高等学校化学学报
CHEMICAL JOURNAL OF CHINESE UNIVERSITIES
综合评述

基于机器学习的负热膨胀研究进展

于坤杰[1], 张　智[1], 高其龙[2]
(1. 郑州大学电气与信息工程学院, 2. 物理学院, 郑州 450001)

摘要　在材料基因工程的背景下, 基于数据驱动的机器学习技术作为一种强大的新型工具, 在材料热膨胀性能研究领域受到了广泛关注. 机器学习能够绕过复杂的实验过程与理论计算, 通过建立描述符与热膨胀性能间的关联, 以较低成本快捷地实现材料热膨胀性能的预测, 有效弥补了传统实验试错法和基于密度泛函理论的方法存在的时间成本高、效率低等不足. 本文简述了机器学习的基本流程与方法, 并重点阐述其在材料热膨胀性能研究中的应用进展. 传统机器学习领域已实现从热膨胀系数单目标预测到引入特征重要性分析解析机制、结合特征选择优化模型, 再到多性能关联多目标预测的逐步深入. 在机器学习原子间势方面, 研究了通过构建原子尺度势能函数驱动分子动力学模拟, 进而揭示热膨胀行为的微观机制. 这些应用加速了负热膨胀材料的设计筛选, 深化了机理理解. 最后, 分析了机器学习在材料热膨胀性能研究中亟待解决的问题, 并据此展望了未来的研究方向与发展趋势.

关键词　数据驱动的机器学习; 热膨胀预测; 负热膨胀材料
中图分类号　O611.6　　**文献标志码**　A　　doi: 10.7503/cjcu20250375

Research Progress on Negative Thermal Expansion Based on Machine Learning

YU Kunjie[1], ZHANG Zhi[1], GAO Qilong[2*]
(*1. School of Electrical and Information Engineering*,
2. School of Physics, *Zhengzhou University*, *Zhengzhou* 450001, *China*)

Abstract　Against the backdrop of materials genetic engineering, data-driven machine learning(ML) techniques, as a powerful new tool, have garnered widespread attention in the field of research on materials' thermal expansion properties. ML can bypass complex experimental processes and theoretical calculations; by establishing correlations between descriptors and thermal expansion properties, it enables rapid prediction of materials' thermal expansion properties at relatively low cost, effectively compensating for the shortcomings of traditional experimental trial-and-error methods and density functional theory(DFT)-based approaches, such as high time costs and low efficiency. This paper outlines the basic processes and methods of machine learning, and focuses on elaborating its application progress in the research on the thermal expansion properties of materials. The field of traditional machine learning has achieved a gradual deepening, starting from the single-target prediction of the coefficient of thermal expansion, moving to the introduction of an analytical mechanism for feature importance analysis and the combination of feature selection to optimize models, and then advancing to the multi-target prediction associated with multiple properties. In the aspect of machine learning-based interatomic potentials, studies drive molecular dynamics simulations by constructing atomic-scale potential energy functions, thereby revealing the microscopic mechanism of thermal

收稿日期: 2025-12-12. 网络首发日期: 2026-01-17.
联系人简介: 高其龙, 男, 博士, 教授, 主要从事固体无机化学合成、晶体结构、负热膨胀及晶格动力学计算方面的研究.
E-mail: qilonggao@zzu.edu.cn
基金项目: 国家自然科学基金(批准号: 22471246, 22071221)资助.
Supported by the National Natural Science Foundation of China(Nos.22471246, 22071221).

expansion behavior. These applications have accelerated the design and screening of negative thermal expansion materials and deepened the understanding of relevant mechanisms. Finally, the paper analyzes the urgent issues to be resolved in the application of ML to the research on materials' thermal expansion properties, and proposes future research directions and development trends accordingly.

Keywords Data-driven machine learning; Thermal expansion prediction; Negative thermal expansion material

热膨胀是材料受热时体积发生变化的普遍现象，固体中热膨胀及热膨胀系数的不匹配会给器件和仪器带来诸多严重问题，如性能恶化、疲劳、暂时性或永久性损伤及开裂等，因此，热膨胀性能是影响高性能器件实际应用与工作寿命的关键因素[1]. 负热膨胀(Negative thermal expansion, NTE)材料[2]作为一类特殊功能材料，在温度升高时其体积会发生反常收缩，这能有效补偿常规材料的正热膨胀，大幅提高系统在温度变化环境下的尺寸稳定性和工作精度，在电子学、光学、生物医学及航空航天部件等温度变化剧烈的应用中至关重要[3].

负热膨胀是指材料的维度随温度升高而反常减小的现象. 这一特性主要源于化学键的非简谐效应，并与电子、自旋和声子等多种自由度的相互作用密切相关，从而形成了负热膨胀材料的多样性. 例如，在$Fe_{65}Ni_{35}$[4]，$Mn_3(Cu_{1-x}Ge_x)N$[5]和$La(Fe, Si, Co)_{13}$[6]等磁性材料中，NTE行为主要归因于磁体积效应；而在$BiNiO_3$[7]和$LaCu_3Fe_4O_{12}$[8]等钙钛矿材料中，电荷转移可引发显著的体积收缩；对于$PbTiO_3$基化合物，其异常热膨胀行为则来源于声子非简谐效应和铁电效应间的竞争[9]. 这些材料的负热膨胀特性均与其电子结构密切相关. 此外，在电子与晶格自由度耦合较弱的开放框架结构化合物(如ZrW_2O_8[10]，ScF_3[11]和$Ag_3[Co(CN)_6]$[12])中，低频声子模对NTE行为起着主导作用. 经过近三十年的研究，对材料热膨胀性能的调控及其物理本质的理解已取得显著进展，但目前研究仍主要集中于若干典型体系. 未来仍需进一步探索新型负热膨胀材料，并加深对其微观机理的认识.

2020年，Gao等[13]在研究框架结构负热膨胀材料的结构柔性与负热膨胀关系的基础上，提出了“平均原子晶格体积(Average atomic volume, AAV)”的概念，以指导新型负热膨胀体系的发现与设计. 该工作给出了“平均原子晶格体积”的关键晶体学指标，即0.016 nm^3. 当化合物的AAV大于该阈值时，通常表现出负热膨胀行为；反之则呈现正热膨胀. 借助这一概念，研究团队发现了$ScCo(CN)_6$[14]，$TiCo(CN)_6$[15]和$NaB(CN)_4$等具有立方相的新型负热膨胀材料. 进一步地，通过建立各向异性负热膨胀材料中AAV与热膨胀系数的关系，又识别出$REFe(CN)_6$(RE= La, Nd, Sm, Gd, Ho和Lu)系列NTE化合物[16~18]，以及Ta_2WO_8[19]，Nb_2WO_8[20]和$Ta_2Mo_2O_{11}$[21]等轴向负热膨胀化合物. 近期，受AAV概念的启发，Zhao等用更大的MoO_4或WO_4单元取代了PO_4四面体，开发出了一类新型的NaSICON类型化合物，其通用化学式为$A^IB^{II}C^{III}M_3O_{12}$($A^I$=Li, Na, K, Rb; B^{II}=Mg, Ni, Zn, Mn; C^{III}=Sc, Cr, In; M=Mo, W)[22~24]. 这种方法旨在提高材料的AAV值，从而增强框架的灵活性并增强NTE特性. 这一设计思路不仅实现了对材料热膨胀行为的精准调控，也解决了传统$A_2M_3O_{12}$体系易发生相变的难题. 实际上，针对材料热膨胀及相变行为的调控研究已逐步展开. 例如，Chen等[25]提出“电荷相互作用指数(Charge interaction index, CII)”，用于热膨胀系数的定量精准设计；Yuan等[26]则引入了“平均有效电负性(Averaged effective electronegativity, AEE)”概念，以实现对材料相变温度的预测.

长期以来，对材料热膨胀机制的探究主要分为实验研究与理论计算两大核心方向. 其中，基于量子力学第一性原理的密度泛函理论(Density functional theory, DFT)[27]计算，通过求解材料的电子结构，理论上可精确计算晶格常数随温度(或准简谐近似下的声子谱)的变化，进而推导热膨胀系数，已成为预测材料热力学性质的核心理论工具. 例如，Weck等[28]利用DFT中的密度泛函微扰理论预测了$Zr_2(WO_4)(PO_4)_2$的热力学性质，包括热膨胀系数和Gruneisen参数；Xu等[29]通过DFT计算发现$FeZr_2$合金中高频光学声子模式主导的巨型单轴NTE，揭示了金属键与声子振动的耦合机制；Pan等[30]则结合高压合成与DFT计算，研究了$PbTiO_{3-x}S_x$中S掺杂对晶格畸变和热膨胀性能的影响，验证了阴离子调控的可行性. 然而，DFT计算在材料热膨胀性能的高通量预测与筛选中，存在计算周期长、成本高及处理复杂体系能力有限等显著瓶颈[31].

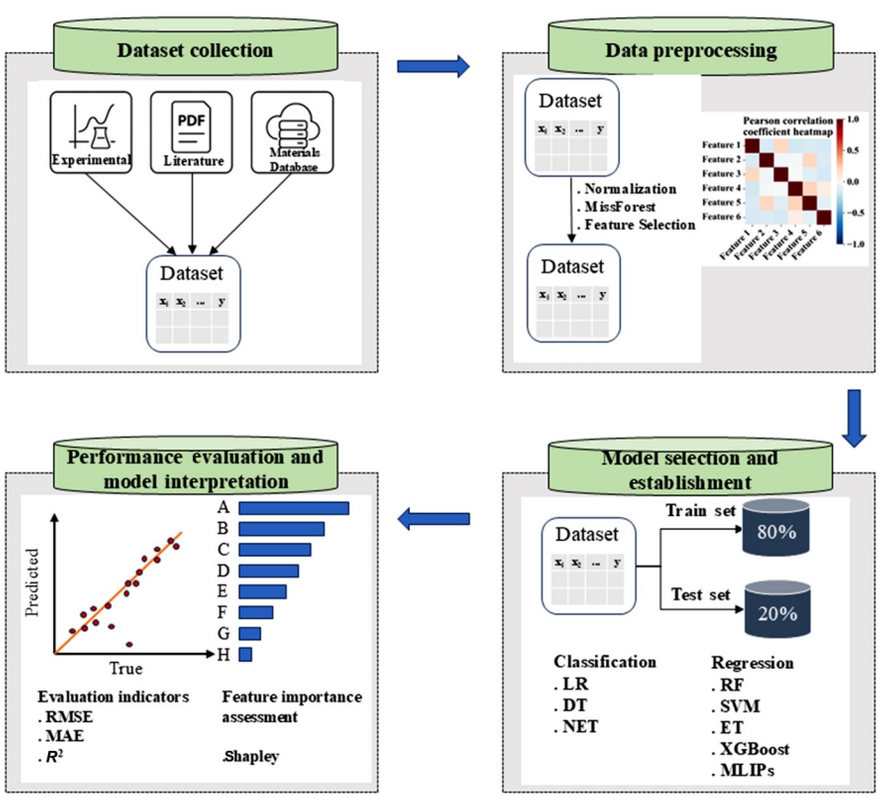
近年来，数据驱动的机器学习(Machine learning，ML)[32]技术的迅猛发展，为突破上述挑战提供了新范式．机器学习方法能从已有的材料数据(包括实验测试数据和高精度计算数据)中，自动学习并发现材料热膨胀性能与其化学组成、晶体结构和原子描述符等特征之间的复杂映射关系，这一核心优势使其在材料热膨胀性能研究中展现出巨大潜力[33]．

机器学习的优势体现在多个方面：(1) 预测速度显著提升，训练良好的模型对新材料热膨胀性能的预测速度比DFT计算快数个数量级(可达ms或s级)，实现近乎实时的预测[34]；(2) 计算成本大幅降低，模型训练虽需一定计算资源，具体取决于数据规模与算法的复杂度(例如，基于小样本的传统回归模型训练仅需数秒至数分钟，而复杂的深度神经网络或原子间势训练可能需数小时至数天)，但预测过程开销极小，仅需秒级即可完成对成千上万个候选结构的筛选，极大减少了对昂贵高性能计算资源的依赖[35]；(3) 擅长挖掘复杂关联，能捕捉高维特征空间中复杂的非线性关系，有助于发现传统物理模型或人为主观分析可能忽略的、影响热膨胀性能的潜在因素和规律[36]．

鉴于此，将机器学习技术与材料科学相结合，为准确、高效、低成本地进行材料热膨胀性能预测，加速新型高性能热膨胀调控材料的研发进程开辟了极具前景的道路．本综合评述将系统总结和评述这一新兴交叉领域的最新研究进展，包括关键方法及代表性成果，分析了当前面临的挑战并探讨了未来的发展方向．

1　机器学习在研究材料热膨胀性能中的一般流程与方法

在材料热膨胀性能的研究中，机器学习方法的应用遵循一定的流程与规范，其核心是通过数据驱动的方式挖掘材料特征与热膨胀性能之间的内在关联，为材料的性能预测与设计提供支持．该流程示意图见Scheme 1，主要包括数据集的收集、数据预处理、模型选择与建立、性能评估与模型解释几个关键环节．



Scheme 1　Schematic diagram of the process of machine learning to study the thermal expansion properties of materials

1.1 数据集的收集

数据集的收集是机器学习应用于材料热膨胀性能研究的基础，其来源主要包括实验测试、文献资料和公共数据库[37]. 实验测试数据通过热膨胀仪和X射线衍射仪等设备获取，包含材料在不同温度下的热膨胀系数、晶体结构变化等直接反映热膨胀特性的信息，具有较高的真实性和可靠性. 从文献资料中提取的材料热膨胀相关数据涵盖材料的成分、制备工艺及性能表现等内容，能够补充实验数据的不足. 公共数据库存储了大量无机材料的晶体结构、物理化学性质等数据，其中包含众多材料的热膨胀性能信息，为研究提供了丰富的基础数据来源.

1.2 数据预处理

原始收集到的材料热膨胀数据往往存在噪声、缺失值及量纲不一致等问题，需要进行预处理以提高数据质量. 归一化[38]是常用的处理手段，通过将不同量级的特征(如晶格常数和键长等)转换到统一区间(如[0,1])，避免因特征量级差异对模型训练产生干扰. 特征选择则是从众多特征中筛选出与热膨胀性能密切相关的特征[39]，如采用皮尔逊相关系数法，通过计算各特征与热膨胀性能(如热膨胀系数)之间的相关系数，保留相关系数绝对值较大的特征，以此减少冗余信息，降低模型复杂度[40]. 对于存在缺失值的数据，可采用均值填充、中位数填充或基于模型预测的方法进行补充. 例如，Tang等[41]提出了一种名为MissForest的数据插补技术，基于随机森林算法对缺失数据进行预测填充；对于噪声数据，可通过滤波、离群点检测等方式进行修正或剔除，以确保数据集的准确性和完整性[42].

1.3 模型选择与建立

经过数据预处理后，便进入模型选择与建立阶段. 针对材料热膨胀研究中的不同任务需求(如材料类型的定性分类或性能参数的定量预测)，需选择合适的算法. 值得注意的是，除部分特定算法外，大多数现代机器学习算法兼具分类与回归功能. 在线性模型方面，逻辑回归(Logistic regression，LogR)[43]常用于二分类任务，能输出材料属于特定热膨胀类型的概率，计算高效且易于解释；而线性回归(Linear regression，LinR)[44]则适用于特征与性能存在明显线性关联的场景，可快速构建热膨胀系数的预测模型.

基于树的集成算法在处理非线性关系时表现出色，且广泛适用于分类和回归任务. 其中，决策树(Decision tree，DT)[45]以树状结构直观呈现规则，决策过程清晰；随机森林(Random forest，RF)[46]通过集成多棵决策树(或回归树)，能捕捉多元掺杂、晶体缺陷等复杂特征间的非线性交互关系，预测精度较高；极端树(Extra trees，ET)作为RF的变体，通过增加随机性进一步增强抗过拟合能力. 梯度提升(Gradient boosting，GB)[47]是一种迭代集成算法. 它通过连续构建决策树，让每一棵新树专门修正前序模型的预测误差，从而逐步提升整体预测精度. 极端梯度提升(eXtreme gradient boosting，XGBoost)[48]是GB的高效优化版本. 它在保留迭代修正误差的基础上，引入了正则化机制以防止过拟合，并优化了计算效率，因此在处理大规模数据集时泛化能力更强. 此外，支持向量机(Support vector machine，SVM)[49]和神经网络(Neural networks，NN)[50]也是通用的强力工具. SVM(包括用于分类的SVC和用于回归的SVR)在小样本、高维数据场景下优势显著，通过核函数有效拟合复杂非线性关系；NN则凭借多层非线性变换能力，在处理高维复杂特征(如多元复合材料的成分与结构组合特征)时表现突出. 同时基于矩张量势(Moment tensor potentials，MTPs)的机器学习原子间势(Machine learning interatomic potentials，MLIPs)[51]也在此领域发挥重要作用. 它通过学习原子的局部环境特征(如原子坐标及配位数等)与相互作用能的关系，构建原子尺度的势能函数，能够从微观层面模拟材料的热膨胀行为，进而预测宏观的热膨胀性能，为揭示热膨胀机理提供了有力工具[52].

图1示出了3种常用于研究材料热膨胀性能的机器学习算法示意图. 在模型建立过程中，需将预处理后的数据集划分为训练集和测试集，利用训练集对模型进行训练，通过调整模型参数(如决策树的深度、SVM的核函数参数及XGBoost的正则化系数等)优化模型性能，最终通过测试集验证模型的泛化能力，确保其在实际材料热膨胀研究中的可靠性.

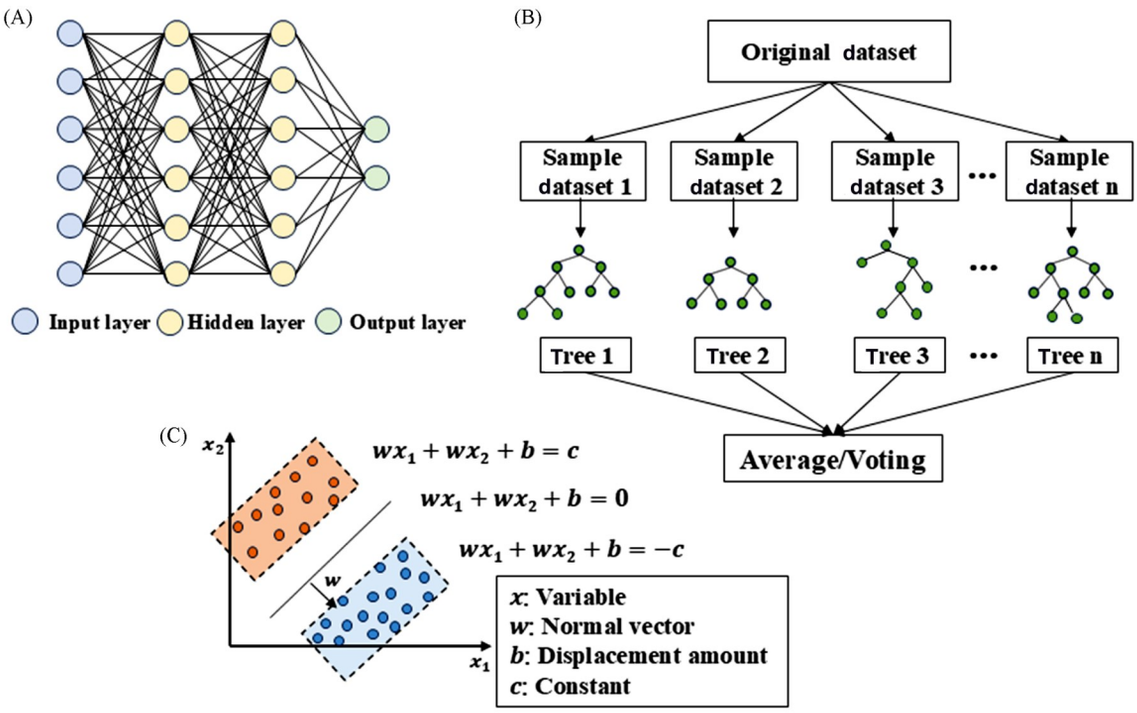


Fig. 1 Common machine learning algorithms of neural network model(A), random forest model(B) and support vector machine model(C)

1.4 性能评估与模型解释

模型建立后，需采用合适的指标对其性能进行评估. 对于回归模型，常用均方根误差(Root mean square error，RMSE)、平均绝对误差(Mean absolute error，MAE)和决定系数(R^2)等指标衡量预测值与真实值的偏差；对于分类模型，准确率、混淆矩阵、ROC曲线和$F1$值等指标可用于评估分类效果[39]. 此外，交叉验证方法[53](如k折交叉验证)能有效避免因数据集划分不当导致的评估偏差，提高评估结果的可靠性.

模型解释旨在阐明模型的预测逻辑，增强研究的可信度. 通过特征重要性分析(如随机森林的特征重要性评分)，可以识别出对材料热膨胀性能影响较大的关键特征(如特定的晶体结构参数)，为理解材料的结构与性能之间的关系提供深刻见解. 部分模型(如决策树)具有天然的可解释性，其决策过程可直观展示；对于黑箱模型(如NN)，可借助SHAP值[54]工具进行解释，揭示输入特征与输出结果之间的关联[55].

2 机器学习在材料热膨胀性能研究中的应用与进展

本章对机器学习在材料热膨胀性能研究中的应用展开论述，主要涵盖两个方向：一类是以传统机器学习算法为核心，通过提取材料的化学组成、晶体结构等宏观特征，构建其与热膨胀系数等性能参数间的映射关系，从而实现性能的快速预测与新材料的高效筛选；另一类则聚焦于机器学习原子间势的构建与应用，通过拟合原子尺度的相互作用数据得到高精度势函数，驱动分子动力学模拟以探究热膨胀行为的微观机制. 二者从不同层面为材料热膨胀性能的研究提供了有力工具，下文将分别阐述其具体应用与研究进展.

2.1 传统机器学习在材料热膨胀性能研究中的应用与进展

传统机器学习通过挖掘材料的宏观特征与热膨胀性能之间的潜在关联，实现了从性能预测到机制解析、再到新材料设计的逐步深入. 其应用演进呈现出清晰的技术脉络，可按研究维度的拓展顺序分为4个阶段，依次为热膨胀系数的单目标预测、引入特征重要性分析的机制解析、结合特征选择的模型优化以及多性能关联的多目标预测. 下文将分别阐述各阶段的技术方法与研究进展.

2.1.1 热膨胀系数的单目标预测 传统机器学习技术在材料热膨胀性能预测领域的应用已取得显著进展，早期研究聚焦于热膨胀系数(CTE)的精准预测，通过构建材料成分、结构与CTE之间的统计关联，实现对材料热膨胀性能的快速评估．如Ayyasamy等[56]利用SVM模型的预测值和标准差估计预期CTE及其不确定性．以$RE_2Si_2O_7$晶体结构(三轴距离、三角距离、体积和晶胞中的原子数)的描述符为输入，以实验测量的CTE数据为输出．对15种$RE_2Si_2O_7$化合物进行训练，将8种独立的$RE_2Si_2O_7$化合物用于测试．随后通过实验验证了$Sm_2Si_2O_7$的结构和CTE预测结果，验证了模型的准确性．Gracheva等[57]提出了两个基于数据驱动的预测模型．这两个机器学习模型均以106个实验样本为训练数据．通过10折交叉验证评估，SMILES-X模型的均方根误差为$(2.65\pm0.09)\times10^{-5}$ K^{-1}，决定系数为0.62±0.03；RF模型相应值分别为$(2.58\pm0.09)\times10^{-5}$ K^{-1}和0.64±0.03，且两种模型对实验室制备的样本的预测结果与实验值一致性良好．

与上述聚焦于CTE精准预测的研究不同，后续研究进一步拓展了机器学习在材料热膨胀领域的应用边界，将其与新材料的设计、发现及实验验证相结合[54]．Zhang等[58]开发了一种用于识别和表征开放框架结构中结构单元连接模式的算法，并构建了该系统热膨胀性能的描述符集，该描述符集由连接方式和元素信息组成．开发的描述符在ML算法的辅助下，可以有效地学习从文献报道的实验数据(246个样本)中收集的数据集中的热膨胀行为．如图2(A)所示，训练后的6个机器学习模型能够准确区分热膨胀行为(PTE/NTE)，效果较好的模型分类准确率可达到92%，图2(B)显示了高精度模型的ROC曲线和混淆矩阵．他们根据材料的连通性特征构建了一个连接因子，如图2(C)和(D)所示，连接因子可以有效区分PTE和NTE材料，总体趋势是因子越大，表示NTE的可能性越高．这种趋势表明，框架材料中结构单元的连通性与其热膨胀行为之间存在密切的相关性．此外，模型还预测了6种新的

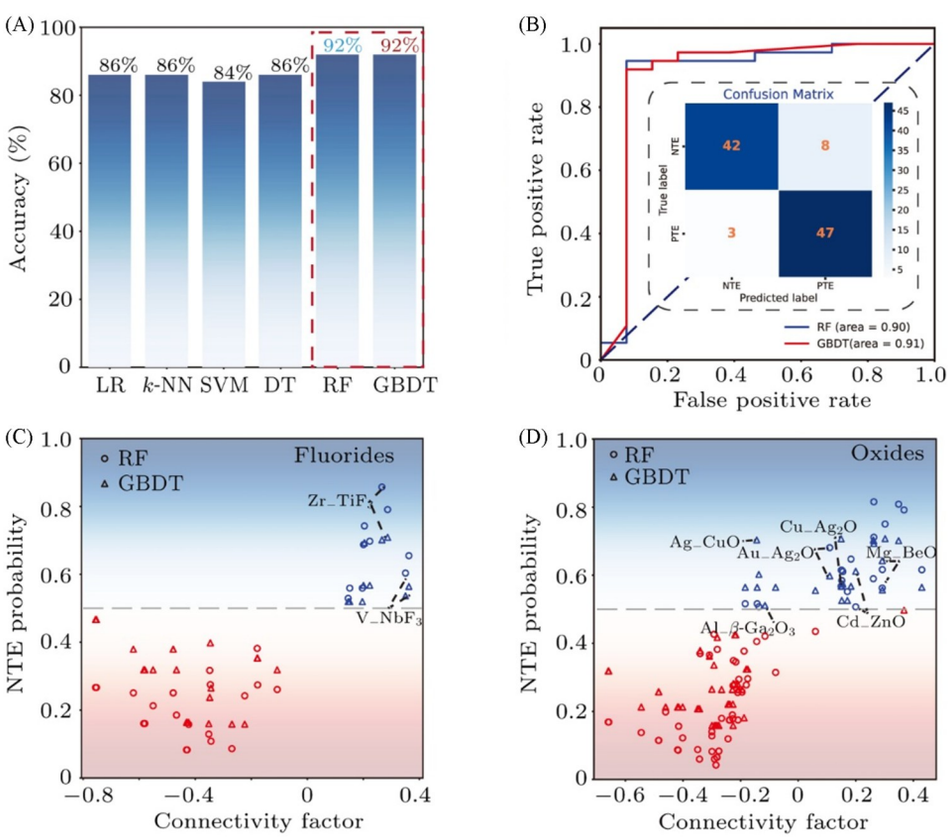


Fig. 2 Accuracy of six trained binary classification models(A), ROC curves and confusion matrices for the high-accuracy RF and GBDT models(B), new NTE candidates predicted for fluorides(C) and oxides(D), and ranked by connectivity factor using the trained RF and GBDT models[58]

Copyright 2023, Chinese Physical Society and IOP Publishing Ltd.

热力学稳定的NTE材料，并通过第一性原理计算进行了理论验证.

为解决镍基高温合金研发中难高效预测CTE、且难兼顾低热膨胀与优异机械性能的问题. Tang等[59]开发了一个SVM回归模型来预测镍基高温合金的CTE. 从通过遗传算法和SVM回归模型生成的候选样品池中，选择了10个样品并生产出来，结果显示其中7个样品表现出优异的CTE性能. 值得注意的是，发现具有最低CTE的合金与现有数据库中的最佳结果相比，显著降低了16.7%. 与传统的低热膨胀高温合金相比，最终选定的合金既表现出低热膨胀性，又表现出优异的机械性能.

2.1.2 引入特征重要性分析的机制解析 在CTE单目标预测的基础上，研究者通过引入特征重要性分析，量化不同输入特征对热膨胀性能的贡献度，从而揭示影响材料热膨胀行为的关键因素. 如Nilsen等[60]采用线性回归和随机森林模型预测混凝土的CTE，通过特征重要性分析发现，粗骨料类型和细骨料类型是影响预测结果的最关键特征，其权重显著高于其它因素，这一结论为混凝土材料的热膨胀性能调控提供了明确方向. Tripathi等[61]构建了用于预测硅酸盐玻璃CTE的RF机器学习回归模型. 玻璃 成分(Al_2O_3、B_2O_3、Na_2O、MgO、CaO、BaO和SiO_2)的摩尔分数被视为预测CTE的模型输入特征. 图3(A)示出了模型预测的CTE与测试集上实际实验测得的CTE. 多数模型预测的CTE值明显聚集在理想拟合线附近，表明模型预测准确性较好. 图3(B)示出了输入特征在预测CTE中的相对重要性. 模型表明，Na_2O的摩尔分数在预测CTE方面比其它成分具有显著更高的特征重要性.

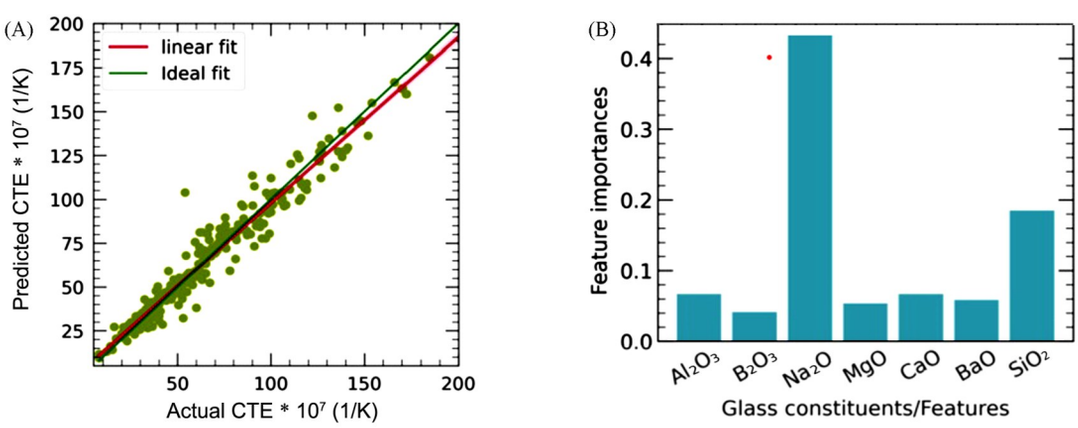


Fig. 3 Relationship between predicted CTE and actual CTE(A) and relative importance of input features in predicting CTE(B)[61]

Copyright 2022, Elsevier.

材料的热膨胀行为由多种特征协同作用决定，不同类型特征对热性能的影响各异，解析其具体作用是理解热膨胀机制的关键. 但传统实验方法难以系统量化各类特征的贡献，因其需精准控制变量以隔离单一特征影响，且存在周期长、成本高及难以覆盖大量特征组合的局限. 针对这一局限性，Yue等[62]将特征按类型分类分析，针对金属有机骨架(MOF)的体积热膨胀系数(α_v)预测，构建包含几何、化学和拓扑特征的输入体系，通过RF模型训练后得到预测结果(见图4)，化学特征和拓扑特征的组合对于准确预测CTE非常重要. 有趣的是，在结合所有3个特征类别(几何、化学和拓扑)的混合描述符上训练的性能最佳的RF模型优于CGCNN模型. 这种分类解析方法为区分不同维度特征对热膨胀性能的影响权重提供了新思路，深化了对材料热膨胀机制的理解. 如图4所示，尽管RF模型在结合几何、化学和拓扑特征后表现优于CGCNN，但两者的绝对预测误差仍然较大(RF的MAE为31.92×10^{-6} K^{-1}，CGCNN的MAE为37.53×10^{-6} K^{-1}). 造成这一现象的原因主要涉及描述符对微观机理的捕捉能力. 首先，针对CGCNN表现欠佳的问题，Yue等[62]分析认为，CGCNN的图卷积算法主要基于原子间的键连接进行编码，具有旋转和平移不变性，这导致它难以区分具有相同配位数但不同空间拓扑排列的结构(例如，平面正方形与四面体构型). 其次，整体预测精度受限也反映了热膨胀性能预测的固有难度. 热膨胀本质上源于晶格的非简谐振动，属于复杂的动力学性质. 而当前的机器学习模型主要依赖静态的几何或化学描述符，难以完全映射出材料在变温过程中的动态结构演化规律. 这也提示未来的研究需进

一步开发能够包含声子信息或动力学特征的物理驱动描述符，以突破现有模型的精度瓶颈.

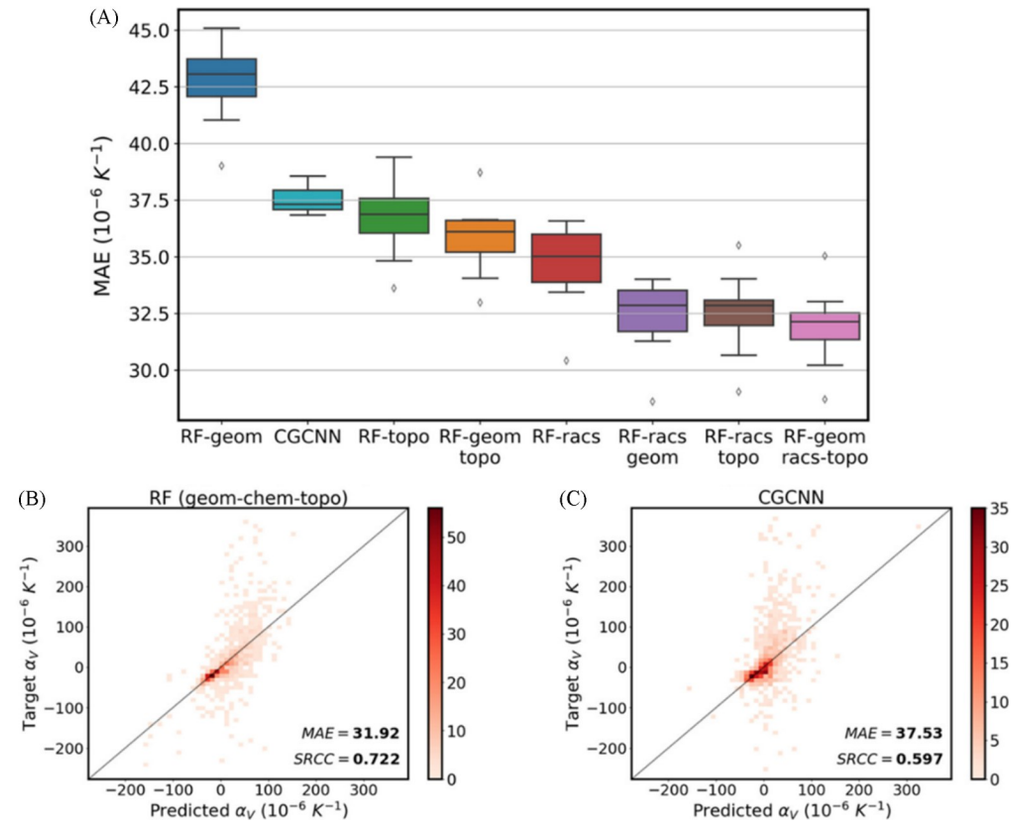


Fig. 4 MAE of models predicting CTE using three different feature combinations(A), parity plot of the best-performing RF model predicting CTE(B) and parity plot of the CGCNN model predicting CTE(C)[62]

Copyright 2024, American Chemical Society.

值得注意的是，特征重要性分析的最新进展已从筛选化学成分等宏观特征，发展到识别具有明确物理意义的微观结构描述符，从而更直接地揭示热膨胀的物理机制. 如，Tian等[63]针对二维晶体线性热膨胀系数的分类与预测问题，创新性地采用特征筛选与物理可解释解析策略，构建涵盖力学核心特征(面内刚度E_2D、面外弯曲刚度D)的输入体系，通过机器学习模型训练与验证得到相关结果. 其中，E_2D与D这两个力学特征的组合及衍生描述符(S_1=1/D，S_2=1/E_2D^2)对于准确分类二维晶体PTE和NTE、精准预测二维晶体线性热膨胀系数至关重要. 更重要的是，研究者通过Grüneisen理论进一步阐释了其物理内涵：面外弯曲刚度与主导负热膨胀的外平面声学声子模式的贡献呈负相关，而面内刚度则反映了化学键本征膨胀导致的正面内应变. 该工作不仅通过可解释的机器学习模型实现了高性能分类预测，更关键的是将模型的输出与材料的基本变形模式(面内拉伸与面外弯曲)直接关联，深刻揭示了二维材料热膨胀行为源于两种模式相互竞争的内在物理机制，为理解与设计热膨胀性能提供了理论依据.

2.1.3 结合特征选择的模型优化 在特征重要性分析的基础上，研究者进一步引入特征选择方法. 通过结合特征选择算法的应用，有效精简了模型的输入维度. 这不仅显著提升了预测效率，更关键的是，通过剔除冗余信息，使得决定负热膨胀性能的核心物理化学机制得以凸显. 如McGuinness等[64]针对$A_xA'_{1-x}B_yB'_{1-y}O_3$型钙钛矿氧化物的热膨胀系数预测，构建了机器学习模型. 研究以A位和B位阳离子的物理化学属性(如电负性、晶体半径和体积模量等)为初始输入特征，共包含902个特征. 如图5所示，采用GA和iPLS的特征选择方法，将初始902个特征减少到259个，显著提高了PLS和SVM模型的性能. SVM回归模型经优化后交叉验证均方根误差低至1.54×10^{-6} K^{-1}，证实了大规模特征筛选在复杂氧化物体系中的有效性. Yang等[65]聚焦二元晶体的CTE预测，采用皮尔逊相关系数法筛选出9个与晶体结构密切相关的关键特征，结合优化的反向传播神经网络，明确平均键长贡献率达27%，且次要特

征对精度提升不可或缺.

鉴于CTE是高熵陶瓷材料设计、性能调控及工程应用中的关键热物理参数，且多维度特征数据集下难以精准筛选主导影响因素，Kim等[66]基于3360个高熵陶瓷数据集(涵盖热力学、弹性及热物理性质)，构建了NET和RF模型以预测CTE. 研究通过计算特征之间的皮尔逊相关系数，排除特征间相关系数高于±0.9的特征，最终确定16个特征作为表征因子. 通过逐个删除输入特征并计算模型性能的变化，确定了影响CTE的主导因素. 在NN与RF模型中，体积模量、杨氏模量和内聚能均表现出较高的特征重要性. 使用这3种最重要的特征进行三维比较，如图6所示，发现测试集和训练集的趋势相似，利用最佳模型和排名前三的特征因子，CTE值的预测准确率超过97%，凸显了特征选择在简化模型同时保留核心物理信息的优势.

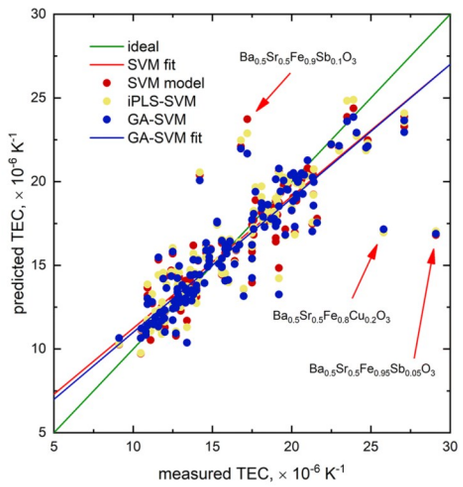


Fig. 5 Comparison of SVM models generated using different feature selection methods[64]

Copyright 2023, the Royal Society of Chemistry.

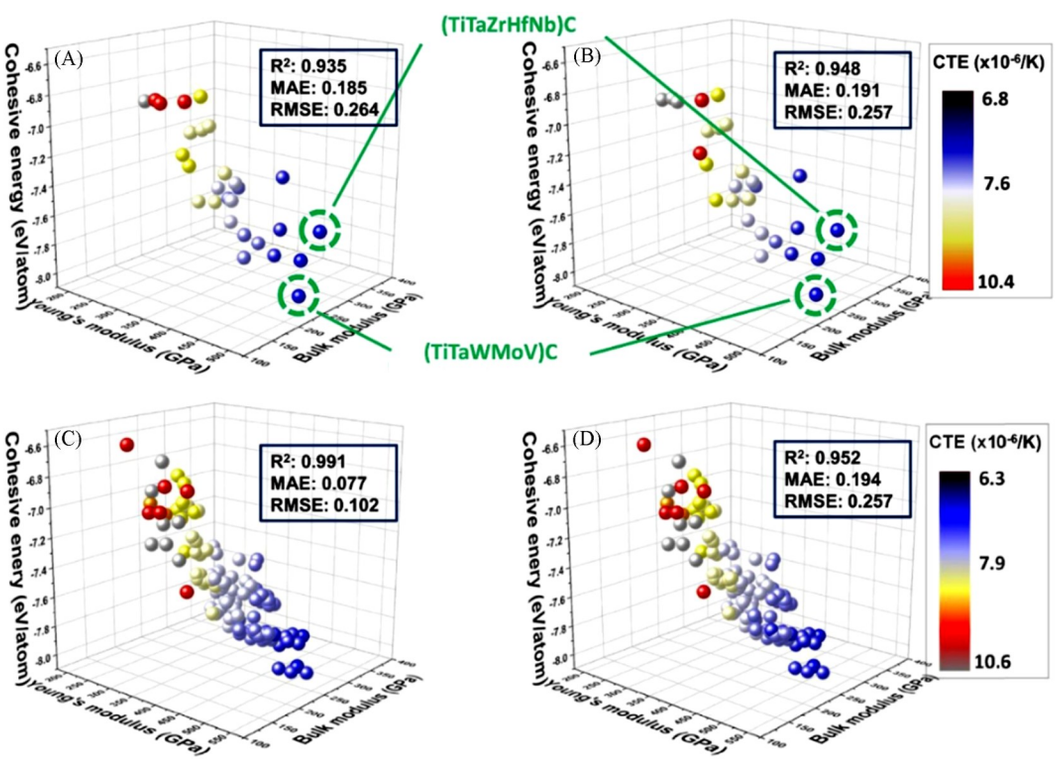


Fig. 6 3D plots of the RF test set(A), the NN test set(B), the RF training set(C), and the NN training set generated using three parameters(bulk modulus, Young's modulus, and cohesive energy)(D)[66]

The CTE magnitude in the 3D plots is indicated by the color scale shown on the right.

Copyright 2024, Elsevier.

为解决钙钛矿阳极材料在实际应用中与氧化钇稳定氧化锆的CTE适配性难题，同时突破传统新材料筛选效率低、特征维度冗余难以精准把控的局限，Wang等[67]针对钙钛矿阳极材料开展研究，为了保留对CTE有显著影响的特征，从原子半径、电负性和晶体结构等特征中通过皮尔逊相关系数筛选出38个关键特征，图7示出了特征选择后的皮尔逊相关系数热图. 红色越深，偏相关性越强. 采用6种机器学习算法进行训练，结果表明，最终ET回归模型为最优模型，其测试集R^2为0.84，RMSE为

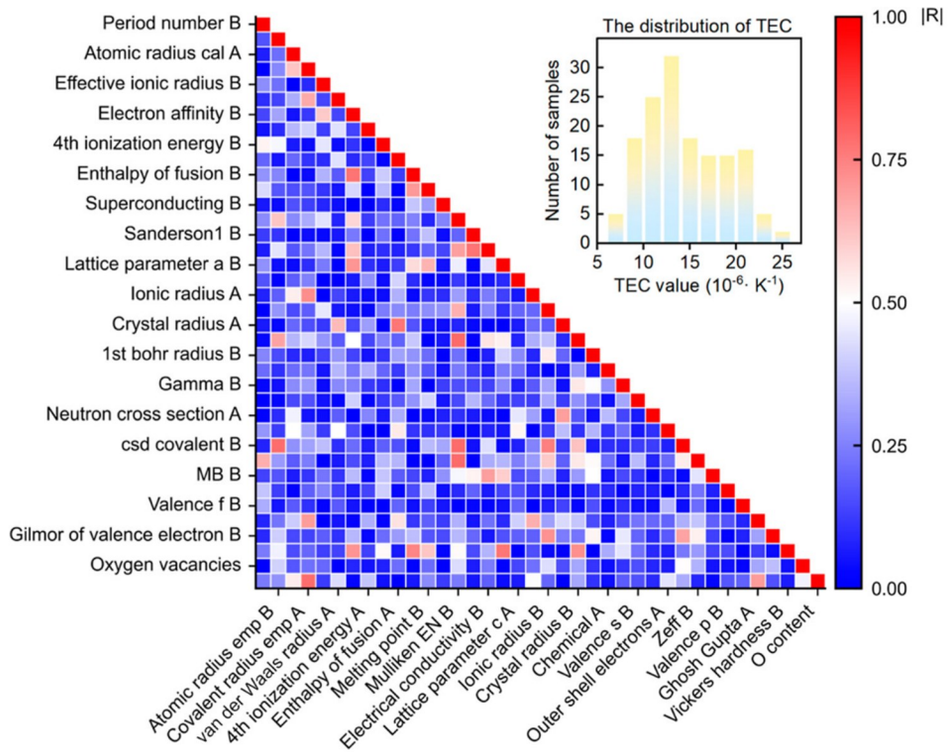


Fig. 7 CTE distribution in the dataset and Pearson correlation heatmap of the selected 38 feature variables[67]

Copyright 2025, American Chemical Society.

1.54×10^{-6} K^{-1}, MAE 为 1.06×10^{-6} K^{-1}. 最终, 基于该模型从 13095 种钙钛矿氧化物中筛选出 507 种与氧化钇稳定的氧化锆相容的候选材料, 体现了特征选择在指导新材料开发中的实用价值.

2.1.4 多性能关联的多目标预测 在材料热膨胀性能的研究中, 仅预测 CTE 的单目标预测虽为理解材料热膨胀特性提供了基础, 但难以满足实际应用对材料综合性能的需求. 材料的热膨胀行为往往与 NTE 对应的温度范围和力学性能等其它关键特性存在内在关联, 且实际应用中需同时考量多种性能指标. 针对这一问题, 多目标机器学习策略的引入主要旨在解决两类问题: 一是通过多目标优化平衡材料相互制约的多项性能指标, 以实现综合性能的最佳匹配; 二是利用不同性能参数间的内在关联, 通过多任务学习等手段辅助提升模型对主要目标的预测精度. 在平衡材料多项性能指标方面, 研究者致力于在维持优异热膨胀性能的同时, 兼顾材料的力学稳定性或工作温区. 镍基高温合金作为航空航天发动机、燃气轮机等极端高温环境下的核心结构材料, 其性能直接决定了装备的使用寿命与安全可靠性[68]. 这类材料需同时满足低热膨胀系数(以减少热应力损伤)和高蠕变寿命(以抵抗持续载荷下的变形失效)的严苛要求, 二者的协同优化是提升材料高温使用性能的关键. Zhu 等[69]针对镍基高温合金的综合性能优化, 构建了兼顾蠕变寿命与 CTE 的多目标预测与优化框架. 通过互信息算法和三步特征选择确定关键特征后, 采用 GB 回归模型分别构建蠕变寿命和 CTE 预测模型, 如图 8 所示, 测试集 R^2 分别为 0.88 和 0.85, 并结合遗传算法实现多目标优化. 该工作设计出蠕变寿命提升 12.13% 且 CTE 降低 6.92% 的新型合金, 实现了力学与热学性能的协同提升.

为解决实验发现 NTE 材料难度大、同时难以大规模筛选潜在 NTE 材料并同步预测其热膨胀系数与适用温度范围的问题, Cai 等[70]针对 NTE 材料的筛选, 采用多步机器学习策略, 不仅预测材料是否具备 NTE 特性及其 CTE 数值, 还同步预测了 NTE 发生的温度范围. 该工作从数据库中筛选出 57 种具有 100% NTE 概率且温区适宜的候选材料[如 $Tb_2(WO_4)_3$ 等], 并通过第一性原理计算进行了验证, 为寻找高性能 NTE 材料提供了从定性筛选到定量综合评估的完整方案. Kireeva 等[71]聚焦热膨胀系数与相变温度的关联预测, 筛选了 17 个描述符, 采用 XGBoost 回归算法实现双目标预测. 十折交叉验证后结果如图 9(A)和(B)所示, 二者的预测 R^2 均达 0.81(CTE 的 RMSE 为 1.170, 相变温度的 RMSE 为 82.289).

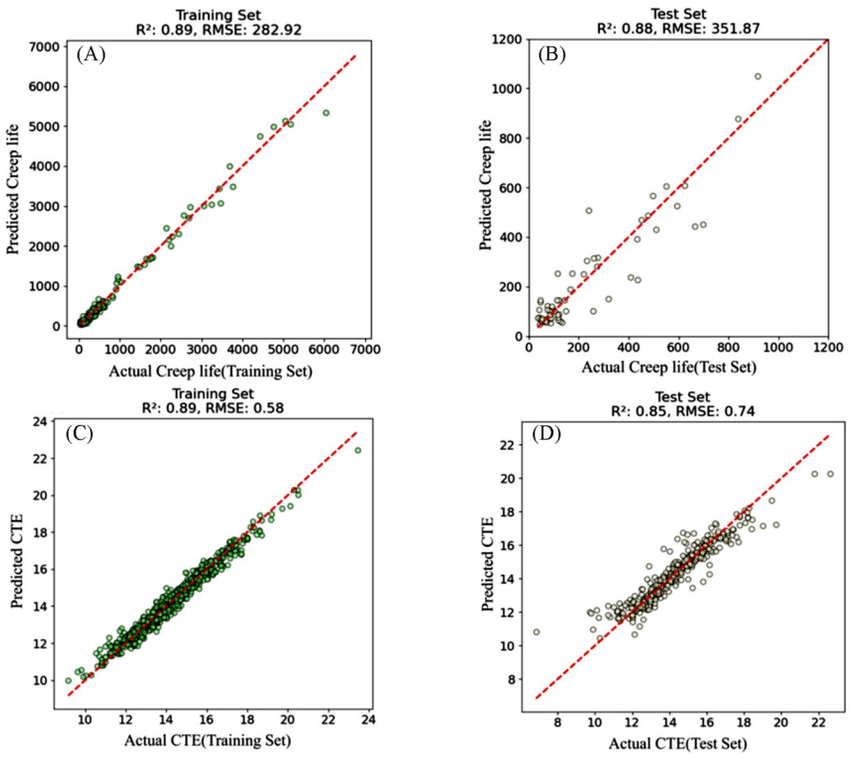


Fig. 8 Relationships between predicted values and experimental calculated values of creep life training set(A), creep life test set(B), CTE training set(C) and CTE test set(D)[69]

Copyright 2025, Springer Nature.

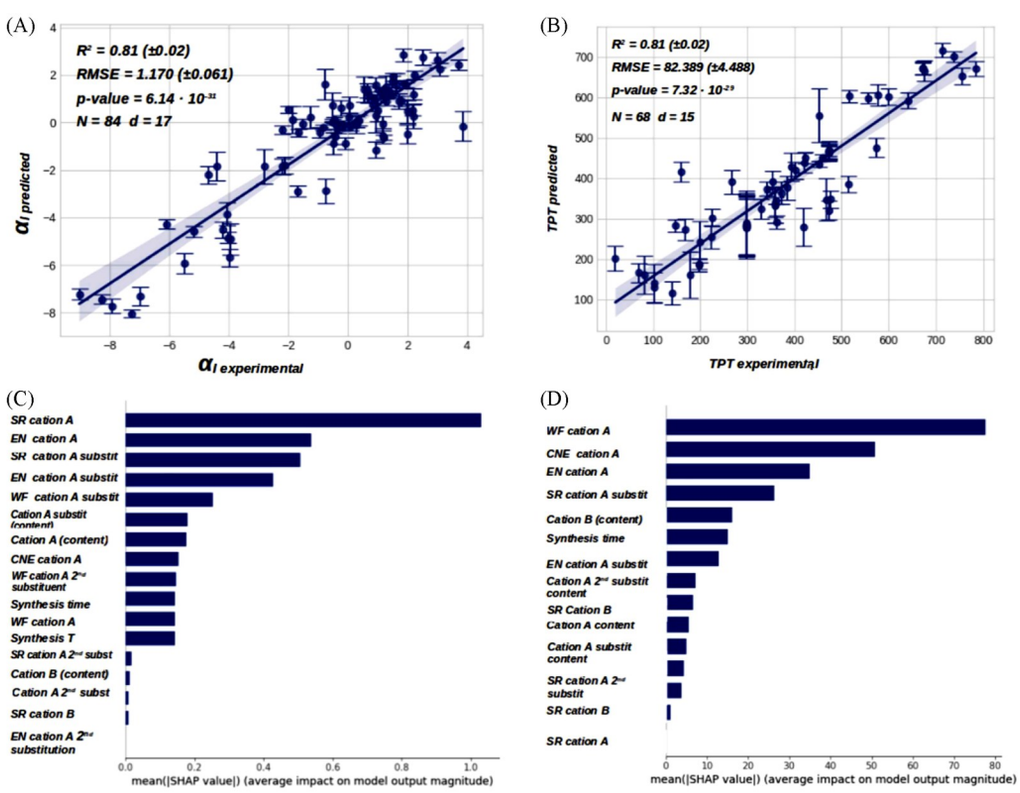


Fig. 9 Relationships between experimental and predicted CTE values(A) and phase transition temperature(B), and SHAP value explanation of the importance of features in predicting CTE(C) and phase transition temperature(D)[71]

Copyright 2024, Elsevier.

图9(C)和(D)通过SHAP值解释了预测CTE和相变温度的特征重要性. 结果表明, A阳离子的Shannon离子半径和电负性等对CTE预测至关重要, 而其功函数和有效核电荷等则主导相变温度预测, 为在保证结构稳定性的前提下调控热膨胀性能提供了理论指导.

在改善模型预测能力方面, 多目标或多任务学习被用于挖掘数据间的潜在联系, 以解决小样本或高维特征下的预测难题. Zeng等[72]提出了一种结合数据增强预处理的多任务学习算法, 用于处理负热膨胀材料中小规模不平衡数据和多目标预测. 该研究将NTE的起始温度、终止温度与CTE作为关联任务同时进行训练, 利用任务间的共享信息显著提升了模型的泛化能力, 其对CTE的预测精度(R^2达到0.898)优于传统的支持向量回归和随机森林模型. 此外, Peng等[73]通过引入阳离子多面体描述符, 建立了晶格参数随温度和成分变化的关联模型. 这种方法本质上利用了晶格常数随温度演变的内在物理规律, 有效解决了复杂氧化物在宽温域下热膨胀行为预测难的问题, 体现了利用物理关联增强模型预测可靠性的思路.

多目标预测的核心在于通过机器学习模型同时解析材料的多种关联性能, 不仅揭示了性能间的耦合机制, 更为材料的高效筛选与定向设计提供了全面的理论支撑, 推动了从单一性能评估向综合性能优化的跨越[74].

2.2 机器学习原子间势在材料热膨胀性能研究中的应用与进展

机器学习原子间势驱动的原子模拟与性能机理研究侧重于利用MLIPs替代传统经验势或第一性原理计算, 驱动高精度、大规模的分子动力学(Molecular dynamics, MD)模拟. 其核心在于在原子尺度上揭示NTE产生的动态结构演化、局域原子行为及内在物理机制, 为理解现象本质提供微观基础.

传统研究二维材料热膨胀的方法存在局限: 从头算分子动力学(AIMD)精度高但计算成本极高, MD依赖经验势函数, 对新型材料精度不足. 针对这一问题, Mortazavi等[75]开发了基于MTPs的MLIPs, 通过“被动训练”策略拟合低成本AIMD轨迹(含能量、力学及结构等数据), 构建了兼顾精度与效率的势函数. 输入为AIMD模拟生成的碳基纳米片的原子构型、能量及力学数据. 将训练好的MTPs进行经典MD模拟, 在50~1000 K温度范围内, 通过监测投影面积(A)随温度(T)的变化[式(1)]计算CTE(∂).

$$\partial = \frac{1}{A} \cdot \frac{\mathrm{d}A}{\mathrm{d}T} \tag{1}$$

该方法计算效率比AIMD高1~2个数量级, 为大规模研究新型2D材料的热膨胀提供了可行工具.

联苯单层作为一种由四元、六元和八元碳环排列构成的新型2D碳同素异形体, 其热膨胀特性的研究对理解非苯环碳材料的独特物理行为具有重要意义[76]. Mortazavi等[77]采用MTPs训练策略, 通过MD模拟计算50~1500 K的CTE, 实验结果证实了联苯单层具有高负热膨胀特性. 图10示出了不同温

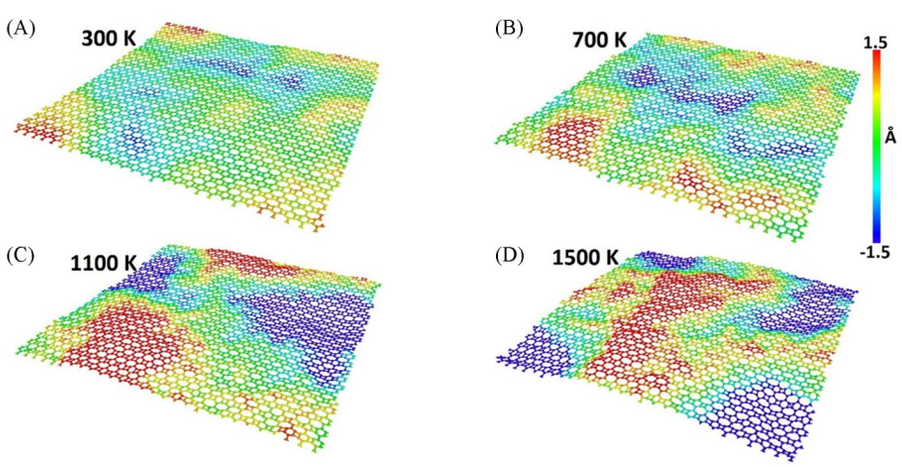


Fig. 10 Atomic configurations of a biphenyl monolayer at different temperatures[77]

The color coding indicates the displacement of atoms relative to the center of mass of the system along the out-of-plane direction.

Copyright 2022, Elsevier.

度下结构平衡后的原子构型. 可以看出, 随着温度升高, 局部褶皱逐渐扩展至整个结构, 使联苯单层沿平面方向发生收缩; 这一现象持续至极高温度, 正是该碳网络具有显著高负热膨胀系数的原因. 此外, MLIPs为联苯单层热膨胀特性的大规模模拟提供了高效工具, 既保证了模拟精度, 又显著提升了计算效率.

石墨炔作为一类含*sp*-*sp*^{2}杂化碳的多孔二维材料, 其独特的拓扑结构赋予了其优异的热学与力学性能. 针对实验合成的γ-石墨炔及5种衍生物, Mortazavi等[78]基于MTPs的MD模拟, 计算了50~1000 K范围内的CTE, 这6种石墨炔的CTE随温度的变化趋势见图11. 结果表明, α-、β-、γ-、λ-、CEY-和δ-石墨烯单层的室温热膨胀系数分别估计为−6.7, −4.1, −2.0, −2.4, −3.6和−3.9×10^{-5} K^{-1}, 比石墨烯的CTE要高1个数量级. 石墨炔的超高负热膨胀源于多孔结构赋予的强面外柔性, 即温度升高时, 多孔区域更易形成密集褶皱, 导致投影面积大幅收缩.

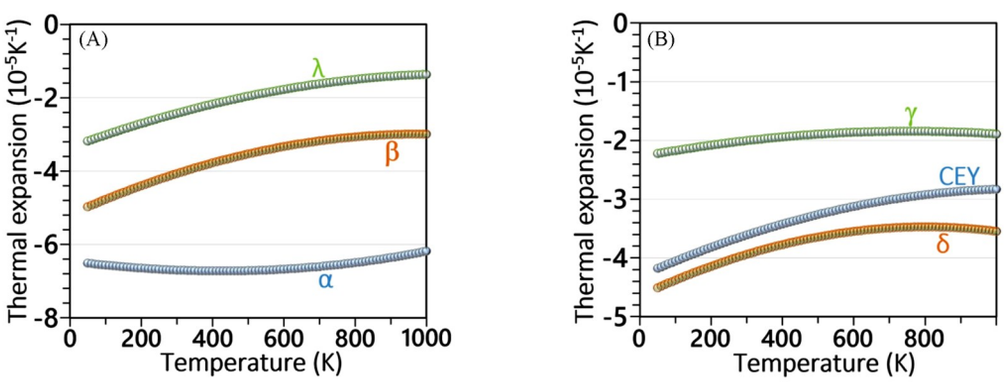


Fig. 11 Changes in CTE of six types of graphyne with temperature[78]

Copyright 2022, Elsevier.

基底作为二维材料实际应用中的关键环境因素, 其对热膨胀特性的影响尚未被系统阐明. Rajabpour等[79]研究了石墨烯、幻石墨烯、C_3N和BC_3单层这4种常见二维结构的悬浮和支撑的热膨胀系数. 该研究进行了MD模拟, 其中原子相互作用由高精度机器学习原子间势定义. 所得结果表明, 通过增加单层和基底之间的范德华相互作用强度(从2 meV增加到8 meV), 石墨烯和幻石墨烯的CTE从负值增加到正值; 而C_3N和BC_3结构的负值仍然保持不变. 分子动力学轨迹分析表明, 基底可以显著减少平面外褶皱的形成, 从而影响CTE值. 研究结果为理解基底对二维材料复杂热膨胀响应的影响提供了参考.

为了在宏观和微观尺度上协同研究材料的热膨胀性能, 传统机器学习和机器学习原子间势可以结合使用. Gao等[80]提出了一种结合ML与MLIPs的两步筛选策略, 用于高效预测新型NTE和零热膨胀框架材料. 第一步构建包含354种材料(194种氧化物、102种氰化物和58种氟化物)的训练集, 通过matminer工具提取元素电负性、孔隙率和键长/键角等特征, 采用递归特征消除优化. 选用LinR、SVM和RF等机器学习模型来完成分类和回归任务, 其中RF模型表现最佳(轴向CTE预测R^2最高达0.825, RMSE<15×10^{-6} K^{-1}), 分类精度达87%以上. 第二步基于MLIPs(MatterSim模型)计算弹性常数, 排除7658种力学不稳定材料, 再结合Phonopy计算声子谱, 排除121种声子软化材料, 最终保留535种稳定材料. Gao等通过特征工程与MLIPs的结合, 从54067种框架材料中筛选出181种候选NTE材料, 并揭示了其热膨胀行为的结构-力学起源. NTE材料需满足孔隙率大于20%, 高孔隙率增强横向原子振动(主导NTE), 而纵向振动导致PTE. 轴向CTE绝对值与杨氏模量呈反比关系, 低模量方向更易发生形变, 导致显著热膨胀.

为了更直观地梳理机器学习在负热膨胀领域的研究现状, 基于本文所引用的研究工作, 分别从文献中实际采用的描述符类型(表1)和涉及的材料体系(表2)两个维度进行了归纳总结.

Table 1 Classification of commonly used descriptors for the thermal expansion properties of materials in machine learning research

Descriptor type	Examples of specific features	Research object	Ref.
Chemical/Compositional Features	Elemental molar fraction, atomic/ionic radius, electronegativity, effective nuclear charge, work function, atomic mass, etc.	Glass composition;	[61]
		Perovskite electronegativity/radius;	[64]
		A-site cation properties	[71]
Geometric/Structural Features	Lattice constants, bond lengths, bond angles, unit cell volume, number of cation polyhedra, porosity, coordination number, etc.	Unit cell parameters;	[56]
		Polyhedral features;	[73]
		Porosity	[70]
Mechanical/Physical Features	Bulk modulus, Young's modulus, cohesive energy, in-plane stiffness, out-of-plane bending stiffness, Debye temperature, etc.	High-entropy ceramic moduli/energy;	[66]
		2D material stiffness;	[63]
		Elastic constants	[80]
Topological/Graph Features	Connectivity factor, persistent homology features, SMILES codes, structural unit connection modes	Connection modes;	[58]
		MOF topological features;	[62]
		SMILES	[57]

Table 2 Main material systems where machine learning has been applied to negative thermal expansion research

Material systems	Typical representatives	Research focus
Framework oxides	$RE_2Si_2O_7$, ZrW_2O_8, $PbTiO_3$, cubic oxides, perovskites(ABO_3)	Single-objective CTE prediction; correlation of polyhedral features; feature selection optimization.
Fluorides	ScF_3, REO_3-type fluorides	Structural flexibility and NTE classification prediction; screening for new materials.
Metals/Alloys	Ni-based superalloys, high-entropy alloys(HEAs), $FeZr_2$	Balancing mechanical properties(creep) and thermal expansion; prediction of phase transition temperatures.
2D Materials	Graphene, Graphyne, biphenylene monolayer, C_3N	Simulating atomic rippling mechanisms using MLIPs; analysis of substrate effects.
Porous Framework Materials	HKUST-1, UiO-66, prussian blue analogues(PBAs)	Analysis of complex topological structures; effect of porosity on thermal expansion.
Amorphous/Others	Silicate glasses, concrete, high-entropy ceramics	Composition-property regression prediction; analysis of aggregate influence.

3 总结与展望

本文综合评述了机器学习在负热膨胀研究中的应用进展. 从早期的热膨胀系数单目标预测, 到引入特征重要性分析解析机制, 再到结合特征选择的模型优化及多性能协同预测, 机器学习已逐步成为该领域不可或缺的研究工具. 特别是MLIPs的发展, 为原子尺度下揭示热膨胀的动力学微观机制提供了高效手段. 尽管数据驱动方法已在加速新材料发现方面展现出巨大潜力, 但当前研究仍面临数据稀缺及通用描述符物理可解释性弱等瓶颈. 单纯依赖晶体结构参数或化学计量比作为输入, 往往需要庞大的训练集才能拟合出复杂的构效关系, 这在高质量热膨胀数据匮乏的现状下极具挑战. 鉴于此, 我们认为未来的研究不应仅局限于算法层面的优化, 而应向“领域知识与数据驱动深度融合”的新范式跨越. 基于本课题组前期的研究实践, 我们认为未来应重点关注以下方向: (1) 构建基于物理判据的领域专用描述符, 突破“小数据”预测瓶颈; (2) 将深厚的领域理论转化为机器学习可识别的高维特征, 是提升模型预测精度与泛化能力的关键. 例如, 本课题组前期提出的“AAV”揭示了框架结构柔性与NTE行为的几何阈值规律, “CII”量化了化学键强对热膨胀系数的贡献, AEE”则与材料的相变温度存在内禀关联.

将上述具有明确物理意义的参数转化为机器学习的物理驱动描述符, 将产生显著的协同效应: (1) 物理规则辅助的高效筛选: 将AAV等物理判据作为机器学习分类任务的前置过滤器或关键特征,

可快速从海量晶体结构数据库中剔除不具备NTE潜力的结构，大幅收敛搜索空间，降低后续高精度计算的成本．(2) 小样本下的高鲁棒性预测：CII和AEE等参数高度浓缩了材料的化学键合与电子结构信息．将其引入回归模型，能够帮助算法在样本量有限的情况下，更敏锐地捕捉热膨胀系数及相变温度的变化规律，显著提升模型在未知体系中的泛化能力．(3) 模型与理论的互馈迭代：通过机器学习的可解释性工具(如SHAP分析)量化AAV、CII等物理描述符在预测中的权重，不仅能提升模型的透明度，还能反过来验证和修正物理模型，形成“物理理论指导AI建模，AI挖掘反哺物理理论”的良性闭环．

综上所述，将领域知识融入机器学习全流程，构建包含AAV、CII及AEE等物理描述符的专用数据库与预测模型，将是从根本上加速新型负热膨胀材料设计与开发的核心路径．

参 考 文 献

[1] Liang E., Sun Q., Yuan H., Wang J., Zeng G., Gao Q., *Front. Phys.*, **2021**, *16*(5), 53302
[2] Takenaka K., *Mater. Trans.*, **2024**, *65*(3), 243—252
[3] Dubey D., Mirhakimi A. S., Elbestawi M. A., *J. Manuf. Mater. Process.*, **2024**, *8*(1), 40
[4] Crisan V., Entel P., Ebert H., Akai H., Johnson D. D., Staunton J., *Phys. Rev. B*, **2002**, *66*(1), 014416
[5] Takenaka K., Takagi H., *Appl. Phys. Lett.*, **2005**, *87*(26), 261902
[6] Huang R., Liu Y., Fan W., Tan J., Xiao F., Qian L., Li L., *J. Am. Chem. Soc.*, **2013**, *135*(31), 11469—11472
[7] Azuma M., Chen W. T., Seki H., Czapski M., Olga S., Oka K., Mizumaki M., Watanuki T., Ishimatsu N., Kawamura N., *Nat. Commun.*, **2011**, *2*(1), 347
[8] Long Y., Hayashi N., Saito T., Azuma M., Muranaka S., Shimakawa Y., *Nature*, **2009**, *458*(7234), 60—63
[9] Pan Z., Fang Y. W., Nishikubo T., Hu L., Kawaguchi S., Azuma M., *Chem. Mater.*, **2022**, *34*(6), 2798—2803
[10] Mary T. A., Evans J., Vogt T., Sleight A., *Science*, **1996**, *272*(5258), 90—92
[11] Greve B. K., Martin K. L., Lee P. L., Chupas P. J., Chapman K. W., Wilkinson A. P., *J. Am. Chem. Soc.*, **2010**, *132*(44), 15496—15498
[12] Goodwin A. L., Calleja M., Conterio M. J., Dove M. T., Evans J. S., Keen D. A., Peters L., Tucker M. G., *Science*, **2008**, *319*(5864), 794—797
[13] Gao Q., Wang J., Sanson A., Sun Q., Liang E., Xing X., Chen J., *J. Am. Chem. Soc.*, **2020**, *142*(15), 6935—6939
[14] Gao Q., Sun Y., Shi N., Milazzo R., Pollastri S., Olivi L., Huang Q., Liu H., Sanson A., Sun Q., *Scr. Mater.*, **2020**, *187*, 119—124
[15] Gao Q., Shi X., Venier A., Carnera A., Huang Q., Wu H., Chen J., Sanson A., Liang E., *Inorg. Chem.*, **2020**, *59*(20), 14852—14855
[16] Gao Q., Sun Q., Venier A., Sanson A., Huang Q., Jia Y., Liang E., Chen J., *Sci. China Mater.*, **2022**, *65*(2), 553—557
[17] Gao Q., Jiao Y., Sanson A., Liang E., Sun Q., *Chin. Chem. Lett.*, **2023**, *34*(5), 107564
[18] Jiao Y., Gao Q., Sanson A., Liang E., Sun Q., Chen J., *Inorg. Chem.*, **2022**, *61*(20), 7813—7819
[19] Luo Y., Qiao Y., Gao Q., Wang J., Guo J., Ren X., Chao M., Sun Q., Jia Y., Liang E., *Inorg. Chem.*, **2021**, *60*(23), 17758—17764
[20] Gao Q., Luo Y., Qiao Y., Guo J., Liang E., *Ceram. Int.*, **2022**, *48*(23), 34707—34712
[21] Gao Y., Wang C., Gao Q., Guo J., Chao M., Jia Y., Liang E., *Inorg. Chem.*, **2020**, *59*(24), 18427—18431
[22] Hu T., Qiao Y., Li M., Fan W., Guo J., Xu M., Gao Q., *Appl. Phys. Lett.*, **2025**, *126*(20), 202201
[23] Wang Q., Qiao Y., Zhao K., Zhang P., Zhao H., Guo J., Shi X., Liang E., Gao Q., *Acta Mater.*, **2024**, *281*, 120358
[24] Zhao H., Qiao Y., Zhao K., Wang Q., Zhang P., Guo J., Chao M., Liang E., Gao Q., *Small*, **2024**, *20*(42), 2403000
[25] Chen X., Gao Q., Zhao K., Qiao Y., Sanson A., Sun Q., Guo J., Kawaguchi S., Liang E., Chen J., *Chem. Sci.*, **2025**, *16*(35), 16331—16338
[26] Yuan H., Wang C., Gao Q., Zeng G., Guo J., Chao M., Kawaguchi S., Jia Y., Liang E., *Mater. Horiz.*, **2021**, *8*(9), 2562—2568
[27] Vasilyev D., Gorev V., *Physica B*, **2025**, *699*, 416850
[28] Weck P. F., Kim E., Gordon M. E., Greathouse J. A., Dingreville R., Bryan C. R., *ACS Omega*, **2018**, *3*(11), 15780—15788
[29] Xu M., Li Q., Song Y., Xu Y., Sanson A., Shi N., Wang N., Sun Q., Wang C., Chen X., *Nat. Commun.*, **2023**, *14*(1), 4439
[30] Pan Z., Liang Z., Wang X., Fang Y. W., Ye X., Liu Z., Nishikubo T., Sakai Y., Shen X., Liu Q., *Mater. Horiz.*, **2024**, *11*(21), 5394—5401
[31] Schleder G. R., Padilha A. C., Acosta C. M., Costa M., Fazzio A., *J. Phys. Mater.*, **2019**, *2*(3), 032001
[32] Pollice R., dos Passos Gomes G., Aldeghi M., Hickman R. J., Krenn M., Lavigne C., Lindner-D'Addario M., Nigam A., Ser C. T., Yao Z., *Acc. Chem. Res.*, **2021**, *54*(4), 849—860
[33] Gao C., Min X., Fang M., Tao T., Zheng X., Liu Y., Wu X., Huang Z., *Adv. Funct. Mater.*, **2022**, *32*(1), 2108044
[34] Schmidt J., Shi J., Borlido P., Chen L., Botti S., Marques M. A., *Chem. Mater.*, **2017**, *29*(12), 5090—5103
[35] Tawfik S. A., Isayev O., Stampfl C., Shapter J., Winkler D. A., Ford M. J., *Adv. Theory Simul.*, **2019**, *2*(1), 1800128

[36] Banik S., Balasubramanian K., Manna S., Derrible S., Sankaranarayananan S. K., *Comput. Mater. Sci.*, **2024**, *236*, 112847
[37] Zhu L., Zhou J., Sun Z., *J. Phys. Chem. Lett.*, **2022**, *13*(18), 3965—3977
[38] Singh D., Singh B., *Pattern Recognit.*, **2022**, *122*, 108307
[39] Wang J., Xu P., Ji X., Li M., Lu W., *Materials*, **2023**, *16*(8), 3134
[40] Han G., Sun Y., Feng Y., Lin G., Lu N., *ES Mater. Manuf.*, **2021**, *14*(18), 20—35
[41] Tang F., Ishwaran H., *Stat. Anal. Data Min.*, **2017**, *10*(6), 363—377
[42] Qayyum F., Kim D. H., Bong S. J., Chi S. Y., Choi Y. H., *Materials*, **2022**, *15*(4), 1428
[43] Bonacini L., Aoyote Pedraza I. L., Senni A. P., Tronco M. L., *IEEE Lat. Am. Trans.*, **2022**, *20*(12), 2471—2477
[44] Maulud D., Abdulazeez A. M., *J. Appl. Sci. Technol. Trends*, **2020**, *1*(2), 140—147
[45] Selvaratnam B., Oliynyk A. O., Mar A., *Inorg. Chem.*, **2023**, *62*(28), 10865—10875
[46] Salman H. A., Kalakech A., Steiti A. Babylon., *J. Mach. Learn.*, **2024**, *2024*, 69—79
[47] Zheng B., Gu G. X., *ACS Appl. Nano Mater.*, **2021**, *4*(3), 3167—3174
[48] Liang H., Jiang K., Yan T. A., Chen G. H., *ACS Omega*, **2021**, *6*(13), 9066—9076
[49] Valkenborg D., Rousseau A. J., Geubbelmans M., Burzykowski T., *Am. J. Orthod. Dentofacial Orthop.*, **2023**, *164*(5), 754—757
[50] Liu X., Tian S., Tao F., Yu W., *Compos. Part B Eng.*, **2021**, *224*, 109152
[51] Mortazavi B., Zhuang X., Rabczuk T., Shapeev A. V., *Mater. Horiz.*, **2023**, *10*(6), 1956—1968
[52] Anstine D. M., Isayev O., *J. Phys. Chem. A*, **2023**, *127*(11), 2417—2431
[53] Bhagat M., Bakariya B., *IJSAT*, **2025**, *16*(1), 1—4
[54] Oviedo F., Ferres J. L., Buonassisi T., Butler K. T., *Acc. Mater. Res.*, **2022**, *3*(6), 597—607
[55] Zhong X., Gallagher B., Liu S., Kailkhura B., Hiszpanski A., Han T. Y. J., *npj Comput. Mater.*, **2022**, *8*(1), 204
[56] Ayyasamy M. V., Deijkers J. A., Wadley H. .N., Balachandran P. V., *J. Am. Ceram. Soc.*, **2020**, *103*(8), 4489—4497
[57] Gracheva E., Lambard G., Samitsu S., Sodeyama K., Nakata A., *Sci. Technol. Adv. Mater. Methods*, **2021**, *1*(1), 213—224
[58] Zhang Y., Mu H., Cai Y., Wang X., Zhou K., Tian F., Fu Y., Zhang L., *Chin. Phys. B*, **2023**, *32*(5), 056302
[59] Tang Y., Zhu G., Tan Q., Kong D., Cao Y., Wang R., *Mater. Lett.*, **2024**, *372*, 137047
[60] Nilsen V., Pham L. T., Hibbard M., Klager A., Cramer S. M., Morgan D., *Constr. Build. Mater.*, **2019**, *220*, 587—595
[61] Tripathi B. M., Sinha A., Mahata T., *Mater. Today Proc.*, **2022**, *67*, 326—329
[62] Yue Y., Mohamed S. A., Jiang J., *J. Chem. Inf. Model.*, **2024**, *64*(13), 4966—4979
[63] Tian S., Zhou K., Yin W., Liu Y., *Nat. Commun.*, **2024**, *15*(1), 6977
[64] McGuinness K. P., Oliynyk A. O., Lee S., Molero-Sanchez B., Addo P. K., *Phys. Chem. Chem. Phys.*, **2023**, *25*(46), 32123—32131
[65] Yang H., Gao C., Jiang D., Zhong D., Ma Y., Li Y., Xing L., Zhao H., Yang L., Li Z., *J. Adv. Ceram.*, **2025**, *14*(8), 9221117
[66] Kim M., Kim J., Kim H., Kim J., *Int. J. Refract. Met. Hard Mater.*, **2024**, *122*, 106738
[67] Wang K., Lei M., Zhou X. G., Chen D., Zhu Y. A., *Ind. Eng. Chem. Res.*, **2025**, *64*(21), 10508—10521
[68] Ma Q., Li X., Xin R., Liu E., Gao Q., Sun L., Zhang X., Zhang C., *J. Mater. Res. Technol.*, **2023**, *26*, 4168—4178
[69] Zhu Y., Peng W., Du D., Dong A., Sun B., Jiang P., *J. Mater. Sci.*, **2025**, *60*, 9743—9757
[70] Cai Y., Wang C., Yuan H., Guo Y., Cho J. H., Xing X., Jia Y., *Mater. Horiz.*, **2024**, *11*(12), 2914—2925
[71] Kireeva N., Tsivadze A. Y., *J. Alloys Compd.*, **2024**, *990*, 174356
[72] Zeng Y., Cao H., Ouyang Q., Qian Q., *Mater. Today Commun.*, **2021**, *27*, 102314
[73] Peng J., Gunda N. H., Bridges C. A., Lee S., Haynes J. A., Shin D., *Comput. Mater. Sci.*, **2022**, *210*, 111034
[74] Xu P., Ma Y., Lu W., Li M., Zhao W., Dai Z., *J. Mater. Inf.*, **2025**, *5*(2), 26
[75] Mortazavi B., Rajabpour A., Zhuang X., Rabczuk T., Shapeev A. V., *Carbon*, **2022**, *186*, 501—508
[76] Yang G., Hu Y., Qiu Z., Li B. L., Zhou P., Li D., Zhang G., *Appl. Phys. Lett.*, **2023**, *122*(8), 082202
[77] Mortazavi B., Shapeev A. V., *FlatChem*, **2022**, *32*, 100347
[78] Mortazavi B., Zhuang X., *FlatChem*, **2022**, *36*, 100446
[79] Rajabpour A., Mortazavi B., *Condensed Matter*, **2022**, *7*(4), 67
[80] Gao Z., Zhou X., Xu M., Xu Y., Tian F., *Acta Mater.*, **2025**, *300*, 121493

(Ed.: H, K, M)