种神经网络等。因此,机器学习技术展现出巨大应用潜力。文中结合近年来机器学习研究成果,综述机器学习在化工领域的应用现状,聚焦流程优化与控制、材料筛选与设计、智能化控制、分子模拟与量子化学性质预测[7-8]及数据建模;总结各种机器学习算法的实施策略。该研究结果可为化工过程的智能化研究提供理论参考。

1 机器学习方法

机器学习通过建立数学框架来处理化工数据,进而从数据中提取关联规律。目前,机器学习已经成为解决复杂系统建模问题的关键工具[9]。根据数据标签的使用方式,可将机器学习分为3类,即监督学习、无监督学习和半监督学习。其中,监督学习利用输入和输出对应关系来训练模型,在化工领域被广泛应用,如通过回归分析预测产率、转化率和物理性质;通过分类算法进行故障诊断。无监督学习主要挖掘数据的潜在结构,通常需对实验数据进行聚类与降维分析。在化工领域,由于获取实验数据的成本较高,带有标签的样本数量较少,半监督学习融合了少量带标签数据和大量无标签数据,在数据稀缺情况下具有独特应用价值。

1.1 常见模型与特征表示

监督学习算法包括支持向量机、随机森林和神经网络等模型,这些模型利用已标注的数据进

收稿日期:2026-05-07
基金项目:国家自然科学基金资助项目(22578187,22508419);国家自然科学基金联合基金重大项目(U25D9013)
作者简介:刘馨羽(1999—),女,博士研究生,主要从事油气资源转化与理论计算研究,(电子信箱)liuxinyu99ykcookie@163.com。
*通信联系人:吴沛文(1990—),男,副教授,博士,主要从事油气资源转化与定向分离研究,(电子信箱)wupeiwen@ujs.edu.cn。
引用格式:刘馨羽,唐旻萌,李宏平,等.机器学习在化工过程优化中的应用[J].宁夏大学学报(自然科学版中英文),2026,47(5):421-429.

行训练,被广泛用于化工过程预测、产品质量控制和复杂装置故障诊断[10]。无监督学习能够挖掘高维数据中的内在结构和隐藏模式,且无需人工标注标签,常被用于数据集的聚类分析与异常检测。强化学习基于智能体与环境交互、奖励反馈机制自主学习,能够动态优化复杂的顺序决策,这使其特别适合解决非线性化工过程控制中时滞和实时反馈调整难题[11]。深度学习是机器学习的重要分支,具有强大的非线性表征能力,在图像与模式识别中占据主导地位。近年来,研究者[5,12-13]通过深度学习解析化学反应机制,加速了新材料的逆向发现进程。图 1 展示了这些模型的方法分类。

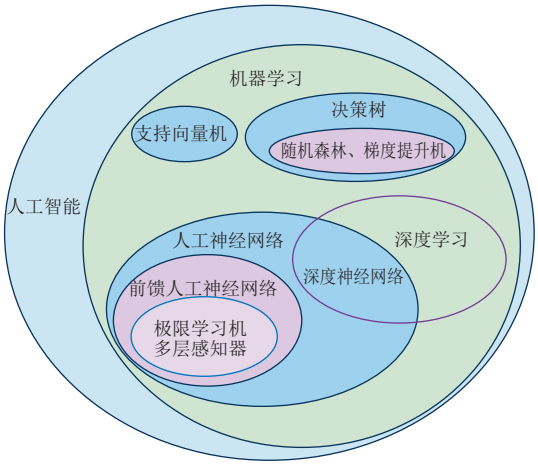


图 1　几种常见的机器学习方法分类[5]

Fig. 1　Classification of several common machine learning methods [5]

数据驱动模型的性能在很大程度上取决于输入特征的质量及其物理意义[14-15]。在化工与材料研究中,必须将微观分子结构或宏观工艺参数等信息转变为计算机可高效处理的数值描述符,具体转化路径见图 2。目前,主流的表征方法包括分子指纹、库仑矩阵和分子图等,其中,库仑矩阵可以捕获静电相互作用;分子图保留了拓扑物理对称性。例如,Liu 等[16]基于密度泛函理论得到高精度量子化学描述符。针对不同的应用领域和预测任务,相关研究[17-22]基于特定的特征空间,采用多种前沿模型进行研究;这些模型的架构特点与适用场景见表 1。近年来,物理信息神经网络受到广泛关注。普通神经网络主要通过样本数据学习输入-输出映射关系,而物理信息神经网络(physics-informed neural networks, PINNs)将质量守恒、能量守恒、组分守恒、反应动力学方程以及初始(边界)条件等物理约束引入损失函数,使模型预测结果受到机理规律约束。在连续搅拌釜反应器、传热反应耦合过程、精馏塔动态建模和多相传递过程等化工场景中,PINNs 常被用于浓度、温度和不可测状态变量的估算[23]。尤其在实验数据不足、传感器布点有限或机理模型与实际工况存在偏差条件下,该方法融合一阶原理模型与生产数据,可为软测量、过程控制和异常诊断提供更具物理一致性的建模工具[24-25]。

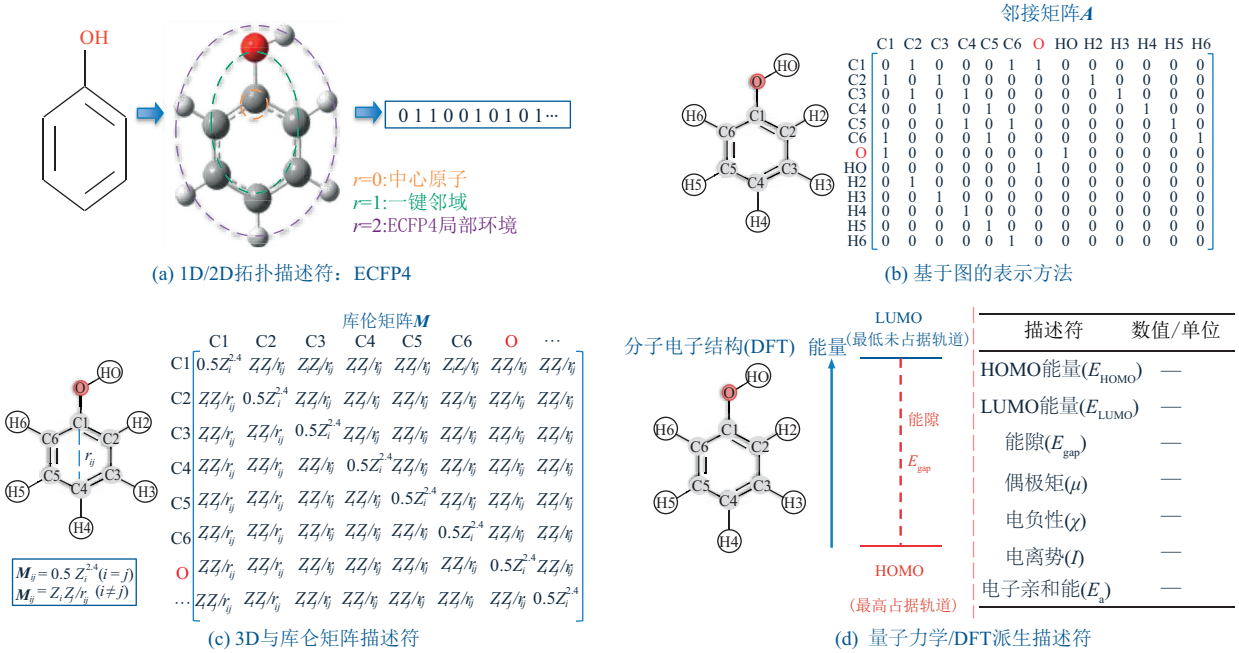


图 2　苯酚分子的多类型描述符示意图

Fig. 2　Schematic diagram of multiple types of descriptors for phenol molecules

表 1　常用机器学习模型及其在化工领域的应用特性

Tab. 1　Common machine learning models and their application characteristics in the chemical engineering field

模型	核心特点	典型化工应用场景
支持向量机	通过核函数处理高维数据,在小样本场景下泛化能力优异	纯物质的物理性质估算与定性分类判断
随机森林	属于集成学习算法,抗噪能力强,能评估特征贡献度	催化剂的活性筛选与工艺参数的重要性分析
人工神经网络	具有强大非线性拟合能力,适合处理大规模多模态数据	复杂反应动力学建模与流程动态模拟
图神经网络	直接分析分子拓扑结构,有效捕获原子间的空间关联关系	新材料筛选与性能预测
强化学习	通过环境交互优化策略,实现长期收益最大化	反应器自动控制与生产调度优化
物理信息神经网络	将质量、能量守恒等物理定律作为约束融入损失函数	稀缺数据下传递过程与反应机理建模

1.2　数据处理与模型评估

化工过程数据通常来源于传感器监测与实验室检测,容易受到电磁干扰、仪表精度和人为操作影响,普遍存在噪声干扰、异常值及缺失问题。因此,采用拉依达准则或中位数绝对偏差法剔除异常值,再通过均值填充、拉格朗日插值或多重插补方法处理缺失数据,以维持数据集的完整性和统计特性[26-27]。此外,不同变量间的量纲差异巨大,如压力、温度和浓度等参数的数值可能相差几个数量级,导致模型训练过程中梯度下降缓慢或权重分配失衡。相关研究[28-29]对特征数据进行归一化或标准化处理,使所有输入特征处于相同数量级,大大提升了算法的收敛速度和预测精度。由于化工过程数据具有高维冗余特性,在特征选择阶段,采用皮尔逊相关系数、互信息法或主成分分析法提取关键描述符。这种降维处理方法不仅降低了模型的计算复杂度,还剔除了无关干扰信息,从根本上缓解了维度灾难问题[30]。在模型验证阶段,为准确评估模型的泛化能力,Ahmed 等[31]将数据分为训练集、验证集和测试集。由于实验数据样本量较小,已有研究[32-33]采用 K 折交叉验证方法,通过多次循环采样降低数据划分带来的随机偏差。同时,为了对模型进行超参数优化,Melo 等[34]采用网格搜索或贝叶斯优化算法,求解最优学习率、正则化系数和神经元数量。这种优化平衡了模型的偏差与方差,避免了过拟合或欠拟合问题。通过上述系统化的数据处理与评估流程,可确保所构建的机器学习模型在未知复杂工况下,依然具备稳健的预测能力与较高的工程应用价值。

2　机器学习的应用

2.1　流程优化与控制

化工生产流程优化涉及多个变量的相互作用,其中,变量包括反应器操作条件、原料配比和空间温度分布,优化的核心目标是提高产物收率、降低能耗并保障装置运行安全。传统优化方法通常依赖经验公式或简化的物理模型,而现代机器学习技术能够精确监测和动态控制复杂工业过程。

在故障检测与诊断领域,深度学习算法可有效处理高维、动态和非线性特征[35]。例如,Chen 等[36]结合卷积神经网络和长短期记忆网络建立一种优化模型,并将该模型应用于田纳西-伊斯曼标准化工仿真过程。与传统方法相比,优化模型能够提取动态故障特征,进而实现对多种复杂工艺故障的准确识别与预警。在核心化工装备中,纯数据驱动的故障诊断系统具有巨大潜力,可替代传统模型的残差分析方法,大大提高对传感器失效或工艺干扰的预警速度[37]。此外,数字孪生技术将实时数据流与物理模型结合,为复杂催化反应的系统研究,提供了全新平台。同时,该平台集成了在线监测、性能预测和智能控制功能[38]。

化工生产面临复杂的决策和过程控制难题,而强化学习可通过智能体与环境交互的试错寻优机制解决这类问题。传统控制方法在非线性和不确定环境条件下存在局限,而强化学习是克服这些局限的关键工具。最新研究将强化学习引入化工分层控制体系,实现了连续生产调度、设定值跟踪及安全运行优化[39]。研究结果显示,在连续搅拌釜反应器、模拟移动床和青霉素生产过程中,模型学习方法在调节控制问题上,通常具有较高的数据利用效率和较低的控制代价波动;基于价值学习的强化学习方法,在闭环辨识和经济优化任务中表现出竞争力[40]。因此,强化学习更适合作为复杂非线性化工过程控制的补充框架,而不是针对模型预测控制的简单替代。

图 3 展示了一个多相过程优化案例[41]。Kordkheili 等[41]采用人工神经网络,预测工业规模柴油加

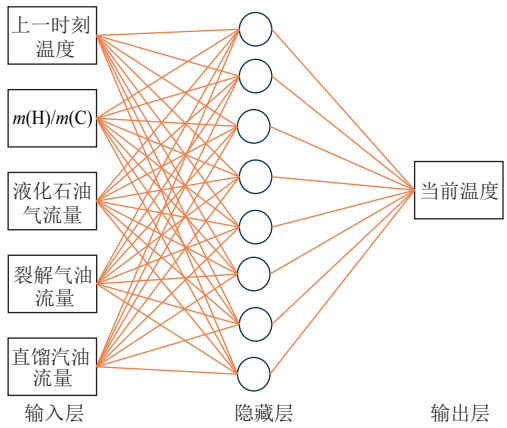


图3 预测工业规模柴油加氢脱硫反应器温度分布的人工神经网络(ANN)结构示意图[41]

Fig. 3 Schematic diagram of the artificial neural network (ANN) for predicting temperature distribution in industrial-scale diesel hydrodesulfurization reactors[41]

氢脱硫反应器的温度分布情况,分析各个输入参数对床层温度预测的重要性,并基于预测温度、反应特性和影响参数来优化反应条件。纯数据驱动的强化学习在面对未知干扰时不够稳定,而最新的控制先验强化学习框架,融合了经典比例积分微分(PID)控制器的抗干扰能力与深度强化学习的数据建模优势,能够解决这个问题。这种混合方法提高了样本利用效率,显著增强了模型在实际化工系统中的稳健性[42]。

此外,化工生产普遍具有明显的时间延迟特性。时间序列模型与非线性模型预测控制的结合应用,已成为当前研究热点。最近研究[43]通过构建一种混合网络架构,其中包含用于提取空间特征的卷积神经网络、挖掘时间依赖关系的长短期记忆网络,实现对间歇反应器温度轨迹的高精确跟踪控制。这种数据驱动的闭环策略,能够实时捕获放热反应的动态波动,进而精确调节冷却液的流量。该方法既保障了系统安全,又提高了过程的自适应能力和热量管理效率。近期,Selvamurugan等[44]改进了非线性模型预测控制架构,并向该架构中加入强化学习机制,以动态更新控制权重。研究结果显示,该混合策略在多变量反应器实验中表现出优异性能,可在预测精度、控制性能与计算效率之间实现良好平衡。

2.2 材料筛选与设计

机器学习正在革新化工和材料科学领域的传统试错研发模式,显著缩短了新材料的筛选与设计过程。传统研究高度依赖大量实验和实践经验,而机器学习可快速分析实验数据和材料特性,从而发现并设计出潜力巨大的新材料(图4)。Ge等[45]采用图卷积神经网络与密度泛函理论计算相结合的方法,从数千种候选材料中快速完成高性能镍基催化剂的逆向筛选。随着计算框架不断演进,定制化设计思路已从单金属载体拓展至复杂多元体系。Xiong等[46]通过提取元素特征构建描述符,并融合d带中心、吸附能等物理化学结构与属性分析,发现机器学习模型不仅能够精准预测用于二氧化碳还原的低过电位铜基三金属组分,还能将金属有机框架等复杂体系的大规模筛选速度提升数个数量级,并揭示体系微观结构与性能之间的定量关系[47]。

随着虚拟高通量筛选技术不断发展,人工智能正推动新材料研发向主动逆向设计与实验闭环转变。扩散模型等生成式架构打破了现有化学空间的限制,能够根据目标催化性能,直接生成高保真度的复杂高熵合金和新型催化剂组分[48]。这些理论预测与自动化机器人平台的深度融合,加快了预测结果向真实材料转化的速度。当前,自驱动实验已成为破解材料实际应用难题的核心途径。例如,Szymanski等[49]开发的A-Lab系统,能够自主接收模型预测到的未知无机晶体结构数据,通过自然语言处理技术提取合成配方,并控制机械臂依次完成粉末混合、焙烧与原位表征,在极短时间内实现数十种新材料的无人化合成。这种将材料从头设计、机器合成与数据反馈高度集成的全流程闭环体系,推动材料设计步入全自动人机协作时代[50]。

这种主动设计范式不断发展,从解析固相微观结构扩展至预测液相体系的宏观物理化学性质,如流体属性(电导率、黏度和热传导率)。Lemaoui等[51]采用多模型交叉验证策略,预测庞大离子液体和低共熔溶剂库。实验结果显示,该策略的预测结果可靠。新型流体合成实验也证实了该预测结果。在实际新能源应用场景中,量子化学与智能算法相结合的方法,能够跨越材料物理性质的试错盲区,从海量候选材料中直接锁定锂离子电池所需的高性能电解质,进而大幅度缩短高适配性聚合物复合器件的开发周期[52]。机器学习的系统性研究覆盖研发全过程,可实现微观催化位点定制和宏观流体性质预测。

2.3 智能化控制

在化工生产的过程控制与研发中,机器学习可实现智能化闭环控制,即在执行实验的同时,通过实时数据反馈,自动优化后续策略(图5)[53]。Koscher等[53]基于机器学习与自动化硬件构建了一个集成化平台。该平台通过生成式模型自动筛选候选分子结构,再通过计算机辅助合成规划,为候

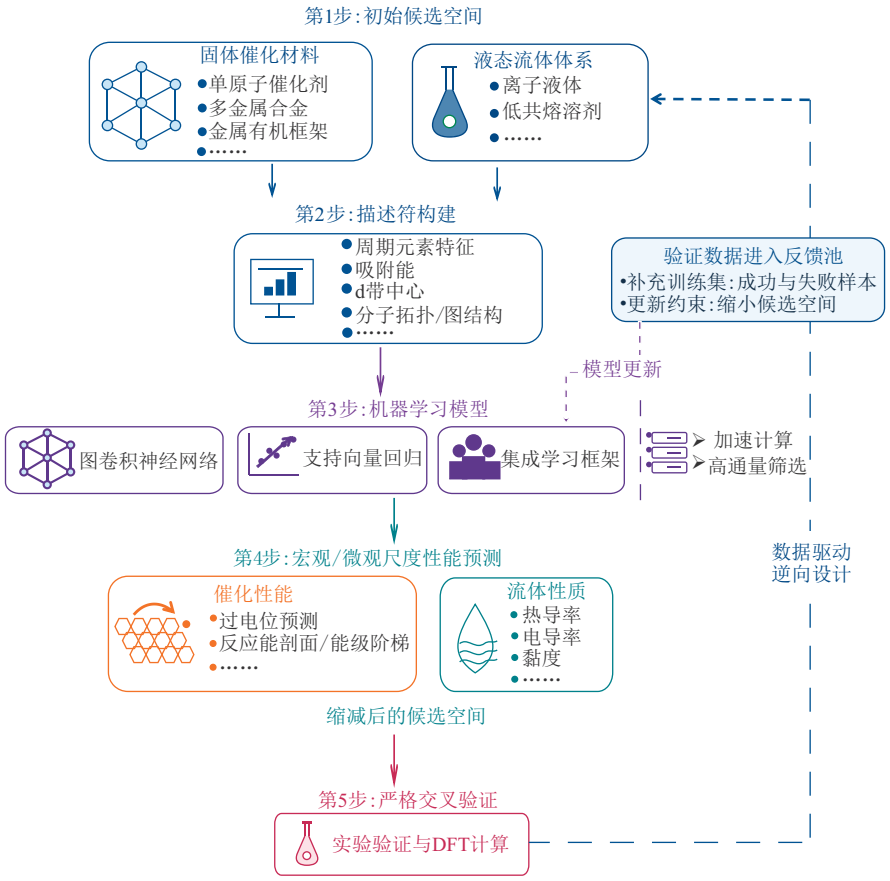


图 4　机器学习驱动的材料筛选与逆向设计闭环流程图

Fig. 4　Closed-loop flowchart of machine learning-driven material screening and inverse design

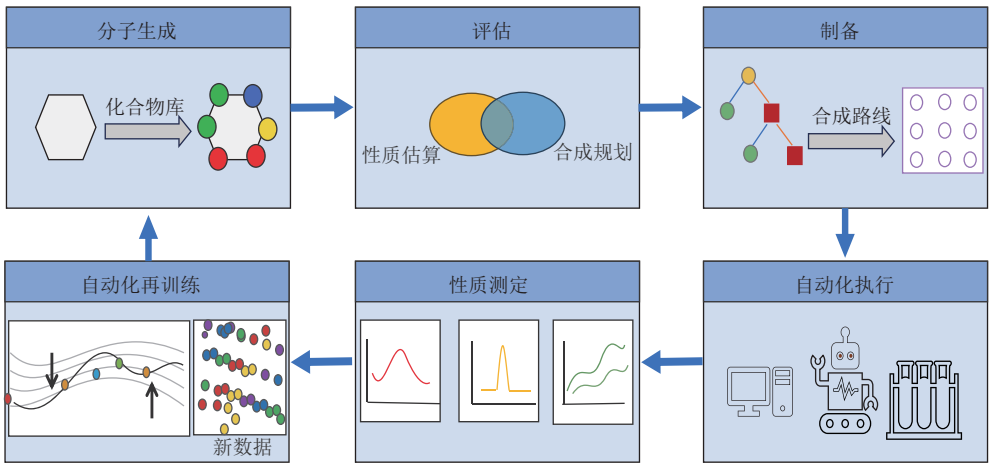


图 5　智能化集成平台关键工作流程[53]

Fig. 5　Key workflow of the intelligent integrated platform[53]

选物分子设计最优合成路径,并将选定的反应路径转化为具体的机器控制与表征工作流。该平台通过调度智能液体处理机器人与在线分析仪器,实现了多线程自动化实验;能够将测试结果实时反馈给算法,从而不断更新预测模型。这种智能化控制系统,不仅实现了化学研发过程中的多线程自主试错,还从根本上打破了传统研发中成本高、探索周期长的瓶颈。

随着大语言模型(large language models, LLMs)的快速发展,高度依赖预设脚本的自动化平台正加速向具备逻辑推理能力的嵌入式人工智能系统演进。最新的突破性研究成果展示了名为“Coscientist”的自主智能体系统。该系统能够理解人类输入的宏观自然语言指令,自主从互联网检索文献,编写并修正控制机械臂和液相泵的底层代码,最终驱动高通量设备完成反应循环,并实现自

我校正[54]。

算法模型实现了数据处理的认知升级。底层物理实验平台融合机器视觉、云计算和非线性优化算法,加速了实验平台向复杂体系和工业级放大应用的进程。由于聚合物纳米颗粒的流体合成过程十分复杂,Knox等[55]在该合成中引入了基于机器学习的闭环控制框架。研究结果显示,该框架能够同时分析单体转化率、相对分子质量分布和颗粒多分散性指数,从而实现全局自主寻优与动态在线调控。同时,智能化闭环控制框架适用于间歇式微量合成,已经拓展至连续流化学领域,且表现出极大工业应用价值。Volk等[56]结合主动学习与微流控技术开发了新型自动化系统,并通过原位光谱监测实现了与硬件的实时交互。研究结果显示,在连续运行状态下,该系统可自主调节停留时间、温度等连续变量,从而实现对多维参数的扫描与稳态控制。总而言之,数据驱动的自动化控制范式,在新材料发现和精细化工生产中展现出巨大应用潜力,推动了传统化学合成向全自动和自校正方向转变[57]。

2.4 分子模拟与量子化学性质预测

实际化工生产和催化过程具有显著动态特征,涉及复杂的溶剂效应、温度变化及化学键断裂与重组。传统研究主要采用从头算分子动力学分析原子尺度动态演变,但其计算成本较高。而且该方法只能模拟数百个原子体系,模拟时间尺度为数皮秒,导致难以真实反映宏观界面催化或大分子反应现象。近年来,以机器学习为代表的全新方法,突破了传统方法的计算限制[58-61]。尤其是等变图神经网络架构的引入,使算法严格保持三维空间旋转、平移等物理对称不变性,不仅极大地提升了模型的数据学习效率,还从根本上确保模型在模拟复杂化学键断裂与生成过程中的物理稳定性与计算准确度[62]。现代机器学习力场可学习大量高精度量子力学数据,不仅确保了密度泛函理论级别的计算精度,还将计算速度提升了几个数量级。

物理模拟正在从静态真空构型向显式溶剂和真实温度下的全动态模拟转变,有助于对复杂化工反应机理的理解。在非均相催化和电化学领域,Yu等[63]融合主动学习框架和机器学习力场,在完全显式溶剂模型下实现长程分子动力学模拟。针对铁-氮-碳单原子催化剂界面,通过长程动态演化方法,研究氧还原反应的限速步骤。该过程揭示了固、液、气三相界面上水分子的氢键网络作用,即通过质子耦合电子转移机制降低反应自由能垒。与此同时,动态模拟的应用范围正迅速扩展至生物大分子和复杂流体体系,其中,含静电嵌入机器学习势的多尺度框架,能够精确、高效模拟酶内高度极化与带电过渡态的转化过程,有效避免了传统混合量子力学-分子力学模拟中的精度损失[64]。Kovács等[65]研发出具有良好迁移性的通用型短程力场,显示针对大分子有机物、液相体系及分子晶体的全动态物理模拟正逐步实现标准化。

机器学习力场不仅在空间环境和复杂体系应用方面不断拓宽边界,还在模拟时间尺度和数据规模上实现了跨越,这为探索高能垒罕见反应事件和宏观热力学相变提供了可能。在实际化工体系中,关键化学键的重组通常需要跨越较高能垒,而传统计算方法难以自动观测到这些缓慢过程。David等[66]将深度势能与增强采样技术相结合,从微秒级时间尺度揭示了液相中肽键形成的竞争性动态网络结构。该方法可同时捕获全局构象和局部化学键解离变化,进而完善了基于静态单路径假设的传统反应机理理论。在大尺度模拟研究中,基于局域化的等变网络架构研究取得重要进展。该架构将具有量子精度的反应动力学模拟拓展至上亿原子体系,缩小了微观化学反应与宏观材料力学性质间的差距[67]。固态电池和多相催化等体系研究需要精准追踪电荷转移过程,而预训练基础模型满足这一需求:将原子磁矩作为电荷约束引入图网络,进而实现跨越元素周期表的零样本泛化能力[20]。这些通用模型兼具宏观视角和量子力学的严谨性,标志着计算化学的动态模拟正在步入大模型时代,从而为复杂化工过程的精准设计提供数据支撑。

3 结论与展望

机器学习已深度融入化学与材料科学研究的各个环节。其中,数据驱动方法在弥补传统机理模型不足方面展现出巨大优势。在化工流程中,结合深度学习与强化学习的混合架构,有效解决了系统非线性和时滞难题,显著提升了对工业装置的动态控制与故障预警能力。在材料研发领域,生成式模型与自驱动实验室高度融合,形成全流程闭环,加速了新材料的逆向设计与物理合成过程。此外,机器学习力场突破了传统计算化学的时空尺度限制;预训练基础模型与等变网络架构相结合,实现了宏观尺度全动态物理模拟。

机器学习在化工领域的实际应用,仍需关注模

型的可靠性、可解释性和跨场景泛化能力。化工过程普遍具备高温、高压及强耦合及动态非线性等特征,单纯依赖数据驱动模型,难以完全覆盖极端工况和异常扰动场景。因此,未来需要加强模型的不确定性评估、安全约束控制和异常工况识别能力,进而提高模型在实际应用中的可信度[5]。物理信息神经网络为机理模型与数据模型融合提供了有效路径,但在复杂多相流、强非线性反应和工业尺度等边界条件下,其训练稳定性与泛化性能仍有待提升[68]。机器学习力场虽然显著提高了分子模拟效率,但在长程相互作用、电荷转移和复杂反应路径外推等方面,仍需更多高质量数据和多尺度验证实验支撑[65]。

未来化工智能化发展方向,并不是单一算法的替代式应用,而是机理模型、数据驱动模型、自动化实验平台和安全控制策略的协同融合。随着高质量数据集、主动学习、物理约束建模和智能实验平台的持续发展,机器学习有望进一步推动化工研发与制造向更加高效、安全和自校正方向发展。

参考文献:

[1] Devarakonda V S, Sun Wei, Tang Xun, et al. Recent advances in reinforcement learning for chemical process control [J]. Processes, 2025, 13(6):1791.

[2] Lavor V, de Come F, dos Santos M T, et al. Machine learning in chemical engineering: Hands-on activities [J]. Education for Chemical Engineers, 2024, 46: 10-21.

[3] Gao Hanyu, Zhu Litao, Luo Zhenghong, et al. Machine learning and data science in chemical engineering [J]. Industrial & Engineering Chemistry Research, 2022, 61(24): 8357-8358.

[4] Xiouras C, Cameli F, Quillo G L, et al. Applications of artificial intelligence and machine learning algorithms to crystallization [J]. Chemical Reviews, 2022, 122 (15): 13006-13042.

[5] Dobbelaere M R, Plehiers P P, van de Vijver R, et al. Machine learning in chemical engineering: Strengths, weaknesses, opportunities, and threats [J]. Engineering, 2021, 7(9): 1201-1211.

[6] 李顺,彭浩然,符秀丽. 数据驱动的高熵碲化物组分高通量理论设计与性能预测[J]. 宁夏大学学报(自然科学版中英文),2026,47(3):263-272.

[7] 申奥,马梅,高乐,等. 人工智能在材料发现中的应用[J]. 宁夏大学学报(自然科学版中英文),2026,47(2): 116-157.

[8] Menon D, Ranganathan R. A generative approach to materials discovery, design, and optimization [J]. ACS Omega, 2022, 7(30): 25958-25973.

[9] Torres A I, Ferreira J, Pedemonte M. Machine learning and process systems engineering for sustainable chemical processes: A short review [J]. Current Opinion in Green and Sustainable Chemistry, 2025, 51 (349):100982.

[10] Yu Beiwei, Zhang Liqin, Ye Xiaoxia, et al. State-of-the-art review on various applications of machine learning techniques in materials science and engineering [J]. Chemical Engineering Science,2025,306:121147.

[11] Paz Ramos M A, Busboom A. Systematic review of reinforcement learning in process industries: A contextual and taxonomic approach [J]. Applied Sciences, 2025,15(24):12904.

[12] Li Wei, Wang Guoqiang, Ma Jing. Deep learning for complex chemical systems [J/OL]. National Science Review, 2023, 10(12): nwad335. https://doi. org/10. 1093/nsr/nwad335.

[13] Duan Chenru, Du Yuanqi, Jia Haojun, et al. Accurate transition state generation with an object-aware equivariant elementary reaction diffusion model [J]. Nature Computational Science,2023,3(12):1045-1055.

[14] Wang Hanchen, Fu Tianfan, DU Yuanqi, et al. Scientific discovery in the age of artificial intelligence [J]. Nature, 2023, 620(7972): 47-60.

[15] Schweidtmann A M, Esche E, Fischer A, et al. Machine learning in chemical engineering: A perspective [J]. Chemie Ingenieur Technik, 2021, 93(12): 2029-2039.

[16] Liu Chengyou, Sun Yan, Davis R, et al. ABT-MPNN: An atom-bond transformer-based message-passing neural network for molecular property prediction [J]. Journal of Cheminformatics, 2023, 15(1):29.

[17] Reiser P, Neubert M, Eberhard A, et al. Graph neural networks for materials science and chemistry [J]. Communications Materials, 2022, 3(1):93.

[18] Karniadakis G E, Kevrekidis I G, Lu Lu, et al. Physics-informed machine learning [J]. Nature Reviews Physics, 2021, 3(6): 422-440.

[19] Merchant A, Batzner S, Schoenholz S S, et al. Scaling deep learning for materials discovery [J]. Nature, 2023, 624(7990): 80-85.

[20] Deng Bowen, Zhong Peichen, Jun K, et al. CHGNet as a pretrained universal neural network potential for charge-informed atomistic modelling [J]. Nature Machine Intelligence, 2023, 5(9): 1031-1041.

[21] Bran M A, Cox S, Schilter O, et al. Augmenting large language models with chemistry tools [J]. Nature Machine Intelligence, 2024, 6(5): 1-11.

[22] Mohr F, Arnese-Feffin E, Barolo M, et al. Smart

process analytics for process monitoring [J]. Computers & Chemical Engineering, 2025, 194:108918.

[23] Nai Dingqi, Li Huayu, Grover M A, et al. Modeling batch crystallization under uncertainty using physics-informed machine learning [J]. Industrial & Engineering Chemistry Research, 2026, 65(18): 9563-9577.

[24] Velioglu M, Zhai Song, Rupprecht S, et al. Physics-informed neural networks for dynamic process operations with limited physical knowledge and data [J]. Computers & Chemical Engineering, 2025, 192: 108899.

[25] Moayedi F, Chandrasekhar A, Rasmussen S, et al. Physics-informed neural networks for process systems: Handling plant-model mismatch [J]. Industrial & Engineering Chemistry Research, 2024, 63(31): 13650-13659.

[26] Ji Cheng, Sun Wei. A review on data-driven process monitoring methods: Characterization and mining of industrial data [J]. Processes, 2022, 10(2):335.

[27] López-Flores F J, Ramírez-Márquez C, González-Campos J B, et al. Machine learning for predicting and optimizing physicochemical properties of deep eutectic solvents: Review and perspectives [J]. Industrial & Engineering Chemistry Research, 2024, 64(6): 3103-3117.

[28] Kappatou C D, Odgers J, Garcia-Munoz S, et al. An optimization approach coupling preprocessing with model regression for enhanced chemometrics [J]. Industrial & Engineering Chemistry Research, 2023, 62(15): 6196-6213.

[29] Qiu Zhen, Zhuang Zechao, Jia Xue, et al. Machine learning-driven advances in water treatment [J]. Journal of Environmental Chemical Engineering, 2025, 13(6):120441.

[30] Mowbray M, Vallerio M, Perez-Galvan C, et al. Industrial data science: A review of machine learning applications for chemical and process industries [J]. Reaction Chemistry & Engineering, 2022, 7(7): 1471-1509.

[31] Ahmed A, del Rio-Chanona E A, Mercangoz M. Comparative study of machine learning and system identification for process systems engineering dynamics [J]. Industrial & Engineering Chemistry Research, 2025, 64(8): 4450-4478.

[32] Jin Miao, Liao Qinzhuo, Patil S, et al. Hyperparameter tuning of artificial neural networks for well production estimation considering the uncertainty in initialized parameters [J]. ACS Omega, 2022, 7(28): 24145-24156.

[33] du Toit T D F, Zhou Yuxing, Deringer V L. Hyperparameter optimization for atomic cluster expansion potentials [J]. Journal of Chemical Theory and Computation, 2024, 20(22): 10103-10113.

[34] Melo A, Câmara M M, Pinto J C. Data-driven process monitoring and fault diagnosis: A comprehensive survey [J]. Processes, 2024, 12(2):251.

[35] Özdemir P, Yildirim R. Ml@ChemE: Past, present, and future of machine learning in chemical engineering [J/OL]. ChemBioEng Reviews, 2025, 12(4): e70012. https://doi.org.10.1002/cben.70012.

[36] Chen Honghua, Cen Jian, Yang Zhuohong, et al. Fault diagnosis of the dynamic chemical process based on the optimized CNN-LSTM network [J]. ACS Omega, 2022, 7(38): 34389-34400.

[37] Du Pu, Abdel Jabbar N M, Wilhite B A, et al. Fault diagnosis in chemical reactors with data-driven methods[J]. Industrial & Engineering Chemistry Research, 2025, 64(11): 6060-6076.

[38] Etezadi R, Tsotsis T T. Digital twin approach toward the design, optimization, and control of a catalytic system for high-purity hydrogen production from biogas [J]. Industrial & Engineering Chemistry Research, 2025, 64(45): 21588-21598.

[39] Szatmári K, Chován T, Németh S, et al. How to support decision making with reinforcement learning in hierarchical chemical process control? [J]. Chemical Engineering Journal Advances, 2025, 22:100753.

[40] OH T H. Quantitative comparison of reinforcement learning and data-driven model predictive control for chemical and biological processes [J]. Computers & Chemical Engineering, 2024, 181(2):108558.

[41] Kordkheili M S, Rarshad F. Artificial neural network and semi-empirical modeling of industrial-scale gasoil hydrodesulfurization reactor temperature profile [J]. Mathematics and Computers in Simulation, 2023, 206: 198-215.

[42] Bloor M, Ahmed A, Kotecha N, et al. Control-informed reinforcement learning for chemical processes [J]. Industrial & Engineering Chemistry Research, 2025, 64(9): 4966-4978.

[43] Selvamurugan A, Ganesan P K, Nayak S S, et al. CNN-LSTM-based nonlinear model predictive controller for temperature trajectory tracking in a batch reactor [J]. ACS Omega, 2024, 9(47): 47203-47212.

[44] Selvamurugan A, Kumar A V, Palan H R, et al. Reinforcement learning-based nonlinear model predictive controller for a jacketed reactor: A machine learning concept validation using jetson orin [J]. ACS Omega, 2025, 10(28): 30864-30878.

[45] Ge Xiaohu, Yin Jun, Ren Zhouhong, et al. Atomic design of alkyne semihydrogenation catalysts via active learning [J]. Journal of the American Chemical Soci-

ety, 2024, 146(7): 4993-5004.

[46] Xiong Bo, Liu Jing, Yang Yingju, et al. Machine learning-assisted screening of Cu-based trimetallic catalysts for electrochemical conversion of CO_2 to Co [J]. Energy & Fuels, 2023, 38(3): 2074-2083.

[47] Xu Liangliang, Zhou Jian, Bandarenka A, et al. Data-driven electrocatalyst discovery: Recent trends in machine learning approaches and descriptor-based design principles[J]. Accounts of Materials Research, 2026, 7(3): 294-308.

[48] Yang Jiaqi, Ye Kailong, Xie Shaohu, et al. Diffusion model-guided inverse design of bimetallic catalysts for ammonia decomposition [J]. Journal of the American Chemical Society, 2026, 148(1): 537-546.

[49] Szymanski N J, Rendy B, Fei Yuxing, et al. An autonomous laboratory for the accelerated synthesis of novel materials [J]. Nature, 2023, 624 (7990): 86-91.

[50] Tom G, Schmid S P, Baird S G, et al. Self-driving laboratories for chemistry and materials science [J]. Chemical Reviews, 2024, 124(16): 9633-9732.

[51] Lemaoui T, Darwish A S, Almustafa G, et al. Machine learning approach to map the thermal conductivity of over 2,000 neoteric solvents for green energy storage applications [J]. Energy Storage Materials, 2023, 59: 102795.

[52] Li Kai, Wang Jifeng, Song Yuanyuan, et al. Machine learning-guided discovery of ionic polymer electrolytes for lithium metal batteries [J]. Nature Communications,2023,14(1):2789.

[53] Koscher B A, Canty R B, mc Donald M A, et al. Autonomous, multiproperty-driven molecular discovery: From predictions to measurements and back [J/OL]. Science, 2023, 382(6677):eadi1407. https://doi. org/10. 1126. science. adi1407.

[54] Boiko D A, Macknight R, Kline B, et al. Autonomous chemical research with large language models [J]. Nature, 2023, 624(7992): 570-578.

[55] Knox S T, Wu K E, Islam N, et al. Self-driving laboratory platform for many-objective self-optimisation of polymer nanoparticle synthesis with cloud-integrated machine learning and orthogonal online analytics [J]. Polymer Chemistry, 2025, 16(12): 1355-1364.

[56] Volk A A, Epps R W, Yonemoto D T, et al. Alphaflow: Autonomous discovery and optimization of multi-step chemistry using a self-driven fluidic lab guided by reinforcement learning [J]. Nature Communications, 2023, 14(1):1403.

[57] Abolhasani M, Kumacheva E. The rise of self-driving labs in chemical and materials sciences [J]. Nature Synthesis, 2023, 2(6): 483-492.

[58] Yalamanchi K K, Monge-Palacios M, van Oudenhoven V C O, et al. Data science approach to estimate enthalpy of formation of cyclic hydrocarbons [J]. The Journal of Physical Chemistry A, 2020, 124 (31): 6270-6276.

[59] Kabylda A, Frank J T, Suárez-Dou S, et al. Molecular simulations with a pretrained neural network and universal pairwise force fields [J]. Journal of the American Chemical Society, 2025, 147(37): 33723-33734.

[60] Poltavsky I, Charkin-Gorbulin A, Puleva M, et al. Crash testing machine learning force fields for molecules, materials, and interfaces: Model analysis in the tea challenge 2023 [J]. Chemical Science, 2025, 16 (8): 3720-3737.

[61] Seute L, Hartmann E, Stühmer J, et al. Grappa: A machine learned molecular mechanics force field [J]. Chemical Science, 2025, 16(6): 2907-2930.

[62] Batzner S, Musaelian A, Sun Lixin, et al. E(3)-equivariant graph neural networks for data-efficient and accurate interatomic potentials [J]. Nature Communications, 2022, 13(1): 2453.

[63] Yu Qinghan, Li Pai, Ni Xing, et al. Dynamics and kinetics exploration of the oxygen reduction reaction at the Fe-N_4/C-water interface accelerated by a machine learning force field [J]. Chemical Science, 2025, 16 (8): 3620-3629.

[64] Gradisteanu V, Chan E W, Hedges L, et al. Simulating enzyme catalysis with electrostatically embedded machine learning potentials [J]. Chemical Science, 2026,17:8542-8556.

[65] Kovács D P, Moore J H, Browning N J, et al. Mace-off: Short-range transferable machine learning force fields for organic molecules [J]. Journal of the American Chemical Society, 2025, 147(21): 17598-17611.

[66] David R, Tunon I, Laage D. Competing reaction mechanisms of peptide bond formation in water revealed by deep potential molecular dynamics and path sampling [J]. Journal of the American Chemical Society, 2024, 146(20): 14213-14224.

[67] Musaelian A, Batzner S, Johansson A, et al. Learning local equivariant representations for large-scale atomistic dynamics [J]. Nature Communications, 2023, 14(1): 579.

[68] Wu Zhiyong, Wang Huan, He Chang, et al. The application of physics-informed machine learning in multiphysics modeling in chemical engineering [J]. Industrial & Engineering Chemistry Research, 2023, 62(44): 18178-18204.

(下转第 440 页)