

文章编号: 1006-9941(XXXX)XX-0001-11

数据驱动策略在含能材料设计及性能预测中的应用

王海风, 王康才, 刘 渝

(中国工程物理研究院化工材料研究所, 四川 绵阳 621999)

摘要: 以数据驱动和人工智能为代表的科技与产业变革正在深刻影响材料科学领域, 也为含能材料的创新带来了前所未有的机遇与挑战。机器学习作为一种新兴技术, 为含能材料的分子设计与合成提供了全新的研发范式, 有望解决效率低下、成本高昂、周期冗长等含能材料研发中长期存在的瓶颈问题。尽管已有部分成功案例被报道, 但机器学习在含能分子“设计→筛选→合成→性能验证”全周期研究中的应用, 相较于其他新材料领域仍处于相对不成熟的阶段。研究综述了机器学习辅助含能材料开发的研究现状, 重点总结了机器学习在含能分子设计、单一性能预测及多性能同步预测中的应用案例。然而, 依托机器学习辅助设计合成具有特定性能的含能材料依然充满了挑战。未来应着力推进含能材料数据质量控制与标准化体系的构建、可解释机器学习模型的开发以跨学科交叉融合体系的建立, 从而进一步推动高性能含能材料的高效创制。

关键词: 数据驱动; 机器学习; 含能材料; 性能预测

中图分类号: TJ55; O64

文献标志码: A

DOI: 10.11943/CJEM2025076

0 引言

含能材料是一类受到外界刺激时会迅速释放出大量化学能和气体的亚稳态物质, 属于国防科技的关键核心材料^[1-2], 其性能的微小提升, 足以推动武器装备实现大幅发展。尽管各军事强国纷纷投入大量人力物力致力于新型含能材料的研制工作, 但含能材料的发展速度依然极其缓慢^[3-5]。目前广泛使用的含能材料, 如 TNT、RDX 等, 均已有 100 多年的历史^[6]。新型含能分子的传统设计与合成手段存在过程繁琐冗长、研发效率较低、成本偏高等弊端, 无法满足高性能含能材料的发展需求, 极大地限制了武器装备的更新换代。

近年来, 数据驱动策略发展迅猛, 已被广泛用于辅助新材料的设计及合成, 可极大缩短材料研发周期, 为材料科学的发展提供了新范式^[7-9]。此外, 与传统的分

子设计和理论计算方法相比, 数据驱动策略在材料性能预测方面展现出显著优势。基于现有含能材料的性能数据, 建立机器学习模型, 并采用连续变量标签训练的监督回归方法, 能够通过新型含能材料的结构对其性能进行预测, 具有高准确性和高效率的特点^[10-11]。采用数据驱动策略辅助新材料研制的关键环节在于数据库的搭建以及机器学习模型的构建。其中, 数据资源主要来源于文献数据、软件程序计算数据, 以及从 CCDC 和 PubChem 等数据库中获取的格式化数据等^[9-10]。机器学习则是通过数据处理和模型拟合实现模式识别与预测, 其模型构建主要包括算法选择和误差评估^[10-11]。常用的算法包括线性回归、支持向量机 (SVM)、决策树、核岭回归 (KRR) 和神经网络 (NN) 等^[12-13], 误差评估参数则主要包括决定系数 (R^2)、相关系数 (r)、均方误差 (MSE) 和平均绝对误差 (MAE) 等^[14]。

数据驱动策略的引入, 不仅能够提高含能材料的研发效率, 还可以为实现特定功能含能材料的定向合成提供支持, 从而推动国防科技的高速发展。近年来, 各国科研人员在基于数据驱动策略辅助的含能材料设计合成领域取得了阶段性的研究进展与成果。本综述系统地梳理了数据驱动策略在含能材料设计合成领域

收稿日期: 2025-04-23; 修回日期: 2025-05-22

网络出版日期: 2025-05-27

基金项目: 国家自然科学基金 (22375189)

作者简介: 王海风 (1997-), 女, 博士, 主要从事新型含能材料设计与合成研究。e-mail: whf202133@163.com

通信联系人: 王康才 (1984-), 男, 副研究员, 主要从事新型含能材料设计与合成研究。e-mail: wangkangcai@caep.cn

引用本文: 王海风, 王康才, 刘渝. 数据驱动策略在含能材料设计及性能预测中的应用[J]. 含能材料, DOI:10.11943/CJEM2025076.

WANG Hai-feng, WANG Kang-cai, LIU Yu. Application of Data-driven Strategies in Energetic Material Design and Performance Prediction[J]. *Chinese Journal of Energetic Materials (Hanneng Cailiao)*, DOI:10.11943/CJEM2025076.

的应用,全面概述了机器学习辅助的材料设计、高通量筛选与性能预测等方面的相关研究进展,深入探讨了未来含能材料发展面临的挑战,并对未来含能材料的发展方向提出了展望。

1 机器学习在含能分子及共晶设计中的应用

采用传统“试错法”开发新型含能分子往往效率低下,难以满足现代军事与工业领域对高性能含能分子的迫切需求^[15]。同时,含能分子的合成与表征过程极为复杂繁琐,使得整个研发周期被大幅拉长,研发成本也居高不下^[16]。因此,在新型含能分子设计之初即要追求其热稳定性、能量密度和感度等多个关键性能指标之间相对平衡,以使其具备优异的综合性能。通过机器学习技术预先精准评估材料性能,能够显著提高研发效率,为含能分子的高效发展开辟新的机遇与可能。

现有熔铸载体炸药(如TNT)因存在能量较低且毒性大等问题,给其应用带来了诸多挑战。为此,中国工程物理研究院化工材料研究所张庆华等^[17]于2021年通过将机器学习技术与实验相结合,实现了新

型熔铸炸药的高效设计与合成。他们的具体研究思路为:(1)高通量生成含能分子:基于排列组合的启发式枚举方法,采用高通量手段,批量生成大量含能分子结构,为后续筛选提供丰富的候选对象。(2)多模型性能预测:构建五种机器学习模型,分别用于预测材料的密度、熔点、分解温度、爆轰速度和爆轰压力等关键性能指标,从而快速筛选出具有理想综合性能的含能分子。(3)实验验证与优化:通过实验手段,对筛选出的优选目标材料进行合成与性能测试,进一步验证模型预测的准确性,并优化材料性能^[17]。该研究从设计的海量候选分子中筛选出136个具有高能量密度、低熔点和高分解温度的候选材料,最终通过实验手段合成了其中8个新化合物(MC-1~MC-8,结构及性能如图1所示),并对其爆轰速度、爆轰压力、熔点、分解温度和机械感度等关键性能进行了验证^[17]。结果表明,MC-1~MC-8的熔点均低于110℃,分解温度均超过170℃,熔点与分解温度之间的温差接近或大于100℃,表现出优异的热稳定性与加工性能。其中,MC-1和MC-5的综合性能尤为突出:MC-1的熔点为75.9℃,分解温度为208.7℃,爆轰速度为8.139 km·s⁻¹,爆轰压力为28.1 GPa;MC-5的熔点为88.2℃,分解温

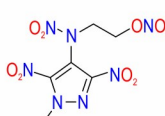
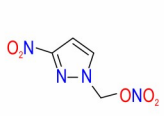
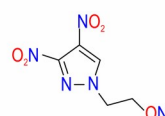
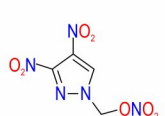
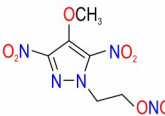
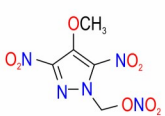
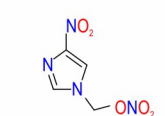
MC-1  D_v: 8.139 km·s⁻¹ P: 28.1 GPa T_m: 75.9 °C T_d: 177.3 °C IS: 8 J FS: 96 N	MC-2  D_v: 8.020 km·s⁻¹ P: 26.9 GPa T_m: 68.4 °C T_d: 161.4 °C IS: >20 J FS: 144 N	MC-3  D_v: 8.165 km·s⁻¹ P: 28.8 GPa T_m: 65.2 °C T_d: 192.3 °C IS: >20 J FS: 96 N	MC-4  D_v: 8.705 km·s⁻¹ P: 33.8 GPa T_m: 97.1 °C T_d: 169.8 °C IS: >20 J FS: 160 N
MC-5  D_v: 7.912 km·s⁻¹ P: 26.3 GPa T_m: 88.2 °C T_d: 189.7 °C IS: >20 J FS: 96 N	MC-6  D_v: 8.238 km·s⁻¹ P: 29.4 GPa T_m: 90.2 °C T_d: 153.3 °C IS: >20 J FS: 96 N	MC-7  D_v: 8.175 km·s⁻¹ P: 29.0 GPa T_m: 106.9 °C T_d: 165.6 °C IS: >20 J FS: 96 N	MC-8  D_v: 7.843 km·s⁻¹ P: 25.2 GPa T_m: 76.2 °C T_d: 167.9 °C IS: >20 J FS: 96 N

图1 化合物MC-1~MC-8的分子结构、爆轰速度、爆轰压力、熔点、分解温度及机械感度^[17]

Fig.1 Molecular structure, detonation velocity, detonation pressure, melting point, decomposition temperature and mechanical sensitivity of MC-1~MC-8^[17]

度为 219.8 °C,爆轰速度为 7.912 km·s⁻¹,爆轰压力为 26.3 GPa。这些化合物不仅密度和爆轰性能均优于 TNT,而且具有优异的机械感度性能,展现出良好的安全性和广阔的应用前景^[17]。由此可见,机器学习技术在含能材料设计与合成中具有巨大的应用潜力。

为解决传统“试错法”在研制新型高性能含能材料时效率较低,以及开发具有特定性能的含能材料难度较大的问题,2022 年张庆华^[18]等开发了一个结合机器学习模型和高通量分子生成的系统,用于快速筛选具有指定性能的新型含能材料,并成功合成了具有优异综合性能的新型含能材料 ICM-104。该研究工作的具体思路如下:(1)数据收集与增强:从 CCDC 数据库中精心筛选出 365 个具有非石墨堆积特性的晶体数据,以及 22 个类石墨堆积晶体数据,之后借助 SMILES 枚举技巧对数据进行增强处理,巧妙地将样本总数扩充至 4000 多个,为后续模型训练提供了丰富的数据基础。(2)特征提取与模型构建:运用 RDKit 库提取自定义描述符和电拓扑指纹特征,这些特征能够精准捕捉含能材料的结构与性能之间的内在联系。在此基础上,采用核岭回归(KRR)算法对性质模型进行训练,确保模型具备高度的准确性和泛化能力。(3)分子生成与高效筛选:通过枚举方法批量生成 25,112 个可能的含能分子结构。利用训练有素的机器学习模型,对这些分子的性质进行精准预测与筛选。通过设定合理的筛选标准,从海量分子中精准筛选出 99 个最具潜力的候选分子,显著提高了研发效率,减少了不必要的实验尝试。(4)实验合成与性能验证:从筛选出的候选分子中,选择最具前景的分子 ICM-104 进行实验合成,并对其各项性能指标进行全面测试。测试结果显示,ICM-104 的密度为 1.825 g·cm⁻³,热分解温度高达 326 °C,爆速达到 8551 m·s⁻¹,爆压为 29.8 GPa,撞击

感度为 35 J,摩擦感度为 360 N,这些性能指标与模型预测结果高度一致^[18]。ICM-104 的合成与性能验证,不仅展示了机器学习技术在特定性能含能材料的研制中具有巨大潜力,还为含能材料的未来发展提供了一种全新的研发范式。

2022 年,针对设计合成能量与 CL-20 相当的高能材料这一极具挑战性的任务,西安近代化学研究所刘英哲^[19]等通过组合库设计与高通量筛选相结合的方式,利用机器学习技术显著提升 CHONF 类高能材料的研发效率。该团队构建了一个包含约 10 万种分子的庞大组合库,并以氧氟平衡、爆速、撞击感度和合成难度为关键标准,开展高通量筛选工作,旨在推动高能密度含能材料的发展。通过对这些标准的严格筛选,该团队成功锁定了综合性能排名前 10 的含能分子(结构如图 2 所示),并对其进行了详细的量子化学计算。结果发现,这些候选含能分子的密度均高于 1.8 g·cm⁻³,且大多数分子的爆速高于 CL-20,同时其生成焓也与 CL-20 相当,表明候选分子在能量密度和爆轰性能方面具有显著优势。该研究通过组合库设计和高通量筛选的方法,成功筛选出了一批具有高爆速的 CHONF 类含能分子。这种方法不仅显著提高了含能材料的设计效率,还为后续分子设计和实验研究提供了极具潜力的候选分子。

2023 年,中国工程物理研究院化工材料研究所张朝阳团队^[20]报道了名为“EM Database v1.0”的含能材料(EMs)数据平台,旨在为含能材料的数据存储和共享提供基准信息,通过数据驱动的方法加速新型含能材料的发现。该平台结合了约 10 万种化合物的量子化学(QC)计算数据和约 1 万种化合物的实验数据,涵盖了从几何构型、电子结构到爆轰性质等多种关键信息,可通过环/链骨架和官能团组合生成含能分子。

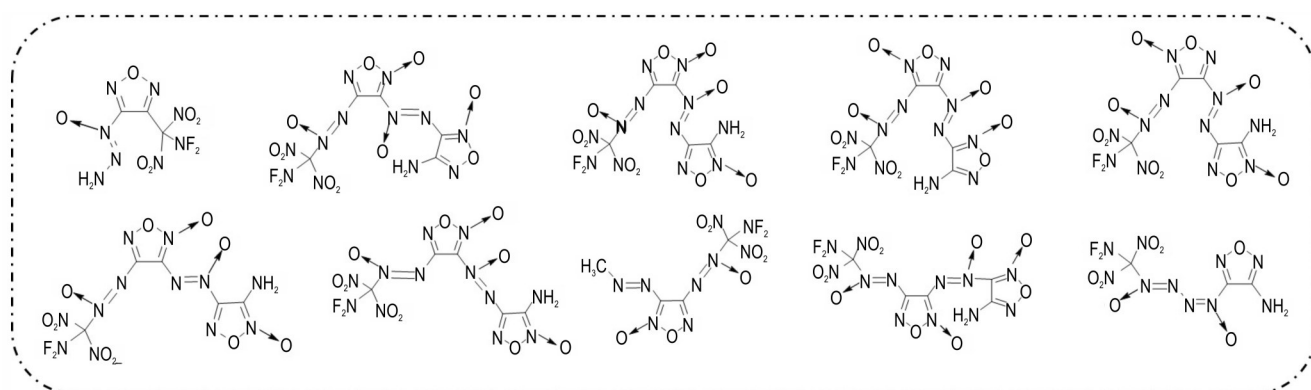


图 2 基于机器学习设计的 CHONF 高能材料的分子结构^[19]

Fig.2 Molecular structures of high energetic materials designed by machine learning based on CHONF^[19]

EM Database v1.0通过标准化的数据收集和处理流程,确保了数据质量,促进了含能材料研究的规范化和可重复性,并且可以结合了电势(ESP)相关经验方程和K-J经验方程,用于校正和预测性质。该平台不仅为研究人员提供了大规模数据支持,还通过高精度的性质预测模型减少了实验成本和时间,推动了新型含能材料从分子设计到实验验证的全流程研究,有望显著加速含能材料的开发和应用。

机器学习技术在含能材料设计领域的另一应用场景是含能共晶。近年来,共晶工程在含能材料领域展现出极高的研究和应用价值,但传统实验方法在系统性筛选含能共晶组分时存在耗时、费力且资源消耗巨大的问题。鉴于此,四川大学蒲雪梅团队^[21]创新性地提出采用机器学习技术,从理论上预测含能共晶组分,从而显著提高了含能共晶的研制效率。基于这一思路,蒲雪梅团队构建了一个包含7871个样本的庞大且高质量的数据集,并提出了一种名为Co-Crystal Graph Network(CCGNet)的灵活图神经网络(GNN)模型。CCGNet模型的设计巧妙,包含消息传递阶段和读出阶段。在消息传递阶段,通过多个CCGBlocks实现节点嵌入的高效传播与动态更新;在读出阶段,引入全局注意力机制优化特征空间,从而提升模型对复杂含能共晶体系的表征能力。为了全面评估CCGNet模型的性能,蒲雪梅团队将其与七种主流竞争模型(包括支持向量机(SVM)、随机森林(RF)、图卷积网络(GCN)、Graph-CNN、enn-s2s、DNN-des和DNN-FP)进行了严格对比。所有模型均在十折交叉验证集上进行训练与评估。研究结果显示,CCGNet在十折交叉验证集上的平衡准确率(BACC)高达98.54%,显著优于其他竞争模型(最高为94.46%)。在独立测试中,针对含能共晶体系,经过转移学习优化后,CCGNet的BACC分别提升至97.83%(TNT)和97.22%(CL-20),进一步验证了其强大的泛化能力和预测精度^[21]。基于CCGNet的精准设计及预测结果,蒲雪梅团队成功合成了一种全新的含能共晶——CL-20/1-甲基-4-硝

基吡唑。这一成果不仅展示了机器学习技术在含能共晶设计中的巨大潜力,还为含能共晶的高效研制提供了可靠的理论依据和创新的实验方法。

这些工作意味着含能材料的研发模式从传统的“试错法”向智能化、高效化以及理论驱动的方向迈进,展示了机器学习技术在新型含能分子及共晶设计中的巨大潜力,为新型含能材料的研发开辟了新的路径,有望推动含能材料领域迈向新的发展阶段。

2 含能材料单一性能预测模型的构建

密度、热稳定性、能量和感度是含能材料的重要性能参数,这些性能参数的预测能够直接提高含能材料研发速率。近年来,各国科研人员采用机器学习技术对含能材料性能的预测做了大量研究工作,并取得了不错的研究进展,推动了含能材料的发展。

从K-J方程可以看出,含能材料密度与能量正相关,密度的大小直接影响含能材料的爆速和爆压^[22]。机器学习作为一种新兴的性能预测技术,能够利用材料结构与密度之间的内在关系,基于数据库参数直接高效预测材料的密度,且预测误差较小。2021年,西南科技大学杨春明等^[23]依托CCDC数据库中的2002种中性硝基化合物的晶体学数据,采用随机森林(RF)、图神经网络(GNN)和支持向量机(SVM)三种算法建立模型,预测含能材料的密度。研究结果表明:GNN模型仅使用分子图作为输入即可实现含能材料密度的高准确度预测。同年,Nguyen等^[24]从CSD中筛选数据,组建了包含10521个分子的数据库,并基于支持向量机(SVM)、随机森林(RF)、偏最小二乘回归(PLSR)和消息传递神经网络(MPNN)四种不同的机器学习算法训练密度预测模型。该工作中,Nguyen等^[24]分别采用Extended 3D Fingerprint、RDKit中的2D分子描述符集以及RDKit中带有原子和键描述符的图表示法用于密度预测。如表1所示,MPNN拥有最高的计算准确度,在大多数密度区间内的误差相对稳定,小于0.05 g·cm⁻³。

表1 不同方法预测高能炸药密度对应的R²和RMSE值^[24]

Table 1 R² and RMSE values of different featurization and method combinations for HE-related density prediction^[24]

Feature	Input information	Feature processing	Method	R ² 1)	RMSE 2)
E3FP	Atomic, 3D positions	precomputed	SVR	0.683	0.085
RDKit (molecular)	Physicochemical/mathematical	precomputed	RF	0.878	0.053
RDKit (molecular)	Physicochemical/mathematical	precomputed	PLSR	0.900	0.048
RDKit (atom/bond)	Atomic/bond, molecule graph	learned	MPNN	0.914	0.044

Note: 1) R² is Coefficient of Determination; 2) RMSE is root mean squared error.

热稳定性是含能材料的重要物化性能参数之一,大量含能材料因为热稳定性较差而不能应用在武器装备中。如果能够在合成前,提前预测材料的热稳定性,那么将极大地提高含能材料的研发效率。然而,含能材料热分解机制涉及复杂的物理和化学过程,传统的计算方法耗时且计算成本高,且对于不同类型材料的普适性较差^[25]。2023年,Wu等^[26]构建了含1022个含能分子的数据集,并基于此开发了具有较高热分解温度预测准确性的梯度提升回归(GBR)模型,其在测试集上的平均绝对误差为27.7℃。通过对异常结果的分析,该研究认为结合分子相互作用相关特征可能会进一步提高模型的准确性^[26]。同年,西安近代化学研究所刘英哲等^[27]建立了包含885种含能化合物热分解温度的数据库,采用代表性回归算法(包括核岭回归(KRR)、最小绝对收缩和选择算子(LASSO)、随机森林(RF)、极端随机树、自适应提升(AdaBoost)、梯度提升(GB)和极端梯度提升(XGBoost))预测含能材料的热分解温度。该研究将数据集以4:1的比例随机分为训练集和测试集,使用网格搜索进行超参数调优,并采

用5折交叉验证来减少过拟合,结果表明RF模型的预测结果最为准确。然而,考虑到较大的样本量会导致数据集信息更加复杂,使得预测难度和误差增加,该研究以MACCS分子指纹图谱为标准,将数据集中的化合物分为四类,然后使用RF算法对每类化合物进行建模分析,其预测结果与实验结果的对比如图3所示,表明含能化合物的热分解与分子组成、电子分布、化学键性质和取代基类型等因素有关。针对数据集的复杂性,该团队于2024年建立了1091种化合物的5个通用描述符集,并将它们与9种算法相结合,构建了一套平均绝对误差范围为41~29 K的预测模型,其中使用RDKit描述符构建的模型表现最好,平均绝对误差为29.38 K^[28]。然而,强调分子性质的描述符集不足以提高某些化合物的预测精度,仍需要探索更高水平的描述符,尤其是晶体属性的描述符^[28]。

生成焓是计算含能材料爆速和爆压不可或缺的重要参数。早在2004年,深圳大学王芳等^[29]就采用神经网络算法,预测了多硝基芳香族化合物的生成焓。2021年,Mathieu^[30]利用两个深度学习模型预测了材

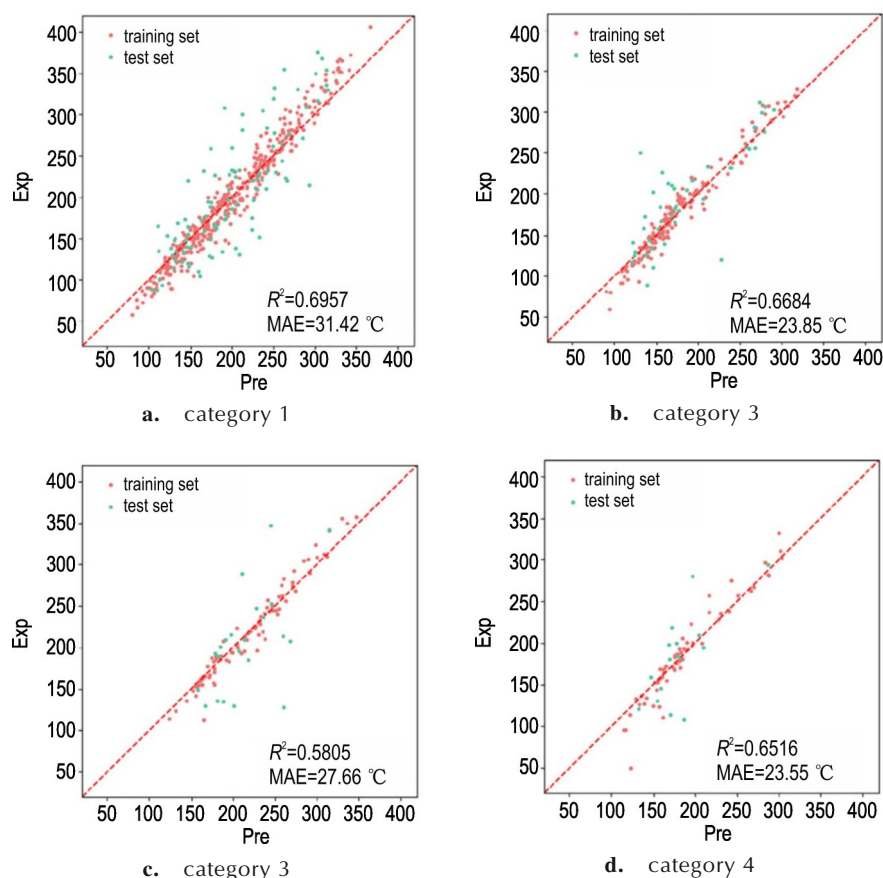


图3 由RF模型导出的四种热分解温度和实验结果的对比图^[27]

Fig.3 Comparison of the experimental results of thermal decomposition temperatures of the four categories derived from the RF model^[27]

料电子能量和振动频率,发现可以利用标准原子等价方案结合预测的能量和频率来计算有机化合物的生成焓。然而,目前严重缺乏有关含能材料生成焓的数据库。为了从相对较小的数据库有效地提取所需的物理化学性质,Chen等^[31]提出了空间矩阵描述符的概念,并基于451个含能分子的较小数据库,开发了体积占用空间矩阵(VOM)和热贡献空间矩阵(HCM)两种新的特征化方法,训练后其对含能材料晶体密度和生成焓的预测精度能够达到 $0.035\text{ g}\cdot\text{cm}^{-3}$ 和 $10\text{ kcal}\cdot\text{mol}^{-1}$,接近Kim等^[32]提出的优秀预测水平。然而,当前的空间矩阵描述符仅考虑了体积占据和热贡献的单体(1-body)和双体(2-body)指数,如果能够引入更高阶的多体效应以及更多物理化学符号,则有望构建出更强大的机器学习模型。

爆轰性能是评估含能材料能量水平的重要参数,其的精准预测能够在极大程度上提高含能材料的研发效率。然而,现有理论计算方法预测材料爆轰性能时,需要同时获得其密度、生成焓和化学组成等数据。这为含能材料爆轰性能的预测带来了极大的挑战。2019年,Chandrasekaran等^[33]开发了一种基于小数据库的ANN爆速预测模型,该模型考虑了材料的生成焓、密度、物理状态和元素组成等对预测精度的影响,实现了在小数据集上高准确度地预测含能材料的爆速。2022年,Wang等^[34]提出了一种基于 $25\text{ }\mu\text{s}$ 左右高速激光诱导冲击波图像的机器学习模型,用于预测27种常规单质炸药的爆轰性能。该研究通过自主搭建的高速纹影成像系统,获取不同样品在约 $25\text{ }\mu\text{s}$ 延迟下的特征激光诱导冲击波图像,并将其作为机器学习的原始输入数据,最终通过方向梯度直方图(HOG)特征提取和支持向量回归(SVR)算法建立了爆速、爆热、爆容、爆压和爆温等参数的高精度预测模型,所有模型训练集的平均相对误差均 $<5\%$,其中爆速定量分析模型测试集的最大相对误差仅为 3.47% 。然而,深度学习在混合炸药爆轰性能预测、炸药状态方程参数的预测以及爆炸特性在预测炸药配方中的应用仍较少。鉴于此,Yang等^[35]基于人工神经网络(ANN)模型和一维卷积神经网络(CNN1D)模型,建立了三种改进的深度学习模型,通过将预测值与参考值进行比较测试模型对真实配方炸药的预测精度,结果表明,对爆压和爆速的预测模型误差分别在 10% 和 3% 以内。

机械感度是评估含能材料安全性的关键核心参数,大量含能材料因机械感度较高而未能在武器装备中获得应用^[36-37]。机械感度的精准预测有助于高安

全性含能材料的筛选,减少不必要的实验合成,提升研发效率及安全性。然而,含能材料的机械感度受材料结构、晶体形貌及环境等多方面因素的影响。现有研究表明,很难实现含能材料机械感度的精准预测。例如,Keshavarz等^[38]采用人工神经网络对含能材料机械感度的预测进行了深入研究,然而,其得出训练数据和测试数据的均方根误差分别为 $41\text{ }\mu\text{m}$ 和 $56\text{ }\mu\text{m}$ 。2025年,张炳儒等^[39]基于多种机器学习模型包括随机森林(RF)、极端梯度提升(XGBoost)、反向传播神经网络(BPNN)、支持向量机(SVM)、k-means聚类算法、卷积神经网络(CNN)、视觉Transformer(ViT),通过分析处理爆炸声信号,能够智能识别炸药的撞击感度。其研究表明,传统机器学习模型RF、XGBoost、BPNN和SVM在原始数据集上的准确率均高于 99.5% ,在条件生成对抗网络(cGAN)增强数据上最终可达到 100% ;而k-means聚类算法初始准确率为 98.5% ,在增强数据上准确率呈现先上升后下降的趋势;深度学习模型CNN和ViT因受限于小样本环境和轻微的过拟合问题,在原始数据上的准确率分别为 98.1% 和 98.4% ,在增强数据上达到 98.4% 和 98.9% ,还有一定的提升空间。

上述实例证明机器学习凭借其强大的数据处理和模式识别能力,为含能材料单一性能的高效预测提供了新途径。在密度、热稳定性、生成焓、爆轰性能及感度预测方面,研究人员开发的模型展现了高准确度和低误差,为新型含能材料的设计和合成提供了有力支持。然而,受制于数据库的匮乏、多参数耦合及影响因素复杂等问题,机器学习模型的精度仍有较大提升空间。总体而言,机器学习在含能材料性能预测中展现出巨大潜力,但仍需进一步优化算法、拓展数据集和深化理论研究,以实现更精准、高效的预测,推动含能材料研发的快速发展。

3 含能材料多种性能同步预测的前沿研究与挑战

含能材料作为国防科技的关键核心材料,其应用潜力评估具有显著的“短板”效应,需要同时满足不同方面的多项性能要求,这为高性能含能材料的发展带来了巨大的挑战^[40]。如果能够实现含能材料多种性能同步预测,将在极大程度上提升含能材料的研发效率,推动含能材料的高效发展。基于此,大量从事机器学习技术研究的科研人员致力于开发一种通过“一次运行”来预测多个性能值的模型。

2020年, Casey等^[41]以氧平衡为参数, 从 Generated Databases (GDB) 数据库中筛选出 26265 个潜在含能分子, 利用高斯和猎豹软件计算了这些分子的电子和能量属性, 并将结果作为卷积神经网络的训练数据。基于此, 该研究利用三维卷积神经网络 (CNN) 预

测了含能材料的爆轰温度、爆轰速度、爆轰压力、生成焓和密度等 (图 4), 结果表明他们所用模型的 R^2 超过了 0.9, 达到了理想的误差范围。然而, 由于 GDB 数据库中含能化合物的数量有限, 这些模型对于含能材料的适用性还有待验证^[40]。

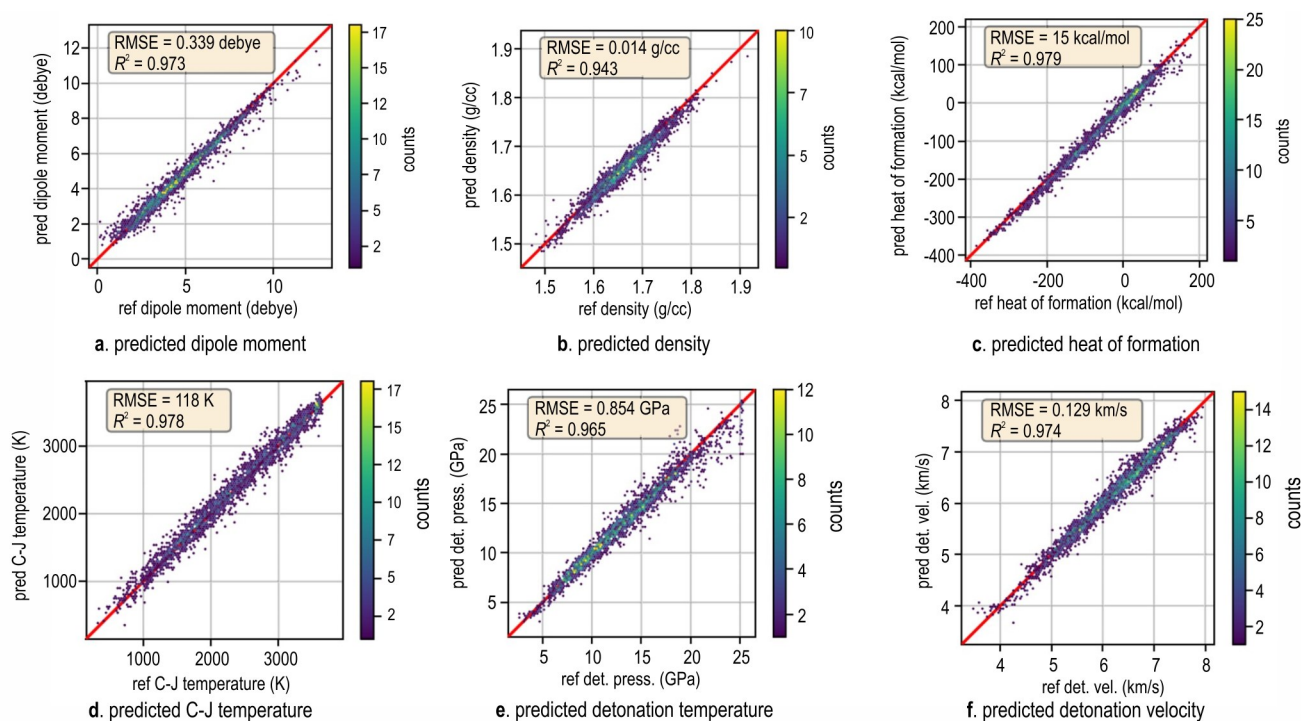


图 4 CNN 预测的含能材料性能数据^[40]

Fig.4 The performances of energetic materials predicted by CNN^[40]

Huang等^[42]通过高通量量子力学计算建立了包含 153 种高能量密度材料性能的数据库, 并按 4:1 的比例将这些数据分为训练数据和测试数据。基于该数据库, 采用核岭回归 (KRR)、随机森林 (RF)、极端梯度提升回归树 (XGBoost)、自适应增强 (AdaBoost) 回归器和多层感知器 (MLP) 进行模型训练, 并将预测结果与实验值进行了比较, 研究结果表明, XGBoost 模型在几乎所有性能的预测中都表现出色。此外, 该研究收集了 109 个分解温度 (T_d) 的实验值以及 612 个关于爆速 (D)、爆压 (P)、爆热 ($Q_{\max}$) 和晶格能量 (LE) 的计算数据, 还收集了 203 个实验值来验证模型的预测 (图 5), 发现除分解温度外, 其他所有性能在测试集上预测结果的决定系数都超过了 0.8。

2021 年, Hou 等^[43]基于包含 436 个含 N 分子结构和爆轰性能 (包括 TNT、CL-20、HMX 和 RDX 等典型炸药) 的数据库, 选用 MATLAB 软件构建了具有 10 个隐藏神经元和 1 个输出层的神经网络模型 (图 6), 仅需 0.038 秒即可预测所有含能分子的爆速、爆压和密度,

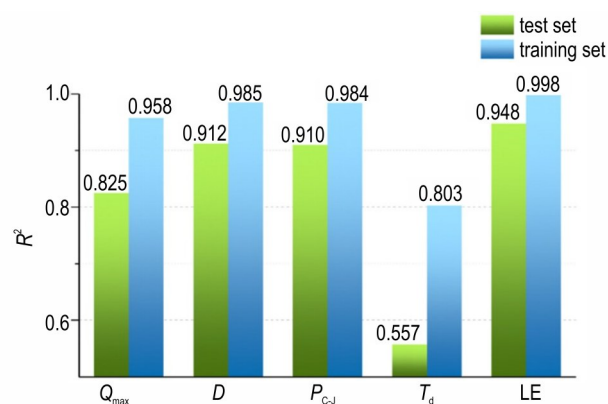
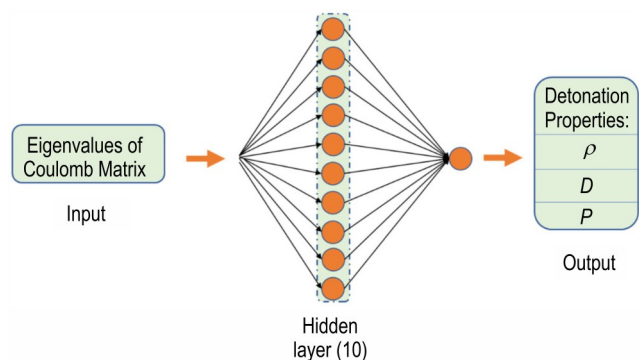


图 5 使用 XGBoost 进行性能预测的决定系数分布^[41]

Fig.5 Coefficient of determination distributions of the performance prediction using XGBoost^[41]

展示出机器学习的高通量特性。该研究通过建立合适的神经网络降低了含能材料爆速、爆压和密度的预测误差, 其平均绝对误差分别为 $0.3456 \text{ km} \cdot \text{s}^{-1}$ 、 1.4933 GPa 和 $0.0259 \text{ g} \cdot \text{cm}^{-3}$, 这表明他们的神经网络模型表现出色, 能够同时准确预测含能分子的爆速、爆压和密度。

图6 神经网络模型示意图^[42]Fig.6 Schematic diagram for the model of the neural network^[42]

上述实例通过构建不同类型的机器学习模型,初步实现了含能材料多项性能的同步预测,证实了机器学习在处理复杂含能材料性能数据方面和多性能同步预测方面的有效性和巨大潜力,为具有优异综合性能的含能材料研发提供了新策略。然而,目前相关研究案例较少,数据集规模和多样性不足,限制了模型的普适性和精准度。此外,不同性能参数间的复杂相互关联尚未被充分挖掘与整合,影响了预测结果的可靠性。未来研究应聚焦于探索多性能参数的耦合,开发更复杂且高效的模型架构,以突破现有瓶颈,实现含能材料多性能同步预测的高效与精准,为高性能含能材料的研发提供更有力的支持。

4 总结与展望

本综述梳理了近年来关于机器学习辅助含能材料设计及性能预测的最新研究进展。机器学习技术的发展提高了含能材料的研发效率。依托含能材料领域现有研究基础,采用枚举法及计算辅助设计,能够实现含能分子的高通量生成。通过机器学习可高效预测含能材料的性能,快速筛选出更具潜力的目标含能分子。之后通过先进的有机合成方法,有望实现目标分子的现实合成。整体来看,上述含能材料高效设计合成策略具有较高的可行性。然而,从现有研究进展来看,采用该策略的研究中仅有较少案例获得了目标含能分子。由此可见,依托机器学习辅助设计合成具有特定性能的含能材料依然充满了挑战,未来可考虑以下方向的进一步发展:

(1)控制数据质量并构建标准化体系。含能材料性能数据库的构建是机器学习技术在该领域应用的基础。然而,目前含能材料实验数据分散、性能测试方法多样,导致难以精准把控数据质量,从而严重制约了机

器学习技术在该领域的应用。因此,建立跨机构数据共享机制、统一性能测试标准、开发数据增强技术,是推动机器学习技术在含能材料领域广泛应用的基础和保障。

(2)开发可解释机器学习模型。机器学习技术采用的是黑箱模型,这限制了含能材料结构与性能间内在关系的深入解析。只有发展可解释的机器学习方法(如符号回归),深入认识材料结构与性能之间的构效关系,提高机器学习模型的准确性,才能进一步推动机器学习技术在含能材料领域的发展。

(3)建立跨学科交叉融合体系。目前通过机器学习设计了大量具有优异性能的潜在含能分子,但通过实验合成获得的实物样品案例较少。这主要是由于设计者和实验者之间存在的认知差异所导致的,可见依靠单一学科知识很难形成“设计-预测-验证”研发模式的闭环。因此,机器学习辅助含能材料设计与合成,需结合计算科学、材料学、有机化学及先进实验表征等多学科知识。加强跨学科交叉融合是推动机器学习辅助含能材料高效研发的重要措施。

参考文献:

- [1] 董海山, 高能量密度材料的发展及对策[J]. 含能材料, 2004, 12(Z1): 1-12.
DONG Hai-shan, The development and countermeasure of high energy density materials[J]. *Chinese Journal of Energetic Materials (Hanneng Cailiao)*, 2004, 12(Z1): 1-12.
- [2] ZLOTIN S, CHURKOV A, EGOROV M, et al. Advanced energetic materials: Novel strategies and versatile applications[J]. *Mendeleev Communications*, 2021, 31(6): 731-749.
- [3] GAO H, ZHANG Q, SHTEEV J. Fused heterocycle-based energetic materials (2012-2019)[J]. *Journal of Materials Chemistry A*, 2020, 8: 4193-4216.
- [4] LI H, ZHANG L, PETRUTIK N, et al. Molecular and crystal features of thermostable energetic materials: Guidelines for architecture of “bridged” compounds[J]. *ACS Central Science*, 2020, 6(1): 54-75.
- [5] ZHANG J, TAN B, ZHANG Q, et al. Strategies for molecular construction and performance regulation of heat-resistant energetic materials: An overview[J]. *Progress in Natural Science: Materials International*, 2024, 33(6): 1132-1157.
- [6] LI C, YU Q, ZHANG J. Design, synthesis, and properties of melt-cast explosive carriers: A review[J]. *Journal of Materials Science*, 2024, 59: 14172-14184.
- [7] YUAN W, HE L, TAO G, et al. Materials-genome approach to energetic materials[J]. *Accounts of Materials Research*, 2021, 2(9): 692-696.
- [8] 刘锐, 刘建, 唐岳川, 等. 人工智能辅助含能分子设计的应用与展望[J]. 含能材料, 2024, 32(4): 408-421.
LIU Rui, LIU Jian, TANG Yue-chuan, et al. Applications and prospects of AI-assisted design of energetic molecules[J]. *Chinese Journal of Energetic Materials (Hanneng Cailiao)*, 2024,

- 32(4): 408–421.
- [9] 王润文, 杨春明, 刘建. 高通量计算与深度学习相结合的稠环含能化合物设计[J]. 含能材料, 2022, 30(12): 1226–1236. WANG Run-wen, YANG Chun-ming, LIU Jian. Exploring novel fused-ring energetic compounds via high-throughput computing and deep learning[J]. *Chinese Journal of Energetic Materials (Hanneng Cailiao)*, 2022, 30(12): 1226–1236.
- [10] ZANG X, ZHOU X, BIAN H, et al. Prediction and construction of energetic materials based on machine learning methods[J]. *Molecules*, 2023, 28(1): 322.
- [11] TIAN X, SONG S, CHEN F, et al. Machine learning-guided property prediction of energetic materials: Recent advances, challenges, and perspectives[J]. *Energetic Materials Frontiers*, 2022, 3: 177–186.
- [12] VANGRSSEL F, PERRY E, MOHAN S, et al. Natural language processing for knowledge discovery and information extraction from energetics corpora[J]. *Propellants, Explosives, Pyrotechnics*, 2023, 48(11): e202300109.
- [13] JUAN Y, DAI Y, YANG Y, et al. Accelerating materials discovery using machine learning[J]. *Journal of Materials Science & Technology*, 2021, 79: 178–190.
- [14] CHEN C, ZUO Y, YE W, et al. A critical review of machine learning of energy materials[J]. *Advanced Energy Materials*, 2020, 10(8): 1903242.
- [15] WANG Y, HU L, PANG S, et al. Nitroimino as an energetic group in designing energetic materials for practical use, a tautomerism from nitroamino[J]. *Journal of Materials Chemistry A*, 2023, 11: 13876–13888.
- [16] MURAVYEV N, WOZNIAK D, PIERCEY D. Progress and performance of energetic materials: Open dataset, tool, and implications for synthesis[J]. *Journal of Materials Chemistry A*, 2022, 10: 11054–11073.
- [17] SONG S, CHEN F, WANG Y, et al. Accelerating the discovery of energetic melt-castable materials by a high-throughput virtual screening and experimental approach[J]. *Journal of Materials Chemistry A*, 2021, 9: 21723–21731.
- [18] SONG S, WANG Y, CHEN F, et al. Machine learning-assisted high-throughput virtual screening for on-demand customization of advanced energetic materials[J]. *Engineering*, 2022, 10(3): 99–109.
- [19] WEN L, WANG B, YU T, et al. Accelerating the search of CHONF-containing highly energetic materials by combinatorial library design and high-throughput screening[J]. *Fuel*, 2022, 310: 122241.
- [20] HUANG X, QIAN W, LIU J, et al. EM Database v1.0: A benchmark informatics platform for data-driven discovery of energetic materials[J]. *Energetic Materials Frontiers*, 2024, 5: 267–273.
- [21] JIANG Y, YANG Z, GUO J, et al. Coupling complementary strategy to flexible graph neural network for quick discovery of coformer in diverse co-crystal materials[J]. *Nature Communications*, 2021, 12: 5950.
- [22] KAMLET M, JACOBS S. Chemistry of Detonations. I. A Simple Method for Calculating Detonation Properties of C—H—N—O Explosives[J]. *Journal of Chemical Physics*, 1968, 48: 23–35.
- [23] YANG C, CHEN J, WANG R, et al. Density prediction models for energetic compounds merely using molecular topology[J]. *Journal of Chemical Information and Modeling*, 2021, 61(6): 2582–2593.
- [24] NGUYEN P, LOVELAND D, KIM J, et al. Predicting energetics materials' crystalline density from chemical structure by machine learning[J]. *Journal of Chemical Information and Modeling*, 2021, 61(5): 2147–2158.
- [25] LIU Z. Review and prospect of thermal analysis technology applied to study thermal properties of energetic materials[J]. *Fire-PhysChem*, 2021, 1(3): 129–138.
- [26] WU J, SONG S, TIAN X, et al. Machine learning-based prediction and interpretation of decomposition temperatures of energetic materials[J]. *Energetic Materials Frontiers*, 2023, 4(4): 254–261.
- [27] ZHANG Z, CAO Y, CHEN C, et al. Machine learning-assisted quantitative prediction of thermal decomposition temperatures of energetic materials and their thermal stability analysis[J]. *Energetic Materials Frontiers*, 2024, 5(4): 274–282.
- [28] ZHANG Z, CHEN C, CAO Y, et al. Descriptors applicability in machine learning-assisted prediction of thermal decomposition temperatures for energetic materials: Insights from model evaluation and outlier analysis[J]. *Thermochimica Acta*, 2024, 735: 179717.
- [29] 王芳, 刘剑洪, 田德余, 等. 用人工神经网络法预估芳香族多硝基化合物生成焓[J]. 含能材料, 2004, 12(4): 207–213. WANG Fang, LIU Jian-hong, TIAN De-yu, et al. Prediction of the enthalpy of formation for aromatic polynitro compounds with artificial neural network[J]. *Chinese Journal of Energetic Materials (Hanneng Cailiao)*, 2004, 12(4): 207–213.
- [30] MATHIEU D. Molecular energies derived from deep learning: application to the prediction of formation enthalpies up to high energy compounds[J]. *Molecular Informatics*, 2021, 41(5): 2100064.
- [31] CHEN C, LIU D, DENG S, et al. Accurate machine learning models based on small dataset of energetic materials through spatial matrix featurization methods[J]. *Journal of Energy Chemistry*, 2021, 63: 364–375.
- [32] KIM C, CHO S, KIM C, et al. Prediction of densities for solid energetic molecules with molecular surface electrostatic potentials[J]. *Journal of Computational Chemistry*, 2008, 29(11): 1818–1824.
- [33] CHANDRASEKARAN N, OOMMEN C, KUMAR V, et al. Prediction of detonation velocity and N—O composition of high energy C—H—N—O explosives by means of artificial neural networks[J]. *Propellants, Explosives, Pyrotechnics*, 2019, 44(5): 579–587.
- [34] WANG X, HE Y, CAO W, et al. Fast explosive performance prediction via small-dose energetic materials based on time-resolved imaging combined with machine learning[J]. *Journal of Materials Chemistry A*, 2022, 10: 13114–13123.
- [35] YANG Z, RONG J, ZHAO Z. Study on the prediction and inverse prediction of detonation properties based on deep learning[J]. *Defence Technology*, 2023, 24: 18–30.
- [36] ZHANG J, FENG Y, BO Y, et al. One step closer to an ideal insensitive energetic molecule: 3, 5-diamino-6-hydroxy-2-oxide-4-nitropyrimidone and its derivative[J]. *Journal of the American Chemical Society*, 2021, 143(32): 12665–12674.
- [37] YU Q, SINGH J, STAPLES R, et al. Assembling Nitrogen-rich, thermally Stable, and insensitive energetic materials by polycyclization[J]. *Chemical Engineering Journal*, 2022, 431(3):

- 133235.
- [38] KESHAVARZ M, JAAFARI M. Investigation of the various structure parameters for predicting impact sensitivity of energetic molecules via artificial neural network[J]. *Propellants, Explosives, Pyrotechnics*, 2006, 31(3): 216–225.
- [39] 张炳儒, 李明, 文玉史, 等. 基于音频信号的含能材料撞击感度机器学习识别[J]. 含能材料, 2025, 33(2): 136–147.
ZHANG Bing-ru, LI Ming, WEN Yu-shi, et al. Machine learning recognition of impact sensitivity of energetic materials based on acoustic signals[J]. *Chinese Journal of Energetic Materials (Hanneng Cailiao)*, 2025, 33(2): 136–147.
- [40] ZLOTIN S, DALINGER I, MAKHOVA N. Nitro compounds as the core structures of promising energetic materials and versatile reagents for organic synthesis [J]. *Russian Chemical Reviews*, 2020, 89(1): 1–54.
- [41] CASEY A, SON S, BILIONIS I, et al. Prediction of energetic material properties from electronic structure using 3D convolutional neural networks [J]. *Journal of Chemical Information and Modeling*, 2020, 60(10): 4457–4473.
- [42] HUANG X, LI C, TAN K, et al. Applying machine learning to balance performance and stability of high energy density materials [J]. *iScience*. 2021, 24(3): 102240.
- [43] HOU F, MA Y, HU Z, et al. Machine learning enabled quickly predicting of detonation properties of N-containing molecules for discovering new energetic materials [J]. *Advanced Theory and Simulations*, 2021, 4(6): 2100057.

Application of Data-driven Strategies in Energetic Material Design and Performance Prediction

WANG Hai-feng, WANG Kang-cai, LIU Yu

(Institute of Chemical Materials, China Academy of Engineering Physics (CAEP), Mianyang 621999, China)

Abstract: Technological and industrial transformations driven by data science and artificial intelligence are profoundly impacting the field of materials science, presenting both unprecedented opportunities and significant challenges for the innovation of energetic materials. As an emerging technology, machine learning offers novel research paradigm for the molecular design and synthesis of energetic materials. It is expected to solve the long-standing bottlenecks such as low efficiency, high cost, and lengthy development cycles. Although some successful cases have been reported, the application of machine learning across the full research cycle of energetic molecules—design, screening, synthesis, and performance validation—remains in a relatively immature stage compared with the application in other advanced materials domains. This review outlines the current research status of machine learning-assisted development of energetic materials, summarizes its applications in molecular design, single-property prediction, and multi-property simultaneous prediction. Nonetheless, the use of machine learning in design and synthesis of energetic materials with targeted properties remains fraught with challenges. Future efforts should prioritize the control of data quality and the construction of standardization frameworks, the development of interpretable machine learning models, and the establishment of interdisciplinary integration platforms, further promoting the efficient creation of high-performance energetic materials.

Key words: data-driven; machine learning; energetic materials; performance prediction

CLC number: TJ55;O64

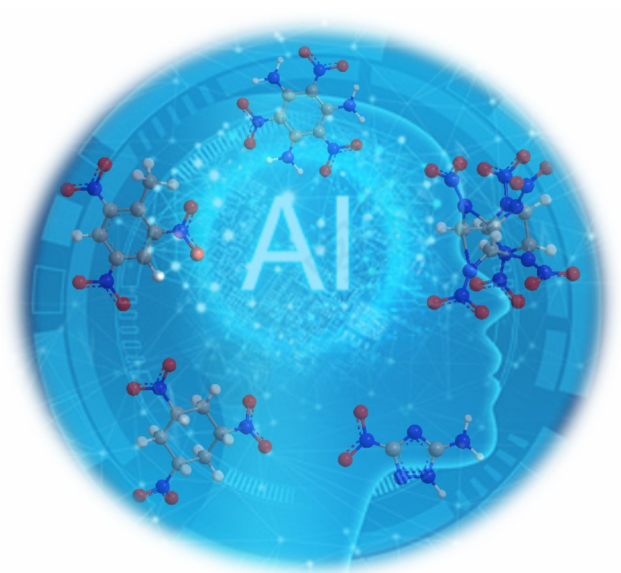
Document code: A

DOI: 10.11943/CJEM2025076

Grant support: the National Natural Science Foundation of China (No. 22375189)

(责编: 姜梅)

图文摘要:



Data-driven strategy is an efficient approach to develop new materials. The machine learning plays a key role in design, high-throughput screening, and performance prediction of energetic materials, providing a way for the efficient creation of high-performance energetic materials.